ADVISORY COMMITTEE BRIEFING MATERIALS: AVAILABLE FOR PUBLIC RELEASE

Bone, Reproductive, and Urologic Drugs Advisory Committee Meeting October 29, 2019

MAKENA®
(hydroxyprogesterone caproate injection)

NDA 021945 / S-023

AMAG Pharmaceuticals, Inc.
1100 Winter Street
Waltham, MA 02451

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviation	Definition
17P	Hydroxyprogesterone caproate injection, 250 mg/mL
ACOG	American College of Obstetricians and Gynecologists
AE	Adverse event
API	Active pharmaceutical ingredient
ASQ	Ages and Stages Questionnaire
BMI	Body mass index
BPD	Bronchopulmonary dysplasia
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
FDA	Food and Drug Administration
GA	Gestational age
GMP	Good Manufacturing Practices
HPC	Hydroxyprogesterone caproate
IM	Intramuscular
ITT	Intent-to-Treat
IVH	Intraventricular hemorrhage
LMP	Last menstrual period
MFMU	Maternal Fetal Medicine Unit
MPC	Maternal pregnancy complication
NDA	New Drug Application
NEC	Necrotizing enterocolitis
NICHD	National Institute of Child Health and Human Development
NICU	Neonatal intensive care unit
PK	Pharmacokinetic
PP	Per Protocol
pPROM	Preterm premature rupture of membranes
PROLONG	Progestin's Role in Optimizing Neonatal Gestation (PROLONG)
PTB	Preterm birth
PT	Preferred term
RDS	Respiratory Distress Syndrome
RRR	Relative risk reduction
SAE	Serious adverse event
SD	Standard deviation
SMFM	Society for Maternal-Fetal Medicine
SPTB	Spontaneous preterm birth
TEAE	Treatment emergent adverse event
US	United States

1. EXECUTIVE SUMMARY

1.1. Overview

Preterm birth (PTB) is a major public health concern in the United States (US). 17P (a synthetic progestin containing the active pharmaceutical ingredient 17α-hydroxyprogesterone caproate), which includes Makena and the recently approved generic formulations, is FDA-approved therapy to reduce recurrent PTB.

The purpose of this Advisory Committee meeting is to discuss the findings from the post-approval confirmatory trial for Makena, which failed to meet its co-primary endpoint. The discussion will focus on better understanding two studies with similar study designs, yet conflicting results.

Study 002 (hereafter referred to as the Meis Study) was the basis for FDA conditional approval of 17P in 2011, and demonstrated consistent and statistically significant efficacy across multiple endpoints. This landmark study was conducted by the National Institute of Child Health and Human Development, Maternal-Fetal Medicine Unit, and enrolled patients entirely in the US.

As part of the conditional approval of Makena, a confirmatory study (Study 003, or "PROLONG") was required. The PROLONG study, conducted predominantly outside the US, as previously mentioned, did not meet its co-primary efficacy objective. However a favorable maternal and fetal safety profile of 17P was reaffirmed, as there were no new or unexpected safety findings, and no clinically meaningful differences in the safety profile across treatment groups.

Key differences in baseline levels of risk for recurrent PTB between the PROLONG and Meis studies limit the applicability of the PROLONG efficacy data to the US population. Nevertheless, the strong efficacy data from the Meis study, previous supporting clinical trial data in the US, and trends favoring treatment benefit for 17P in post-hoc analyses focused on patients enrolled in the US, coupled with a favorable safety profile, support the continued use of 17P.

1.2. Preterm Birth Prevalence and Prevention

PTB, defined as birth before the 37th week of gestation, is a serious health concern, and is recognized as the leading cause of neonatal mortality and morbidity in the US [ACOG 2012]. One of the most significant risk factors for spontaneous singleton PTB is a patient's history of PTB. Women who have had a prior PTB have a 2.5-fold greater risk for subsequent PTB than women without a prior history of PTB [Iams et al 1998; Mercer et al 1999]. Approximately 3.3% of pregnant women, or 130,000 annually, have a history of prior singleton spontaneous PTB.

Infants born prematurely have increased risks of mortality and morbidity throughout childhood, especially during the first year of life. Premature birth is the number one cause of death of children under 5 years old worldwide. Infants who do survive premature birth often suffer long-term health problems and potential for long-term physical and cognitive disabilities.

According to the Centers for Disease Control and Prevention, ~10% of liveborn births, or nearly 400,000, each year are born prematurely. Rates of PTB are highest in the areas of the country with the greatest disparities in health care, particularly in minorities and poor communities.

Approximately 30% of women who deliver preterm had a history of a prior singleton spontaneous PTB [Gallagher et al 2018].In addition to prior PTB, there are additional known risk factors. Studies have reported that Black women are twice as likely as White women to have preterm deliveries and three times as likely to have very preterm deliveries (<32 weeks), which are the most vulnerable to mortality and long-term morbidities [Carmichael et al 2014; McKinnon et al 2016]. While the rate of PTB in the US is lower than the estimated global rate, the US ranked among the top ten countries in total number of PTBs, and remains among the highest in developed countries. In 2010, the World Health Organization ranked the US as 131st out of 184 countries in regard to rates of PTB.

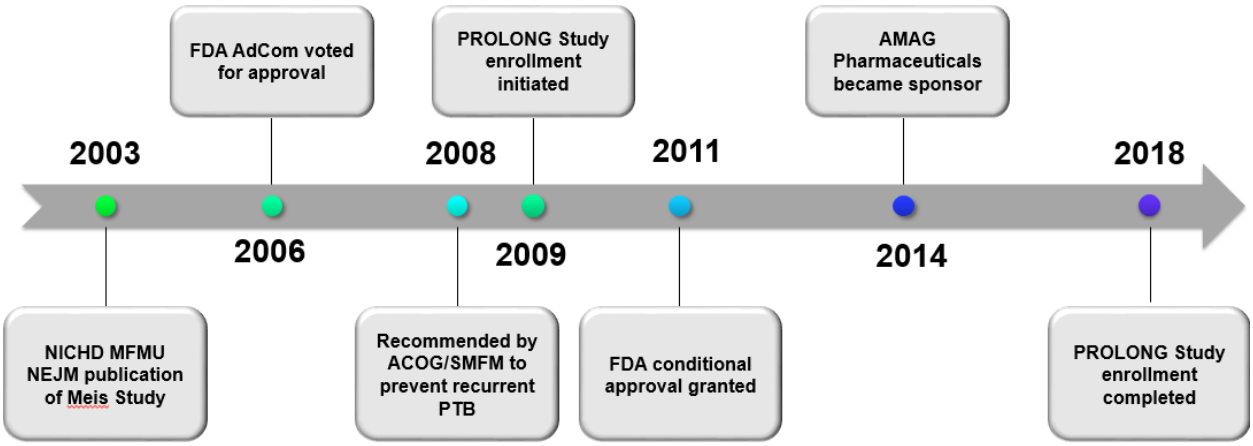
Progesterone agents have demonstrated effectiveness in the prevention PTB in randomized trials [Keirse 1990; Meis and Aleman 2004] which are thought to support gestation by reducing inflammation and inhibiting uterine activity. Hydroxyprogesterone caproate (HPC), or "17P", has demonstrated efficacy in randomized clinical trials to prevent pre-term birth in women with a prior spontaneous singleton pregnancy. In addition, a number of controlled studies support the use of 17P for this same patient population [Levine 1964; Papiernik-Berkhauser 1970; Johnson et al 1975; Yemini et al 1985; Suvonnakote 1986, Meis et al 2003, Saghafi et al 2011]. Vaginal progesterone has also been studied for the reduction of PTB in women with a history of spontaneous PTB, however, vaginal progesterone is not FDA-approved to prevent PTB in women with a prior spontaneous PTB or an incidental short cervix.

Progestogens, including 17P, have been recommended for use in treatment guidelines issued by professional societies. In 2008, the American College of Obstetricians and Gynecologists (ACOG) Committee on Obstetric Practice and Society for Maternal-Fetal Medicine (SMFM) issued a joint opinion that progesterone be used to prevent recurrent preterm birth [ACOG 2008]. In 2012, ACOG and the Society for Maternal-Fetal Medicine (SMFM) issued separate guidelines regarding the management of women at risk for PTB. In the SMFM guideline, an algorithm recommends the use of vaginal progesterone for women with an incidental short cervix and the use of 17P for women with histories of spontaneous PTB. The ACOG guideline was more general and stated only that "progesterone supplementation should be offered" to women with histories of spontaneous PTB [Practice Bulletin 2012].

1.3. Makena

A summary of the regulatory history for Makena is depicted in Figure 1 .

Figure 1: Regulatory Timeline



Abbreviations: NEJM=New England Journal of Medicine

1.3.1. Approval

Makena® was approved by FDA under the accelerated approval provisions of Subpart H of 21 CFR Part 314 in February 2011 (New Drug Application [NDA] 21945). Under Subpart H, FDA may grant approval based on an effect on a surrogate endpoint that is reasonably likely to predict a drug's clinical benefit.

"Makena is a progestin indicated to reduce the risk of PTB in women with a singleton pregnancy who have a history of singleton spontaneous PTB. The effectiveness of Makena is based on improvement in the proportion of women who delivered <37 weeks of gestation. There are no controlled trials demonstrating a direct clinical benefit, such as improvement in neonatal mortality and morbidity."

The Meis study was the pivotal study that served as the basis for approval. As part of the accelerated approval (granted based large unmet need for condition with no other treatment option), FDA required a confirmatory efficacy study be performed in order to demonstrate neonatal benefit as a primary outcome. During the review process, FDA recognized the difficulty of conducting a study once the drug was approved and adopted based on the recommendations of clinical guidelines supporting its use in this patient population. As a result FDA required that at least 5% of the patients be enrolled prior to approval of Makena, and that at least 10% of the patients be enrolled from North America. As such, the confirmatory study began in 2009, and once the North America enrollment requirement was met in 2011, Makena received FDA approval.

The confirmatory trial (PROLONG) was designed in conjunction with the FDA. FDA required that clinical efficacy be confirmed using the co- primary endpoints of PTB rates at less than 35 weeks and and rates of incident cases of neonatal morbidity/mortality with predefined criteria. FDA also wanted additional safety data to better understand the incidence of early fetal loss.

1.3.2. Availability of 17P

Prior to the approval of Makena, 17P was available to patients only through pharmacy compounding. Unlike pharmaceutical manufacturers, compounding pharmacies do not have to demonstrate the safety and efficacy of compounded products or adhere to FDA Good Manufacturing Practices (GMPs). In 2011, the original sponsor of Makena (KV Pharmaceuticals) obtained samples of compounded 17P and the active pharmaceutical ingredient used by pharmacists to compound 17P, and identified that compounded versions of 17P did not meet the purity and potency specifications designated for Makena [Chollet and Jozwiakowski 2012].

In addition to lack of comparability, there are significant potential risks associated with pharmacy compounding products. A stark reminder of these potential safety concerns that can arise from the lack of regulation around purity, potency and sterility of drug products, occurred in the Fall of 2012 when a fungal meningitis outbreak was traced to contaminated compounded drugs formulated and distributed by the New England Compounding Center (NECC). There were 76 deaths were attributed to these substandard sterile injectable drugs produced by the NECC, with over 700 patients being gravely sickened [FDA 2017; Raymond 2017].

The key issue is the lack of standard quality oversight of compounded products from a GMP perspective. Whenever this process is lacking or deficient, there is the potential for untoward effects and unnecessary harm to patients. Without FDA-approved forms of 17P (Makena, plus the four generic products available), pharmacy compounding may be the only available source of this injectable drug for pregnant women.

1.4. Overview of Clinical Studies

An overview of the key adequate and well-controlled safety and efficacy studies comprising the Makena clinical development program is provided in Table 1.

Table 1: Overview of Key Clinical Studies

	Meis	**PROLONG**
Year	1999 to 2002	2009 to 2018
Sites	19 sites, US Only	93 sites, 9 countries
Randomization	2:1	2:1
Study Drug	17P 250 mg/mL or vehicle	17P 250 mg/mL or vehicle
Dose	1 dose/week through 36^6 weeks gestation or delivery	1 dose/week through 36^6 weeks gestation or delivery
Study Population	Women 16 to 20 weeks gestation with history of spontaneous preterm delivery	Women 16 to 20 weeks gestation with history of spontaneous preterm delivery
Sample Size	17P: N=310 Vehicle: N=153	17P: N=1130 Vehicle: N=578
Primary Endpoint(s)	• PTB <37 weeks	• PTB <35 weeks • Neonatal Composite Index
Key Secondary Endpoints	• PTB <35 and <32 weeks • Neonatal morbidity/mortality	• PTB <37 and <32 weeks • Fetal/early infant death

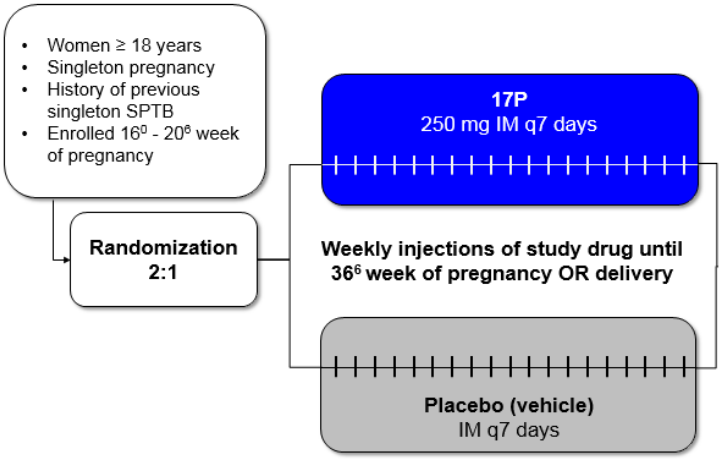
1.5. Meis: Pivotal Trial Results

The Meis study was conducted from 1999 to 2002 by the National Institutes of Child Health and Human Development (NICHD) through the Maternal Fetal Medicine Units Network (MFMU). The study was a US-only, double-blind, randomized placebo-controlled trial in pregnant women with a documented history of spontaneous preterm delivery. Women were enrolled at 19 clinical centers in the US, primarily located in inner city academic institutions with a high proportion of minorities.

A dose of 250 mg IM was selected based on earlier clinical trials designed to determine if 17P could prevent premature delivery [LeVine 1964; Johnson et al 1975; Yemini et al 1985].

The design of Meis is provided in Figure 2**.**

Figure 2: Meis Study Schematic



In 2002, the prespecified stopping criterion (p=0.015) for efficacy was met at the second interim analysis and the Data Monitoring Committee recommended stopping the trial prior to enrolling the proposed 500 patients. Stopping criteria were in place to assure that once efficacy was established the drug could be made available to all appropriate patients.

1.5.1. Efficacy

Patients randomized to the two treatment groups were comparable in mean age, race, body mass index (BMI) prior to pregnancy, marital status, years of education, and substance use during pregnancy. The majority of patients were Black (approximately 59%), with a mean age of 26.2 years. The mean pre-pregnancy BMI was approximately 26.6 kg/m^2. Approximately 50% of patients in the study were married, and approximately 22% smoked, approximately 8% consumed alcohol, and 3% used illicit drugs during the study pregnancy. Compared to the vehicle group, the 17P patients had significantly fewer previous preterm deliveries, fewer previous spontaneous preterm deliveries, and a lower percentage of patients with >1 previous preterm delivery.

1.5.1.1. Primary Efficacy Endpoint Analysis: Recurrent Preterm Birth

The risk of delivering prior to 37^0 weeks gestation in the Meis study was significantly reduced in the 17P group (37.1% vs 54.9%; p=0.0003) (Table 2).

Table 2: Percentage of Patients with Delivery <37^0 Weeks of Gestation (Meis)

Data Source	17P n (%)	Vehicle n (%)	Nominal p-value[a]	Treatment difference [95% CI[b]]
ITT Population	115 (37.1)	84 (54.9)	0.0003	-17.8% [-28%, -7%]
Only available data	111 (36.3)	84 (54.9)	0.0000	-18.6% [-29%, -8%]

Source: FDA Background Gestiva (August 2, 2006), Table 4.
Note: ITT population was all randomized patients (17P N=310; Vehicle N=153). The 4 patients with missing outcome data were classified as having a preterm birth of <37^0 weeks (i.e., treatment failure). "Only available data" does not include the 4 patients in the 17P group with missing outcome data.
[a] Chi-square test. Adjusting for interim analyses, p-values should be compared to 0.035 rather than the usual 0.05.
[b] CI adjusted for the 2 interim analyses and the final analysis. To preserve the overall Type I error rate of 0.05, a p-value boundary of 0.035 was used for the adjustment (equivalent to a 96.5% confidence interval).

1.5.1.2. Secondary Endpoint Analyses

1.5.1.2.1. Preterm Birth <35 and <32 Weeks Gestational Age

Despite the fact that the study was not powered to determine statistically significant differences in births at <35^0 and <32^0 weeks gestation, 17P demonstrated clinically important reductions in the number of births before 35^0 weeks (p=0.032) and before 32^0 weeks gestation (p=0.046) (Table 3).

Table 3: Percentage of Patients with Delivery <35^0 and <32^0 Weeks of Gestation (Meis)

Pregnancy Outcome	17P (N=310) n (%)	Vehicle (N=153) n (%)	Nominal p-value[a]
Delivery <35^0	67 (21.6)	47 (30.7)	0.032
Delivery <32^0	39 (12.6)	30 (19.6)	0.046

Source: FDA Background Gestiva (August 2, 2006), Table 6.
Data presented are from the ITT population (i.e., all randomized patients). The 4 patients with missing outcome data were classified as having a preterm birth <37^0 weeks (i.e., treatment failure).
[a] Adjusting for interim analyses, p-values should be compared to 0.035 rather than the usual 0.05.

At the <37^0, <35^0, and <32^0 weeks gestation, the percentage of deliveries was numerically lower in the 17P treatment arm (Table 4). There was no difference between treatment groups for the percentages of deliveries <28^0 weeks.

Table 4: Percentage of Patients with Delivery <37^0, 35^0, 32^0, and 28^0 Weeks of Gestation (Intent-to-Treat Population - Meis)

Time of Delivery (Gestational Age)	17P N=310 %	Vehicle N=153 %	Treatment difference[a] [95% CI[b]]
<37^0 weeks	37.1	54.9	-17.8% [-28%, -7%]
<35^0 weeks	21.6	30.7	-9.1% [-18%, 0.3%]
<32^0 weeks	12.6	19.6	-7.05 [-14%, 0.8%]
<28^0 weeks	10.0	10.5	-0.5% [-6.9, 5.9]

Source: FDA Background Gestiva (August 2, 2006), Table 7.
[a] Chi-square test.
[b] CI based on a t-test are adjusted for the 2 interim analyses and the final analysis. To preserve the overall Type I error rate of 0.05, a p-value boundary of 0.035 was used for the adjustment (equivalent to a 96.5% confidence interval).

1.5.1.2.2. Neonatal Morbidity and Mortality

A prespecified key secondary endpoint was the incidence rate of having a qualifying event in the composite neonatal morbidity index. The neonatal composite index included neonates with death, respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), grade 3 or 4 IVH, proven sepsis, or NEC) was lower in the 17P group, but the between group difference was not statistically significant (11.9% vs 17.2%; p=0.119)

The study was not powered to detect statistically significant differences between 17P and vehicle treatments in neonatal mortality or morbidities, however, reductions were observed with 17P in the rates of NEC, any grade of IVH, and the need for supplemental oxygen.

Although the overall rate of neonatal deaths was lower in the 17P arm versus vehicle, it was observed that miscarriages (defined as spontaneous loss of fetus from 16^0 to 19^6 weeks gestation) were numerically higher in the 17P arm, as were stillbirths (defined as birth of an infant ≥20 weeks gestation who died prior to delivery) (Table 5). In the vehicle group, the incidence of neonatal death was twice the rate of the 17P group, however the between group difference was not statistically significant due to the small sample size (p=0.116). Two other NICHD MFMU studies were subsequently conducted; when miscarriage and stillbirth are reviewed in the totality of these studies, the rates were similar between 17P and vehicle [Rouse et al 2007, Caritis et al 2009].

Table 5: Miscarriages, Stillbirths, and Neonatal Deaths (Meis)

Pregnancy Outcome	17P (N=306) n (%)	Vehicle (N=153) n (%)	Nominal p-value[a]
Total Deaths	**19 (6.2)**	**11 (7.2)**	**0.689**
Miscarriages <20 weeks gestation	5 (1.6)	0	0.175
Stillbirth	6 (2.0)	2 (1.3)	0.725
Antepartum stillbirth	5 (1.6)	1 (0.6)	---
Intrapartum stillbirth	1 (0.3)	1 (0.6)	---
Neonatal deaths	8 (2.6)	9 (5.9)	0.116

Source: FDA Background Gestiva (August 2, 2006), Table 8.
[a] No adjustment for multiple comparisons.

1.5.2. Safety

The most common type of adverse event (AE) reported during the Meis study was injection site reactions, which was expected considering that patients received weekly 1 mL IM injections. Pain, swelling, itching, and nodule formation were among the most common reactions regardless whether the solution being injected was 17P or vehicle. However, there was a significantly higher incidence of swelling at the injection site in the 17P group than vehicle (17.1% vs. 7.8%; p=0.007). Nevertheless, few women (1.7%) discontinued the study due to injection site reactions.

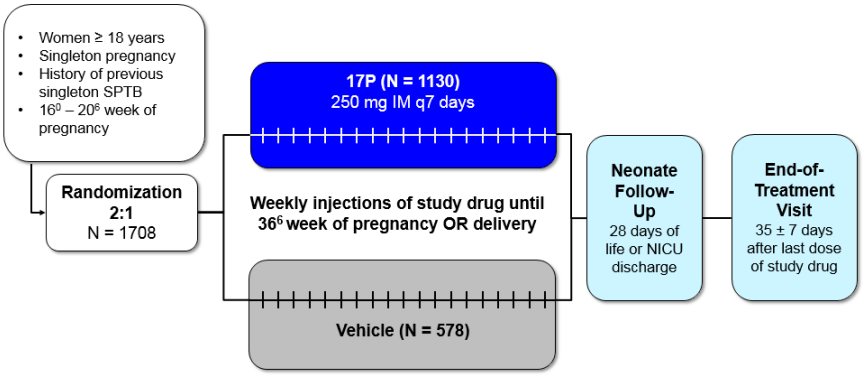
The incidence of pregnancy complications, such as preeclampsia, gestational diabetes, or clinical chorioamnionitis, as well as the incidence of serious adverse events (SAEs), was not different between the 17P and vehicle groups. SAEs reported were predominately miscarriages, stillbirths, and neonatal deaths, which were not unexpected events in the high-risk patient population, and were considered by the Investigator to be unrelated to study drug.

1.6. PROLONG: Trial Results

PROLONG was an international, double-blind, randomized, placebo-controlled trial in pregnant women with a documented history of spontaneous preterm delivery conducted from 2009 through 2018. PROLONG was approximately four times the size of the Meis trial, and was powered to detect a 30% and 35% difference between treatments in the co-primary endpoints, PTB <35 weeks gestation and neonatal composite index, respectively.

The design of PROLONG is provided in Figure 3.

Figure 3: Study Schematic (PROLONG)



PROLONG began in 2009, and once the North America enrollment requirement was met in 2011, Makena received FDA approval. Following approval of Makena, recruitment and enrollment in the US became increasingly difficult. Additional sites were then opened in Ukraine and Russia, as these countries had previously been the top enrollers in Europe.

Women were enrolled at 93 clinical centers in 9 countries. Russia and Ukraine accounted for 61% of study patients, and the US had 23%. The remaining 16% of patients were enrolled in Hungary, Spain, Bulgaria, Canada, Czech Republic, and Italy, each enrolling less than 100 patients. Enrollment in PROLONG was completed in 2018.

1.6.1. Efficacy

A total of 1708 patients were randomized 2:1 (1130 to 17P and 578 to Vehicle) and were included in the Intent-to-Treat (ITT) Population.

