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.

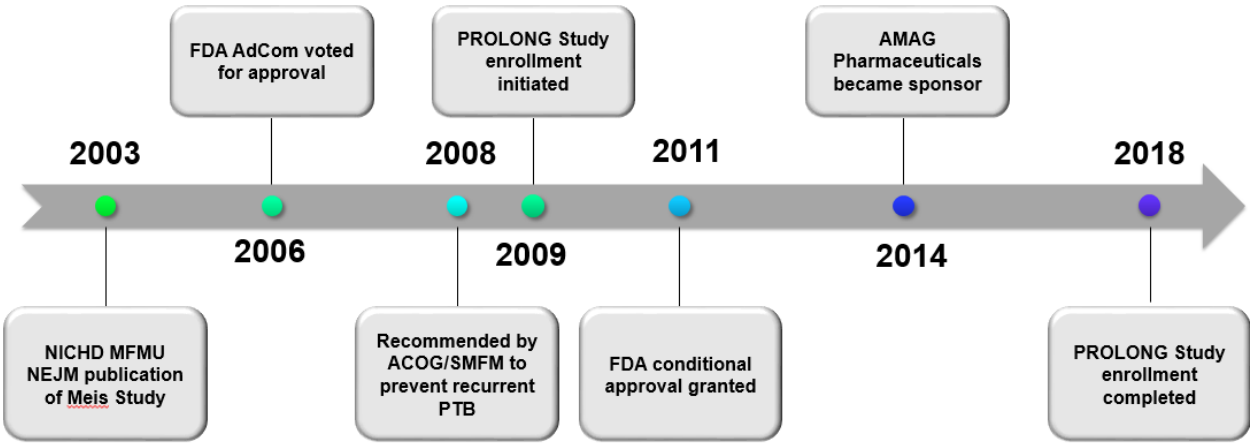
Progesterone agents have demonstrated effectiveness in the prevention PTB in randomized trials [Keirse 1990; Meis and Aleman 2004] which are thought to support gestation by reducing inflammation and inhibiting uterine activity. Hydroxyprogesterone caproate (HPC), or "17P", has demonstrated efficacy in 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

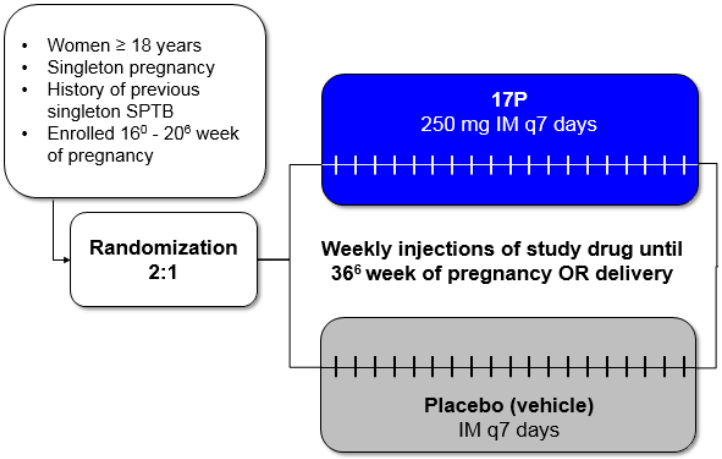
1.5. Meis: Pivotal Trial Results

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.

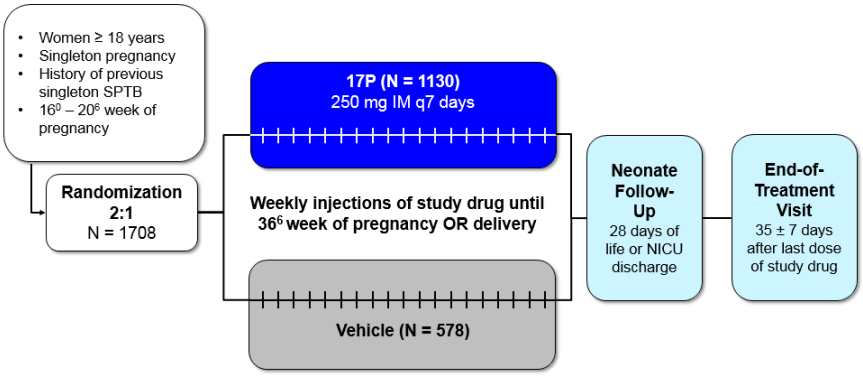
The incidence of pregnancy complications, such as preeclampsia, gestational diabetes, or clinical chorioamnionitis, as well as the incidence of serious adverse events (SAEs), was not different between the 17P and vehicle groups. SAEs reported were predominately miscarriages, stillbirths, and neonatal 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)

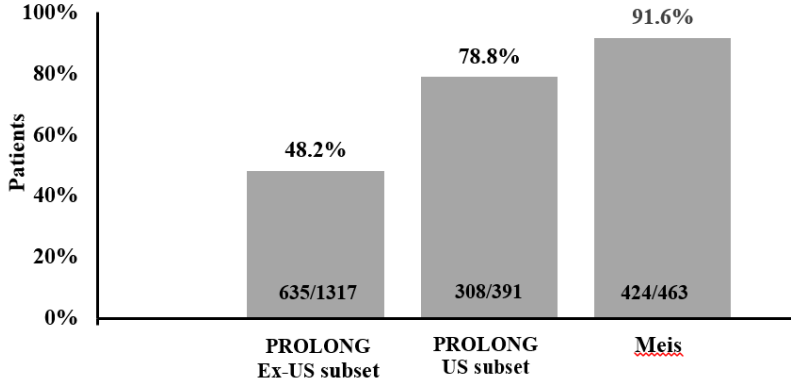
Source: PROLONG Ad Hoc Table 14.1.3.1.9

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.

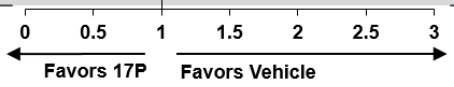
In Meis, which was a higher risk population, a treatment benefit favoring 17P was observed not only with the <37 weeks gestational age, but also at <35 weeks and even at <32 weeks, an important endpoint since it is known that babies born at earlier than 32 weeks have a significant risk of mortality and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

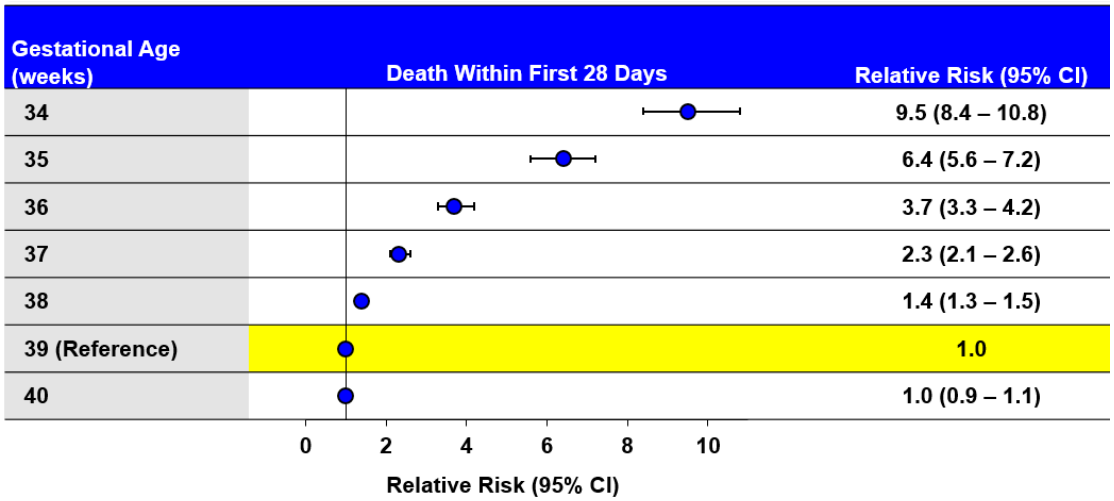
- extremely preterm (<28 weeks);
- very preterm (28 to <32 weeks);
- moderate or late preterm (32 to <37 completed weeks of gestation)

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.