Although the study entry criteria were similar between PROLONG and Meis, there were differences in the patient populations that were enrolled. When comparing demographics and baseline characteristics of patients enrolled in the two studies, the differences across race and other potential surrogates of socioeconomic status were noteworthy, with Meis representing a much higher-risk population. In comparison to Meis, PROLONG patients had lower risk for spontaneous PTB based on the following key features:

- The majority of patients were White (approximately 89%), non-Hispanic or Latino (approximately 91%) with a mean age of 30 years.
- Approximately 90% of patients were married at the time of study entry.
- Substance use during pregnancy was low in PROLONG (~8% smoked, ~3% consumed alcohol, and 1.4% used illicit drugs).
- Approximately 15% of patients in PROLONG reported >1 previous spontaneous preterm delivery (compared to ~35% in Meis).

1.6.1.1. Primary Endpoint Analysis

The study did not meet its co-primary efficacy objectives, which were to demonstrate a reduction in PTB prior to 35^0 weeks gestation and in the neonatal composite index.

Rate of PTB

The overall rate of PTBs prior to 35^0 weeks gestation was lower than anticipated based on the event rates observed in Meis. Rates of PTB $<35^0$ weeks were low in both groups and not statistically different between groups (11.0% for 17P and 11.5% for vehicle; Table 6)

Neonatal Composite Index

No statistically significant differences in the rates of neonatal mortality or morbidity as measured by the neonatal composite index, were noted (5.4% for 17P and 5.2% for vehicle; Table 6).

The incidence of individual components of the neonatal composite were similar between treatment groups (Table 7). RDS accounted for almost all of the infants who met the criteria for this index, and rates across treatment groups were not statistically significantly different, at 4.9% and 4.6% in neonates born to patients in the 17P treatment group and vehicle group, respectively.

Table 6: Primary Efficacy Outcomes (PROLONG)

Primary Efficacy Outcomes	**17P (N=1130)**	**Vehicle (N=578)**
PTB $<35^0$ Weeks Gestation (ITT Population)		
Overall Outcome rate n/N* (%)	**122/1113 (11.0)**	**66/574 (11.5)**
p-value[a]	0.716	
Relative risk (95% CI)	0.95 (0.71, 1.26)	
Neonatal Composite Index (Liveborn Neonatal Population)	**(N=1091)**	**(N=560)**
Neonatal Composite Index – Overall, n (%)[d]	**59 (5.4)**	**29 (5.2)**
p-value[b]	0.840	
Relative risk (95% CI)	1.05 (0.68, 1.61)	

Source: PROLONG CSR Table 14.2.1.1.1 and Table 14.2.1.1.2, PROLONG Ad Hoc Table 14.2.1.1.1.26.
[a] p-value from the Cochran-Mantel-Haenszel test.
[b] p-value from the Cochran-Mantel-Haenszel test.
N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 35^0 weeks in the specified category.
The composite index was defined as a liveborn neonate with any of the following occurring at any time during the birth hospitalization up through discharge from the NICU: neonatal death, Grade 3 or 4 IVH, RDS, BPD, NEC, or proven sepsis.

Table 7: Components of Neonatal Composite Index from NICU Outcomes (Liveborn Neonatal Population - PROLONG)

Individual Components of Neonatal Composite Index	17P (N=1091) n (%)	Vehicle (N=560) n (%)
Neonatal Composite Index – Overall	**59 (5.4)**	**29 (5.2)**
Neonatal death prior to discharge	3 (0.3)	2 (0.4)
Grade 3/4 intraventricular hemorrhage	2 (0.2)	1 (0.2)
Respiratory distress syndrome	54 (4.9)	26 (4.6)
Bronchopulmonary dysplasia	6 (0.5)	1 (0.2)
Necrotizing enterocolitis	2 (0.2)	2 (0.4)
Proven sepsis	5 (0.5)	3 (0.5)

Source: PROLONG CSR Table 15.

1.6.1.1.1. Subgroup Analysis

Subgroup analyses of the primary endpoints were conducted by geographic region and obstetric history.

Geographic Region

The event rates for PTB and the neonatal composite index were 1.5 to 2 times higher at 16 to 18% in the US relative to ex-US regions (10%). The rates of PTB among US patients were the highest of the three top enrolling countries in the study (Russia, Ukraine and US), while the rates in Russia and Ukraine were the lowest. The rates of the neonatal composite index in the regions with the highest enrollments (Russia and Ukraine) were among the lowest observed. This is consistent with the known epidemiology, as well as the substantially different health care delivery systems in these countries, where early intervention to improve prenatal care and reduce neonatal complications is emphasized and universally available [Healthy Newborn Network 2015; Russian Federation: Federal State Statistics Service 2012; UNICEF 2017; USAID 2011].

Obstetric History

Rates of PTB $<35^0$ weeks gestation and neonatal composite index were also examined for differences in obstetrical history including gestational age of qualifying delivery, gestational age of earliest prior PTB, and number of previous preterm deliveries. Results were similar for both treatment groups across subgroups.

1.6.1.2. Key Secondary Endpoint Analyses

1.6.1.2.1. Preterm Birth <37 and <32 Weeks Gestational Age

There were no statistically significant differences in births at $<37^0$ (p=0.567) or $<32^0$ weeks gestation (p=0.698) (Table 8).

Table 8: Percentage of Patients with Delivery <37^0 and <32^0 Weeks of Gestation (Intent-to-Treat Population, PROLONG)

	17P (N=1130) n/N* (%)	**Vehicle (N=578) n/N* (%)**
<32^0 Weeks Gestation	**54/1116 (4.8)**	**30/574 (5.2)**
p-value[a]	0.698	
Relative risk (95% CI)	0.92 (0.60, 1.42)	
<37^0 Weeks Gestation	**257/1112 (23.1)**	**125/572 (21.9)**
p-value[a]	0.567	
Relative risk (95% CI)	1.06 (0.88, 1.28)	

Source: PROLONG Table 14.2.3.2.1 and Table 14.2.3.1.1, PROLONG Ad Hoc Table 14.2.1.1.1.26.
[a] p-value Cochran-Mantel-Haenszel test.
Notes: n=number of patients with delivery <32^0 or 37^0 weeks (as indicated) gestation.
N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 32^0 or 37^0 weeks (as indicated) in the specified category.

1.6.2. Safety

1.6.2.1. Fetal and Early Infant Death (Primary Safety Outcome)

The primary safety objective of PROLONG was to rule out a doubling in the risk of fetal or early infant death in the 17P group compared to vehicle. This objective was included specifically to address the Agency's concern of a potential "safety signal" relative to the numerically higher rate of both miscarriage and stillbirth from the Meis study.

Fetal/early infant death was defined as a spontaneous abortion or miscarriage occurring at 16 weeks 0 days through 19 weeks 6 days; a stillbirth, either antepartum or intrapartum; or a neonatal death, occurring minutes after birth until 28 days of life.

If the upper bound of the CI is less than or equal to 2.0, a doubling in risk of fetal/early infant death can be ruled out. A doubling of risk was selected and agreed upon with FDA based on sample size calculations.

Rates were low and similar between treatment groups (1.68% and 1.90% in the 17P and vehicle groups, respectively) with a relative risk of 0.79 (95% CI 0.37–1.67) (Table 9). Given that the upper bound of the 95% CI is less than 2.0, a doubling in the risk of fetal/early infant death was adequately and firmly excluded.

Table 9: Fetal and Early Infant Death (Intent-to-Treat Population, PROLONG)

Primary Safety Outcome	17P (N=1130) n (%)	Vehicle (N=578) n (%)
Fetal/Early Infant Death	**19 (1.68)**	**11 (1.90)**
Relative Risk (95% CI) [a]	**0.79 (0.37 - 1.67)**	

Source: 17P-ES-003 CSR, Table 14.3.1.1.1.
[a] Relative risk of fetal/early infant death is from the Cochran-Mantel-Haenszel test.
Notes: N=number of patients in the ITT Population in the specified treatment group.
n=number of patients with Fetal/Early Infant Death in the specific category. Fetal/Early Infant Death is defined as neonatal death occurring in liveborns born at less than 24 weeks of gestation, spontaneous abortion/miscarriage or stillbirth

1.6.2.2. Treatment-emergent Adverse Events

The AE profile between the two treatment groups was comparable. There were 57.3% and 57.8% of patients with at least one treatment-emergent AEs (TEAEs) in the 17P and vehicle group, respectively. The majority of TEAEs were mild in intensity, and most were considered unrelated to study drug. There was a low percentage of TEAEs leading to study drug withdrawal (1.0% and 0.9%) in the 17P and vehicle group, respectively, with both groups experiencing similar and low rates of serious adverse events (SAEs; 3.0% and 3.1% in the 17P and vehicle group, respectively).

The most frequently reported TEAEs in either treatment group were anemia (9.2% in 17P and 9.7% in vehicle) and headache (6.0% in 17P and 4.8% in vehicle). Other commonly reported TEAEs in the 17P group included nausea (4.9%) and back pain (4.4%).

1.6.2.3. Maternal Pregnancy Complications (MPC)

There were 27.7% and 28% of patients who experienced at least one MPC in the 17P and vehicle group respectively. The majority of patients who experienced MPC experienced mild events, and most were unrelated to study drug. The most frequently reported MPCs by PT for the 17P group were cervical incompetence (3.0%), gestational diabetes (2.9%), anemia of pregnancy (2.7%), and placental disorder and pre-eclampsia (2.6% each). The incidence of these MPC were similar in the vehicle group.

The number of patients diagnosed with gestational diabetes during PROLONG was low (~4% in both treatment groups), and consistent with the incidence each year in the US (2 to 10% of pregnancies) per Center for Disease Control estimates [CDC 2019].

1.6.2.4. Miscarriage and Stillbirth

Stillbirths were reported for 12 (1.1%) 17P patients and 3 (0.5%) vehicle patients (Table 37). All of the stillbirths were deemed unrelated to study drug by the Investigator. Among the 12 that occurred in the 17P group, 8 were listed as "definitely not related," 3 as "unlikely related", and 1 "not related." Two women in the 17P group who delivered stillbirths reported smoking during pregnancy, one tested positive for cannabinoids, 1 had a large subserous myoma, and another had uncontrolled Type 1 diabetes mellitus with documented nephropathy and retinopathy.

Ten women had a miscarriage: 4 (0.35%) in the 17P group and 6 (1.04%) in the vehicle group.

1.6.2.5. Serious Adverse Events

Overall, 34 (3.0%) 17P patients and 18 (3.1%) vehicle patients experienced serious TEAEs or MPCs. The most frequently reported serious TEAE or MPC for patients treated with 17P were premature separation of placenta (5 patients, 0.4%), placental insufficiency (4 patients, 0.4%), and pneumonia (3 patients, 0.3%); Escherichia coli sepsis, pyelonephritis, and wound infection were each reported by 2 patients in the 17P group. The most frequently reported serious TEAE or MPC for patients treated with vehicle were cholestasis (3 patients, 0.5%), and premature separation of placenta (2 patients, 0.3%).

Two patients each had one serious TEAE/MPC considered possibly related to study treatment (one patient in the 17P group had the TEAE of mild nephrolithiasis considered possibly related and one patient in the vehicle group had the severe MPC of cholestasis considered probably related).

1.6.2.6. Discontinuation due to Adverse Event

In total, 11 (1.0%) 17P patients and 5 (0.9%) vehicle patients experienced a TEAE and/or MPC that led to discontinuation of study medication (predominantly associated with the injection site). None of these events were deemed serious by the study investigator.

1.7. Exploratory Analyses

Unlike the Meis trial, which showed a treatment benefit, treatment with 17P in PROLONG did not decrease rates of PTB or the overall neonatal composite index in the overall study population.

To better understand these discrepant results, exploratory analyses were conducted. These post hoc analyses examined the potential role that differences between the study populations (demographics and patient characteristics associated with baseline risk levels), and differences in health care delivery systems and geography (access to universal health care, emphasis on preventative care) may have had on the results of the study.

1.7.1. Comparison of Demographics

When comparing demographics and baseline characteristics from PROLONG and Meis, the differences across race and other potential surrogates of socioeconomic status that have been linked to higher rates of PTB were noteworthy, with most of those differences driven by the ex-US PROLONG subset population (Table 10). Compared to the US PROLONG subset and Meis, the ex-US PROLONG population represented a cohort with a lower baseline risk for PTB.

- **Prior spontaneous PTB:** In ex-US PROLONG, 11% had more than 1 prior spontaneous PTB, compared to 27% in US PROLONG and 32% in Meis.
- **Race/ethnicity:** In ex-US PROLONG, only 1 patient was Black or African American, compared to 29% in US PROLONG and nearly 60% in Meis. Hispanic or Latinos accounted for approximately 8% of patients in ex-US PROLONG, 14% in US PROLONG, and 15% in Meis.
- **Marital status:** In ex-US PROLONG, 4% of patients were unmarried with no partner, compared to 31% in US PROLONG and 50% in Meis.

- **Substance use:** In ex-US PROLONG, approximately 4% of patients reported any substance use during pregnancy (smoking, alcohol or illicit drugs), compared to 28% in US PROLONG and 26% in Meis.

Table 10: Differences in Race and Socioeconomic Status (Meis and PROLONG)

Demographics/Baseline Characteristics – n (%)	Ex-US PROLONG (N=1317)	US PROLONG (N=391)	Meis (N=463)
>1 previous SPTB	141 (10.7)	107 (27.4)	149 (32.2)
Race/ethnicity			
Black/African American	1 (0.1)	113 (28.9)	273 (59.0)
Hispanic or Latino	101 (7.7)	54 (13.8)	69 (14.9)
Gestational age at randomization			
16-17 weeks	603 (45.8)	138 (35.3)	151 (32.6)
18-20^{6} weeks	714 (54.2)	253 (64.7)	312 (67.4)
Unmarried with no partner	53 (4.0)	120 (30.7)	233 (50.3)
Educational status (≤12 years)	549 (41.7)	197 (50.5)	330 (71.3)
Any substance use during pregnancy	47 (3.6)	111 (28.4)	121 (26.1)
Smoking	44 (3.3)	89 (22.8)	100 (21.6)
Alcohol	6 (0.5)	36 (9.2)	37 (8.0)
Illicit drugs	1 (0.1)	23 (5.9)	15 (3.2)

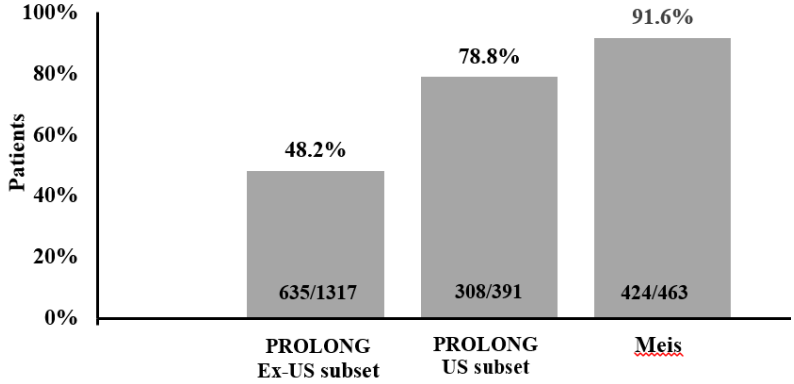
Source: PROLONG Ad Hoc Table 14.1.3.1.9

It is important to note that while US PROLONG patients were more similar to those in Meis, there remain differences related to baseline levels of risk for PTB.

Figure 4 displays a post hoc assessment of select composite risk factors associated with risk of PTB across Meis and PROLONG. The components selected for inclusion (beyond the required entry criteria for at least one prior spontaneous PTB) are >1 prior spontaneous PTB, any substance use, ≤12 years of education, unmarried with no partner, and Black or African American. Importantly, other than a prior history of more than 1 spontaneous PTB, the other components are merely imperfect surrogates of socioeconomic status, an important known predictor of rates of PTB.

The ex-US subset of PROLONG (a low risk population) had a much lower percentage of patients (48.2%) with more than one additional risk factor for PTB compared to the subset of US patients in PROLONG, an intermediate risk population (78.8%) and patients in Meis, a high risk population (91.6%).

Figure 4: Differences in Baseline Risk Factors (Known or Surrogate) Associated with Preterm Birth - Post Hoc (Meis and PROLONG)



Source: PROLONG Ad Hoc Table 14.1.3.1.9
Notes: The composite risk factors (in addition to the required prior spontaneous PTB) included >1 prior spontaneous PTB, substance use, educational status (≤12 years), unmarried with no partner, and Black/African American. Percentages expressed as n/N x 100, where n is the number of patients with at least 1 additional risk factor and N is the number of patients in the cohort.

1.7.2. Comparison of Efficacy Outcomes

Study populations with a greater percentage of high risk patients defined by the previously described composite of risk factors appeared to show improved treatment benefit with 17P compared to those with a lower percentage of those patients as shown in Figure 5.

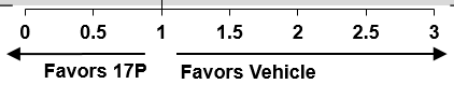
In Meis, which was a higher risk population, a treatment benefit favoring 17P was observed not only with the <37 weeks gestational age, but also at <35 weeks and even at <32 weeks, an important endpoint since it is known that babies born at earlier than 32 weeks have a significant risk of mortality and neonatal complications.

In addition, the intermediate risk population from the US subset of PROLONG also shows trends of a treatment effect favoring 17P beginning to emerge, as this population becomes more similar to Meis. These trends can be seen at <35 weeks and even at <32 weeks, however not at <37 weeks.

In contrast, the lower risk population of patients from the ex-US subset of PROLONG tend to show no trends of 17P treatment benefit compared to vehicle.

Figure 5: Comparison of Maternal Efficacy Endpoints – Post Hoc (Meis and PROLONG)

Study Population	Gestational Age (weeks)	17P (n/N)	Vehicle (n/N)	Relative Risk (95% CI)
Meis	32	39/310	30/153	0.63 (0.41, 0.97)
Meis	35	67/310	47/153	0.70 (0.51, 0.96)
Meis	37	115/310	84/153	0.68 (0.55, 0.83)
Prolong US	32	14/256	12/131	0.58 (0.27, 1.21)
Prolong US	35	40/256	23/131	0.88 (0.55, 1.40)
Prolong US	37	85/256	37/131	1.16 (0.84, 1.61)
Prolong Ex- Us	32	40/860	18/443	1.14 (0.66, 1.97)
Prolong Ex- Us	35	82/857	43/443	0.98 (0.69, 1.39)
Prolong Ex- Us	37	172/856	88/441	1.01 (0.80, 1.27)



Source: PROLONG Ad Hoc Table 14.2.1.1.1.26.

1.8. Discussion

PROLONG did not meet the predefined co-primary objectives. AMAG believes that the results from PROLONG were influenced by differences in the study population from that previously studied in Meis. While the entry criteria of Meis and PROLONG were similar, the study population in PROLONG was different than that of Meis, with the latter comprised of a higher risk population.

Efficacy

When comparing demographics and baseline characteristics from PROLONG and Meis, the differences across race and other potential surrogates of socioeconomic status that have been linked to higher rates of PTB were noteworthy, with most of those differences were driven by the ex-US PROLONG subset population. As a result, key differences in baseline risk associated with PTB even within the PROLONG study population, notably US vs. ex-US subset populations, make the applicability of the efficacy data particularly challenging in the US.

A review of the baseline characteristics of patients who enrolled in PROLONG in the US demonstrates that although they are more similar to Meis than that of the overall PROLONG population, they remain differ from Meis on many of the risk factors thought to be associated with risk of PTB.

A post-hoc investigation into baseline risk factors indicate that, compared to Meis (a high-risk population), the PROLONG US subset was an intermediate risk group for recurrent PTB, with the PROLONG ex-US subset at lower risk. The lower baseline risk for PTB in ex-US PROLONG could be attributed to varying healthcare delivery systems (more preventive than acute care) with universal access in ex-US countries, which represented 75% of the study population (61% from Russia and Ukraine alone). In a number of these countries, there are dedicated programs that target prevention of PTB and adverse fetal outcomes with evidence-based technologies to improve the quality of perinatal care. Often, these programs include comprehensive measures for pregnancy planning, screening, primary prophylaxis, and risk factor

reduction, as well as providing healthcare and treatment of co-morbid conditions prior to pregnancy. In addition, compliance with prenatal care is associated with state-provided financial incentives for new mothers [Healthy Newborn Network 2015; Russian Federation: Federal State Statistics Service 2012; UNICEF 2017; USAID 2011].

Of note, exploratory analyses of PTB rates by baseline risk suggest an increasing treatment benefit associated with 17P with increasing levels of baseline risk for recurrent PTB. Treatment effect was observed at <37, <35, and <32 weeks gestation for the highest risk group (Meis), while the lowest risk group (ex-US PROLONG) showed no effect. Trends favoring 17P emerge in the US PROLONG subset as the population becomes more similar to that of Meis, with increased effect at <35 and <32 weeks, but not at <37 weeks gestation.

In totality, it is possible that differences in baseline risk for PTB underpin the lack of correlation between the efficacy results observed in Meis and PROLONG.

Safety

The key safety outcome of PROLONG was to rule out a doubling of risk of fetal or early infant death in the 17P group relative to vehicle. This endpoint was included specifically to address the Agency's concern of a potential safety signal relative to the numerically higher rate of both miscarriage and stillbirth from the Meis study. The relative risk of 0.79 with an upper bound of the 95% CI of 1.67 excludes that risk.

The favorable maternal and fetal safety profile of 17P was reaffirmed as there were no new or unexpected safety findings, and no clinically meaningful differences in the safety profile across treatment groups. Specifically, there were no clinically meaningful differences in TEAEs across the two treatment groups (17P and vehicle).

Proposed Changes to Prescribing Information

Based on the results from PROLONG, AMAG is proposing to maintain the indication with the current limitations of use and to amend the current prescribing information to include the following updates:

- Section 6 Adverse Reactions: to include pooled (Meis and PROLONG) safety information
- Section 14.1 Clinical Trials to Evaluate Reduction of Risk of Preterm Birth: to include findings from PROLONG. In particular AMAG proposes that it is important to include information that helps place the results from PROLONG in context with those observed from Meis.

1.8.1. Conclusions

Differences in study populations between Meis and PROLONG as it relates to baseline levels of risk associated with PTB contributed to the vastly lower rates of PTB and associated prematurity complications seen in PROLONG. It is relevant to acknowledge that in the nearly 20 years since Meis was initiated and PROLONG was completed, there have been substantial improvements in neonatal care that have increased survival. However, rates of PTB in the US have remained relatively constant over that time period and there remains a significant public health concern regarding PTB. Moreover, women with a prior history of spontaneous PTB, particularly if the

preterm birth is early (<32 week gestation), or if there is a history of more than one prior spontaneous PTB, are at the highest risk for a recurrent PTB.

The totality of clinical data including more than 16 years of clinical use support 17P's positive benefit-risk profile and support its availability for clinicians to make patient-specific prescribing decisions, based upon their clinical judgment and shared decision-making with their patients.

2. PRETERM BIRTH

Summary

- Preterm birth (PTB), defined as birth before the 37th week of gestation, is a serious health concern and is recognized as the leading cause of neonatal mortality and morbidity in the United States (US) [ACOG 2012].
  - Women who have had a prior PTB have a 2.5-fold greater risk for subsequent PTB than women with no prior history of PTB [Iams et al 1998; Mercer et al 1999].
- Infants born prematurely have increased risks of mortality and morbidity throughout childhood, especially during the first year of life. Infants who do survive premature birth often suffer long-term health problems.
- Despite advances in perinatal care, the incidence of PTB remains high in the US, with rates among the highest among industrialized countries [March of Dimes 2015].
  - Approximately 10% of liveborn births each year, or nearly 400,000, are born prematurely
  - The PTB rate in the US worsened for a third consecutive year.
- Preterm birth rates vary significantly by race and geographic location.
  - Black women are twice as likely as White women to have preterm deliveries and three times as likely to have very preterm deliveries (<32 weeks) [Carmichael et al 2014; McKinnon et al 2016].
  - While the rate of PTB in the US is lower than the estimated global rate, the US ranked among the top ten countries in total number of PTBs and remains among the highest in developed countries [Chawanpaiboon et al 2019].

2.1. Preterm Birth: Definitions and Complications

Preterm birth (PTB), defined as birth before the 37th week of gestation, is a serious health concern, and is recognized as the leading cause of neonatal mortality and morbidity in the United States (US) [ACOG 2012]. The World Health Organization (WHO) further subcategorizes PTB on the basis of gestational age:

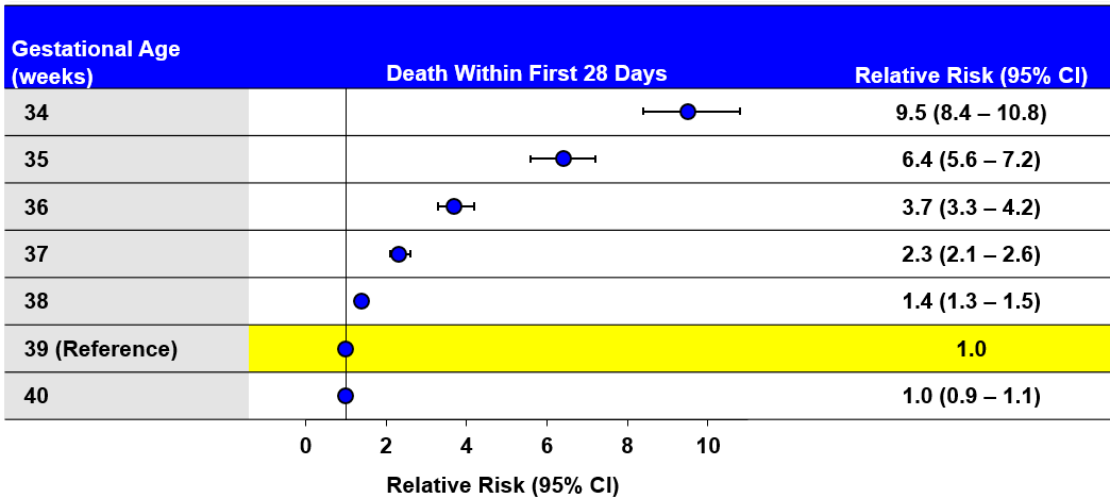
- extremely preterm (<28 weeks);
- very preterm (28 to <32 weeks);
- moderate or late preterm (32 to <37 completed weeks of gestation)

One of the most significant risk factors for spontaneous singleton PTB is a patient's history of PTB. Women who have had a prior PTB have a 2.5-fold greater risk for subsequent PTB than women with no prior history of PTB [Iams et al 1998; Mercer et al 1999].

Infants born prematurely have increased risks of mortality and morbidity throughout childhood, especially during the first year of life. Premature birth is the number one cause of death of children under 5 years old worldwide. Of the estimated 5.43 million deaths of children under the age of 5 in 2017, complications from preterm births accounted for nearly 1 million deaths [WHO 2018]. When using 39 weeks as the reference point of 1.0 for both neonatal and infant

mortality, death within the first 28 days is significantly higher for those babies born at 34, 35 and even 36 weeks of gestation, with the relative risk of neonatal mortality being 9.5 times for a baby born at 34 weeks than that of a baby born at 39 weeks and 3.7 times greater for a baby born at 36 weeks (Figure 6).

Figure 6: Neonatal Mortality Rates by Gestational Age



Source: Reddy et al 2009, Table 2.

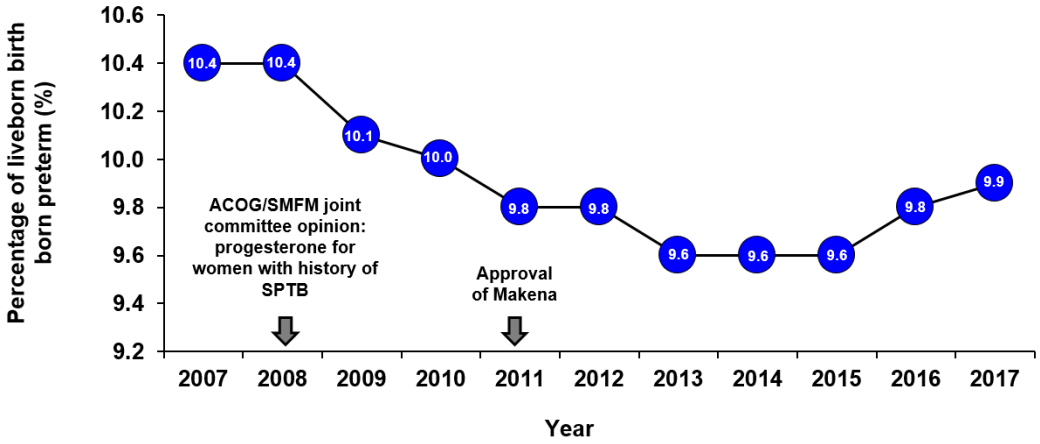
Infants who do survive premature birth often suffer long-term health problems and potential for long-term physical and cognitive disabilities. During the birth hospitalization, late preterm infants are at increased risk for morbidities such as respiratory distress, hypothermia, feeding difficulties, hyperbilirubinemia, and hypoglycemia. After discharge, late preterm infants are at increased risk for rehospitalization, mortality, and other morbidities, including neurologic, respiratory, developmental, and psychiatric/behavioral disorders [Huff et al 2019].

2.2. Prevalence

Despite advances in perinatal care, the incidence of PTB remains high in the US, with rates among the highest among industrialized countries [March of Dimes 2015].

According to the Centers for Disease Control and Prevention, ~10% of liveborn births each year, or nearly 400,000, are born prematurely (Figure 7). Rates of PTB are highest in the areas of the country with the greatest disparities in health care, particularly in minorities and poor communities.

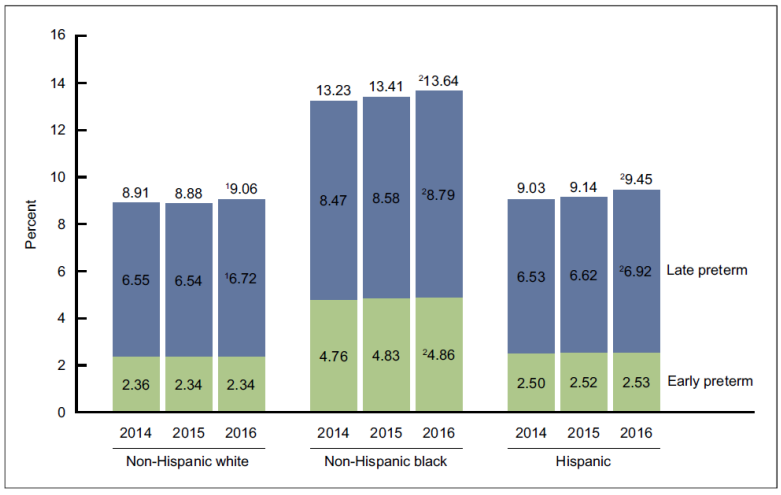
Figure 7: Preterm Birth Rates in United States (2007 through 2017)



Source: Adapted from March of Dimes 2018.
Data from NCHS, National Vital Statistics System, Natality.

Approximately 30% of women who deliver preterm had a history of a prior singleton spontaneous PTB [Gallagher et al 2018]. In addition to prior PTB, there are additional known risk factors. A review of rates of PTB in the US demonstrates a higher PTB rates in non-Hispanic Black women (Figure 8), who are more likely to experience adverse pregnancy outcomes such as PTB, hypertensive disease of pregnancy, and small-for-gestational age birth [Grobman et al 2018]. Other studies have reported that Black women are twice as likely as White women to have preterm deliveries and three times as likely to have very preterm deliveries (<32 weeks), which are the most vulnerable to mortality and long-term morbidities [Carmichael et al 2014; McKinnon et al 2016]. In 2009, reported PTB rates were as high as 17.5% in Black Americans, compared to just 10.9% in White Americans [Martin et al 2011].

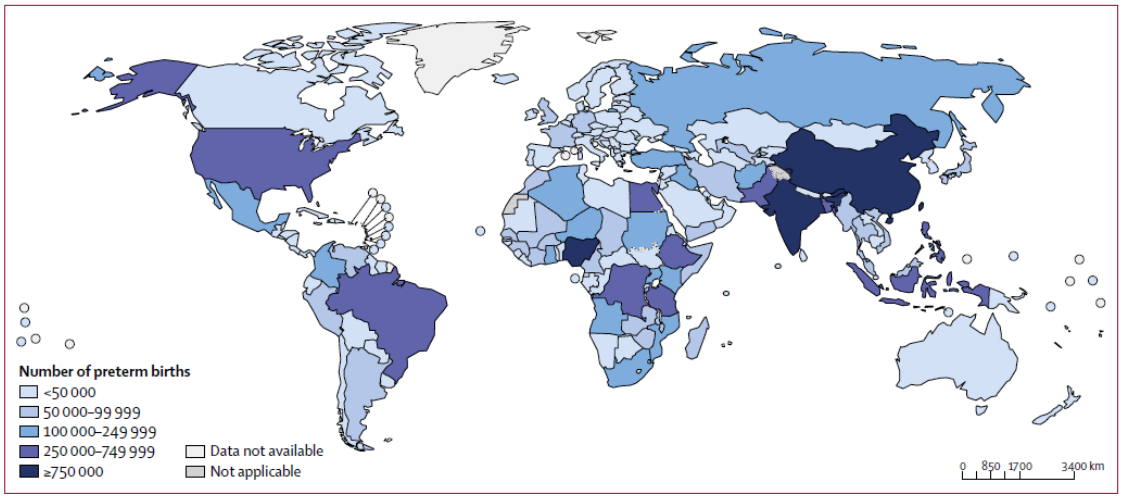
Figure 8: Preterm Birth Rates in the United States by Race and Ethnicity (2014 to 2016)



Source: Martin and Osterman 2018, Figure 3
[1] Significant increase from 2014 and 2015 ($p<0.05$).
[2] Significantly increasing linear trend for 2014-2016 ($p<0.05$).
Notes: Preterm is <37 weeks, late preterm is 34-36 weeks, and early preterm is <34 weeks of gestation. Figures may not add to totals because of rounding. Data source from NCHS, National Vital Statistics System, Natality.

In 2014, the estimated global PTB rate was 10.6%, equating to an estimated 14.84 million (12.65 million to 16.73 million) live preterm births [Chawanpaiboon et al 2019]. While the rate of PTB in the US is lower than the estimated global rate, the US ranked among the top ten countries in total number of PTBs (Figure 9), and remains among the highest in developed countries. In 2010, the World Health Organization ranked the US as 131st out of 184 countries in regard to rates of PTB.

Figure 9: Estimated Numbers of Preterm Births Worldwide (2014)



Source: Chawanpaiboon et al 2019, Figure 2.

3. PREVENTION OF PRETERM BIRTH

Summary

- Hydroxyprogesterone caproate (HPC) or "17P", has a history of being prescribed for use in pregnant women dating back approximately 6 decades, supported by 7 controlled studies on the use of HPC for prevention of PTB [Levine 1964; Papiernik-Berkhauser 1970; Johnson et al 1975; Yemini et al 1985; Suvonnakote 1986, Meis et al 2003, Saghafi et al 2011].
- In a large (N=463), controlled clinical study (Meis et al 2003), 17P was shown to:
  - Reduce the incidence of PTB <37^0 weeks of gestation compared with vehicle (p <0.001);
  - reduce the incidence of PTB when defined as <35^0 (p=0.026) or <32^0 (p=0.027) weeks of gestation;
  - Prolong the duration of pregnancy from time of enrollment (p=0.002);
  - Lower the rates of low birth-weight infants (<2500 g), neonates with necrotizing enterocolitis (NEC), neonates having any grade 3 or 4 intraventricular hemorrhage (IVH), neonates requiring supplemental oxygen, and neonates requiring admission to the neonatal intensive care unit (NICU) (p<0.05).
- Progestogens, including 17P, have been recommended for use in treatment guidelines issued by professional societies.
- Clinicians rely on 17P as the only FDA-approved therapy to prevent recurrent PTB.
- Given the adverse consequences associated with PTB, coupled with the increasing incidence of PTB in the US, there is a clear continued medical need for effective prophylaxis agents such as 17P.

3.1. Prophylactic Methods

Prophylactic methods for prevention of PTB, including tocolytic drugs, bed rest, and other interventions such as cerclage, have been shown in most studies to be ineffective [Creasy 1993; Keirse et al 1989]. One of the preventive measures that has shown effectiveness in randomized trials is the use of progesterone agents [Keirse 1990; Meis and Aleman 2004]. Progesterone has been shown to support gestation and to inhibit uterine activity.

3.1.1. Hydroxyprogesterone Caproate

Hydroxyprogesterone caproate (HPC), or "17P", has a history of use in pregnant women dating back approximately 6 decades when it was marketed as Delalutin® (E.R. Squibb & Sons, Inc.). In addition, a number of controlled studies support the use of 17P for prevention of preterm births [Levine 1964; Papiernik-Berkhauser 1970; Johnson et al 1975; Yemini et al 1985; Suvonnakote 1986, Meis et al 2003, Saghafi et al 2011].

In a large (N=463), controlled clinical study conducted by the National Institutes of Child Health and Human Development (NICHD) through the Maternal Fetal Medicine Units Network (MFMU) (Study 17P-CT-002, hereafter referred to as the "Meis study" [Meis et al 2003]), HPC injection, 250 mg/mL (17P) was shown to:

- significantly reduce the rate of recurrent PTB among women at high-risk for PTB;
- reduce the incidence of PTB $<37^0$ weeks of gestation compared with vehicle ($p<0.001$);
- reduce the incidence of PTB when defined as $<35^0$ ($p=0.026$) or $<32^0$ ($p=0.027$) weeks of gestation;
- prolong the duration of pregnancy from time of enrollment ($p=0.002$); and
- lower the rates of low birth-weight infants (<2500 g), neonates with necrotizing enterocolitis (NEC), neonates having any grade 3 or 4 intraventricular hemorrhage (IVH), neonates requiring supplemental oxygen, and neonates requiring admission to the neonatal intensive care unit (NICU) ($p<0.05$).

Additional details regarding the design and results for this study are presented in Section 6.1.

A follow-up study of children born to mothers who participated in the Meis study was conducted. Of 348 eligible surviving children, 278 (80%) were available for evaluation (194 in the 17P group and 84 in the placebo group). The mean age at follow-up was 48 months. The authors reported that they did not detect differences in developmental delays, safety concerns related to overall health or physical development, or genital or reproductive anomalies between children with in-utero exposure to placebo and in-utero exposure to 17P [Northen et al 2007].

Based on data from the Meis study, 17P was approved under the accelerated approval provisions of Subpart H of 21 CFR Part 314 in February 2011 (New Drug Application [NDA] 21945). Under Subpart H, FDA may grant approval based on demonstrating an effect on a surrogate endpoint that is reasonably likely to predict a drug's clinical benefit.

3.1.2. Vaginal Progesterone

Vaginal progesterone has been studied for the reduction of PTB in women with a history of spontaneous PTB. Several large placebo-controlled trials have failed to find a benefit of vaginal progesterone in patients with a history of SPTB [O'Brien et al 2007; Norman et al 2009; Crowther et al 2017). A 2003 Brazilian study [daFonseca et al 2003] using vaginal progesterone in 142 high-risk women (the majority of whom had a history of preterm delivery) reported a reduction in preterm birth; however, questions have been raised regarding the 14 subjects excluded from the statistical analysis [Tita and O'Day 2004]. A small number of studies have been conducted comparing 17P to vaginal progesterone; these studies have varied in their inclusion criteria. A 2017 Society for Maternal-Fetal Medicine (SMFM) statement noted that the largest of the studies, a Saudi Arabian study by Maher et al [Maher et al 2013], was not generalizable to the US and that vaginal progesterone is not an appropriate substitute for 17P in women with a history of SPTB. Vaginal progesterone has also been studied for a different PTB risk factor of short cervical length; while there have been several studies [Fonseca et al 2007; Hassan et al 2011] indicating a benefit (using varying doses, formulation and inclusion criteria), a 2012 FDA Advisory Committee voted to not approve vaginal progesterone for short cervix as the single study cited in support of the application had inconsistent results, with overall efficacy driven by only two ex-US countries (Belarus and South Africa) [Soule 2012].

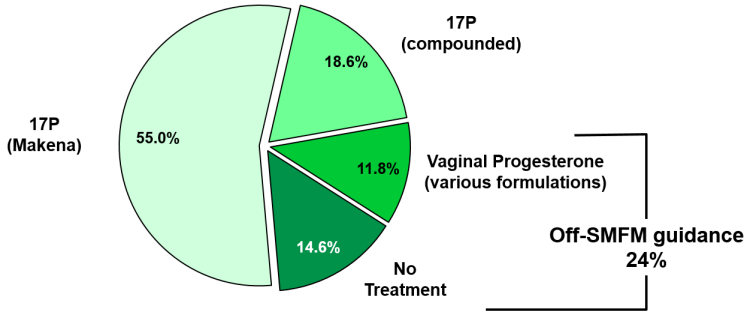
3.1.3. Treatment Guidelines

Progestogens, including 17P, have been recommended for use in treatment guidelines issued by professional societies. In 2008, the American College of Obstetricians and Gynecologists (ACOG) Committee on Obstetric Practice and SMFM issued a joint opinion that progesterone should be offered to patients to prevent recurrent PTB [ACOG 2008].

In 2012, ACOG and SMFM issued separate guidelines regarding the management of women at risk for PTB. In the SMFM guideline, an algorithm recommends the use of vaginal progesterone for women with an incidental short cervix and the use of 17P for women with histories of spontaneous PTB. The ACOG guideline was more general and stated only that "progesterone supplementation should be offered" to women with histories of spontaneous PTB [Practice Bulletin 2012].

Based on a retrospective chart review conducted in 2017, the majority of treatment for the prevention of PTB in women with a history of spontaneous PTB in the US is via branded 17P (Makena) (Figure 10) [Gallagher et al 2018].

Figure 10: Type of Treatment for Prevention of Preterm Birth



Source: Adapted from Gallagher et al 2018, Figure 2.
Note: Proportion of SMFM guidance-eligible patients managed by study physicians in previous 12 months by type of treatment/no treatment option based on retrospective chart review (April to June 2017).

3.2. Compounding of 17P

Prior to the approval of Makena in 2011, 17P was available to patients only through pharmacy compounding. Unlike pharmaceutical manufacturers, compounding pharmacies do not have to demonstrate the safety and efficacy of compounded products or adhere to FDA Good Manufacturing Practices (GMPs). GMPs are legally enforceable regulations that specify how pharmaceutical manufacturing, packaging, labeling, testing, and distribution must be done for FDA-approved medications manufactured domestically or imported into the US in order to ensure their identity, strength, quality, and purity. Manufacturing processes must be validated to consistently meet quality standards. Further, GMPs require an independent quality control unit to oversee the manufacturing, packaging, and testing processes and to reject substandard batches [Gudeman et al 2013]. Only about 2% of compounding pharmacies participate in the industry's voluntary accreditation program [Kliff 2012].

When Makena was approved, there were initial concerns regarding patient access to the FDA approved therapy. In March 2011, FDA issued a statement, noting:

> *"In order to support access to this important drug, at this time and under this unique situation, FDA does not intend to take enforcement action against pharmacies that compound hydroxyprogesterone caproate based on a valid prescription for an individually identified patient unless the compounded products are unsafe, of substandard quality, or are not being compounded in accordance with appropriate standards for compounding sterile products. As always, FDA may at any time revisit a decision to exercise enforcement discretion."* [FDA 2011]

The original sponsor of Makena (KV Pharmaceuticals) subsequently obtained samples of compounded 17P and the active pharmaceutical ingredient (API) used by pharmacists to compound 17P, and identified that compounded versions of 17P did not meet the purity and potency specifications designated for Makena [Chollet and Jozwiakowski 2012].

In June 2012, FDA issued an updated statement pertaining to compounding and Makena; of particular relevance is the following position:

> *"If there is an FDA-approved drug that is medically appropriate for a patient, the FDA-approved product should be prescribed and used. Makena was approved based on an affirmative showing of safety and efficacy. The company also demonstrated the ability to manufacture a quality product. The pre-market review process included a review of the company's manufacturing information, such as the source of the API used in the manufacturing of the drug, proposed manufacturing processes, and the firm's adherence to current good manufacturing practice.*
>
> *Compounded drugs do not undergo the same premarket review and thus lack an FDA finding of safety and efficacy and lack an FDA finding of manufacturing quality. Therefore, when an FDA-approved drug is commercially available, the FDA recommends that practitioners prescribe the FDA-approved drug rather than a compounded drug unless the prescribing practitioner has determined that a compounded product is necessary for the particular patient and would provide a significant difference for the patient as compared to the FDA-approved commercially available drug product."* [FDA 2012]

In addition to lack of comparability, there are significant potential safety risks associated with pharmacy compounding products. A stark reminder of these potential safety concerns that can arise from the lack of regulation around purity, potency and sterility of drug products, occurred in the Fall of 2012 when a fungal meningitis outbreak was traced to contaminated compounded drugs formulated and distributed by the New England Compounding Center (NECC). There were 76 deaths attributed to these substandard sterile injectable drugs produced by the NECC, with over 700 patients being gravely sickened [FDA 2017; Raymond 2017]. This public health catastrophe resulted in the passage of the Drug Quality and Security Act, which has expanded FDA's oversight of pharmacy compounding (traditionally regulated under the practice of pharmacy by individual State Boards of Pharmacy).

The key issue is the lack of standard quality oversight of compounded products from a GMP perspective. Whenever this process is lacking or deficient, there is the potential for untoward effects and unnecessary harm to patients. Without FDA-approved forms of 17P (Makena, plus

the 4 generic products available), pharmacy compounding may be the only available source of this injectable drug for pregnant women.

3.3. Continued Medical Need

Clinicians rely on 17P as the only FDA-approved therapy to prevent recurrent PTB. In 2018, an estimated 59,000 of the 135,000 eligible patients were treated with Makena.

Given the adverse consequences associated with PTB, coupled with the increasing incidence in the US, there is a clear continued medical need for effective prophylaxis agents such as 17P, manufactured in a GMP environment.

4. HYDROXYPROGESTERONE CAPROATE

Summary

- Makena, designated as an orphan drug, was approved by FDA in 2011.
- Makena (HPC injection) is available in single or multi-dose vials for intramuscular (IM) injection; it is can be administered via autoinjector for subcutaneous injection.
- HPC is a synthetic progestin with actions similar to naturally occurring progesterone but unlike progesterone it is not metabolized into estrogen or androgens.
- The exact mechanism by which HPC prevents recurrent PTB is not known but is thought to work by decreasing inflammation and stabilizing the myometrium.
- The FDA-approved indication for 17P (Makena, HPC Injection) is that it is a progestin indicated to reduce the risk of PTB in women with a singleton pregnancy who have a history of singleton spontaneous PTB.
- Following the expiration of the orphan drug exclusivity in February 2018, four generic HPC products have been approved.

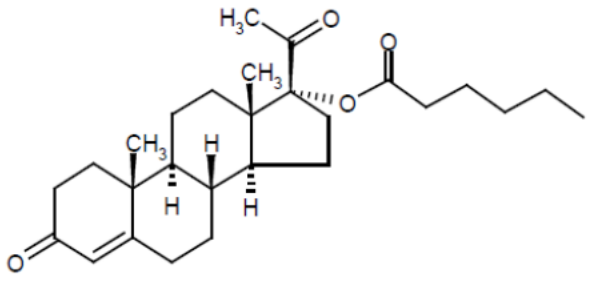
4.1. Makena (HPC Injection)

4.1.1. Product Description

Makena was approved by FDA in 2011 and is a clear, yellow, sterile, non-pyrogenic solution for intramuscular (vials) or subcutaneous (auto-injector) injection. Each 1.1 mL Makena auto-injector for subcutaneous use and each 1 mL single-dose vial for intramuscular use contains HPC USP, 250 mg/mL (25% w/v), in a preservative-free solution containing castor oil USP (30.6% v/v) and benzyl benzoate USP (46% v/v). Each 5 mL multi-dose vial contains HCP USP, 250 mg/mL (25% w/v), in castor oil USP (28.6%) and benzyl benzoate USP (46% v/v) with the preservative benzyl alcohol NF (2% v/v).

The structural formula of HPC is depicted in Figure 11.

Figure 11: Makena Structural Formula



4.1.2. Mechanism of Action

HPC is a synthetic progestin with actions similar to naturally occurring progesterone but unlike progesterone is not metabolized into estrogen or androgens. The exact mechanism by which HPC prevents recurrent PTB is not known but it is thought to work by decreasing inflammation and stabilizing the myometrium.