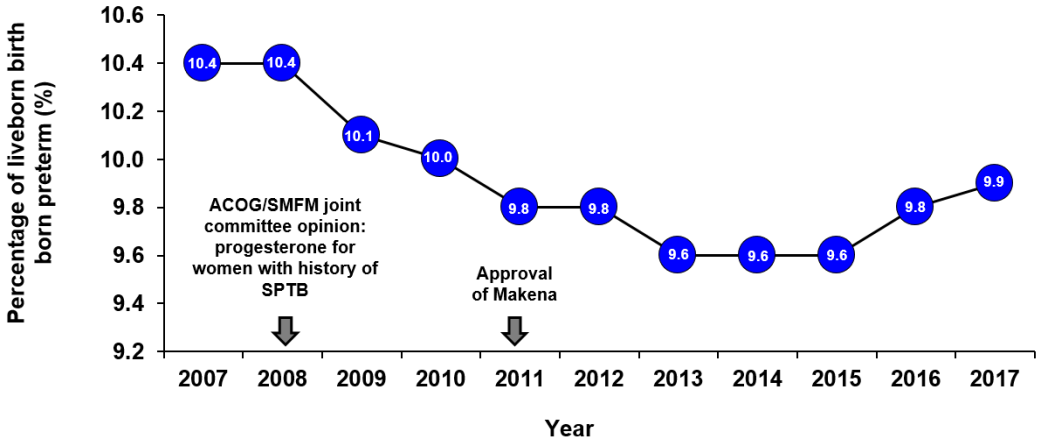
Infants who do survive premature birth often suffer long-term health problems and potential for long-term physical and cognitive disabilities. During the birth hospitalization, late preterm infants are at increased risk for morbidities such as respiratory distress, hypothermia, feeding difficulties, 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.

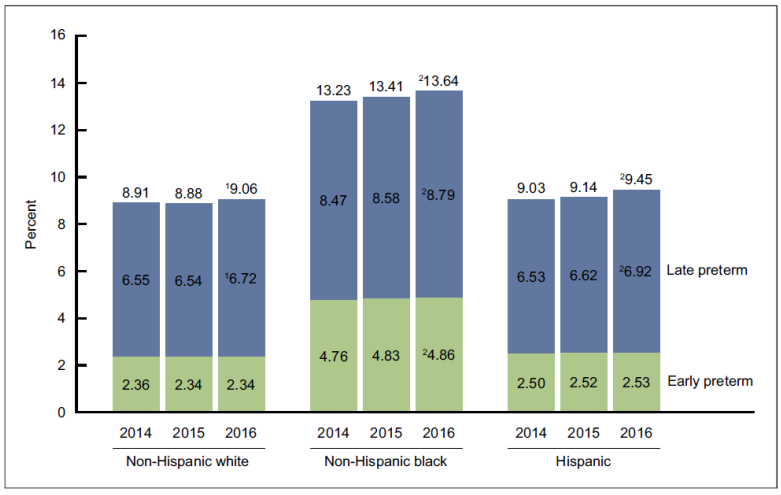
Figure 7: Preterm Birth Rates in United States (2007 through 2017)



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].

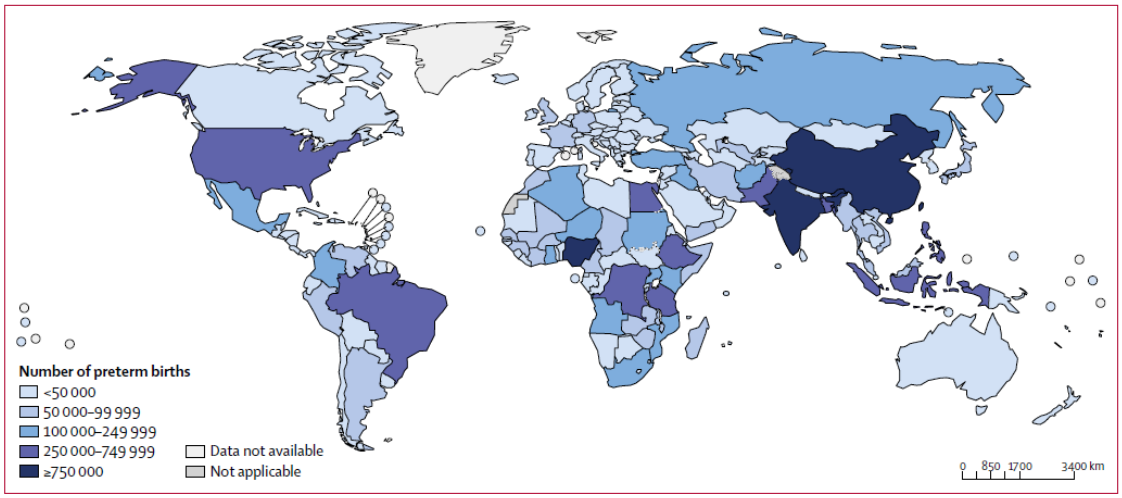
Figure 8: Preterm Birth Rates in the United States by Race and Ethnicity (2014 to 2016)



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].

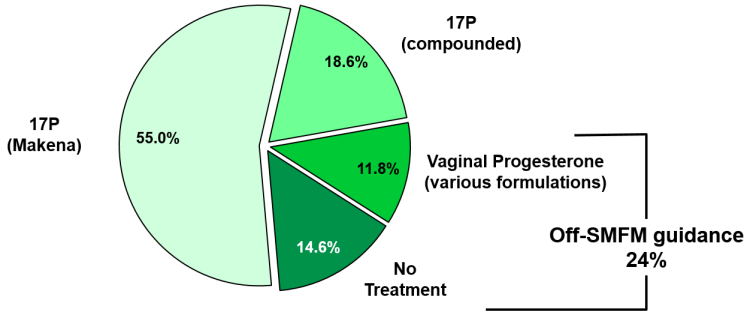
3.1.3. Treatment Guidelines

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.

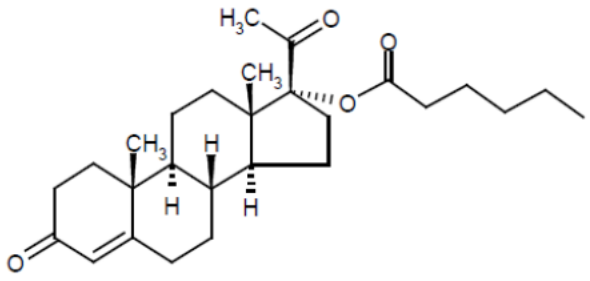
4.1. Makena (HPC Injection)

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.