4.1.3. Indication

"Makena is a progestin indicated to reduce the risk of PTB in women with a singleton pregnancy who have a history of singleton spontaneous PTB. The effectiveness of Makena is based on improvement in the proportion of women who delivered <37 weeks of gestation. There are no controlled trials demonstrating a direct clinical benefit, such as improvement in neonatal mortality and morbidity.

Limitation of use: While there are many risk factors for preterm birth, safety and efficacy of Makena has been demonstrated only in women with a prior spontaneous singleton preterm birth. It is not intended for use in women with multiple gestations or other risk factors for preterm birth."

4.2. Generic HPC

Following the expiration of the orphan drug exclusivity for Makena in February 2018, four generic 17P products have been approved. The first generic product was approved by FDA in June 2018, with three others subsequently approved.

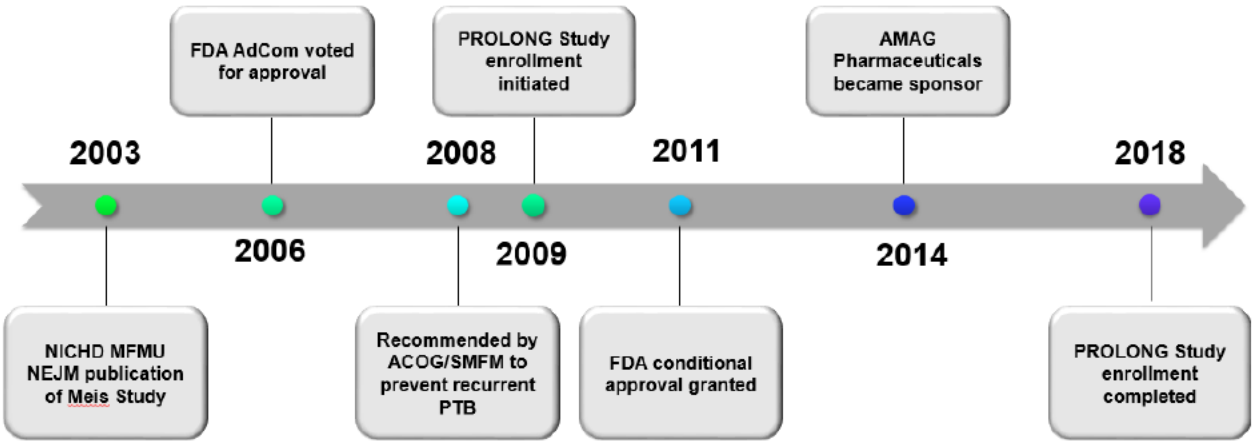
5. REGULATORY HISTORY

Summary

- The Meis trial was conducted by the NICHD and MFMU Network at 19 study centers in the US from 1999 to 2002.
- An FDA Advisory Committee Meeting was held in 2006, with the panel voting unanimously that an additional confirmatory study was required to evaluate safety/efficacy.
- The confirmatory trial (Study 17P-ES-003 or "PROLONG") was initiated in 2009.
- Conditional approval of Makena was granted by FDA in 2011.
- Enrollment in PROLONG was completed in 2018.

A summary of the regulatory history for Makena is depicted in Figure 12.

Figure 12: Makena Regulatory Timeline



Abbreviation: NEJM=New England Journal of Medicine.

The Meis study was a multi-center, double-blind, placebo-controlled trial of pregnant women with a documented history of spontaneous preterm delivery conducted by the NICHD and MFMU Network. The study enrolled 463 patients at 19 clinical centers in the US from 1999 through 2002 [Meis et al 2003]. Treatment with 17P significantly reduced the risk of delivery at <37 weeks of gestation and delivery at <35 weeks of gestation. Patients treated with 17P also had numerically lower rates of delivery at <32 weeks of gestation. Infants of women treated with 17P had lower rates of NEC, IVH, and need for supplemental oxygen.

Recognizing the benefit of having an HPC product manufactured under FDA-regulated GMPs, the NICHD provided Adeza Biomedical access to the clinical data for the purpose of seeking FDA approval of 17P (referred to as "Gestiva" at that time).

5.1. FDA Advisory Committee Meeting (2006)

Following the Gestiva NDA submission in April 2006 which included data from Meis trial as well as follow-up information on infants born to mothers enrolled in that trial, an FDA Advisory Committee Meeting was held in August 2006. Only 5 of 21 panelists felt that a reduction in PTB

prior to 37 weeks gestation was an adequate surrogate endpoint. However, the committee felt that reductions in PTB <35 weeks (yes: 13, no: 8) and <32 weeks (yes: 20, no: 1) were adequate surrogates for neonatal outcomes.

Twelve (12) of the 21 members voted that the Applicant's data provided substantial evidence that 17P treatment prevented preterm birth <35 weeks gestation, and 13 of the 21 members voted that the existing safety data were sufficient to support marketing approval of 17P without the need for additional pre-approval safety data.

All panelists agreed that additional data post-approval was needed to further investigate the safety and efficacy profile of 17P.

5.2. FDA Review of NDA Submission

The original NDA submission for 17P underwent 3 review cycles with FDA.

Cycle 1 (April 2006 to October 2006)

FDA issued an Approvable Letter indicating that future approval under Subpart H would be possible but that additional well-controlled trial(s) would be required to 1) confirm the clinical benefit of 17P, and 2) evaluate the association of 17P treatment with a potential increased risk of second trimester miscarriage and stillbirth. A draft protocol(s) and evidence of feasibility of conducting these trial(s) was required. Additional deficiencies regarding chemistry, manufacturing, and controls and reproductive toxicology were also described in the Approvable Letter.

Cycle 2 (April 2008 to January 2009)

In a Complete Response Letter, FDA stated that "adequate assurance of feasibility could only be addressed by actual initiation of the confirmatory trial".

Cycle 3 (July 2010 to February 2011)

FDA acknowledged the more recent concerns regarding the increased morbidity and mortality of late PTB relative to term births, and recommended that reduction in PTB <37 weeks was an adequate surrogate for clinical benefit.

5.3. Orphan Drug Designation

Orphan status is given to drugs and biologics defined as those intended for the safe and effective treatment, diagnosis or prevention of rare diseases/disorders that affect fewer than 200,000 people in the U.S., or that affect more than 200,000 persons but are not expected to recover the costs of developing and marketing a treatment drug [CFR 21 Part 316]. Orphan drug designation for use of 17P for the prevention of preterm birth in singleton pregnancies was granted on 25 January 2007.

5.4. Confirmatory Study Requirement for Makena

Study 17P-ES-003 (Progestin's Role in Optimizing Neonatal Gestation Trial; hereafter referred to as "PROLONG"), was designed in conjunction with FDA to address the Agency's review of the NDA. In that review and subsequent communication, the FDA requested that efficacy be

established based on both an outcome of PTB and neonatal morbidity/mortality and that the safety endpoint of early fetal loss be examined. Enrollment in PROLONG was initiated in 2009.

During the review process, FDA recognized the difficulty of conducting a study once the drug was approved and adopted due to guidelines supporting its use in this patient population. As a result FDA required that at least 5% of the patients be enrolled prior to approval of Makena, and that at least 10% of the patients be enrolled from North America. After the requisite 10% of patients from North America were enrolled, Makena received approval in 2011.

Given the approval under the accelerated approval pathway, the Indications and Usage section of the label also provides "The effectiveness of [Hydroxyprogesterone Caproate Injection] is based on improvement in the proportion of women who delivered <37 weeks of gestation. There are no controlled trials demonstrating a direct clinical benefit, such as improvement in neonatal mortality and morbidity."

At the time of approval, the Division director commented that:

"Since the time of the meeting, there has been reconsideration of this view, with new recognition of the impact of "late" preterm birth on infant morbidity and mortality. For this reason, the Advisory Committee's overall opinion regarding the merits of a reduction in preterm births at <37 week gestation as an adequate surrogate for a reduction in fetal and neonatal morbidity/mortality is not likely to reflect views currently held by most obstetricians and pediatricians."

However, data that supports the surrogacy of this endpoint to improved neonatal outcomes has been reported. Late PTB (currently defined as occurring 34 to 36 weeks gestation) represents approximately 75% of all PTB. Late preterm births have been increasingly recognized as contributing to both short-term complications and long-term consequences [Moster et al 2008; Reddy et al 2009; Kugelman and Colin 2013]. At 34 weeks gestation, the brain weight is 65% of that of term weight and formation is incomplete [Kugelman and Colin 2013]. Cerebral palsy, mental retardation, psychosocial disorders and other disabilities reported at greater frequency at 34 to 36 weeks compared to >37 weeks [Moster et al 2008]. In addition, neonatal and infant mortality significantly decreases as delivery is closer to 39 to 40 weeks of gestation [Reddy et al 2009].

5.4.1. Postmarketing Commitments

5.4.1.1. PROLONG Study

PROLONG was managed by numerous Sponsors over this period of time (Hologic, KV Pharmaceutical, Lumara Health, and AMAG Pharma USA, Inc.). In 2014, AMAG acquired Lumara Health, who continued to function as a wholly owned subsidiary of AMAG, and from 2016 onward, the study was managed directly by AMAG.

As a result of enrollment challenges for this orphan indication, AMAG submitted two requests to extend the post-marketing requirement timeline (in 2017 and 2019). Enrollment into PROLONG was completed in 2018, and topline data were shared with FDA in early 2019.

Results from PROLONG are provided in Section 6.2.

5.4.1.2. Infant Follow-up Study

A second post-marketing commitment required a clinical follow-up safety study of children born to women who participated in PROLONG. Study 17P-FU-004 is ongoing; participating sites and study staff are blinded to treatment assignment of the subject's mother during PROLONG.

The primary objective of the study is to determine whether there is a difference in developmental status between children, aged 23 to 25 months after adjustment for gestational age, whose mothers received 17P or vehicle while participating in PROLONG.

Although AMAG has been unblinded to PROLONG, it is still blinded to the treatment arm associated with the infant. As of April 1, 2019, a total of 402 child subjects have been consented to participate by their parent(s)/legal guardian(s). Of these, 232 patients have reached 22 months of age and, therefore, their parent(s)/legal guardian(s) have been mailed the Ages and Stages Questionnaire version 3 (ASQ). Of the 232 ASQ's mailed, to date, 183 (78.9%) questionnaires have been returned. Of the 183 received, 42 patients (23%) have scored positive for developmental delay in at least one of the five ASQ domains and have been referred for Bayley Scales of Infant and Toddler Development and neurological exam.

The estimated date for study completion is 4Q2020.

6. CLINICAL DEVELOPMENT PROGRAM

Summary

- The Makena clinical development program was comprised of two key studies:
  - Meis, the pivotal study that served as the basis for approval
    - 19 sites in US (17P N=310; Vehicle: N=153)
    - Enrollment from 1999 to 2002
  - PROLONG, the confirmatory study
    - 93 sites in 9 countries (17P: N=1130; Vehicle: N=578)
    - Enrollment from 2009 to 2018
- Key design elements of both studies:
  - Patients at 16 to 20 weeks of gestation with history of prior PTB
  - Randomized 2:1 to receive weekly IM injections of 17P (250 mg) or vehicle through 36 weeks of gestation or delivery
  - Maternal endpoints of PTB <37 weeks, <35 weeks, and <32 weeks
  - Neonatal morbidity endpoints (death, respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), grade 3 or 4 IVH, proven sepsis, or NEC)

Meis Study

- Risk of PTB $<37^0$ weeks gestation was significantly reduced in the 17P group (37.1% vs 54.9%; p=0.0003).
- 17P also reduced the risk of PTB $<35^0$ weeks gestation (p=0.032) and PTB $<32^0$ weeks gestation (p=0.046).
- The composite neonatal morbidity was numerically lower in the 17P group, but the between group difference was not statistically significant.
- There were no statistical differences in neonatal death rate between the two groups, although the incidence of neonatal death was numerically twice the rate in the vehicle group.

PROLONG

- Study was powered to detect a difference in the co-primary endpoints based on the effect size observed in Meis.
- The study did not meet its co-primary efficacy objectives.
  - Rates of PTBs $<35^0$ weeks gestation were lower than expected (11.0% for 17P and 11.5% for vehicle) and not statistically different (p=0.716).
  - No statistically significant difference in the rates of neonatal mortality or morbidity were noted (5.4% for 17P and 5.2% for vehicle; p=0.840).
- No statistically significant differences between groups were observed in the rates of PTB $<32^0$ weeks (p=0.698) or $<37^0$ weeks gestation (p=0.567).
- Rates of fetal/infant death were low and excluded a doubling of the risk of fetal/early infant death (relative risk 0.79 [95% CI 0.37–1.67].
- Treatment with 17P was generally well tolerated, reaffirming that the safety profile remains acceptable and unchanged.

An overview of the key adequate and well-controlled safety and efficacy studies comprising the Makena clinical development program is provided in Table 11.

Table 11: Overview of Key Clinical Studies

	Meis	**PROLONG**
Year	1999 to 2002	2009 to 2018
Sites	19 sites, US Only	93 sites, 9 countries
Randomization	2:1	2:1
Study Drug	17P 250 mg/mL or vehicle	17P 250 mg/mL or vehicle
Dose	1 dose/week through 36^6 weeks gestation or delivery	1 dose/week through 36^6 weeks gestation or delivery
Study Population	Women 16 to 20 weeks gestation with history of spontaneous preterm delivery	Women 16 to 20 weeks gestation with history of spontaneous preterm delivery
Sample Size	17P: N=310 Vehicle: N=153	17P: N=1130 Vehicle: N=578
Primary Endpoint(s)	• PTB <37 weeks	• PTB <35 weeks • Neonatal Composite Index
Key Secondary Endpoints	• PTB <35 and <32 weeks • Neonatal morbidity/mortality	• PTB <37 and <32 weeks • Fetal/early infant death

In addition to Meis and PROLONG, an initial formulation study (Study 17P-IF-001) was conducted by the NICHD. The study began in February 1998, but treatment was terminated in March 1999 because the active study drug (17P) was recalled by its manufacturer, under the direction of the FDA, due to violations of manufacturing practices potentially affecting the potency of the drug. At the time of termination, only 150 of the proposed 500 patients had been randomized, and no data analysis had been done. Eighty six (86) patients completed the treatment regimen before the study was stopped: 57 on 17P and 29 on Vehicle. Information from this study was considered to be of limited value in supporting either the safety or efficacy of 17P and is not discussed further as it was not part of the initial approval.

6.1. Meis: Pivotal Trial Design and Results

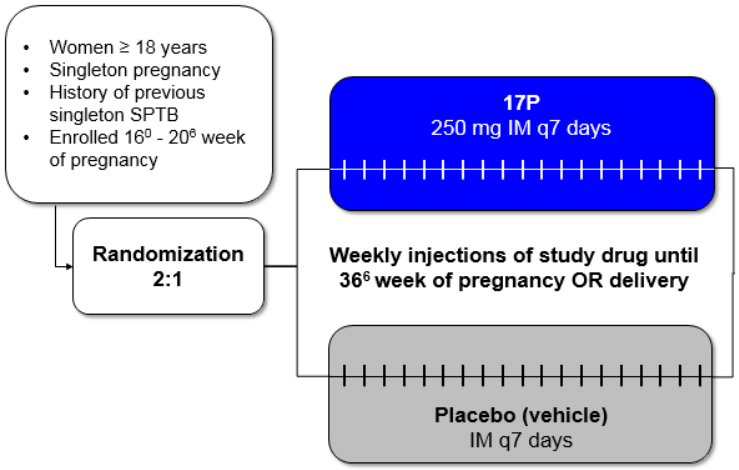
6.1.1. Study Design

The Meis study was conducted by the NICHD through the MFMU from 1999 to 2002. The study was a US-only, double-blind, placebo-controlled trial in pregnant women with a documented history of spontaneous preterm delivery.

The design of the study is depicted in Figure 13. Patients were randomly assigned in a 2:1 ratio, to receive either 17P (250 mg) or vehicle. The vehicle contained all the excipients used in the manufacturing of 17P and contained no active drug. Study drug was administered weekly by IM injection. Weekly study injections continued until delivery or to 36^6 weeks of gestation.

A dose of 250 mg IM was selected based on earlier clinical trials designed to determine if 17P could prevent premature delivery [LeVine 1964; Johnson et al 1975; Yemini et al 1985].

Figure 13: Meis Study Schematic



6.1.1.1. Study Objectives

The primary efficacy outcome was delivery <37^0 weeks. All deliveries occurring from randomization through 36^6 weeks gestation, including miscarriages occurring from 16^0 to 19^6 weeks gestation and elective abortions, were included in the primary outcome.

Secondary objectives of the study were to determine if treatment with 17P:

- reduced the use of tocolytic therapy and/or cervical cerclage.
- reduced neonatal morbidity/mortality.
- reduced the risk of PTB at <35^0 weeks gestation.
- reduced the risk of PTB at <32^0 weeks gestation.
- reduced overall neonatal morbidity based on a composite measure of neonatal morbidity.

6.1.1.2. Statistical Analysis

The primary analysis population was the Intention-To-Treat (ITT), consisting of all randomized patients. Patients with missing outcome data were considered to have delivered at the date last known pregnant.

All statistical comparisons were between 17P and vehicle. Except where explicitly indicated, data were pooled across study centers for all statistical analyses. Patients were analyzed based on the group to which they were randomized.

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

All statistical tests were reported as 2-sided p-values. The final primary efficacy analysis utilized the Type 1 α=0.034 level of statistical significance as required by the O'Brien Fleming

boundary. For all other analyses, no adjustments were made for multiple comparisons and a nominal α=0.05 level of statistical significance was used.

6.1.1.3. Calculation of Gestational Age

Gestational age calculated from the last menstrual period (LMP), date of the first ultrasound (required prior to randomization), and the patient's gestational age at the first ultrasound, derived from the ultrasound measurements. If the LMP date was sure and the ultrasound confirmed the gestational age within a specified number of days, the LMP derived gestational age was used. Otherwise, the ultrasound was used to determine project gestational age.

6.1.2. Study Enrollment

Women were enrolled at 19 clinical centers in the US. In 2002, the prespecified stopping criterion (p=0.015) for efficacy was met at the second interim analysis and the Data Monitoring Committee recommended stopping the trial prior to enrolling the proposed 500 patients. Stopping criteria were in place to assure that once efficacy was established the drug could be made available to all appropriate patients.

6.1.3. Demographics and Baseline Characteristics

In Meis, patients randomized to the two treatment groups were comparable in mean age, race, body mass index (BMI) prior to pregnancy, marital status, years of education, and substance use during pregnancy (Table 12). The majority of patients were Black (approximately 59%), with a mean age of 26.2 years. The mean pre-pregnancy BMI was approximately 26.6 kg/m^2. Approximately 50% of patients in the study were married, and approximately 22% smoked, approximately 8% consumed alcohol, and 3% used illicit drugs during the study pregnancy.

Table 12: Demographic and Baseline Characteristics (Intent-to-Treat Population, Meis)

Characteristic	17P (N=310) n (%)	Vehicle (N=153) n (%)
Age, years		
Mean (SD)	26.0 (5.6)	26.5 (5.4)
Race/ethnic group		
African American	183 (59.0)	90 (58.8)
Caucasian	79 (25.5)	34 (22.2)
Hispanic	43 (13.9)	26 (17.0)
Asian	2 (0.6)	1 (0.7)
Other	3 (1.0)	2 (1.3)
Marital status		
Married or living with partner	159 (51.3)	71 (46.4)
Divorced, widowed, or separated	32 (10.3)	18 (11.8)

Characteristic	17P (N=310) n (%)	Vehicle (N=153) n (%)
Never married	119 (38.4)	64 (41.8)
Pre-pregnancy BMI (kg/m^2)		
Mean (SD)	26.9 (7.9)	26.0 (7.0)
Years of education		
Mean (SD)	11.7 (2.3)	11.9 (2.3)
Substance use during current pregnancy		
Smoking	70 (22.6)	30 (19.6)
Alcohol	27 (8.7)	10 (6.5)
Illicit drugs	11 (3.5)	4 (2.6)

Source: Study 17P-CT-002 Table 11-1.

Obstetrical histories were comparable in the 17P and vehicle groups for gestational age at randomization, gestational age of qualifying delivery, number of previous term deliveries, percentage with previous miscarriages and stillbirths (Table 13). Compared to the vehicle group, the 17P patients had significantly fewer previous preterm deliveries, fewer previous spontaneous preterm deliveries, and a lower percentage of patients with >1 previous preterm delivery.

Table 13: Obstetrical Risk Factors for Preterm Delivery (Intent-to-Treat Population, Meis)

Obstetrical History	17P (N=310) n (%)	Vehicle (N=153) n (%)	p-value
No. of previous preterm deliveries			0.007[a]
Mean (SD)	1.4 (0.7)	1.6 (0.9)	
>1 Previous preterm birth	86 (27.7)	63 (41.2)	0.004[b]
No. of previous SPTB			0.002[a]
Mean (SD)	1.3 (0.7)	1.5 (0.9)	
No. of previous term deliveries			0.665[a]
Mean (SD)	0.8 (1.1)	0.7 (1.0)	
Duration of gestation at randomization, week			0.593[a]
Mean (SD)	18.9 (1.4)	18.8 (1.5)	
Gestational age of qualifying delivery, week			0.208[a]
Mean (SD)	30.6 (4.6)	31.3 (4.2)	
Previous miscarriage	93 (30.0)	57 (37.3)	0.117[b]
Previous stillbirth	31 (10.0)	13 (8.5)	0.604[b]
Infection during pregnancy (before randomization)	98 (31.6)	55 (35.9)	0.351[b]
Corticosteroids during pregnancy (before randomization)	5 (1.6)	8 (5.2)	0.036[c]

Source: Study 17P-CT-002 Table 11-2.
[a] p-value from the Wilcoxon rank sum test.
[b] p-value from the chi-square test.
[c] p-value from the Fisher exact test.

6.1.4. Efficacy

6.1.4.1. Primary Efficacy Endpoint Analysis: Preterm Birth

The risk of delivering prior to 37^0 weeks gestation in the Meis study was significantly reduced in the 17P group (37.1% vs 54.9%; p=0.0003) (Table 14).

Table 14: Percentage of Patients with Delivery <37^0 Weeks of Gestation (Meis)

Data Source	17P n (%)	Vehicle n (%)	Nominal p-value[a]	Treatment difference [95% CI[b]]
ITT Population	115 (37.1)	84 (54.9)	0.0003	-17.8% [-28%, -7%]
Only available data	111 (36.3)	84 (54.9)	0.0000	-18.6% [-29%, -8%]

Source: FDA Background Gestiva (August 2, 2006), Table 4.
Note: ITT population was all randomized patients (17P N=310; Vehicle N=153). The 4 patients with missing outcome data were classified as having a preterm birth of <37^0 weeks (i.e., treatment failure). "Only available data" does not include the 4 patients in the 17P group with missing outcome data.
[a] Chi-square test. Adjusting for interim analyses, p-values should be compared to 0.035 rather than the usual 0.05.
[b] CI adjusted for the 2 interim analyses and the final analysis. To preserve the overall Type I error rate of 0.05, a p-value boundary of 0.035 was used for the adjustment (equivalent to a 96.5% confidence interval).

Because there was an imbalance between the 17P and vehicle groups with regard to the number of previous preterm deliveries, an analysis with adjustment for this variable was performed. The adjusted relative risk of delivery before 37 weeks of gestation in the 17P group as compared with the vehicle group was 0.70 (95% CI, 0.57 to 0.85).

6.1.4.2. Secondary Endpoint Analyses

6.1.4.2.1. Preterm Birth <35 and <32 Weeks Gestational Age

Despite the fact that the study was not powered to determine statistically significant differences in births at <35^0 and <32^0 weeks gestation, 17P demonstrated clinically important reductions in the number of births before 35^0 weeks (p=0.0324) and before 32^0 weeks gestation (p=0.0458) (Table 15).