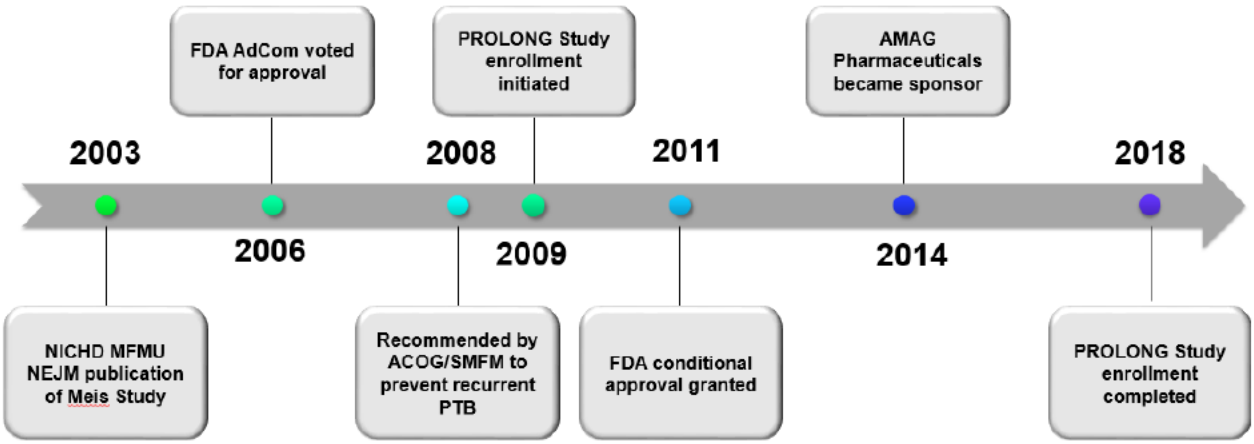
5. REGULATORY HISTORY

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

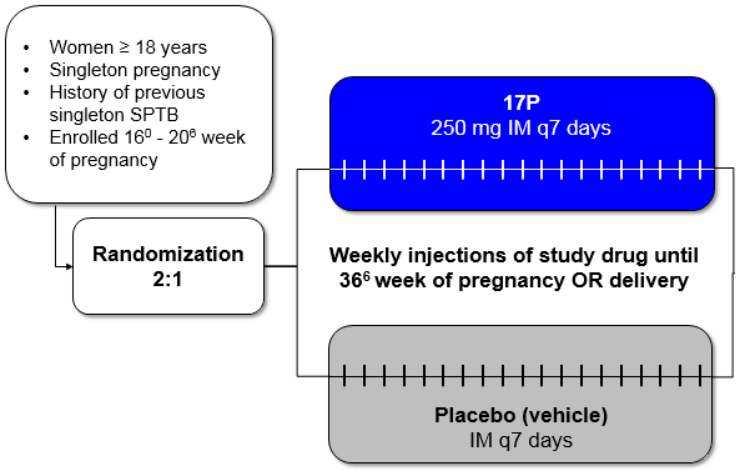
6.1.1. Study Design

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.

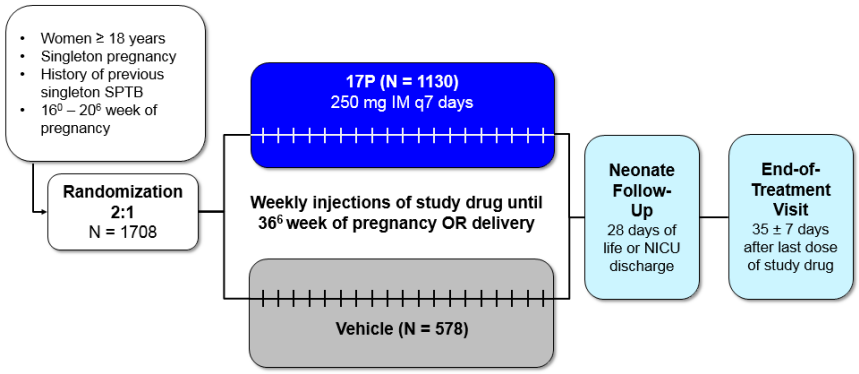
Each patient was randomized in a 2:1 ratio to receive either 17P (250 mg/mL) or vehicle, respectively. Patients received weekly injections of study drug from randomization (16^0 through 20^6 weeks of gestation) through 36^6 weeks of gestation or delivery, whichever occurred first. All injections were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

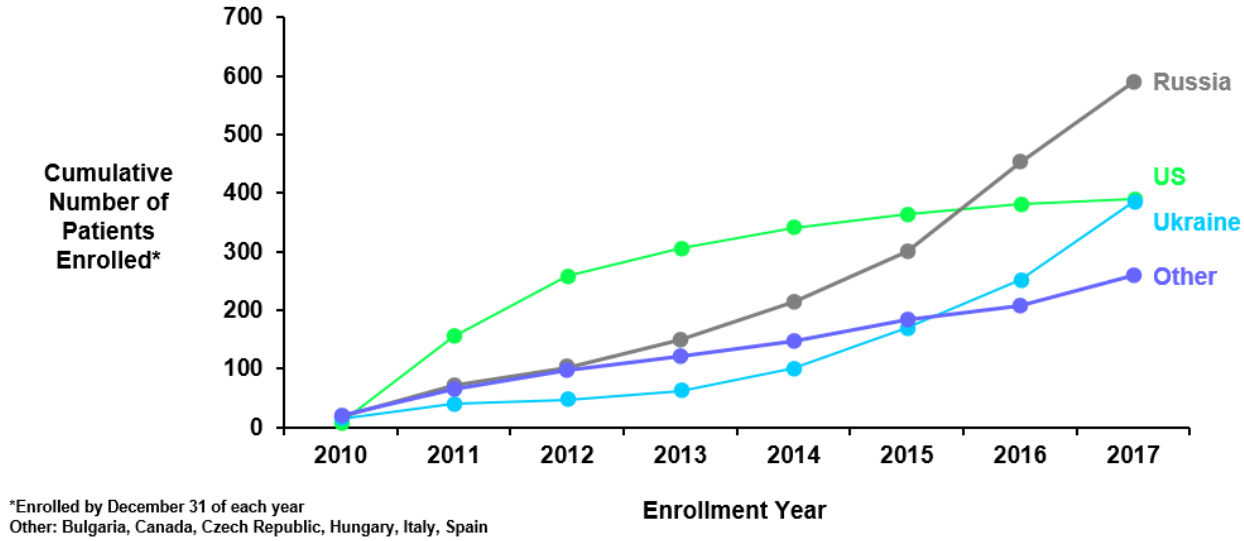
6.2.1.4. Calculation of Gestational Age

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)

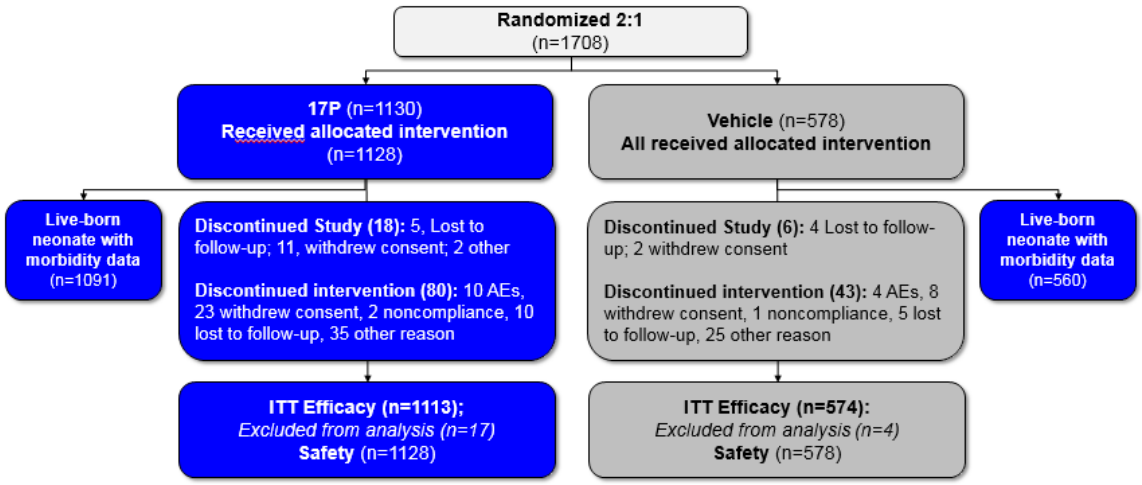
Country	Sites (n)	Patients Receiving Trial Injection (n)	Patients Randomized (n)	Randomized to 17P (n)	Randomized to Vehicle (n)
Overall	**93**	**1740**	**1708**	**1130**	**578**
Russia	12	628	621	414	207
Ukraine	10	424	420	277	143
United States	41	407	391	258	133
Hungary	5	91	91	59	32
Spain	8	85	85	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)

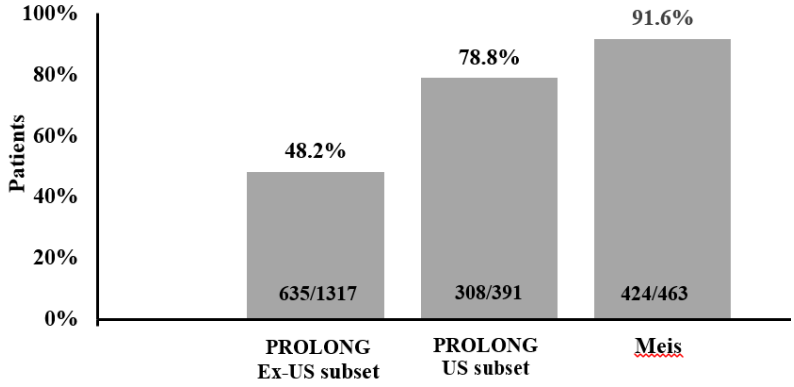
Source: PROLONG Ad Hoc Table 14.1.3.1.10 and Ad Hoc Table 14.1.3.1.11.
[a] Hispanic or Latino included in both race and ethnicity category.
[b] Study 002/PROLONG preterm delivery tables differ. PROLONG % PTB deliveries calculated manually.
NC=not collected.

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.

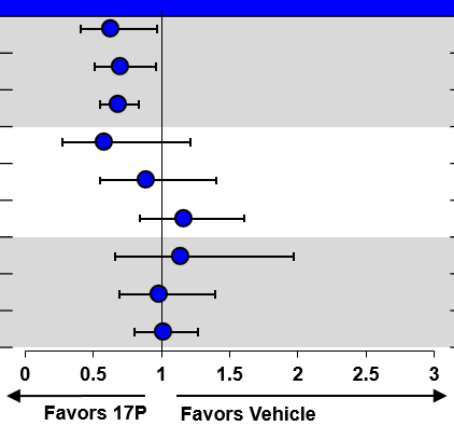
In addition, the intermediate risk population from the US subset of PROLONG also shows trends of a treatment effect favoring 17P beginning to emerge, as this population becomes more similar

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.

FDA Briefing Document
NDA 021945
Hydroxyprogesterone Caproate Injection
(trade name Makena)

Bone, Reproductive, and Urologic Drugs Advisory Committee (BRUDAC) Meeting
October 29, 2019
Division of Bone, Reproductive, and Urologic Products
Office of Drug Evaluation III
Office of New Drugs

Division of Biometrics III
Division of Biometrics VII
Office of Biostatistics
Office of Translational Sciences

Division of Epidemiology II
Office of Surveillance and Epidemiology

Center for Drug Evaluation and Research

DISCLAIMER STATEMENT

The attached package contains background information prepared by the Food and Drug Administration (FDA) for the panel members of the advisory committee. The FDA background package often contains assessments and/or conclusions and recommendations written by individual FDA reviewers. Such conclusions and recommendations do not necessarily represent the final position of the individual reviewers, nor do they necessarily represent the final position of the Review Division or Office. We have brought new information from the new drug application for Makena (17-hydroxyprogesterone caproate) to this Advisory Committee in order to gain the Committee's insights and opinions, and the background package may not include all issues relevant to the final regulatory recommendation and instead is intended to focus on issues identified by the Agency for discussion by the advisory committee. The FDA will not issue a final determination on the issues at hand until input from the advisory committee process has been considered and all reviews have been finalized. The final determination may be affected by issues not discussed at the advisory committee meeting.

Table of Contents

Table of Tables

Table of Figures

INTRODUCTORY MEMORANDUM

To: Bone, Reproductive and Urologic Drugs Advisory Committee

From: Christine P. Nguyen, MD
Deputy Director for Safety

Hylton V. Joffe, MD, MMSc
Director

Division of Bone, Reproductive, and Urologic Products (DBRUP)

Subject: Makena (hydroxyprogesterone caproate injection)
New Drug Application 021945/Supplement 023
Overview of topics to be discussed at the October 29, 2019, advisory committee meeting

---

The FDA is convening this Advisory Committee (AC) meeting to discuss the evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth and improving neonatal outcomes to inform FDA's regulatory decision-making for this product. In 2011, Makena received accelerated approval (a type of approval discussed in greater detail below) based on a reduced risk of recurrent preterm birth (PTB) prior to 37 weeks, a surrogate endpoint that FDA considered reasonably likely to predict clinical benefit to the neonate. Consistent with FDA's accelerated approval framework [21 CFR part 314, subpart H and section 506(c) of the Federal Food, Drug, and Cosmetic Act (FD&C Act)], FDA required the Applicant to conduct a post-approval confirmatory trial to verify and describe the clinical benefit. Completed at the end of 2018, this confirmatory trial did not verify Makena's efficacy on obstetrical or neonatal outcomes. In a supplemental new drug application (sNDA), the Applicant proposes to add findings from this trial to the drug label.

BACKGROUND:

Current clinical practice

Preterm birth, defined as birth prior to 37 weeks of gestation, currently affects approximately 10% of all births and 8% of singleton pregnancies.[1] Premature birth is a significant public health problem because these infants are at an increased risk of neonatal mortality and significant morbidity, as well as long-term physical and developmental impairment. To date, there are no drugs approved for reducing neonatal morbidity or mortality or long-term sequelae of preterm birth.

Progesterone, administered by intramuscular injection or intravaginally, has been used for certain conditions that may increase a pregnant woman's risk of PTB. Current professional practice

[1] https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm (accessed September 19, 2019)

guidelines recommend progesterone treatment starting in the second trimester of pregnancy to reduce the risk of recurrent preterm birth in women with a singleton pregnancy and a prior spontaneous preterm birth (sPTB). The guidelines also recommend vaginal progesterone to reduce the risk of PTB in women without a prior preterm birth and with a shortened cervix in the current pregnancy, although such use is not FDA-approved.[2] Makena is the only pharmacotherapy approved to reduce the risk of recurrent preterm birth. Based on its accelerated approval, Makena's indication states that it is approved to "reduce the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. 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."

Regulatory History of Hydroxyprogesterone Caproate:
The drug substance of Makena, hydroxyprogesterone caproate (HPC), also referred to as 17-HPC, 17-OHPC, or 17P, was approved by FDA in 1956 for conditions generally responding to progestogens, under the tradename Delalutin (HPC) injection 125 mg/mL and 250 mg/ml (NDAs 010347, 016911). This approval was based on safety considerations because it occurred prior to the Kefauver-Harris Amendment of 1962 to the FD&C Act requiring that approved drugs be supported by substantial evidence of effectiveness, in addition to demonstrated safety. Delalutin remained approved for certain gynecologic indications after undergoing the Drug Efficacy Study Implementation review, which determined the efficacy of marketed drugs approved before 1962. At the Applicant's request, FDA withdrew approval of the NDAs for Delalutin in 2000 (not for efficacy or safety reasons) (65 Fed. Reg. 55264, Sept. 13, 2000). FDA has approved generic products of Delalutin that are currently marketed. Note that Delalutin and its generics are not approved for reducing the risk of preterm birth.

Published literature from the 1960s through the 1980s included several clinical studies evaluating the efficacy of HPC for obstetrical uses. Conflicting findings regarding the effectiveness of HPC for the prevention of PTB prompted the National Institute for Child Health and Human Development (NICHD), via the Maternal-Fetal Medicine Units (MFMU) Network, to conduct a multicenter, double-blind, placebo-controlled clinical trial in women with a history of spontaneous preterm singleton birth to assess the efficacy of HPC for preventing recurrent PTB (Study 17P-CT-002, or Trial 002 hereinafter). In June 2003, the trial's findings were published,[3] reporting that HPC 250 mg injection reduced the proportion of women who delivered at less than 37 weeks gestation.

[2] American College of Obstetricians and Gynecologists (ACOG) Practice Bulletin: Prediction and Prevention of Preterm Birth (2012, reaffirmed 2018); Society for Maternal-Fetal Medicine Statement: "The choice of progestogen for the prevention of preterm birth in women with singleton pregnancy and prior preterm birth" (March 2017). While the ACOG Practice Bulletin did not specify the formulation of progesterone for women with a prior sPTB, SMFM recommended treatment with hydroxyprogesterone caproate injection and not vaginal progesterone in this population.

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

Makena's accelerated approval

In 2006, an applicant submitted NDA 021945 seeking marketing approval of HPC injection for the prevention of recurrent PTB. The NDA relied on data from the MFMU Network Trial 002 for primary support of efficacy and safety. At that time, no drug was approved in the U.S. to reduce the risk of PTB. However, HPC was compounded and used widely for the prevention of PTB in women at high risk.