Table 15: Percentage of Patients with Delivery <35^0 and <32^0 Weeks of Gestation (Meis)

Pregnancy Outcome	17P (N=310) n (%)	Vehicle (N=153) n (%)	Nominal p-value[a]
Delivery <35^0	67 (21.6)	47 (30.7)	0.032
Delivery <32^0	39 (12.6)	30 (19.6)	0.046

Source: FDA Background Gestiva (August 2, 2006), Table 6.
Data presented are from the ITT population (i.e., all randomized patients). The 4 patients with missing outcome data were classified as having a preterm birth <37^0 weeks (i.e., treatment failure).
[a] Adjusting for interim analyses, p-values should be compared to 0.035 rather than the usual 0.05.

At the $<37^0$, $<35^0$, and $<32^0$ weeks gestation, the percentage of deliveries was numerically lower in the 17P treatment arm (Table 16). There was no difference between treatment groups for the percentages of deliveries $<28^0$ weeks.

Table 16: Percentage of Patients with Delivery $<37^0$, 35^0, 32^0, and 28^0 Weeks of Gestation (Intent-to-Treat Population - Meis)

Time of Delivery (Gestational Age)	17P N=310 %	Vehicle N=153 %	Treatment difference[a] [95% CI[b]]
$<37^0$ weeks	37.1	54.9	-17.8% [-28%, -7%]
$<35^0$ weeks	21.6	30.7	-9.1% [-18%, 0.3%]
$<32^0$ weeks	12.6	19.6	-7.05 [-14%, 0.8%]
$<28^0$ weeks	10.0	10.5	-0.5% [-6.9, 5.9]

Source: FDA Background Gestiva (August 2, 2006), Table 7.
[a] Chi-square test.
[b] CI based on a t-test are adjusted for the 2 interim analyses and the final analysis. To preserve the overall Type I error rate of 0.05, a p-value boundary of 0.035 was used for the adjustment (equivalent to a 96.5% confidence interval).

6.1.4.2.2. Neonatal Morbidity and Mortality

A prespecified key secondary endpoint was the incidence rate of having a qualifying event in the composite neonatal morbidity index. The neonatal composite index included neonates with death, respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), grade 3 or 4 IVH, proven sepsis, or NEC) was lower in the 17P group, but the between group difference was not statistically significant (11.9% vs 17.2%; p=0.119) (Table 17).

The study was not powered to detect statistically significant differences between 17P and vehicle treatments in neonatal mortality or morbidities, however, reductions were observed with 17P in the rates of NEC, any grade of IVH, and the need for supplemental oxygen.

Although the overall rate of neonatal deaths was lower in the 17P arm versus vehicle, it was observed that miscarriages (defined as spontaneous loss of fetus from 16^0 to 19^6 weeks gestation) were numerically higher in the 17P arm, as were stillbirths (defined as birth of an infant ≥20 weeks gestation who died prior to delivery) (Table 18). The incidence of neonatal death was twice the rate in the vehicle group, but the between group difference was not statistically significant (p=0.116). Two other NICHD MFMU studies were subsequently conducted; when miscarriage and stillbirth are reviewed in the totality of these studies, the rates were similar between 17P and vehicle [Rouse et al 2007, Caritis et al 2009].

Table 17: Neonatal Morbidity for Live Births (Meis)

Morbidity	17P (N=295) n (%)	Vehicle (N=151) n (%)
Transient tachypnea	11 (3.7)	11 (7.3)
Respiratory distress syndrome	29 (9.9)	23 (15.3)
Bronchopulmonary dysplasia	4 (1.4)	5 (3.3)
Persistent pulmonary hypertension	2 (0.7)	1 (0.7)
Ventilator support	26 (8.9)	22 (14.8)
Supplemental oxygen	45 (15.4)	36 (24.2)
Patent ductus arteriosus	7 (2.4)	8 (5.4)
Seizures	3 (1.0)	0
Any intraventricular hemorrhage	4 (1.4)	8 (5.3)
Grade 3 or 4 IVH	2 (0.7)	0
Other intracranial hemorrhage	1 (0.3)	2 (1.3)
Retinopathy of prematurity	5 (1.7)	5 (3.3)
Proven newborn sepsis	9 (3.1)	4 (2.6)
Confirmed pneumonia	3 (1.0)	4 (2.7)
Necrotizing enterocolitis	0	4 (2.7)
Composite Neonatal Morbidity Score[a]	35 (11.9)	26 (17.2)

Source: FDA Background Gestiva (August 2, 2006), Table 10.
[a] The composite neonatal morbidity measure counted any liveborn infant who experienced death, RDS, BPD, grade 3 or 4 IVH, proven sepsis, or NEC.

Table 18: Miscarriages, Stillbirths, and Neonatal Deaths (Meis)

Pregnancy Outcome	17P (N=306) n (%)	Vehicle (N=153) n (%)	Nominal p-value[a]
Total Deaths	**19 (6.2)**	**11 (7.2)**	**0.689**
Miscarriages <20 weeks gestation	5 (1.6)	0	0.175
Stillbirth	6 (2.0)	2 (1.3)	0.725
Antepartum stillbirth	5 (1.6)	1 (0.6)	---
Intrapartum stillbirth	1 (0.3)	1 (0.6)	---
Neonatal deaths	8 (2.6)	9 (5.9)	0.116

Source: FDA Background Gestiva (August 2, 2006), Table 8.
[a] No adjustment for multiple comparisons.

6.1.4.3. Subgroup Analysis

A post-hoc subgroup analysis of results for PTB <32 weeks, and <35 weeks stratified by race was conducted (Table 19). This analysis demonstrated significant reductions in PTB across all gestational ages in Black patients. Additionally, significant reductions in PTB <37 weeks were observed in non-Black patients. Of note, the study was stopped early based on <37 weeks data, and Blacks made up 59% of the study population relative to 41% non-Black patients.

Table 19: Preterm Birth Stratified by Race (Intent-to-Treat Population, Meis)

	17P (N=310) n/N (%)	Vehicle (N=153) n/N (%)	Difference in % (95% CI)
<32^0 Weeks Gestation			
Black	23/183 (12.6)	22/90 (24.4)	-11.9 (-22.0, -1.8)
Non-Black	16/127 (12.6)	8/63 (12.7)	-0.1 (-10.1, 9.9)
<35^0 Weeks Gestation			
Black	39/183 (21.3)	32/90 (35.6)	-14.2 (-25.8, -2.7)
Non-Black	28/127 (22.0)	15/63 (23.8)	-1.8 (-14.5, 11.0)
<37^0 Weeks Gestation			
Black	66/183 (36.1)	47/90 (52.2)	-16.2 (-28.6, -3.7)
Non-Black	49/127 (38.6)	37/63 (58.7)	-20.1 (-35.0, -5.3)

Source: FDA Table 1, FDA Table 2, and FDA Table 3

6.1.5. Safety

The most common type of adverse event (AE) reported during the study was injection site reactions, which was expected considering that patients received weekly 1 mL IM injections. Pain, swelling, itching, and nodule formation were among the most common reactions regardless whether the solution being injected was 17P or vehicle. However, there was a significantly higher incidence of swelling at the injection site in the 17P group than vehicle (17.1% vs. 7.8%; p=0.007). Nevertheless, few women (1.7%) discontinued the study due to injection site reactions.

The incidence of pregnancy complications, such as preeclampsia, gestational diabetes, or clinical chorioamnionitis, as well as the incidence of serious adverse events (SAEs), was not different between the 17P and vehicle groups. SAEs reported were predominately miscarriages, stillbirths, and neonatal deaths, which were not unexpected events in the high-risk patient population, and were considered by the Investigator to be unrelated to study drug.

6.2. PROLONG: Trial Design and Results

As noted above, Meis was a US-only study that demonstrated that treatment with 17P resulted in a statistically significant reduction in PTB (<37 weeks gestation). The endpoint of PTB defined as <37 weeks gestation was considered an adequate surrogate for clinical benefit to support approval of 17P under subpart H regulations with a single trial. A confirmatory trial

(PROLONG) was required, and FDA requested that PTB defined as <35 weeks and an effect on the neonatal composite index be analyzed as co-primary endpoints.

PROLONG was an international, double-blind, randomized, placebo-controlled trial in pregnant women with a documented history of spontaneous preterm delivery conducted from 2009 through 2018.

6.2.1. Study Design

The design of PROLONG is depicted in Figure 14.

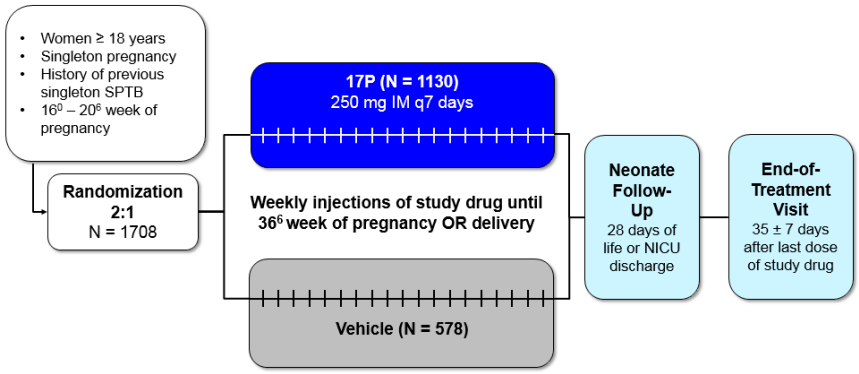
Each patient was randomized in a 2:1 ratio to receive either 17P (250 mg/mL) or vehicle, respectively. Patients received weekly injections of study drug from randomization (16^0 through 20^6 weeks of gestation) through 36^6 weeks of gestation or delivery, whichever occurred first. All injections were administered at the study site.

Randomized patients were to be followed for efficacy outcomes through the date of delivery and for AEs up to the End-of-Treatment Period Visit, defined as 35 ± 7 days after the last dose of study drug. Neonates of randomized patients were followed until Day 28 or the date of discharge from the NICU or equivalent, whichever occurred later. Following delivery, follow-up visits were conducted for both mother and baby.

A prospective, non-interventional infant follow-up study, similar to what was done for Meis, is also being conducted for PROLONG, and is described in Section 5.4.1.2.

Pharmacokinetic (PK) assessments were made based on a sparse sampling of approximately 450 patients (300 active and 150 vehicle), stratified according to BMI to analyze the dose-plasma concentration-time relationship of 17P.

Figure 14: Study Schematic (PROLONG)



6.2.1.1. Study Objectives

There were two co-primary objectives of the study:

- Determine if treatment with 17P injection, 250 mg/mL reduced the rate of PTB $<35^0$ weeks of gestation in women with a singleton pregnancy, aged 18 years or older, with a previous singleton spontaneous preterm delivery.

- Determine if 17P reduced the rate of neonatal mortality or morbidity. Neonatal mortality or morbidity was measured by a composite index comprised of:
  - Neonatal death
  - Grade 3 or 4 IVH
  - RDS
  - BPD
  - NEC
  - Proven sepsis

A key secondary objective of the study was to exclude a doubling of the risk of fetal/early infant death, which was included to address concerns from the original review. Fetal/early infant death was defined as spontaneous abortion/miscarriage (delivery from 16^0 through 19^6 weeks of gestation) or neonatal death (from minutes after birth until 28 days of life) occurring in liveborns born at <24 weeks gestation or stillbirth (antepartum or intrapartum death from 20 weeks gestation through term), in the 17P group compared to the vehicle group.

Additional secondary objectives were to:

- Determine if 17P reduced the rate of PTB $<32^0$ weeks of gestation.
- Determine if 17P reduced the rate of PTB $<37^0$ weeks of gestation.
- Determine if 17P reduced the rate of stillbirth, defined as all stillbirths/fetal deaths/in-utero fetal losses occurring from 20 weeks gestation until term.
- Determine if 17P reduced the rate of neonatal death (from minutes after birth until 28 days of life) occurring in liveborns born at 24 weeks gestation or greater.
- Evaluate the PK/pharmacodynamics of 17P in a subset of pregnant women.

6.2.1.2. Study Population

Study eligibility criteria for PROLONG were based on those used for women in Meis.

Key inclusion criteria included:

- Age ≥18 years
- Singleton gestation
- Project gestational age between 16^0 weeks and 20^6 weeks of gestation at the time of randomization, based on clinical information and evaluation of the first ultrasound
- Documented history of a previous singleton spontaneous preterm delivery, defined as delivery from 20^0 to 36^6 weeks of gestation following spontaneous preterm labor or preterm premature rupture of membranes (pPROM)

Key exclusion criteria included:

- Multifetal gestation
- Known major fetal anomaly or fetal demise (as determined by ultrasound examination between 14^0 through 20^3 weeks of gestation)
- Receipt of a progestin during the current pregnancy AND met one of the following criteria were excluded.
  - Progestin was administered in the 4 weeks preceding the first dose of study medication
  - Patients received HPC
  - Progestin was administered by a route other than oral or intra-vaginal.
- Heparin therapy during current pregnancy or history of thromboembolic disease.
- Maternal medical/obstetrical complications including cerclage, hypertension requiring medication, or seizure disorder
- Presence of a uterine anomaly (except uterine fibroids)
- Prior participation in the trial in a previous pregnancy
- Known hypersensitivity to HPC injection or its components.

6.2.1.3. Statistical Methodology

Analyses were conducted as per the Statistical Analysis Plan, which was approved prior to database lock. All statistical analyses in PROLONG were performed using SAS Version 9.4

6.2.1.3.1. Analysis Populations

Efficacy analyses were conducted using the ITT Population, the Per Protocol (PP) Population, and the Liveborn Neonatal Population. The ITT Population consisted of all randomized patients regardless of whether they received study medication. The efficacy analysis utilized the ITT population which included all randomized patients. No patients were excluded from the efficacy analysis.

The PP Population consisted of all patients who complied with the study protocol. Compliance was based on the following criteria: patient did not have a major protocol deviation potentially affecting efficacy or the evaluation of efficacy as determined by the Sponsor in a blinded review, received the correct blinded study medication for the majority of the duration of study drug receipt, was at least 90% compliant with study medication (based on receipt of study medication through 36^6 weeks of gestation or delivery, whichever occurred first), and had outcome data available.

The Liveborn Neonatal Population consisted of all babies of randomized women who were liveborn and have morbidity data available.

The Safety Population consisted of patients who received any amount of blinded medication.

6.2.1.3.2. Determination of Sample Size

PROLONG was approximately four times the size of the Meis trial and was powered to detect a 30% and 35% treatment difference in the co-primary endpoints (PTB <35 weeks gestation and neonatal composite index).

With 2:1 randomization of 17P and vehicle, a total of 1707 patients were needed to detect a 30% reduction in PTB <35 weeks (from 30% to 21%), giving the study 98% power assuming two-sided type 1 error at 5%. A total of 1665 liveborn infants were needed to detect a 35% reduction in the neonatal composite index (from 17% to 11%), giving 90% power assuming two-sided type 1 error at 5%. Assuming 2.5% of pregnancies result in miscarriage or stillbirth, another 42 women were required (N=1707; 1138 active and 569 vehicle).

Since the outcome measures were co-primary endpoints, the power to detect statistically significant differences between treatments was reduced:

- If outcome measures were independent, power was 88.2%
- If outcome measures were highly correlated (as with Meis), power was 90%.

Assuming 4% fetal/early infant death rate in both treatment arms, a sample size of 1707 provided 82.8% power to rule out a doubling of risk of early fetal/infant death (i.e. the upper bound of the confidence interval for relative risk of 17P compared to vehicle was ≤2.0).

6.2.1.3.3. Interim Analysis

No interim analysis of efficacy was conducted for PROLONG.

6.2.1.3.4. Efficacy Analyses

Primary Efficacy Analyses

Statistically significant differences between the 17P and vehicle treatments in the percentage of patients who delivered $<35^0$ weeks gestation were determined using a Cochran-Mantel-Haenszel (CMH) test stratified by project gestational age at randomization (16^0 weeks – 17^6 weeks gestation and 18^0 weeks – 20^6 weeks gestation).

The number and percentage of neonates in the Liveborn Neonatal Population with the neonatal composite index are presented by project gestational age at randomization stratum and overall for each treatment group. Statistically significant differences between the 17P and vehicle treatment groups were determined using the CMH test stratified by project gestational age at randomization.

Patients with missing delivery data who were known to be pregnant at ≥35 weeks were included in the analysis as not having a PTB<35 weeks. Multiple imputation was used to address other missing data.

Secondary Efficacy Analyses

Statistically significant differences between the 17P and vehicle treatments were determined using the CMH test stratified by project gestational age at randomization. Multiple imputation was used to address missing data for the secondary outcomes as well as was the date last known pregnant as described above for PTB <35 weeks.

6.2.1.3.5. Safety Analyses

Primary Safety Analysis

Analysis of the safety outcome of fetal/early infant death was conducted in the ITT Population. For each gestational age at randomization stratum and overall, the percentage of patients with a fetal/early infant death is provided. The relative risk of fetal/early infant death for the 17P treatment relative to the vehicle treatment was determined using the CMH procedure stratified by project gestational age at randomization stratum. A two-sided 95% CI for the relative risk was constructed using the CMH method adjusted for project gestational age at randomization stratum. If the upper bound of the 95% CI was ≤2.0, a doubling in the risk of fetal/early infant death was ruled out.

6.2.1.3.6. Other Analyses

Study Drug Administration

Dosing information was summarized as the number of injections received and compliance with the expected dosing regimen. Differences between treatment groups in the number of injections and compliance were determined using the Wilcoxon Rank Sum test and for the percentage of patients fully compliant, with the chi-square test.

Gestational Age at Delivery and Neonatal Outcome

A logistic regression model of the neonatal composite index with covariate terms for treatment and gestational age at randomization as a continuous variable was conducted. The odds ratio and 95% CI for the odds ratio for each covariate were calculated.

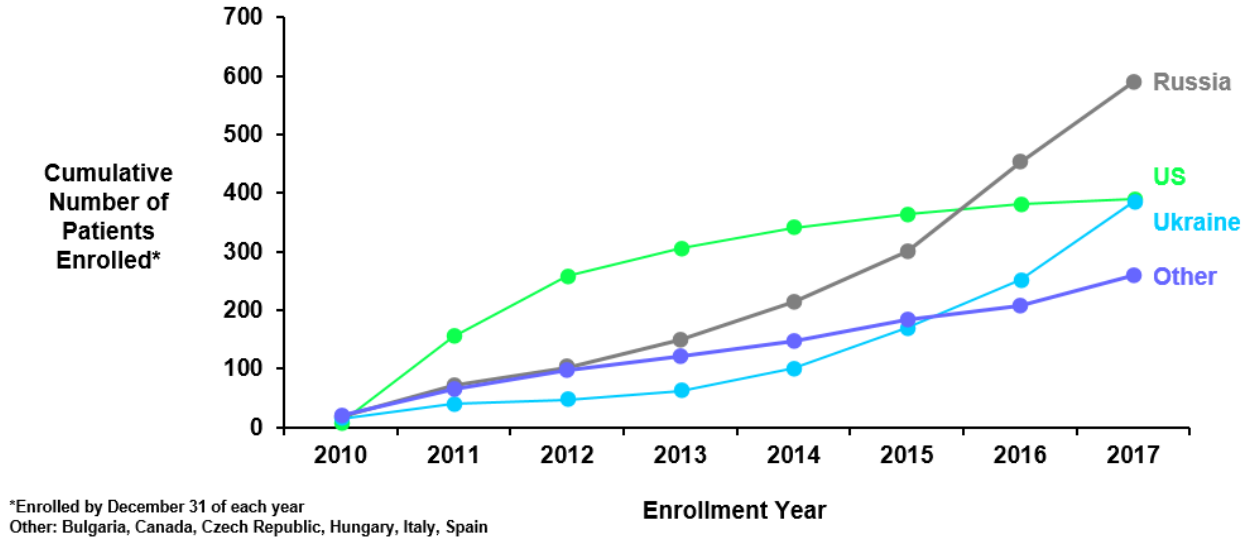
6.2.1.4. Calculation of Gestational Age

Similar to Meis, gestational age in PROLONG was calculated from the patient's menstrual history and measurements obtained at the patient's first ultrasound.

6.2.2. Study Enrollment

Enrollment into PROLONG began in 2009. Following approval of Makena in the US, recruitment in the US became increasingly difficult. Cumulative enrollment rates by year and geographical region showed that, although the overall study enrollment occurred from 2009 to 2018, there was a gradual decline in enrollment rates in the US each year, with nearly 80% of all US patients enrolled by 2013 and nearly 90% by 2014 (Figure 15). By contrast, enrollment rates in Russia and the Ukraine continued to increase with time. It is important to note that both US and ex-US sites were held to the same ICH/GCP standards and ethic committee approvals. Sites in Russia and Ukraine were audited and there were no Major or Critical Findings.

Figure 15: PROLONG Cumulative Enrollment at Year-end (All Countries)



Source: PROLONG CSR, Listing 16.1.1.1.

There were 43 sites in the US that enrolled at least 1 patient in PROLONG. Most of these sites, in contrast to Meis, were in non-urban areas, with 25% of patients residing on military bases.

Table 20 provides an overview of patient enrollment by country. Russia and Ukraine accounted for 61% of study patients, and the US had 23%. The remaining 16% of patients were enrolled in Hungary, Spain, Bulgaria, Canada, Czech Republic, and Italy, each enrolling less than 100 patients.

Table 20: Patient Enrollment by Country (PROLONG)

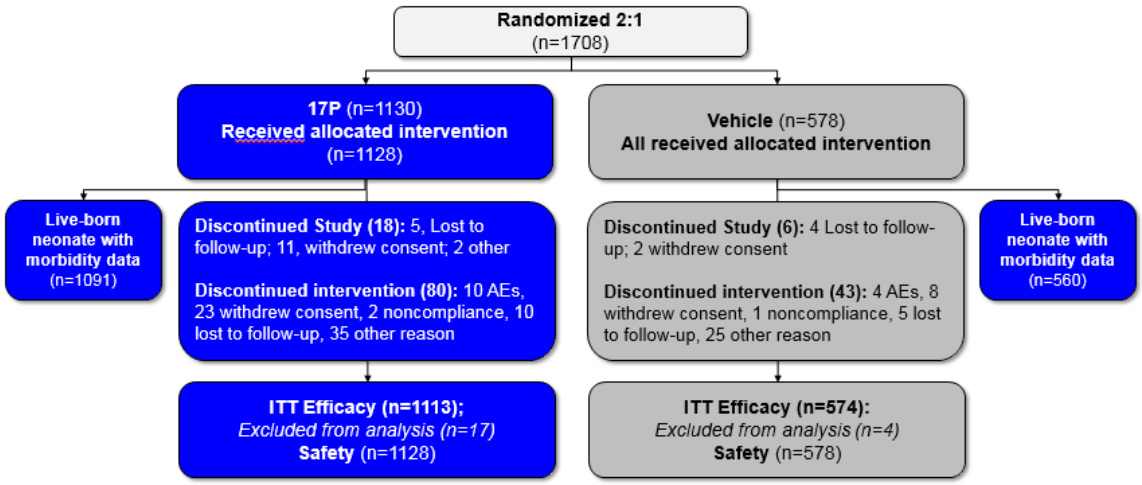
Country	Sites (n)	Patients Receiving Trial Injection (n)	Patients Randomized (n)	Randomized to 17P (n)	Randomized to Vehicle (n)
Overall	**93**	**1740**	**1708**	**1130**	**578**
Russia	12	628	621	414	207
Ukraine	10	424	420	277	143
United States	41	407	391	258	133
Hungary	5	91	91	59	32
Spain	8	85	85	57	28
Bulgaria	6	50	50	33	17
Canada	5	34	31	19	12
Czech	5	15	14	9	5
Italy	1	6	5	4	1

Source: PROLONG CSR Table 14.1.1.1.2.

6.2.3. Disposition

The disposition of patients in PROLONG is presented in Figure 16. A total of 1708 patients were randomized (1130 to 17P and 578 to Vehicle) and included in the ITT Population.

Figure 16: Disposition of Patients (PROLONG)



Source: PROLONG CSR, Figure 1.

A summary of analysis populations is provided in Table 21.

Table 21: Analysis Populations (PROLONG)

	17P n (%)	Vehicle n (%)
Patients randomized (ITT Population)	1130	578
Patients who are protocol compliant (PP Population)	1057 (93.5)	530 (91.7)
Patients excluded from the PP Population:	73 (6.5)	48 (8.3)
Major protocol deviation [a]	29 (2.6)	30 (5.2)
<90% blinded study medication compliance [b]	46 (4.1)	21 (3.6)
No delivery data	18 (1.6)	6 (1.0)
Safety Population	1128 (99.8)	578 (100)
Number of liveborn infants with morbidity data available	1091 (96.5)	560 (96.9)

Source: PROLONG CSR Table 14.1.1.4.
[a] Includes not meeting inclusion/exclusion criteria.
[b] 90% study medication compliance was based on a 10-day cycle.
[c] The Liveborn Neonatal Population consists of all babies of randomized women who were liveborn and have morbidity data available. Excluded are stillbirths (n=16), miscarriages (n=10), elective abortions (n=2), babies for which insufficient data were available to determine liveborn status (n=5) and babies with no morbidity data (n=1).

6.2.4. Demographics and Baseline Characteristics

The treatment groups were comparable across demographic (Table 22), social history (Table 23), and obstetrical characteristics, as well as for social history characteristics (Table 24).

Although the study entry criteria were similar between PROLONG and Meis, the enrolled patient populations differed. When comparing demographics and baseline characteristics of patients enrolled in the two studies, the differences across race and other potential surrogates of socioeconomic status were noteworthy, with Meis representing a much higher-risk population. In comparison to Meis, PROLONG patients had lower risk for spontaneous PTB based on the following key features:

- The majority of patients were White (approximately 89%), non-Hispanic or Latino (approximately 91%) with a mean age of 30 years.
- Approximately 90% of patients were married at the time of study entry.
- Substance use during pregnancy was low in PROLONG (~8% smoked, ~3% consumed alcohol, and 1.4% used illicit drugs).
- Approximately 15% of patients in PROLONG reported >1 previous spontaneous preterm delivery (compared to ~35% in Meis).

Table 22: Demographic and Baseline Characteristics (Intent-to-Treat Population, PROLONG)

	17P (N=1130) n (%)	**Vehicle (N=578) n (%)**
Age (years), n	1130	578
Mean (SD)	30.0 (5.17)	29.9 (5.22)
Ethnicity		
Hispanic or Latino	101 (8.9)	54 (9.3)
Non-Hispanic or Latino	1029 (91.1)	524 (90.7)
Race		
White	1004 (88.8)	504 (87.2)
Black, African American/African heritage	73 (6.5)	41 (7.1)
Native Hawaiian/Pacific Islander	1 (0.1)	0 (0)
Asian	23 (2.0)	22 (3.8)
American Indian or Alaska native	3 (0.3)	0 (0)
Mixed race	8 (0.7)	7 (1.2)
Other	18 (1.6)	4 (0.7)
Pre-pregnancy BMI (kg/m^2), n	1130	577
Mean (SD)	24.3 (7.05)	24.7 (8.65)

Source: PROLONG CSR Table 14.1.3.1.

Table 23: Social History at Baseline (Intent-to-Treat Population, PROLONG)

	17P (N=1130) n (%)	Vehicle (N=578) n (%)
Marital Status		
Married/living with partner	1013 (89.6)	522 (90.3)
Divorced/widowed/separated	31 (2.7)	16 (2.8)
Never married	86 (7.6)	40 (6.9)
Years of Education, n	1129	578
Mean (SD)	13.0 (2.37)	13.0 (2.36)
Substance Use During Current Pregnancy		
Smoking	92 (8.1)	41 (7.1)
Alcohol	24 (2.1)	18 (3.1)
Illicit drugs	16 (1.4)	8 (1.4)

Source: PROLONG CSR Table 14.1.3.2.

Table 24: Obstetrical Risk Factors for Preterm Delivery (Intent-to-Treat Population, PROLONG)

	17P (N=1130) n (%)	Vehicle (N=578) n (%)	p-value[a]
Gestational age at randomization (weeks)[b]			
$<16^0$	6 (0.5)	4 (0.7)	0.051
16^0-17^6	495 (43.8)	236 (40.8)	
18^0-20^6	28 (55.6)	333 (57.6)	
$>20^6$	1 (0.1)	5 (0.9)	
Number of previous preterm deliveries			
Only 1 previous spontaneous preterm delivery	964 (85.3)	494 (85.5)	0.828
>1 previous spontaneous preterm delivery	166 (14.7)	82 (14.2)	
Number of previous miscarriages			
None	644 (57.0)	337 (58.3)	0.873
1	278 (24.6)	139 (24.0)	
>1	208 (18.4)	102 (17.6)	
Number of previous stillbirths			
None	1071 (94.8)	543 (93.9)	0.762
1	55 (4.9)	33 (5.7)	
>1	4 (0.4)	2 (0.3)	
Gestational age of qualifying delivery (weeks)			
20^0-$<28^0$	238 (21.1)	102 (17.6)	0.425
28^0-$<32^0$	202 (17.9)	105 (18.2)	
32^0-$<35^0$	347 (30.7)	187 (32.4)	
35^0-$<37^0$	340 (30.1)	181 (31.3)	

Source: PROLONG CSR Table 14.1.3.3 and PROLONG CSR Erratum Table 14.1.3.4.
[a] p-value is for 17P vs. Vehicle and is from chi-square test or Fisher's exact text for dichotomous variables and the Wilcoxon Rank Sum test for ordinal and continuous variables.
[b] Refers to project gestational age which is the correct gestational age calculated from the patient's menstrual history and measurements obtained at the patient's first ultrasound
[c] Cervical length measurement was not captured for some patients.

6.2.5. Exposure to Study Treatment

Treatment groups were comparable in the mean number of injections received (17.6 and 17.5 injections for patients in the 17P and vehicle groups, respectively; Table 25). More than 96% of patients were considered in full compliance with the injection schedule.

Table 25: Study Medication Administration (Intent-to-Treat Population, PROLONG)

	17P (N=1130)	Vehicle (N=578)	p-value[a]
Number of Injections Received			
N	1128	578	0.991
Mean (SD)	17.6 (3.65)	17.5 (3.81)	
Injection Schedule Compliance (%)[b]			
N	1128	578	0.957
Mean (SD)	96.0 (13.93)	96.4 (13.12)	
Number of patients with Full Compliance[c]	1087 (96.2)	561 (97.1)	0.484
Injection Schedule Compliance (%)			
<80 %	33 (2.9)	17 (2.9)	0.845
80-120 %	44 (3.9)	19 (3.3)	
>120 %	1051 (93.0)	542 (93.8)	

Source: PROLONG CSR Table 14.3.4.
[a] p-value for the Number of Injections Received and Compliance (a) is from the Wilcoxon Rank Sum Test. p-value for Full Compliance (b) and Compliance (c) is from the chi-square test.
[b] Compliance is defined as the number of injections received divided by the number of expected injections (x 100) based on a 7-day injection schedule.
[c] Full compliance is defined as ≥90% compliance based on a 10-day injection schedule.

6.2.6. Efficacy

The study did not meet its co-primary efficacy objectives, which were to demonstrate a reduction in PTB prior to 35^0 weeks gestation and in the neonatal composite index. When comparing demographics and baseline characteristics of patients enrolled in the two studies, the differences across race and other potential surrogates of socioeconomic status were noteworthy, with Meis representing a much higher-risk population.

6.2.6.1. Primary Endpoint Analysis

Rate of PTB

Rates of PTB $<35^0$ weeks were low in both groups and not statistically different between groups (11.0% for 17P and 11.5% for vehicle; Table 26).

Neonatal Composite Index

No statistically significant difference in the rates of neonatal mortality or morbidity as measured by the neonatal composite index, were noted (5.4% for 17P and 5.2% for vehicle; Table 26).

The incidence of individual components of the neonatal composite were similar between treatment groups (Table 27). RDS accounted for almost all of the infants who met the criteria for this index, and rates across treatment groups were not statistically significantly different, at 4.9% and 4.6% in neonates born to patients in the 17P treatment group and vehicle group, respectively

Table 26: Primary Efficacy Outcomes (PROLONG)

Primary Efficacy Outcomes	**17P (N=1130)**	**Vehicle (N=578)**
PTB <35^0 Weeks Gestation (ITT Population)		
Overall Outcome rate n/N* (%)	**122/1113 (11.0)**	**66/574 (11.5)**
p-value[a]	0.716	
Relative risk (95% CI)	0.95 (0.71, 1.26)	
Neonatal Composite Index (Liveborn Neonatal Population)	**(N=1091)**	**(N=560)**
Neonatal Composite Index – Overall, n (%)[d]	**59 (5.4)**	**29 (5.2)**
p-value[b]	0.840	
Relative risk (95% CI)	1.05 (0.68, 1.61)	

Source: PROLONG CSR Table 14.2.1.1.1 and Table 14.2.1.1.2, PROLONG Ad Hoc Table 14.2.1.1.1.26.
[a] p-value from the Cochran-Mantel-Haenszel test.
[b] p-value from the Cochran-Mantel-Haenszel test.
N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 35^0 weeks in the specified category.
The composite index was defined as a liveborn neonate with any of the following occurring at any time during the birth hospitalization up through discharge from the NICU: neonatal death, Grade 3 or 4 IVH, RDS, BPD, NEC, or proven sepsis.

Table 27: Components of Neonatal Composite Index from NICU Outcomes: Liveborn Neonatal Population (PROLONG)

	17P (N=1091) n (%)	**Vehicle (N=560) n (%)**
Neonatal Composite Index – Overall	**59 (5.4)**	**29 (5.2)**
Neonatal death prior to discharge	3 (0.3)	2 (0.4)
Grade 3/4 intraventricular hemorrhage	2 (0.2)	1 (0.2)
Respiratory distress syndrome	54 (4.9)	26 (4.6)
Bronchopulmonary dysplasia	6 (0.5)	1 (0.2)
Necrotizing enterocolitis	2 (0.2)	2 (0.4)
Proven sepsis	5 (0.5)	3 (0.5)

Source: PROLONG CSR Table 14.2.4.1
N=number of babies in the Liveborn Neonatal population in the specified treatment group.

6.2.6.1.1. Assessment for Interaction

Logistic regression analyses of PTB <35^0 weeks gestation and neonatal composite index were conducted to assess whether there was an interaction between treatment and gestational age at the time of randomization. The logistic regression analyses showed no significant interaction between treatment and gestational age at randomization for either primary outcome, indicating a consistent treatment effect regardless of gestational age at randomization.

6.2.6.2. Key Secondary Endpoint Analyses

6.2.6.2.1. Preterm Birth <37 and <32 Weeks of Gestation

There were no statistically significant differences in births at <37^0 (p=0.567) or <32^0 weeks gestation (p=0.698) (Table 28). Rates of PTB were comparable between treatment groups regardless of gestational age at randomization.

Table 28: Percentage of Patients with Delivery <37^0 and <32^0 Weeks of Gestation (Intent-to-Treat Population, PROLONG)

	17P (N=1130) n/N* (%)	**Vehicle (N=578) n/N* (%)**
<32^0 Weeks Gestation	**54/1116 (4.8)**	**30/574 (5.2)**
p-value[a]	0.698	
Relative risk (95% CI)	0.92 (0.60, 1.42)	
<37^0 Weeks Gestation	**257/1112 (23.1)**	**125/572 (21.9)**
p-value[a]	0.567	
Relative risk (95% CI)	1.06 (0.88, 1.28)	

Source: PROLONG Table 14.2.3.2.1 and Table 14.2.3.1.1, PROLONG Ad Hoc Table 14.2.1.1.1.26.
[a] p-value Cochran-Mantel-Haenszel test.
Notes: n=number of patients with delivery <32^0 or 37^0 weeks (as indicated) gestation.
N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 32^0 or 37^0 weeks (as indicated) in the specified category.

Similar rates of spontaneous PTB were observed in each treatment group (Table 29). In addition, the mean gestational age at delivery was comparable for both treatment groups

Table 29: Gestational Age at Delivery (Intent-to-Treat Population, PROLONG)

Gestational Age at Randomization (weeks)[a]	**17P (N=1130)**	**Vehicle (N=578)**
16^0-17^6, n	493	238
Mean (SD)	37.6 (3.6)	37.5 (4.0)
18^0-20^6, n	619	334
Mean (SD)	37.8 (2.7)	37.7 (2.9)
Overall, n	**1112**	**572**
Mean (SD)	**37.7 (3.1)**	**37.6 (3.4)**
p-value[b]	0.952	
p-value[c]	0.981	

Source: PROLONG CSR Table 14.2.4.6.1.
[a] Refers to project gestational age which is the correct gestational age calculated from the patient's menstrual history and measurements obtained at the patient's first ultrasound.
[b] p-value is from the Van Elteren test for continuous variables stratified by gestational age at randomization.
[c] p-value is from the Wilcoxon test for differences in Kaplan-Meier curves.

The treatment groups also had similar maternal delivery characteristics. Most patients had spontaneous labor (71.9% 17P patients and 72.3% vehicle patients). At least one episode of preterm labor was reported for 16.5% 17P patients and 14.5% vehicle patients. Approximately 25% of patients in both treatment groups underwent cesarean section. The median duration of hospitalization was 5.0 days for patients in both treatment groups.

6.2.6.2.2. NICU Outcomes

Table 30 summarizes the NICU outcomes for liveborn neonates. Among the liveborn population of neonates born at ≥24 weeks gestational age, deaths were reported for 3 neonates born to mothers treated with 17P and 2 neonates born to mothers treated with vehicle. In total, 12.4% of neonates born to patients in the 17P treatment group and 10.4% of neonates born to patients in the vehicle group were admitted to the NICU.

Table 30: Infant NICU Outcome (Liveborn Neonatal Population, PROLONG)

	17P (N=1091) n (%)	Vehicle (N=560) n (%)
Components of Neonatal Composite Index		
Neonatal death[a]	3 (0.3)	2 (0.4)
Grade 3/4 intraventricular hemorrhage	2 (0.2)	1 (0.2)
Respiratory distress syndrome	54 (4.9)	26 (4.6)
Bronchopulmonary dysplasia	6 (0.5)	1 (0.2)
Necrotizing enterocolitis	2 (0.2)	2 (0.4)
Proven sepsis	5 (0.5)	3 (0.5)
Other NICU Outcomes[c]		
Any intraventricular hemorrhage	46 (4.2)	19 (3.4)
Transient tachypnea	37 (3.4)	11 (2.0)
Neonatal hypoglycemia	10 (0.9)	5 (0.9)
Confirmed pneumonia	10 (0.9)	2 (0.4)
Retinopathy of prematurity	5 (0.5)	7 (1.3)
Patent ductus arteriosis	4 (0.4)	4 (0.7)
Seizures	5 (0.5)	0 (0)
Persistent pulmonary hypertension	2 (0.2)	2 (0.4)
Other intracranial hemorrhage	3 (0.3)	0 (0)
Grade 3/4/5 retinopathy of prematurity	2 (0.2)	0 (0)
Periventricular leukomalacia	1 (0.1)	0 (0)
Infant NICU Outcome		
All infants admitted (N*)	135 (12.4)	58 (10.4)
Died before final discharge from NICU	3 (2.2)	2 (3.4)
Discharged to home	107 (79.3)	46 (79.3)
Discharged to chronic care facility	6 (4.4)	1 (1.7)
Discharged to non-medical facility (other than home)	2 (1.5)	1 (1.7)
Discharged to step-down unit	15 (11.1)	8 (13.8)
Unknown	2 (1.5)	0 (0.0)
Respiratory Needs		
Number of neonates on ventilator support/ receiving supplemental oxygen	130 (11.9)	54 (9.6)
Number of days of respiratory therapy, n	130	54
Mean (SD)	8.3 (23.8)	10.4 (23.4)
Median	2.0	2.0

Source: PROLONG CSR Table 14.2.2 and Table 14.2.4.1.

[a] Number and percent of neonatal deaths was based on the Liveborn Neonatal Population Born at ≥24 Weeks Gestational Age (N for 17P=1089 and for vehicle=558).

[c] NICU outcomes that were part of the Neonatal Composite Index as well as an NICU outcome are presented here only once as part of the Neonatal Composite Index.

Notes: N=number of babies in the Liveborn Neonatal population in the specified treatment group.
n=number of babies within a specific category. Percentages are calculated as 100 x (n/N) except for the Infant NICU Outcome section in which percentages are calculated as 100 x (n/N*) where N* is the value in the All Infants Admitted row.

6.2.6.3. Subgroup Analysis

6.2.6.3.1. Efficacy by Geographic Region

The event rates for PTB and the neonatal composite index were 1.5 to 2 times higher at 16 to 18% in the US relative to ex-US regions (10%) (Table 31). The rates of PTB among US patients were the highest of the three top enrolling countries in the study (Russia, Ukraine and US), while the rates in Russia and Ukraine were the lowest (Table 32). The rates of the neonatal composite index in the regions with the highest enrollments (Russia and Ukraine) were among the lowest observed. This is consistent with the known epidemiology, as well as the substantially different health care delivery system in these countries, where early intervention to improve prenatal care and reduce neonatal complications is universally available [Healthy Newborn Network 2015; Russian Federation: Federal State Statistics Service 2012; UNICEF 2017; USAID 2011].

Table 31: Primary Efficacy Outcomes by Geographic Region (PROLONG)

Primary Efficacy Outcomes	17P (N=1130)	Vehicle (N=578)
PTB <35^0 Weeks Gestation (ITT Population) (Note 1)		
US Outcome rate n/N* (%)	**40/256 (15.6)**	**23/131 (17.6)**
Relative Risk (95% CI)	**0.88 (0.55, 1.40)**	
Ex-US Outcome rate n/N* (%)	**82/857 (9.6)**	**43/443 (9.7)**
Relative Risk (95% CI)	**0.98 (0.69, 1.39)**	
Russia	27/406 (6.7)	18/206 (8.7)
Ukraine	27/270 (10.0)	14/142 (9.9)
Hungary	11/59 (18.6)	4/32 (12.5)
Spain	8/57 (14.0)	3/28 (10.7)
Canada	5/19 (26.3)	3/12 (25.0)
Bulgaria	4/33 (12.1)	0/17 (0)
Czech Republic	0/9 (0)	1/5 (20.0)
Italy	0/4 (0)	0/1 (0)
Neonatal Composite Index (Liveborn Neonatal Population) (Note 2)	**(N=1091)**	**(N=560)**
US Outcome rate n/N* (%)	**18/252 (7.1)**	**12/126 (9.5)**
Relative Risk (95% CI)	**0.77 (0.39, 1.54)**	
Ex-US Outcome rate n/N* (%)	**41/839 (4.9)**	**17/434 (3.9)**
Relative Risk (95% CI)	**1.27 (0.73, 2.21)**	
Russia	17/401 (4.2)	8/200 (4.0)
Ukraine	13/265 (4.9)	5/140 (3.6)
Canada	4/19 (21.1)	2/12 (16.7)
Spain	3/54 (5.6)	1/27 (3.7)
Hungary	2/57 (3.5)	1/32 (3.1)
Bulgaria	1/30 (3.3)	0/17 (0)
Czech Republic	1/9 (11.1)	0/5 (0)
Italy	0/4 (0)	0/1 (0)

Source: PROLONG CSR Table 14.2.1.5, Table 14.2.1.6, Table 14.2.1.10, and Table 14.2.1.11, PROLONG Ad Hoc Table 14.2.1.1.1.26.

Note 1: N=number of patients in the ITT Population in the specified treatment group.
n=number of patients with delivery <35^0 weeks of gestation in the specified category.
N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 35^0 weeks in the specified category.

Note 2: N=number of babies in the Liveborn Neonatal population in the specified treatment group.
N*=number of babies of patients in the indicated region.
n=number of babies in the specific category. Percentages are calculated as 100 x (n/N*).

Table 32: Preterm Birth by Weeks Gestation for the Three Countries with Largest Enrollments (Intent-to-Treat Population, PROLONG)

Gestation Age at Randomization [a] Outcome Rate	17P (N=1130) n/N* (%)	Vehicle (N=578) n/N* (%)
<32^0 Weeks Gestation		
Russia	13/407 (3.2)	7/206 (3.4)
Ukraine	14/272 (5.1)	6/142 (4.2)
United States	14/256 (5.5)	12/131 (9.2)
<35^0 Weeks Gestation		
Russia	27/406 (6.7)	18/206 (8.7)
Ukraine	27/270 (10.0)	14/142 (9.9)
United States	40/256 (15.6)	23/131 (17.6)
<37^0 Weeks Gestation		
Russia	60/406 (14.8)	35/204 (17.2)
Ukraine	61/269 (22.7)	30/142 (21.1)
United States	85/256 (33.2)	37/131 (28.2)

Source: PROLONG CSR Table 14.2.1.5, Table 14.2.3.1.3, and Table 14.2.3.2.3.
[a] Refers to project gestational age which is the correct gestational age calculated from the patient's menstrual history and measurements obtained at the patient's first ultrasound.
Notes: N=number of patients in ITT Population in the specified treatment group.
n=number of patients with delivery <32^0, 35^0, or 37^0 weeks (as indicated) gestation in the specified category.
N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 32^0, 35^0, or 37^0 weeks (as indicated) in the specified category.

6.2.6.3.2. Efficacy by Obstetric History

Rates of PTB <35^0 weeks gestation and neonatal composite index were also examined for differences in obstetrical history including gestational age of qualifying delivery, gestational age of earliest prior PTB, and number of previous preterm deliveries. Results were similar for both treatment groups across subgroups (Table 33).

Table 33: Primary Efficacy Outcomes by Gestational Age of Qualifying Delivery, Earliest Prior Preterm Birth, and Number of Previous Preterm Deliveries (PROLONG)

Primary Efficacy Outcomes	17P n/N* (%)	Vehicle n/N* (%)
PTB <35^0 Weeks Gestation (ITT Population)	**(N=1130)**	**(N=578)**
Gestational Age of Qualifying Delivery		
20^0-<28^0	29/229 (12.7)	9/101 (8.9)
28^0-<32^0	24/201 (11.9)	20/104 (19.2)
32^0-<35^0	36/344 (10.5)	24/186 (12.9)
35^0-<37^0	32/336 (9.5)	13/180 (7.2)
Gestational Age of Earliest Prior PTB		
20^0-<28^0	40/275 (14.5)	14/125 (11.2)
28^0-<32^0	26/207 (12.6)	20/105 (19.0)
32^0-<35^0	30/336 (8.9)	20/177 (11.3)
35^0-<37^0	26/295 (8.8)	12/165 (7.3)
Number of Previous Preterm Deliveries, n (%)		
1	80/949 (8.4)	51/491 (10.4)
>1	42/164 (25.6)	15/81 (18.5)
Neonatal Composite Index (Liveborn Neonatal Population)[a]	**(N=1091)**	**(N=560)**
Gestational Age of the Qualifying Delivery		
20^0-<28^0	17/221 (7.7)	3/97 (3.1)
28^0-<32^0	14/198 (7.1)	13/102 (12.7)
32^0-<35^0	15/339 (4.4)	9/182 (4.9)
35^0-<37^0	13/330 (3.9)	4/176 (2.3)
Gestational Age of Earliest Prior PTB		
20^0-<28^0	20/265 (7.5)	5/121 (4.1)
28^0-<32^0	13/202 (6.4)	13/103 (12.6)
32^0-<35^0	15/333 (4.5)	8/173 (4.6)
35^0-<37^0	11/291 (3.8)	3/161 (1.9)
Number of Previous Preterm Deliveries, n (%)		
1	43/933 (4.6)	22/478 (4.6)
>1	16/158 (10.1)	7/80 (8.8)

Source: PROLONG CSR Table 14.2.1.2, Table 14.2.1.3, Table 14.2.1.4, Table 14.2.1.7, Table 14.2.1.8, and Table 14.2.1.9.

For PTB <35^0 weeks gestation, n=number of patients with delivery <35^0 weeks of gestation in the specified category and N*=number of ITT patients with non-missing delivery data or who were known to still be pregnant at 35^0 weeks in the specified category.

[a] For neonatal composite index, n=number of babies of patients in the specified category and N*=number of babies of patients in the Liveborn Neonatal Population in the specified category.

6.2.7. Safety

6.2.7.1. Primary Safety Outcome: Fetal and Early Infant Death

The primary safety objective of PROLONG was to rule out a doubling in the risk of fetal or early infant death in the 17P group compared to vehicle. This objective was included specifically to address the Agency's concern of a potential "safety signal" relative to the numerically higher rate of both miscarriage and stillbirth from the Meis study.