After three review cycles and one Advisory Committee meeting, in February 2011, the FDA granted Makena accelerated approval based on reduction in preterm birth prior to 37 weeks, a surrogate endpoint considered to be reasonably likely to predict the clinical benefit of reducing neonatal morbidity or mortality.

Initiated in 1999 and completed in 2002, Trial 002 enrolled 463 women with a singleton pregnancy and at least one prior sPTB from 19 university-based clinical centers in the United States in the MFMU Network. The primary efficacy endpoint was the proportion of pregnant women delivering prior to 37 weeks gestation, with those delivering prior to 35 or 32 weeks as secondary endpoints. The trial showed that Makena (HPC 250 mg) injection administered intramuscularly once weekly starting at 16 weeks 0 days (16^0) to 20 weeks 6 days (20^6) gestation and used through 36^6 weeks gestation or birth reduced the proportion of women who delivered <37 weeks gestation from 55% (placebo) to 37% (Makena). The treatment difference was -17.8% [95% confidence interval (CI): -28%, -7.4%]. This treatment benefit appeared independent of race, number of prior preterm deliveries, and gestational age of the prior preterm birth. The treatment effect was sufficiently persuasive to support drug approval based on the findings of a single adequate and well-controlled trial. The proportions of women delivering at <35 and <32 weeks gestation were also statistically lower among women treated with Makena compared to placebo. The treatment difference was -9.4% (95% CI: -19.0%, -0.4%) for delivery <35 weeks gestation and -7.7% (95% CI: -16.1%, -0.3%) for delivery <32 weeks gestation.

Issues regarding generalizability of Trial 002's findings to the broader U.S. population included (a) approximately 60% of the trial participants being self-identified Blacks, (b) subject recruitment from only academic centers, with 25% of subjects from a single academic center, and (c) the notably high rate of recurrent preterm birth in the placebo arm (55%).[4] As a condition of accelerated approval, the Applicant was required to submit data from a confirmatory efficacy and safety trial to verify the clinical benefits of Makena, and the trial was to be completed with due diligence.

CONFIRMATORY TRIAL (Trial 003)

Prior to approving Makena in 2011, the FDA recognized the challenges of the feasibility of conducting a confirmatory efficacy and safety trial in the United States, given the endorsement of professional practice guidelines and accepted clinical practice of using progesterone for preterm birth. Prior to approval, the FDA required that the Applicant provide evidence that it could successfully complete the confirmatory trial, which must be ongoing at the time of approval, and that at least 10% of subjects be enrolled from the U.S. and Canada. Initiated in 2009 and completed in 2018, this confirmatory trial (Trial 003) was a multicenter, international,

[4] Background recurrent preterm birth rate used to power Trial 002 was 36%, as this was the background rate from the MFMUN uterine monitoring trial in the 1990s.

randomized, double-blind, placebo-controlled study that enrolled women with eligibility criteria like those of Trial 002. The trial's coprimary efficacy endpoints were delivery prior to 35 weeks gestation and a neonatal morbidity/mortality composite index (neonatal composite index).[5] The inclusion of a clinical endpoint (the neonatal composite index) addressed the accelerated approval's regulations of verifying that initial findings based on a surrogate endpoint (gestational age at delivery) lead to direct clinical benefit. Trial 003 randomized a total of 1,708 women from nine countries, with Russia, Ukraine, and the United States enrolling 36%, 25%, and 23% of women, respectively. Data were available for 1651 liveborn neonates. The trial did not demonstrate a statistically significant treatment effect for the coprimary endpoints of proportion of women delivering prior to 35 weeks (11% Makena compared to 12% placebo, p=0.72) or neonatal composite index (5.4% Makena compared to 5.2% placebo, p=0.84). Also, no differences between Makena and placebo were seen in the secondary outcomes related to other gestational ages at delivery (<37 weeks [23% Makena vs. 22% placebo, p=0.57), <32 weeks gestation [4.8% Makena vs. 5.2% placebo, p=0.70]) or for the individual components of the neonatal index.

The Applicant raised concerns that the study populations of Trial 002 (U.S only) and Trial 003 (international, including U.S.) differed substantially and that this may have contributed to the discordant outcomes between the two trials. Therefore, exploratory subgroup analyses and comparisons of Trial 003's U.S. population (003-U.S. subgroup) and non-U.S. patients were undertaken. There were no relevant differences in the treatment effect when analyzed by region (U.S. vs. non-U.S.), even though the non-U.S. subgroup appeared to have a lower risk profile based on demographics, social, and behavioral factors compared to the U.S. subgroup. There was no evidence of interaction between treatment and U.S. vs. non-U.S. region for the coprimary endpoints. In the 003-U.S. subgroup:

- Makena did not improve the neonatal composite index. The treatment effect was -2.2% (95% CI: -8.3, 3.9) when analyzed using the stratified Cochran-Mantel-Haenszel (CMH) method and -0.2% (95% CI: -4.9, 2.8) using another approach known as shrinkage analysis.
- Makena did not reduce the risk of delivery <35 weeks (16% Makena vs. 18% placebo). The treatment difference was -2.2% (95% CI: -10.1, 5.7) using the stratified CMH analytical method; this difference was -0.8% (95% CI: -6.0, 3.5) with shrinkage estimation.
- Point estimates of the proportions of women with delivery occurring <37 weeks (33% Makena vs. 28% placebo, a treatment effect of 4.7% [95% CI: -5%, 14%] by the CMH method) or <32 weeks (5.5% Makena vs. 9.2% placebo, a treatment effect of -3.9% [95% CI: -9.6, 1.7] by the CMH method) showed contradictory trends in the treatment effect.

A comparison among Trial 003 overall, the 003-U.S. subgroup, and Trial 002 populations indicated that a greater proportion of subjects in Trial 002 had certain risk factors for PTB, such as being self-identified Blacks or having > 1 prior sPTB, than the 003-U.S. subgroup or Trial 003 overall. However, exploratory subgroup analyses did not show statistically significant

[5] The neonatal morbidity/mortality composite index includes neonatal death, Grade 3 or 4 intraventricular hemorrhage, respiratory distress syndrome, bronchopulmonary dysplasia, necrotizing enterocolitis, and proven sepsis.

interactions between these risk factors and treatment effect of Makena in Trial 002 or Trial 003. Although these risk factors may have an impact on the PTB rate, there was no evidence in Trial 003 that they impact the treatment effect nor was there consistent convincing evidence of treatment benefit within a specific subpopulation across the two trials.

Published literature on progesterone's effect on preterm birth in women with a prior sPTB
Because findings from Trial 003 were discordant with those of Trial 002, we evaluated published evidence from six randomized, placebo-controlled trials that assessed the effect of progesterone in preterm birth and that included pregnant women with a prior sPTB. These trials studied vaginal progesterone at different doses (90 – 200 mg) in women with various risks for PTB, including a history of sPTB, with different gestational ages at delivery as the primary outcome. The overall evidence based on subgroup analyses in pregnant women with a prior sPTB did not suggest a treatment benefit with progesterone over placebo in reducing the risk of recurrent PTB in these women. These trials and their findings, however, are not directly applicable to Makena; none evaluated injectable HPC in the same target population measuring the same efficacy endpoints as Makena. We also reviewed two recent large meta-analyses. These meta-analyses evaluated progesterone formulations, doses, patient populations, and endpoints dissimilar to those of the trials for Makena and did not reliably inform the treatment effect of Makena for its intended use.