Fetal/early infant death was defined as a spontaneous abortion or miscarriage occurring at 16 weeks 0 days through 19 weeks 6 days; a stillbirth, either antepartum or intrapartum; or a neonatal death, occurring minutes after birth until 28 days of life.

If the upper bound of the CI is less than or equal to 2.0, a doubling in risk of fetal/early infant death can be ruled out. A doubling of risk was selected and agreed upon with FDA based on sample size calculations.

Rates were low and similar between treatment groups (1.68% and 1.90% in the 17P and vehicle groups, respectively) with a relative risk of 0.79 (95% CI 0.37–1.67) (Table 34).Given that the upper bound of the 95% CI is less than 2.0, a doubling in the risk of fetal/early infant death was adequately excluded.

Table 34: Fetal and Early Infant Death (Safety Population, PROLONG)

Primary Safety Outcome	17P (N=1130) n (%)	Vehicle (N=578) n (%)
Fetal/Early Infant Death	**19 (1.68)**	**11 (1.90)**
Relative Risk (95% CI) [a]	**0.79 (0.37 - 1.67)**	

Source: PROLONG CSR, Table 14.3.1.1.1.
[a] Relative risk of fetal/early infant death is from the Cochran-Mantel-Haenszel test.
Notes: N=number of patients in the ITT Population in the specified treatment group.
n=number of patients with Fetal/Early Infant Death in the specific category. Fetal/Early Infant Death is defined as neonatal death occurring in liveborns born at less than 24 weeks of gestation, spontaneous abortion/miscarriage or stillbirth

6.2.7.2. Adverse Events and Maternal Pregnancy Complications (MPC)

Treatment-emergent Adverse Events

The AE profile between the two treatment groups was comparable. There were 57.3% and 57.8% of patients with at least one treatment-emergent AEs (TEAEs) in the 17P and vehicle group, respectively (Table 35). The majority of TEAEs were mild in intensity, and most were considered unrelated to study drug. There was a low percentage of TEAEs leading to study drug withdrawal (1.0% and 0.9%) in the 17P and vehicle group, respectively, with both groups experiencing similar and low rates of serious adverse events (SAEs; 3.0% and 3.1% in the 17P and vehicle group, respectively).

The most frequently reported TEAEs in either treatment group were anemia (9.2% in 17P and 9.7% in vehicle) and headache (6.0% in 17P and 4.8% in vehicle). Other commonly reported TEAEs in the 17P group included nausea (4.9%) and back pain (4.4%).

Table 35: Most Common (≥2% for Either Treatment Group by PT) Treatment Emergent Adverse Events (Safety Population, PROLONG)

System Organ Class Preferred Term	17P (N=1128) n (%)	Vehicle (N=578) n (%)
Patients with at least one TEAE	**653 (57.9)**	**336 (58.1)**
Blood and lymphatic system disorders		
Anaemia	104 (9.2)	56 (9.7)
Anaemia of pregnancy	30 (2.7)	18 (3.1)
Gastrointestinal disorders		
Abdominal pain	40 (3.5)	27 (4.7)
Abdominal pain lower	23 (2.0)	7 (1.2)
Constipation	38 (3.4)	17 (2.9)
Diarrhea	23 (2.0)	13 (2.2)
Dyspepsia	37 (3.3)	25 (4.3)
Nausea	55 (4.9)	26 (4.5)
Vomiting	42 (3.7)	19 (3.3)
General disorders and administration site conditions		
Injection site pain	36 (3.2)	24 (4.2)
Injection site pruritus	42 (3.7)	23 (4.0)
Oedema peripheral	25 (2.2)	11 (1.9)
Infections and infestations		
Nasopharyngitis	39 (3.5)	27 (4.7)
Urinary tract infection	44 (3.9)	23 (4.0)
Vaginal infection	41 (3.6)	21 (3.6)
Vaginitis bacterial	35 (3.1)	22 (3.8)
Vulvovaginal candidiasis	21 (1.9)	12 (2.1)
Metabolism and nutrition disorders		
Gestational diabetes	33 (2.9)	21 (3.6)
Musculoskeletal and connective tissue disorders		
Back pain	50 (4.4)	20 (3.5)
Nervous system disorders		
Dizziness	22 (2.0)	13 (2.2)
Headache	68 (6.0)	28 (4.8)

Table 35 Most Common (≥2% for Either Treatment Group by PT) Treatment Emergent Adverse Events and Maternal Pregnancy Complications (Safety Population, PROLONG) (Continued)

System Organ Class Preferred Term	17P (N=1128) n (%)	Vehicle (N=578) n (%)
Pregnancy, puerperium and perinatal conditions		
Afterbirth pain	48 (4.3)	24 (4.2)
Cervical incompetence	34 (3.0)	16 (2.8)
Placental disorder	28 (2.5)	11 (1.9)
Pre-eclampsia	29 (2.6)	23 (4.0)
Psychiatric disorders		
Insomnia	36 (3.2)	13 (2.2)
Reproductive system and breast disorders		
Shortened cervix	18 (1.6)	15 (2.6)
Skin and subcutaneous tissue disorders		
Pruritus	17 (1.5)	13 (2.2)

Source: NDA 021945 Module 2.7.4 Table 7A-003.
Notes: Version 21.1 of MedDRA was used to code maternal pregnancy complications.
Patients reporting a particular AE (preferred term) or MPC more than once are counted only once by preferred term and System Organ Class.
TEAE were AE occurring on/after randomization through the End of Treatment Period Visit.

Maternal Pregnancy Complications (MPC)

There were 10% and 11.1% of patients who experienced at least one MPC in the 17P and vehicle group respectively (Table 36). The majority of patients who experienced MPC experienced mild events, and most were unrelated to study drug. The most frequently reported MPCs for the 17P group was pre-eclampsia (4.2%) and gestational diabetes (2.9%). The incidence of MPC were similar to that in the vehicle group.

The number of patients diagnosed with gestational diabetes during PROLONG was low, and consistent with the incidence each year in the US (2 to 10% of pregnancies) per Center for Disease Control estimates [CDC 2019].

Table 36: Maternal Pregnancy Complications (Safety Population, PROLONG)

	17P (N=1128) n (%)	Vehicle (N=578) n (%)
Patients with at least one maternal pregnancy complication	**113 (10.0)**	**64 (11.1)**
Gestational diabetes	33 (2.9)	21 (3.6)
Antepartum hemorrhage	5 (0.4)	1 (0.2)
Oligohydramnios	8 (0.7)	11 (1.9)
Preclampsia or gestational hypertension	47 (4.2)	30 (5.2)
Chorioamnionitis	9 (0.8)	2 (0.3)
Premature separation of placenta	16 (1.4)	4 (0.7)
HELLP syndrome	2 (0.2)	0 (0.0)
Eclampsia	1 (0.1)	0 (0.0)

Source: NDA 021945 Module 2.7.4 Table 5A-003.

6.2.7.3. Serious Adverse Events

Overall, 34 (3.0%) 17P patients and 18 (3.1%) vehicle patients experienced serious TEAEs or MPCs. The most frequently reported serious TEAE or MPC for patients treated with 17P were premature separation of placenta (5 patients, 0.4%), placental insufficiency (4 patients, 0.4%), and pneumonia (3 patients, 0.3%); Escherichia coli sepsis, pyelonephritis, and wound infection were each reported by 2 patients in the 17P group. The most frequently reported serious TEAE or MPC for patients treated with vehicle were cholestasis (3 patients, 0.5%), and premature separation of placenta (2 patients, 0.3%).

Two patients each had one serious TEAE/MPC considered possibly related to study treatment (one patient in the 17P group had the TEAE of mild nephrolithiasis considered possibly related and one patient in the vehicle group had the severe MPC of cholestasis considered probably related).

6.2.7.4. Stillbirth and Miscarriage

Stillbirths were reported for 12 (1.1%) 17P patients and 3 (0.5%) vehicle patients (Table 37). All of the stillbirths were deemed unrelated to study drug by the Investigator. Among the 12 that occurred in the 17P group, 8 were listed as "definitely not related," 3 as "unlikely related", and 1 "not related." Two women in the 17P group who delivered stillbirths reported smoking during pregnancy, one tested positive for cannabinoids, 1 had a large subserous myoma, and another had uncontrolled Type 1 diabetes mellitus with documented nephropathy and retinopathy. Ten women had a miscarriage: 4 (0.5%) in the 17P group and 6 (1.3%) in the vehicle group.

Table 37: Stillbirths, Miscarriages, and Early Infant Deaths (Safety Population, PROLONG)

	17P (N=1128) n/N (%)	Vehicle (N=578) n/N (%)	Relative Risk (95% CI)[a]
Fetal/Early Infant Death	**19/1128 (1.7)**	**11/578 (1.9)**	**0.87 (0.42, 1.81)**
Miscarriage	4/866 (0.5)	6/448 (1.3)	0.32 (0.09, 1.14)
Stillbirth	12/1124 (1.1)	3/571 (0.5)	2.07 (0.59, 7.29)
Antepartum stillbirth	4/1124 (0.4)	0/571 (0.0)	-
Intrapartum stillbirth	8/1124 (0.7)	3/571 (0.5)	1.38 (0.37, 5.17)
Early Infant Death	3/1112 (0.3)	2/569 (0.4)	0.73 (0.12, 4.48)

Source: PROLONG Ad Hoc Table 9A-003.
Notes: Fetal/Early Infant Death is defined as spontaneous abortion/miscarriage, stillbirth, or death (from minutes after birth until 28 days of life) occurring in liveborns born at less than 24 weeks gestation.
Miscarriage is defined as delivery from 16 weeks up until 20 weeks of gestation. Includes subjects enrolled prior to 20 weeks 0 days.
Stillbirth is defined as all stillbirths/fetal deaths/in-utero fetal losses occurring from 20 weeks gestation until term (excludes deliveries <20 weeks gestation).
[a] Relative risk for 17P relative to Vehicle (Placebo) and is from the Cochran-Mantel-Haenszel test adjusted for gestational age at randomization.

There was a low percentage of TEAEs (predominantly associated with the injection site) leading to study drug withdrawal (1.0% and 0.9%) in the 17P and vehicle group, respectively (Table 38). None of these events were deemed serious by the study investigator.

Table 38: Treatment Emergent Adverse Events and Maternal Pregnancy Complications Leading to Premature Discontinuation of Study Medication (Safety Population, PROLONG)

Preferred Term	17P (N=1128)	Vehicle (N=578)
Patients with at least one TEAE/MPC leading to discontinuation of study medication	**11 (1.0)**	**5 (0.9)**
Injection site erythema	2 (0.2)	0
Injection site nodule	0	1 (0.2)
Injection site pruritus	0	1 (0.2)
Injection site rash	0	1 (0.2)
Injection site reaction	2 (0.2)	0
Hypothyroidism	1 (0.1)	0
Nausea	1 (0.1)	0
Vomiting	1 (0.1)	0
Cholestasis	0	2 (0.3)
Headache	0	1 (0.2)
Fetal growth restriction	1 (0.1)	0
Pre-eclampsia	0	1 (0.2)
Mood altered	1 (0.1)	0
Shortened cervix	1 (0.1)	0
Vaginal hemorrhage	1 (0.1)	0
Dermatitis allergic	1 (0.1)	0
Dry skin	1 (0.1)	0

Source: NDA 021945 Module 2.7.4 Table 8A-003.
Notes: Version 21.1 of MedDRA was used to code adverse events.
Patients reporting a particular adverse event (preferred term) or MPC more than once are counted only once by preferred term.

6.2.7.5. Safety Conclusions

Results from PROLONG reaffirmed the safety of 17P demonstrated in the Meis study. Importantly, PROLONG excluded any doubling of risk of fetal/early infant death.

There were no new or unexpected safety findings from PROLONG, as 17P demonstrated a safety profile that was comparable to vehicle. 17P was well-tolerated and the majority of patients in PROLONG who experienced TEAEs or MPCs experienced mild events that were unrelated to study drug.

To date the safety information received from the post-marketing setting is consistent with the known safety profile, and no new safety signals have been identified.

6.2.8. Pharmacokinetics

Patients were offered the opportunity to participate in a PK substudy until approximately 450 patients (300 active and 150 vehicle) had been enrolled. PK assessments were made based on sparse sampling, stratified according to pre-pregnancy BMI, to analyze the dose-plasma concentration-time relationship of 17P.

Three blood samples were obtained:

- Before study drug dosing at either Visit 6 or 7 (i.e., Dose 5 or 6).
- Before study drug dosing at either Visit 8 or 9 (i.e., Dose 7 or 8).
- At a separate, non-dosing visit 1 to 6 days after Visit 9, 10, or 11 (i.e., 1 to 6 days after Doses 8, 9, or 10).

The PK analysis, based on a limited number of samples per patient, demonstrated that apparent clearance increased with each of increasing weight and increasing BMI. In turn, systemic exposure to 17P decreased with increasing weight and BMI. However, the magnitude of difference in exposure between the lowest and highest quartiles of BMI was small.

There was no evidence that the PK characteristics of 17P were altered by administration of concomitant medications known to induce or inhibit pathways believed to be involved in the metabolism of 17P. However, the number of patients using relevant concomitant medications was small.

There was also no evidence that the incidence of PTB varied as a function of exposure to 17P. Similarly, there was no evidence that any of seven neonatal outcomes varied as a function of exposure to 17P; however, the incidence of these outcomes was low in both vehicle and 17P treated patients, minimizing the opportunity to assess an exposure-response relationship.

7. EXPLORATORY POST HOC ANALYSES

Summary

- Differences across race and other potential surrogates of socioeconomic status linked to higher rates of PTB were noteworthy between Meis and PROLONG, with most of those differences driven by the ex-US PROLONG subset.
- Compared to the US PROLONG subset and Meis, the ex-US PROLONG population represented a cohort with a lower baseline risk for PTB:
  - Lower percentage with prior spontaneous PTB (11% ex-US PROLONG, 27% US PROLONG, 32% in Meis).
  - Fewer Black patients (1 Black patient ex-US PROLONG, 29% US PROLONG, 60% in Meis).
  - Lower percentage of unmarried patients (4% ex-US PROLONG, 31% US PROLONG, 50% in Meis).
  - Lower percentage of patients with any substance use during pregnancy (4% ex-US PROLONG, 28% US PROLONG, and 26% in Meis).
- The ex-US and US PROLONG subsets had patient populations with lower risk for future PTBs than that of Meis.
  - Nearly 92% of patients in Meis had at least one additional risk factor for PTB (beyond 1 previous spontaneous PTB), compared to 79% in US PROLONG and 48% in ex-US PROLONG.
- A treatment benefit associated with 17P was correlated with increasing levels of baseline risk for recurrent PTB.
  - Meis, the highest risk population, had a treatment benefit favoring 17P at <37, <35, and <32 weeks gestation.
  - No treatment effect favoring 17P was observed in the ex-US PROLONG subset, a decidedly lower risk study population.
  - In the US PROLONG subset, a more intermediate and higher risk population, trends of a treatment effect favoring 17P begin to emerge at <35 weeks and <32 weeks.

PROLONG was the largest trial to date to study the effects of 17P in women with prior spontaneous PTB. Unlike the Meis trial, which showed a treatment benefit, treatment with 17P in PROLONG did not decrease rates of PTB or the overall neonatal composite index.

To better understand these discrepant results, exploratory analyses were conducted. These post hoc analyses examined the potential role that differences between the study populations (demographics and patient characteristics associated with baseline risk levels), and differences in health care delivery systems and geography (access to universal health care, emphasis on preventative care) may have had on the results of the study.

7.1. Comparison of Study Demographics

When comparing demographics and baseline characteristics from PROLONG and Meis, the differences across race and other potential surrogates of socioeconomic status that have been linked to higher rates of PTB were noteworthy, with most of those differences driven by the

ex-US PROLONG subset population (Table 39). Compared to the US PROLONG subset and Meis, the ex-US PROLONG population represented a cohort with a lower baseline risk for PTB.

- **Prior spontaneous PTB:** In ex-US PROLONG, 11% had more than 1 prior spontaneous PTB, compared to 27% in US PROLONG and 32% in Meis.
- **Race/ethnicity:** In ex-US PROLONG, only 1 patient was Black or African American, compared to 29% in US PROLONG and nearly 60% in Meis. Hispanic or Latinos accounted for approximately 8% of patients in ex-US PROLONG, 14% in US PROLONG, and 15% in Meis.
- **Marital status:** In ex-US PROLONG, 4% of patients were unmarried with no partner, compared to 31% in US PROLONG and 50% in Meis.
- **Substance use:** In ex-US PROLONG, approximately 4% of patients reported any substance use during pregnancy (smoking, alcohol or illicit drugs), compared to 28% in US PROLONG and 26% in Meis.

Table 39: Demographics and Baseline Characteristics – Post Hoc (Meis and PROLONG)

Variable	PROLONG (Overall)		PROLONG (Ex-US)		PROLONG (US Only)		Meis	
	17P (N=1130)	Vehicle (N=578)	17P (N=872)	Vehicle (N=445)	17P (N=258)	Vehicle (N=133)	17P (N=310)	Vehicle (N=153)
Age, years (mean ±SD)	30.0 ± 5.2	29.9 ± 5.2	30.5 ± 5.1	30.9 ± 4.9	28.1 ± 5.1	26.7 ± 5.1	26.0 ± 5.6	26.5 ± 5.4
Race, n (%)								
Black or African	73 (6.5)	41 (7.1)	1 (0.1)	0	72 (27.9)	41 (30.8)	183 (59.0)	90 (58.8)
White	1004 (88.8)	504 (87.2)	834 (95.6)	420 (94.4)	170 (65.9)	84 (63.2)	79 (29.6)	34 (26.8)
Hispanic or Latino	101 (8.9)	54 (9.3)	70 (8.0)	31 (7.0)	31 (12.0)	23 (17.3)	43 (13.9)[a]	26 (17.0)[a]
>1 previous SPTB	166 (14.7)	82 (14.2)	95 (10.9)	46 (10.3)	71 (27.5)	36 (27.1)	86 (27.7)[b]	63 (41.2)[b]
Gestational age of qualifying delivery, weeks	31.3 ± 4.35	31.6 ± 4.16	30.9 ± 4.40	31.3 ± 4.21	32.5 ± 3.92	32.5 ± 3.86	30.6 ± 4.6	31.3 ± 4.2
Married or living with partner	1013 (89.6)	522 (90.3)	833 (95.5)	431 (96.9)	180 (69.8)	91 (68.4)	159 (51.3)	71 (46.4)
BMI before pregnancy	24.3 ± 7.1	24.7 ± 8.7	23.4 ± 4.47	23.3 ± 4.39	27.4 ± 11.76	29.3 ± 15.29	26.9 ± 7.9	26.0 ± 7.0
Years of education	13 ± 2.4	13 ± 2.4	13.1 ± 2.40	13.1 ± 2.39	13.0 ± 2.25	12.5 ± 2.22	11.7 ± 2.3	11.9 ± 2.3
Any substance use during pregnancy - n (%)	105 (9.3)	51 (8.8)	36 (4.1)	11 (2.5)	69 (26.7)	40 (30.1)	85 (27.4)	36 (23.5)
Smoking	92 (8.1)	40 (6.9)	34 (3.9)	10 (2.2)	58 (22.5)	31 (23.3)	70 (22.6)	30 (19.6)
Alcohol	23 (2.0)	18 (3.1)	4 (0.5)	2 (0.4)	20 (7.8)	16 (12.0)	27 (8.7)	10 (6.5)
Illicit drugs	15 (1.3)	8 (1.4)	1 (0.1)	0	15 (5.8)	8 (6.0)	11 (3.5)	4 (2.6)

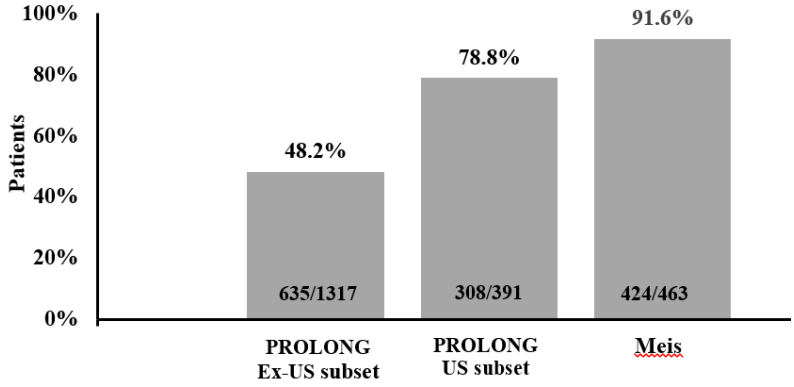
Source: PROLONG Ad Hoc Table 14.1.3.1.10 and Ad Hoc Table 14.1.3.1.11.
[a] Hispanic or Latino included in both race and ethnicity category.
[b] Study 002/PROLONG preterm delivery tables differ. PROLONG % PTB deliveries calculated manually.
NC=not collected.

It is important to note that while US PROLONG patients were more similar to those in Meis, there remain differences related to baseline levels of risk for PTB.

Figure 17 displays a post hoc assessment of select composite risk factors associated with risk of PTB across Meis and PROLONG. The components selected for inclusion (beyond the required entry criteria for at least one prior spontaneous PTB) are >1 prior spontaneous PTB, any substance use, ≤12 years of education, unmarried with no partner, and Black or African American. Importantly, other than a prior history of more than 1 spontaneous PTB, the other components are merely imperfect surrogates of socioeconomic status, an important known predictor of rates of PTB.

The ex-US subset of PROLONG (a low risk population) had a much lower percentage of patients (48.2%) with more than one additional risk factor for PTB compared to the subset of US patients in PROLONG, an intermediate risk population (78.8%) and patients in Meis, a high risk population (91.6%).

Figure 17: Differences in Baseline Risk Factors (Known or Surrogate) Associated with Preterm Birth - Post Hoc (Meis and PROLONG)



Source: PROLONG Ad Hoc Table 14.1.3.1.9.
Notes: The composite risk factors (in addition to the required prior spontaneous PTB) included >1 prior spontaneous PTB, substance use, educational status (≤12 years), unmarried with no partner, and Black/African American. Percentages expressed as n/N x 100, where n is the number of patients with at least 1 additional risk factor and N is the number of patients in the cohort.

7.2. Comparison of Efficacy Outcomes

Study populations with a greater percentage of high risk patients defined by the previously described composite of risk factors appeared to show improved treatment benefit with 17P compared to those with a lower percentage of those patients as shown in Figure 18.

In Meis, which was a higher risk population, a treatment benefit favoring 17P was observed not only with the <37 weeks gestational age, but also at <35 weeks and even at <32 weeks, an important endpoint since it is known that babies born at earlier than 32 weeks have a significant risk of mortality and neonatal complications.

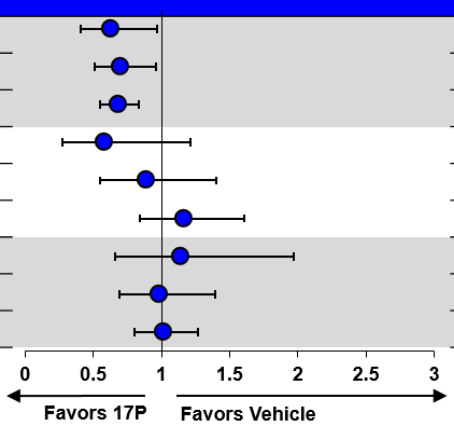
In addition, the intermediate risk population from the US subset of PROLONG also shows trends of a treatment effect favoring 17P beginning to emerge, as this population becomes more similar

to Meis. These trends can be seen at <35 weeks and even at <32 weeks, however not at <37 weeks.

In contrast, the lower risk population of patients from the ex-US subset of PROLONG tend to show no trends of 17P treatment benefit compared to vehicle.