Accelerated approval and evidentiary standards for drug approval
When appropriate, the accelerated approval pathway allows for earlier approval of a drug to treat a serious condition and fill an unmet medical need based on a surrogate endpoint that is reasonably likely to predict clinical benefit but is not itself a direct measure of clinical benefit. The Applicant is required to conduct trial(s) after receiving accelerated approval to confirm the expected clinical benefit. If the confirmatory trial(s) shows that the drug provides clinical benefit, then the conditions initially attached to accelerated approval are generally terminated. (See 21 CFR 314.560.) If the confirmatory trial(s) fail to demonstrate such benefit, FDA may withdraw approval of the drug in accordance with section 506(c)(3) of the FD&C Act and 21 CFR 314.530. With accelerated approval, there is less certainty at the time of approval that the drug will ultimately be shown to improve how patients feel, function or survive; however, this pathway provides earlier patient access than would otherwise be possible to an approved drug that is reasonably likely to confer clinical benefit for a serious condition with an unmet need. In the case of Makena, FDA granted accelerated approval based on the reduction in preterm birth seen in Trial 002; however, confirmatory Trial 003 did not verify clinical benefit on adverse neonatal outcomes to infants born prematurely.

For FDA approval, including accelerated approval, the drug must meet the regulatory standard of "substantial evidence" of effectiveness and the benefits must outweigh the risks. Generally, FDA interprets substantial evidence of effectiveness as evidence of effectiveness from two or more adequate and well-controlled trials. A single positive trial, even if well-designed and well-conducted, may have undetected systemic biases or may reflect a chance finding, increasing the risk of concluding that a drug is effective when in fact it is not. The requirement for at least two adequate and well-controlled trials ensures independent substantiation of experimental findings and strengthens a conclusion of effectiveness. Nonetheless, when appropriate, FDA has the authority and flexibility to conclude that there is substantial evidence of effectiveness based on a

single adequate and well-controlled trial. In the case of Makena, FDA determined that Trial 002 was adequate, well-controlled and very persuasive and concluded that this single trial provided substantial evidence of an effect on a surrogate endpoint (effectiveness for reduction in the risk of recurrent preterm birth). It is important to note, however, that at the time this determination was made in 2011, there were no other adequate and well-controlled trials with Makena, and that had there been such additional trial(s), FDA would have considered those data when deciding whether there was substantial evidence of effectiveness.

There are two important scientific and regulatory implications for Makena:

- Accelerated approval: A drug approved under the accelerated approval pathway based on a surrogate endpoint reasonably likely to predict clinical benefit must undergo a confirmatory trial postapproval to verify clinical benefit (i.e., an improvement in how patients feel, function or survive). In the case of Makena, confirmatory Trial 003 did not demonstrate a reduction in adverse neonatal outcomes from preterm birth; therefore, the clinical benefit of Makena remains unverified.

- Substantial evidence of effectiveness: Trial 003 also did not confirm an effect of Makena on gestational age of delivery, the surrogate endpoint used in Trial 002 to support accelerated approval. This raises the question as to whether Makena's accelerated approval is still supported by substantial evidence of effectiveness for the reduction in recurrent preterm birth.

AREAS OF FOCUS FOR ADVISORY COMMITTEE

Based on the above considerations, the key issues are whether there remains substantial evidence of effectiveness of Makena on preterm birth, the unconfirmed clinical benefit of Makena on neonatal outcomes, and implications for Makena's marketing status. Makena received accelerated approval based on findings from Trial 002, which showed a reduction in the proportion of women with preterm delivery <37 weeks compared to placebo, a surrogate endpoint considered reasonably likely to predict clinical benefit. However, Trial 003, an adequate and well-controlled, well-conducted and appropriately powered confirmatory trial, did not show a reduction in preterm birth with Makena compared to placebo, nor did it demonstrate a reduction in neonatal morbidity/mortality. Under accelerated approval regulations, FDA may withdraw the approval of Makena if the Applicant fails to provide confirmatory evidence of efficacy and safety. To place this discussion in the appropriate context, we ask that the Advisory Committee members consider:

- The applicability of the findings of Trial 003 to the U.S. population
- Factors, if any, that may account for the differences in outcomes between Trial 002 and Trial 003
- Whether there continues to be substantial evidence that Makena reduces the risk of recurrent preterm birth in the context of two adequate and well-controlled trials with discrepant efficacy findings on this surrogate endpoint
- If a new confirmatory trial is required, the design of such a trial, including the comparator arm, dose(s) of study medication, location (U.S./North America or international), efficacy endpoints and importantly, the feasibility and likelihood of successfully completing such a trial in a timely manner

- If Makena were to be withdrawn from the market because of lack of efficacy, the likely consequences and their potential impact on public health.

We look forward to a thorough and reasoned discussion of these complex, important matters. Thank you in advance for the vital public health contribution you are making through your participation in this meeting.

Draft Points to Consider:

1. Discuss the effectiveness of Makena, including:
    a. The effects of Makena on recurrent preterm birth in Trial 003, and your interpretation of the discrepant preterm birth results between Trial 002 and Trial 003;
    b. The effects of Makena on neonatal morbidity and mortality;
    c. Relevance of the findings in Trial 003 to the U.S. population and current clinical practice.

2. If a new efficacy trial were to be conducted, discuss the study design, including control, dose(s) of study medication, efficacy endpoints and the feasibility of completing such a trial.

3. Discuss the potential consequences of withdrawing Makena on patients and clinical practice.

4. Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal outcomes?

    Provide rationale for your vote.

5. Based on the findings from Trial 002 and Trial 003, is there substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth?

    Provide rationale for your vote.

6. FDA approval, including accelerated approval, of a drug requires substantial evidence of effectiveness, which is generally interpreted as clinically and statistically significant findings from two adequate and well-controlled trials, and sometimes from a single adequate and well-controlled trial. For drugs approved under the accelerated approval pathway based on a surrogate endpoint, the Applicant is required to conduct adequate and well-controlled postapproval trial(s) to verify clinical benefit. If the Applicant fails to conduct such postapproval trial(s) or if such trial(s) do not verify clinical benefit, FDA may, following an opportunity for a hearing, withdraw approval.

    Should FDA:

    A. Pursue withdrawal of approval for Makena
    B. Leave Makena on the market under accelerated approval and require a new confirmatory trial
    C. Leave Makena on the market without requiring a new confirmatory trial

    Provide rationale for your vote and discuss the following:

    - Vote (A) (withdraw approval) may be appropriate if you believe the totality of evidence does not support Makena's effectiveness for its intended use.

- Discuss the consequences of Makena removal (if not previously discussed in Discussion point 3)

- Vote (B) (require a new confirmatory trial) may be appropriate if you believe the totality of evidence supports Makena's effectiveness in reducing the risk of recurrent preterm birth, but that there is no substantial evidence of effectiveness on neonatal outcomes. Vote (B) would also reflect a belief that a new confirmatory trial is necessary and feasible.

    - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth, based on the surrogate endpoint of gestational age at delivery.

    - Also discuss key study elements, including study population, control, dose(s), and efficacy endpoints of the new confirmatory trial (if not previously discussed in Discussion point 2) and approaches to ensure successful completion of such a trial.

- Vote (C) (leave Makena on the market without a new confirmatory trial) may be appropriate if you believe Makena is effective for reducing the risk of recurrent preterm birth and that it is not necessary to verify Makena's clinical benefit in neonates.

    - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth and why it is not necessary to verify Makena's clinical benefits in neonates.

1. Background

1.1. The Condition and Treatment Options

1.1.1. Preterm Birth

Preterm birth (PTB), defined as delivery between 20 and 37 completed weeks of gestation, is a significant public health concern. Preterm birth may be spontaneous (birth following a spontaneous process, such as preterm labor or preterm premature rupture of membranes) or indicated (delivery initiated by the healthcare provider for maternal or fetal health). According to the Centers for Disease Control and Prevention, in 2017, the U.S. PTB rate was 9.9% overall and 8.1% in singleton pregnancies; the incidence was highest in black women (13.9%) compared to white or Hispanic women (9.1% and 9.6%, respectively).[6] The CDC reported that the rate of preterm birth in the U.S. declined from 2007 (10.4%) to 2014 (9.6%), mostly because of a decline in teenage pregnancy, but has increased from 2014 until 2017 (9.9%). The latter trend is mostly due to an increase in the rate of late preterm birth (delivery 34-36 weeks gestation), while rates for early preterm birth (less 34 weeks) have remained unchanged from 2015. The World Health Organization estimates the global PTB rate to be 10.6%, which is similar to the rate of 11.2% in North America, but there are differences across geographic regions, ranging from 8.7% in Europe to 13.4% in North Africa.[7] In 2015, PTB accounted for 17% of infant deaths [8] and surviving children often suffer developmental delay or long-term neurologic impairment. In 2016, complications of PTB were the leading cause of death globally in children younger than 5 years of age, accounting for approximately 16% of all deaths in this age group, and 35% of deaths among neonates.[9] In general, the risk of adverse outcomes in the preterm neonate decreases with increasing gestational age at delivery.