Figure 18: Comparison of Maternal Efficacy Endpoints – Post Hoc (Meis and PROLONG)

Study Population	Gestational Age (weeks)	17P (n/N)	Vehicle (n/N)	Relative Risk (95% CI)
Meis	32	39/310	30/153	0.63 (0.41, 0.97)
Meis	35	67/310	47/153	0.70 (0.51, 0.96)
Meis	37	115/310	84/153	0.68 (0.55, 0.83)
Prolong US	32	14/256	12/131	0.58 (0.27, 1.21)
Prolong US	35	40/256	23/131	0.88 (0.55, 1.40)
Prolong US	37	85/256	37/131	1.16 (0.84, 1.61)
Prolong Ex- Us	32	40/860	18/443	1.14 (0.66, 1.97)
Prolong Ex- Us	35	82/857	43/443	0.98 (0.69, 1.39)
Prolong Ex- Us	37	172/856	88/441	1.01 (0.80, 1.27)



Source: PROLONG Ad Hoc Table 14.2.1.1.1.26.

7.3. Integrated Safety (PROLONG and Meis)

In an effort to continue to fully characterize the safety profile of Makena, an integrated safety analysis was conducted, using two data cohorts from PROLONG and Meis:

1. All patients treated across both studies (17P: N=1438; Vehicle: N=731)
2. US patients only (17P: N=567; Vehicle; N=286)
    - The safety profile of the US only group was consistent with that of the overall integrated dataset and is not discussed further in this document.

MedDRA version 8.0 was used to code AEs in Meis, and Version 21.1 was used for PROLONG.

7.3.1. Common Adverse Events

Similar proportions of patients experienced at least 1 TEAE during the study (56.8% of patients in each treatment group). The most commonly reported TEAE was injection site pain, which occurred in ~10% of patients in each treatment group (Table 40).

Table 40: Incidence of Treatment-Emergent Adverse Events Occurring in at least 2% of Patients in Either Treatment Group by System Organ Class and Preferred Term (Safety Population- PROLONG and Meis Combined)

System Organ Class Preferred Term	17P (N=1438)	Vehicle (N=731)
Patients with at least one TEAE	**817 (56.8)**	**415 (56.8)**
Blood and lymphatic system disorders		
Anaemia	104 (7.2)	56 (7.7)
Anaemia of pregnancy	30 (2.1)	18 (2.5)
Gastrointestinal disorders		
Abdominal pain	43 (3.0)	31 (4.2)
Constipation	40 (2.8)	18 (2.5)
Diarrhoea	30 (2.1)	14 (1.9)
Dyspepsia	37 (2.6)	25 (3.4)
Nausea	73 (5.1)	33 (4.5)
Vomiting	52 (3.6)	24 (3.3)
General disorders and administration site conditions		
Injection site nodule	32 (2.2)	12 (1.6)
Injection site pain	144 (10.0)	74 (10.1)
Injection site pruritus	60 (4.2)	28 (3.8)
Injection site swelling	58 (4.0)	14 (1.9)
Infections and infestations		
Nasopharyngitis	39 (2.7)	27 (3.7)
Urinary tract infection	44 (3.1)	23 (3.1)
Vaginal infection	41 (2.9)	21 (2.9)
Vaginitis bacterial	35 (2.4)	22 (3.0)
Metabolism and nutrition disorders		
Gestational diabetes	33 (2.3)	22 (3.0)
Musculoskeletal and connective tissue disorders		
Back pain	54 (3.8)	21 (2.9)
Nervous system disorders		
Headache	72 (5.0)	28 (3.8)
Pregnancy, puerperium and perinatal conditions		
Afterbirth pain	48 (3.3)	24 (3.3)
Cervical incompetence	34 (2.4)	16 (2.2)

System Organ Class Preferred Term	17P (N=1438)	Vehicle (N=731)
Pre-eclampsia	29 (2.0)	23 (3.1)
Psychiatric disorders		
Insomnia	38 (2.6)	14 (1.9)
Reproductive system and breast disorders		
Shortened cervix	18 (1.3)	15 (2.1)
Skin and subcutaneous tissue disorders		
Pruritus	41 (2.9)	22 (3.0)
Urticaria	43 (3.0)	17 (2.3)

Source: NDA 021945 Module 2.7.4 Table 7A.
N=number of patients in the Safety Population in the specified treatment group.
n=number of patients in the specific category. Percentages are calculated as 100 x (n/N).
Patients reporting a particular AE (PT) more than once are counted only once by PT and System Organ Class.

7.3.2. Serious Adverse Events

In the overall pooled population, less than 4% of patients experienced a serious TEAE (17P 3.5%, vehicle 2.9%) (Table 41). Stillbirth, spontaneous abortion, and premature separation of placenta were the most frequently reported SAE in the 17P group. Fetal/early infant deaths, stillbirths, and miscarriages are described further in the sections that follow.

There were no maternal deaths reported in either study.

Table 41: Incidence of Serious Treatment-Emergent Adverse Events Occurring in at least 2 Patients in Either Treatment Group by Preferred Term (Safety Population- PROLONG and Meis Combined)

Preferred Term	17P (N=1438)	Vehicle (N= 731)
Patients with at least one Serious TEAE	**50 (3.5)**	**21 (2.9)**
Stillbirth	6 (0.4)	2 (0.3)
Abortion spontaneous	5 (0.3)	0 (0.0)
Premature separation of placenta	5 (0.3)	2 (0.3)
Placental insufficiency	4 (0.3)	1 (0.1)
Pneumonia	3 (0.2)	0 (0.0)
Endometritis	2 (0.1)	1 (0.1)
Escherichia sepsis	2 (0.1)	0 (0.0)
Pyelonephritis	2 (0.1)	1 (0.1)
Wound infection	2 (0.1)	0 (0.0)
Cholestasis	0 (0.0)	3 (0.4)

Source: NDA 021945 Module 2.7.4 Table 6A.
N=number of patients in the Safety Population in the specified treatment group.
n=number of patients in the specific category. Percentages are calculated as 100 x (n/N).
Patients reporting a particular AE (PT) more than once are counted only once by PT.
Maternal pregnancy complications are included as TEAEs where applicable.

7.3.2.1. Fetal and Early Infant Deaths

In the overall pooled population, the incidence of fetal death was low and similar in both treatment arms (relative risk 1.01 [95% CI 0.57, 1.79]) (Table 42).

Table 42: Fetal and Early Infant Death (Safety Population- PROLONG and Meis Combined)

Fetal/Early Infant Death[a] by Gestational Age at Randomization		17P (N=1438)	Vehicle (N=731)
16 - <18 Weeks	n[b]/N[c] (%)	17/605 (2.8)	9/287 (3.1)
18 - <21 Weeks	n/N (%)	17/833 (2.0)	8/444 (1.8)
Fetal/Early Infant Death	n/N (%)	34/1438 (2.4)	17/731 (2.3)
Relative Risk[d]	RR (95% CI)	1.01 (0.57, 1.79)	

Source: NDA 021945 Module 2.7.4 Table 1A.

[a] Fetal/Early Infant Death is defined as spontaneous abortion/miscarriage, stillbirth, or death (from minutes after birth until 28 days of life) occurring in liveborns born at less than 24 weeks gestation.

[b] n=number of patients within a specific category. Percentages are calculated as 100 x (n/N).

[c] N=number of patients in the Safety Population in the specified treatment group. The safety population consists of all patients who received any amount of study medication.

[d] Relative risk of fetal/early infant death for 17P relative to vehicle (placebo) and is for the Cochran-Mantel-Haenszel test adjusted for gestational age at randomization.

7.3.2.2. Stillbirths and Miscarriages

In the overall pooled population, miscarriage and stillbirth were infrequent and similar between the treatment groups (Table 43). Stillbirths were reported in 1.3% of 17P patients and 0.7% vehicle-treated patients. Fifteen women had a miscarriage: 9 in the 17P group and 5 in the vehicle group.

Table 43: Stillbirths, Miscarriages, and Early Infant Deaths (Safety Population – PROLONG and Meis Combined)

	17P (N=1438) n/N (%)	Vehicle (N=731) n/N (%)	Relative Risk (95% CI)[a]
Fetal/Early Infant Death	34/1438 (2.4)	17/731 (2.3)	1.01 (0.57, 1.79)
Miscarriage	9/1075 (0.8)	6/555 (1.1)	0.73 (0.26, 2.04)
Stillbirth	18/1429 (1.3)	5/724 (0.7)	1.86 (0.69, 4.99)
Antepartum stillbirth	9/1429 (0.6)	1/724 (0.1)	4.67 (0.58, 37.31)
Intrapartum stillbirth	9/1429 (0.6)	4/724 (0.6)	1.16 (0.36, 3.76)
Early Infant Death	7/1411 (0.5)	6/720 (0.8)	0.58 (0.20, 1.73)

Source: Ad Hoc Table 9A.

Notes: Fetal/Early Infant Death is defined as spontaneous abortion/miscarriage, stillbirth, or death (from minutes after birth until 28 days of life) occurring in liveborns born at less than 24 weeks gestation.

Miscarriage is defined as delivery from 16 weeks up until 20 weeks of gestation. Includes subjects enrolled prior to 20 weeks 0 days.

Stillbirth is defined as all stillbirths/fetal deaths/in-utero fetal losses occurring from 20 weeks gestation until term (excludes deliveries <20 weeks gestation).

[a] Relative risk for 17P relative to Vehicle (Placebo) and is from the Cochran-Mantel-Haenszel test adjusted for gestational age at randomization.

8. DISCUSSION

PROLONG did not meet the predefined co-primary objectives. AMAG believes that the results from PROLONG were influenced by differences in the study population from that previously studied in Meis. While the entry criteria of Meis and PROLONG were similar, the study population in PROLONG was different than that of Meis, with the latter comprised of a higher risk population.

Efficacy

When comparing demographics and baseline characteristics from PROLONG and Meis, the differences across race and other potential surrogates of socioeconomic status that have been linked to higher rates of PTB were noteworthy, with most of those differences were driven by the ex-US PROLONG subset population. As a result, key differences in baseline risk associated with PTB even within the PROLONG study population, notably US vs. ex-US subset populations, make the applicability of the efficacy data particularly challenging in the US.

A review of the baseline characteristics of patients who enrolled in PROLONG in the US demonstrates that although they are more similar to Meis than that of the overall PROLONG population, they remain differ from Meis on many of the risk factors thought to be associated with risk of PTB.

A post-hoc investigation into baseline risk factors indicate that, compared to Meis (a high-risk population), the PROLONG US subset was an intermediate risk group for recurrent PTB, with the PROLONG ex-US subset at lower risk. The lower baseline risk for PTB in ex-US PROLONG could be attributed to varying healthcare delivery systems (more preventive than acute care) with universal access in ex-US countries, which represented 75% of the study population (61% from Russia and Ukraine alone). In a number of these countries, there are dedicated programs that target prevention of PTB and adverse fetal outcomes with evidence-based technologies to improve the quality of perinatal care. Often, these programs include comprehensive measures for pregnancy planning, screening, primary prophylaxis, and risk factor reduction, as well as providing healthcare and treatment of co-morbid conditions prior to pregnancy. In addition, compliance with prenatal care is associated with state-provided financial incentives for new mothers [Healthy Newborn Network 2015; Russian Federation: Federal State Statistics Service 2012; UNICEF 2017; USAID 2011].

Of note, exploratory analyses of PTB rates by baseline risk suggest an increasing treatment benefit associated with 17P with increasing levels of baseline risk for recurrent PTB. Treatment effect was observed at <37, <35, and <32 weeks gestation for the highest risk group (Meis), while the lowest risk group (ex-US PROLONG) showed no effect. Trends favoring 17P emerge in the US PROLONG subset as the population becomes more similar to that of Meis, with increased effect at <35 and <32 weeks, but not at <37 weeks gestation.

In totality, it is possible that differences in baseline risk for PTB underpin the lack of correlation between the efficacy results observed in Meis and PROLONG.

Safety

The key safety outcome of PROLONG was to rule out a doubling of risk of fetal or early infant death in the 17P group relative to vehicle. This endpoint was included specifically to address the

Agency's concern of a potential safety signal relative to the numerically higher rate of both miscarriage and stillbirth from the Meis study. The relative risk of 0.79 with an upper bound of the 95% CI of 1.67 excludes that risk.

The favorable maternal and fetal safety profile of 17P was reaffirmed as there were no new or unexpected safety findings, and no clinically meaningful differences in the safety profile across treatment groups. Specifically, there were no clinically meaningful differences in TEAEs across the two treatment groups (17P and vehicle).

Proposed Changes to Prescribing Information

Based on the results from PROLONG, AMAG is proposing to maintain the indication with the current limitations of use and to amend the current prescribing information to include the following updates:

- Section 6 Adverse Reactions: to include pooled (Meis and PROLONG) safety information
- Section 14.1 Clinical Trials to Evaluate Reduction of Risk of Preterm Birth: to include findings from PROLONG. In particular AMAG proposes that it is important to include information that helps place the results from PROLONG in context with those observed from Meis.

8.1. Conclusions

Differences in study populations between Meis and PROLONG as it relates to baseline levels of risk associated with PTB contributed to the vastly lower rates of PTB and associated prematurity complications seen in PROLONG. It is relevant to acknowledge that in the nearly 20 years since Meis was initiated and PROLONG was completed, there have been substantial improvements in neonatal care that have increased survival. However, rates of PTB in the US have remained relatively constant over that time period and there remains a significant public health concern regarding PTB. Moreover, women with a prior history of spontaneous PTB, particularly if the preterm birth is early (<32 week gestation), or if there is a history of more than one prior spontaneous PTB, are at the highest risk for a recurrent PTB.

The totality of clinical data including more than 16 years of clinical use support 17P's positive benefit-risk profile and support its availability for clinicians to make patient-specific prescribing decisions, based upon their clinical judgment and shared decision-making with their patients.

9. REFERENCES

ACOG Committee Opinion. Use of progesterone to reduce preterm birth. Obstet Gynecol. Nov 2003;102(5 Pt 1):1115-1116.

ACOG Committee Opinion number 419 October 2008 (replaces no. 291, November 2003). Use of progesterone to reduce preterm birth. Obstet Gynecol. Oct 2008;112(4):963-965.

ACOG practice bulletin no. 127: Management of preterm labor. Obstet Gynecol. Jun 2012;119(6):1308-1317.

Caritis SN, Rouse DJ, Peaceman AM, Sciscione A, et al. Prevention of preterm birth using 17 alpha-hydroxyprogesterone caproate: a randomized controlled trial. Obstet Gynecol. Feb 2009;113(2):285-292.

Carmichael SL, Cullen MR, Mayo JA, Gould JB, Loftus P, Stevenson DK, Wise PH, Shaw GM. Population-level correlates of preterm delivery among Black and White Women in the US. PLos ONE 2014;9(4): e94153. Doi:10.1371/journal.pone.0094153

Chawanpaiboon S, Vogel JP, Moller AB, Lumbiganon P, Petzold M, Hogan D, Landoulsi S, Jampathong N, Kongwattanakul K, Laopaiboon M, Lewis C, Rattanakanokchai S, Teng DN, Thinkhamrop J, Watananirun K, Zhang J, Zhou W, Gülmezoglu AM. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Health. 2019;7:E37–46.

CDC. Gestational Diabetes. https://cdc.gov/diabetes/basics/gestational.html. Accessed 20August2019.

Code of Federal Regulations. Title 21 Part 316 Orphan Drugs. https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=718f6fcbc20f2755bd1f5a980eb5eecd&mc=true&n=sp21.5.316.c&r=SUBPART&ty=HTML#se21.5.316_120. Accessed 21September2019.

Chollet JL, Jozwiakowski MJ. Quality investigation of hydroxyprogesterone caproate active pharmaceutical ingredient and injection. Drug Dev Industr Pharm. 2012; 1-10.

Creasy RK. Preterm birth prevention: where are we? Am J Obstet Gynecol. Apr 1993;168(4):1223-1230.

Crowther CA, Ashwood P, McPhee AJ, et al; PROGRESS Study Group. Vaginal progesterone pessaries for pregnant women with a previous preterm birth to prevent neonatal respiratory distress syndrome (the PROGRESS Study): a multicentre, randomised, placebo-controlled trial. PLoS Med 2017;14(09):e1002390.

da Fonseca EB, Bittar RE, Carvalho MHB, Zugaib M. Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: a randomized placebo-controlled double-blind study. Am J Obstet Gynecol 2003;188(02):419–424.

FDA. FDA Statement on Makena. March 30, 2011.

FDA. Questions and answers on updated FDA statement on compounded versions of hydroxyprogesterone caproate (the active ingredient in Makena). June 2012. http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm310215.htm.

FDA. FDA's Human Drug Compounding Progress Report. January 2017.

Fonseca EB, Celik E, Parra, M, Singh, M, Nicolaides KH. Progesterone and the Risk of Preterm Birth among Women with a Short Cervix. NEJM 2007;357:462-9.

Gallagher JR, Gudeman J, Heap K, Vink J, Carroll S. Understanding if, how, and why women with prior spontaneous preterm births are treated with progestogens: a national survey of obstetrician practice patterns. Am J Perinatol Rep 2018; 8:e315-324.

Gudeman J, Jozwiakowski M, Chollet J, Randell M. Potential risks of pharmacy compounding. Drugs R D 2013. Doi:10.1007/s40268-013-0005-9.

Grobman WA, Parker CB, Willinger M, Wing DA, Silver RM, Wapner RJ, Simhan HN, Parry S, Mercer BM, Haas DM, Peaceman AM, Hunter S, Wadhwa P, Elovitz M, Foroud T, Saade G, Reddy UM. Racial Disparities in Adverse Pregnancy Outcomes and Psychosocial Stress. Obstetrics & Gynecology. 2018;131(2):328-335.

Hassan SS, Romero R, VidyadhariD, et al; PREGNANT Trial. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, doubleblind, placebo-controlled trial. Ultrasound Obstet Gynecol 2011; 38(01):18–31.

Healthy Newborn Network. 2015. Accessed December 2017. https://www.healthynewbornnetwork.org/numbers/

Huff K, Rose RS, Engle WA. Late preterm infants: morbidities, mortality, and management recommendations. Pediatr Clin N Am. 2019;66: 387–402.

Iams JD, Goldenberg RL, Mercer BM, et al. The Preterm Prediction Study: recurrence risk of spontaneous preterm birth. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Am J Obstet Gynecol. May 1998;178(5):1035-1040.

Keirse MJ, Grant A, King J. Preterm Labour. In: Chalmers L, Enkin M, Keirse MJ, eds. Effective Care in Pregnancy and Childbirth. New York: Oxford University Press; 1989:694-745.

Keirse MJ. Progestogen administration in pregnancy may prevent preterm delivery. Br J Obstet Gynaecol. Feb 1990;97(2):149-154.

Kliff S, Sun L. Critics charge that drug compounding pharmacies require government oversight. Washington Post. October 18, 2012. http://www.washingtonpost.com/national/health-science/critics-charge-that-drug-compounding-pharmacies-require-government-oversight/2012/10/18/7b726132-17d8-11e2-9855-71f2b202721b_story.html. Accessed 19August2019.

Kugelman A, Colin AA. Late preterm infants: near term but still in a critical development time period. Pediatrics. 2013;132:741–751.

Johnson JW, Austin KL, Jones GS, Davis GH, King TM. Efficacy of 17alpha-hydroxyprogesterone caproate in the prevention of premature labor. N Engl J Med. Oct 2 1975;293(14):675-680.

Levine L. Habitual Abortion. A controlled study of progestational therapy. Western Journal of Surgery, Obstetrics, and Gynecology. Jan-Feb 1964;72:30-36.

Maher MA, Abdelaziz A, Ellaithy M, Bazeed MF. Prevention of preterm birth: a randomized trial of vaginal compared with intramuscular progesterone. Acta Obstet Gynecol Scand 2013;92(02):215–222.

March of Dimes. Prematurity Campaign 2014 Progress Report. 2015. https://www.marchofdimes.org/materials/2014-march-of-dimes-prematurity-campaign-progress-report.pdf. Accessed 16August2019.

March of Dimes. Prematurity Report Card. 2018. https://www.marchofdimes.org/mission/prematurity-reportcard-tv.aspx. Accessed 5August2019.

Martin JA, Hamilton BE, Ventura P, Osterman M, Kirmeyer S, Mathews MS, Wilson E. Births: Final Data for 2009. National Vital Statistic Reports 2011. 60.

Martin JA, Hamilton BE, Osterman MJK, Driscoll AK, Mathews TJ. Births: final data for 2015. National Vital Statistics Reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System. 2017;66(1):1-70.

Martin JA, Hamilton BE, Osterman MJK, Driscoll AK, Drake P. Births: final data for 2016. National Vital Statistics Reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System. 2018;67(1).

Martin JA, Osterman MJK. Describing the increase in preterm births in the United States, 2014-2016. NCHS Data Brief, no 312, Hyattsville, MD; National Center for Health Statistics. 2018

McKinnon B, Yang S, Kramer MS, Bushnik T, Sheppard AJ, Kaufman JS. Comparison of black-white disparities in preterm births between Canada and the United States. CMAJ 2016;188:E19-26.

Meis PJ, Klebanoff M, Thom E, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. N Engl J Med. Jun 12 2003;348(24):2379-2385.

Meis PJ, Aleman A. Progesterone treatment to prevent preterm birth. Drugs. 2004;64(21):2463-2474.

Mercer BM, Goldenberg RL, Moawad AH, et al. The preterm prediction study: effect of gestational age and cause of preterm birth on subsequent obstetric outcome. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Am J Obstet Gynecol. Nov 1999;181(5 Pt 1):1216-1221.

Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. N Engl J Med 2008;359:262-73.

Norman JE, Mackenzie F, Owen P, et al. Progesterone for the prevention of preterm birth in twin pregnancy (STOPPIT): a randomised, double-blind, placebo-controlled study and metaanalysis. Lancet 2009;373(9680):2034–2040.

Northen AT, Norman GS, Anderson K, et al. Follow-up of children exposed in utero to 17 alphahydroxyprogesterone caproate compared with placebo. Obstetrics and gynecology. 2007;110(4):865-872.

O'Brien JM, Adair CD, Lewis, DF, Hall DR, DeFrance EA, Fusey S, et al. Progesterone vaginal gel for the reduction of recurrent preterm birth: primary results from a randomized, double-blind, placebo-controlled trial. Ultrasound Obstet Gynecol 2007; 30: 687-696.

Papiernik-Berkhauser E. Double blind study of an agent to prevent pre-term delivery among women at increased risk. Edition Schering. 1970;Serie IV(fiche 3):65-68.

Practice bulletin no. 130: prediction and prevention of preterm birth. Obstet Gynecol. Oct 2012;120(4):964-973.

Raymond N. US prosecuters seek 35 years in prison for meningitis outbreak exec. June 23, 2017. https://www.reuters.com/article/us-massachusetts-meningitis-idUSKBN19E1GR. Accessed 5September2019.

Reddy UM, Ko C-W, Raju TNK, Willinger M. Delivery indications at late-preterm gestations and infant mortality rates in the United States. Pediatrics. 2009:124;234-240.

Rouse DJ, Caritis SN, Peaceman AM, Sciscione A, et al. A trial of 17-alpha-hydroxyprogesterone caproate to prevent prematurity in twins. N Engl J Med. Aug 2 2007;357:454-461.

Russian Federation: Federal State Statistics Service. 2012. Accessed December 2017. http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/en/figures/population/.

Saghafi N, Khadem N, Mohajeri T, Shakeri MT. Efficacy of 17α-hydroxyprogesterone caproate in prevention of preterm delivery. J. Obstet. Gyn Res. 2011:1-4.

Soule L. FDA Meeting of the Advisory Committee for Reproductive Health Drugs. January 20, 2012.

Suvonnakote T. Prevention of pre-term labour with progesterone. J Med Assoc. Thailand. 1986;69(10):537-542.

Tita ATN and O'Day MP. Letter to the editor: Prophylactic progesterone to prevent preterm birth. Am J Obstet Gyn. 2004. pp 1799-1800.

UNICEF. 2017. 'Table 1. Basic Indicators'. https://data.unicef.org/wp-content/uploads/2017/12/SOWC-Statistical-Tables-2017.xlsx.

USAID. 2011. "Ukraine Maternal and Infant Health Project (MIMP-II) 2006-2011 Report.", 1-18.

WHO. Global Health Observatory (GHO) Data. 2018. http://apps.who.int/gho/data/view.main.CM1300N?lang=en. Accessed 27August2019.

Yemini M, Borenstein R, Dreazen E, et al. Prevention of premature labor by 17 alpha-hydroxyprogesterone caproate. Am J Obstet Gynecol. Mar 1 1985;151(5):574-577.