While the burden of PTB is clear, the causes of PTB are less so, and identifying women who will give birth preterm is challenging. Spontaneous PTB represents a syndrome and its causes are multifactorial. Risk factors for PTB include uterine distension (seen in multifetal pregnancies and polyhydramnios), dysfunction of the cervix (reduced mechanical competence, either resulting from genetic mutations in components of collagen that is required for integrity of the cervix or from repeated surgeries on the cervix), infection of the lower genital tract, and other factors (such as cigarette smoking, inadequate maternal weight, and illicit drug use). The contribution of these factors to PTB, however, is not well-characterized. However, an accepted major risk factor is short cervical length (typically defined as <25 mm observed prior to 24 weeks gestation). Regarding the risk of recurrent PTB, one of the strongest risk factors is a history of a preterm birth, which increases the risk of PTB by about 1.5 to 2-fold. Additionally, the number of prior PTBs and the gestational age of the prior PTBs impact the recurrence risk.

---

[6] National Vital Statistics Reports, Vol 67, No. 8, November 7, 2018. https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_08-508.pdf

[7] Chawanpaiboon S, Vogel JP, Moller A-B, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systemic review and modelling analysis. Lancet Glob Health 2019;7(1): e37-46.

[8] CDC – Division of Reproductive Health, National center for Chronic Disease Prevention and Health Promotion. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm

[9] UN Inter-Agency Group for Child Mortality Estimation. Levels and trends in child mortality: Report 2017. New York: United Nations Children's Fund, 2017.

Nonetheless, two-thirds of PTBs occur among women with no identifiable risk factors, causality of PTB has been difficult to determine, and the pathogenesis remains poorly understood.[10]

1.1.2. Treatment to Reduce the Risk of Recurrent Preterm Birth

In January 2003, Trial 002 was presented by the NICHD as the first abstract at the Society for Maternal-Fetal Medicine Meeting. The positive findings from this trial immediately gained extensive media attention, leading to the wide use of compounded HPC to reduce the risk of recurrent PTB. Following the June 2003 publication of Trial 002 in the New England Journal of Medicine, the American College of Obstetricians and Gynecologists (ACOG) Committee on Obstetric Practice endorsed the use of progesterone only in women with a documented history of a previous spontaneous birth at less than 37 weeks of gestation. In its most recent Practice Bulletin (published 2012, reaffirmed 2018), ACOG recommends progesterone (without specifying the formulation of progesterone) starting in the second trimester in women with a singleton pregnancy and a prior sPTB. ACOG also recommends vaginal progesterone in women with a singleton pregnancy with a shortened cervix and without a prior sPTB. In 2003, the Society for Maternal-Fetal Medicine (SMFM) recommended treatment with either HPC injection or vaginal progesterone for women with a prior spontaneous PTB to prevent the recurrence of PTB; this recommendation was reaffirmed in 2008.[11] Based on published findings of several clinical trials, the SMFM in 2012 revised the guideline to recommend that HPC 250 mg IM weekly be given, starting at 16 to 20 weeks of gestation until 36 weeks or birth, to women with a singleton gestation whose prior sPTB occurred between 20-$36^{6/7}$ weeks gestation.[12] In 2017, SMFM reaffirmed its 2012 recommendation and added that vaginal progesterone should not be considered a substitute for HPC in these patients.[13] As noted previously, Makena is the only FDA-approved treatment for PTB.

1.2. Regulatory Background

1.2.1. Regulatory Standards of Drug Approval

1.2.1.1. Accelerated Approval

Under the accelerated approval pathway [21 CFR part 314, subpart H, and 506(c) of the FD&C Act], FDA may grant marketing approval for a new drug based on adequate and well-controlled clinical trials establishing that the drug has an effect on a surrogate endpoint that is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict

[10] PRETERM BIRTH CAUSES, CONSEQUENCES, AND PREVENTION. Committee on Understanding Premature Birth and Assuring Healthy Outcomes. Board on Health Sciences Policy. Richard E. Behrman and Adrienne Stith Butler, Editors. INSTITUTE OF MEDICINE OF THE ACADEMIES. THE NATIONAL ACADEMIES PRESS. Washington, D.C. Copyright 2007 by the National Academy of Sciences.

[11] Society for Maternal-Fetal Medicine Publications Committee: Use of progesterone to reduce preterm birth. ACOG Committee opinion number 419, October 2008 (replaces no. 291, November 2003) Obstet Gynecol, 112 (2008), pp. 963-965.

[12] Society for Maternal-Fetal Medicine Publications Committee, with assistance of Vincenzo Berghella. Progesterone and preterm birth prevention: translating clinical trials data into clinical practice. Am J Obstet Gynecol, 206 (2012), pp. 376-386.

[13] The choice of progestogen for the prevention of preterm birth in women with singleton pregnancy and prior preterm birth Society for Maternal-Fetal Medicine (SMFM) Publications Committee, 2017

clinical benefit. A measurement of clinical benefit directly assesses how a patient feels, functions, or survives. Because gestational age at delivery does not directly measure how a neonate feels, functions, or survives, it is considered a surrogate endpoint, but one that we determined to be a reasonably reliable predictor of the clinical benefit for the neonate. In general, two major concerns with surrogate endpoints are (1) it may not be a true predictor of the clinical benefit and (2) it may not provide a quantitative measure of benefit. Thus, approval under this regulation requires that the Applicant study the drug further to verify and describe its clinical benefit. The confirmatory trials must be adequate and well-controlled and be conducted with due diligence. These trials are usually already ongoing at the time of accelerated approval to ensure their timely completion.

For drugs approved under the accelerated approval pathway, the regulations also outline the conditions that may prompt FDA to withdraw approval:

(1) A postmarketing clinical study fails to verify clinical benefit;
(2) The Applicant fails to perform the required postmarketing study with due diligence;
(3) Use after marketing demonstrates that postmarketing restrictions are inadequate to assure safe use of the drug product;
(4) The Applicant fails to adhere to the postmarketing restrictions agreed upon;
(5) The promotional materials are false or misleading; or
(6) Other evidence demonstrates that the drug product is not shown to be safe or effective under its conditions of use.
(See 21 CFR 314.530)

1.2.1.2. Substantial Evidence of Effectiveness

For FDA approval, including accelerated approval, a drug must meet the regulatory standard of "substantial evidence" of effectiveness for the intended use and the benefits must outweigh the risks.[14] Traditionally, FDA has interpreted substantial evidence of effectiveness as clinically and statistically significant findings from at least two adequate and well-controlled trials. A single positive trial, even if well-conducted, may have biases or may reflect a chance finding, increasing the risk of concluding that a drug is effective when in fact it is not. The requirement for at least two adequate and well-controlled trials ensures independent substantiation of experimental findings and strengthens a conclusion of effectiveness. Nonetheless, when appropriate, FDA has the authority and flexibility to conclude that there is substantial evidence of effectiveness based on a single adequate and well-controlled trial. Conclusions based on two high-quality trials will generally be more secure than those based on a single comparably persuasive study. Therefore, reliance on a single trial is generally limited to situations where a second trial is not feasible (e.g., rare diseases) or ethical (e.g., when one trial has demonstrated a clinically meaningful effect on mortality, irreversible morbidity, or prevention of a serious disease). Characteristics of a single trial that could support a conclusion of substantial evidence of effectiveness include a large multicenter trial with consistency across study subsets, multiple studies within a single study, multiple endpoints involving different events, and statistically very persuasive findings.

[14] FDA Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products, May 1998.

1.3. Trial 002 and Approval of Makena

1.3.1. Trial 002

In 1999, the National Institute of Child Health and Human Development initiated a multicenter, double-blind, randomized, placebo-controlled clinical trial through its Maternal-Fetal Medicine Units Network to evaluate the efficacy and safety of HPC injection. The study randomized pregnant women with at least one documented prior sPTB of a singleton fetus to either HPC or placebo in a 2:1 ratio. Eligible subjects were at a gestational age between 16^0 weeks and 20^6 weeks at randomization. Pregnancies with multifetal gestation and known major fetal anomaly (as documented by an ultrasound examination after 14 weeks gestation) were excluded. Women who had progesterone treatment prior to randomization were also excluded, as were women experiencing maternal medical complications (e.g., hypertension requiring medication, seizure disorder) or obstetrical complications. The subjects received HPC 250 mg weekly injections or placebo vehicle beginning on the day of randomization through 36^6 weeks gestation or delivery, whichever occurred first. The primary efficacy endpoint was the proportion of delivery prior to 37^0 weeks gestation in the intent-to-treat (ITT) population.

A total of 463 women were randomized to receive either HPC (N=310) or placebo (N= 153). The two study groups were similar with respect to age, race or ethnicity, body mass index prior to pregnancy, marital status, education, and substance use during pregnancy; 59% of the subjects were African American. Of the 463 women randomized, 418 (90.3%) completed dosing through 36^6 weeks or birth, including 279 (90.0%) in the HPC group and 139 (90.8%) in the placebo group. The efficacy results for gestational age at delivery are shown in Table 1.

Table 1: Proportion of Subjects in Each Treatment Arm Who Delivered at <37 Weeks, <35 Weeks, and <32 Weeks Gestational Age (Trial 002)

Delivery outcome	HPC* %	Placebo %	Treatment Difference and 95% Confidence Interval**
<37 weeks	37.1	54.9	-17.8% [-28.0%. -7.4%]
<35 weeks	21.3	30.7	-9.4% [-19.0%, -0.4%]
<32 weeks	11.9	19.6	-7.7% [-16.1%, -0.3%]

*Four HPC-treated subjects were lost to follow-up. They were counted as deliveries at their gestational ages at time of last contact (184, 220, 343, and 364 weeks).
**Adjusted for interim analysis.
Source: FDA-approved Makena prescribing information

Pregnancy after the time of randomization was maintained for an average of six days longer in the HPC group (131 vs. 125 days), with the mean gestational age at delivery being one week greater (36.2 vs. 35.2 weeks for HPC and placebo subjects, respectively).

Makena's effect on reducing recurrent preterm birth appeared independent of race, number of previous preterm deliveries, and gestational age of previous preterm birth. The proportion of women who delivered at <37 weeks in the placebo group appeared notably high (55%). See Table 2.

Table 2: Percentages of Subjects With Delivery <37 Weeks by Gestational Age of Previous Birth, Race, and Number of Previous Preterm Deliveries (Trial 002)

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

Data based on ITT Population (all randomized subjects). The 4 subjects with missing outcome data were classified as having a preterm birth <37^0 weeks (i.e., treatment failure).
Abbreviations: n = number of subjects in a specific category who delivered study pregnancy at <37^0 weeks gestation; N = total number of subjects overall in a specific category
Source: Table 11-4, Final Report for Study 17-CT-002

This trial was terminated by the Data and Safety Monitoring Board prior to enrolling the planned 500 subjects because the pre-specified stopping criteria for the primary efficacy endpoint of delivery < 37 weeks gestation were attained at an interim analysis.

Data on the individual components that subsequently constituted the neonatal composite index were prospectively collected. The analysis of a composite index, developed by the Applicant at the request of the FDA, was conducted post-hoc, after the initial submission of the NDA in 2006, to evaluate adverse outcomes in live births and as supportive evidence of Makena's benefit on reducing the risk of recurrent preterm delivery. The neonatal composite index was based on the number of neonates who died or experienced respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), grade 3 or 4 intraventricular hemorrhage (IVH), proven sepsis, or necrotizing enterocolitis (NEC). Although the proportion of neonates who experienced one or more events was numerically lower in the Makena arm than placebo (12% vs. 17%, P=0.7), the number of adverse outcomes was limited and the difference between arms was not statistically significant. The same neonatal composite index was prospectively evaluated as a coprimary endpoint for Trial 003.

1.3.2. Approval of Makena

Following the publication of results from Trial 002 in 2003, Adeza Biomedical[15] obtained access to the NICHD data and began discussion with the FDA regarding submission of a new drug application (NDA) based on Trial 002.

[15] The NDA ownership was subsequently transferred to several entities, including Hologics, KV Pharmaceutical, Lumara Health, Inc., and AMAG. Hereafter, all are referred to as "the Applicant."

During the first review cycle of the NDA, FDA brought Makena to the Advisory Committee on Reproductive Health Drugs (the Committee) for discussion in August 2006. As noted previously, the primary endpoint of Trial 002 was the rate of PTB prior to 37 weeks gestation; however, 16 of 21 Committee members found that PTB <37 weeks was not an adequate surrogate for reduction in fetal/neonatal mortality and neonatal morbidity. Thirteen of the 21 Committee members voted that PTB <35 weeks was an adequate surrogate, and 12 members voted that the data submitted provided substantial evidence that Makena prevents PTB at <35 weeks. However, the Committee overwhelmingly voted (19 no, 2 yes) that the submitted data did not provide substantial evidence of benefit on neonatal mortality or morbidity, based on the results of the neonatal morbidity/mortality composite index.[16]

FDA did not approve the application in 2006.[17] The primary deficiency was that efficacy based on a single trial that relied on a surrogate endpoint (deemed by most Committee members to be an inadequate surrogate of neonatal morbidity and mortality) was not sufficiently robust to support approval. FDA determined that further study was needed to provide confirmatory evidence of the drug's efficacy in terms of direct clinical benefit on neonatal outcomes or through an established surrogate such as the rate of preterm birth prior to 35 and 32 weeks gestation. To address this deficiency, the FDA recommended that the Applicant submit a draft protocol and evidence of the feasibility of conducting an additional adequate and well-controlled trial to verify and describe further the clinical benefit of preventing recurrent PTB, as stated under the accelerated approval regulations.

In the second review cycle that began in 2008, the Applicant provided a protocol for a postapproval confirmatory trial for an accelerated approval, and another protocol for an infant follow-up study. During the review, the American College of Obstetricians and Gynecologists (ACOG) issued a revised Committee Opinion on Use of Progesterone to Reduce Preterm Birth.[18] In contrast to the 2003 Committee Opinion,[19] which stated:

> *"When progesterone is used, it is important to restrict its use to only women with a documented history of previous spontaneous birth at less than 37 weeks of gestation because unresolved issues remain, such as optimal route of drug delivery and long-term safety of the drug."*

The 2008 Committee Opinion stated:

> *"Progesterone supplementation for the prevention of recurrent preterm birth should be offered to women with a singleton pregnancy and a prior spontaneous preterm birth due to spontaneous preterm labor or premature rupture of membranes."*

---

[16] Cross-Discipline Team Leader Review dated February 3, 2011. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000CrossR.pdf

[17] Approvable Letter, dated October 20, 2006.

[18] ACOG Committee on Obstetric Practice. Use of progesterone to reduce preterm birth. No. 419, October 2008.

[19] ACOG Committee on Obstetric Practice. Use of progesterone to reduce preterm birth. No. 291, November 2003.

FDA interpreted this new Opinion as establishing a *de facto* standard of care for women with a previous spontaneous PTB. FDA was concerned that this opinion could adversely impact recruitment of subjects into a placebo-controlled trial. Although the trial protocol (including study design, planned sample size, primary and secondary objectives, and proposed analysis plan) was deemed satisfactory, FDA declined to approve the application again in 2009, requesting that the Applicant provide adequate documentation that it would be feasible to conduct and successfully complete the confirmatory trial. FDA stated that "adequate assurance of feasibility of [the confirmatory trial] can only be addressed by actual initiation of the trial." Further, noting that one clinical site (University of Alabama at Birmingham) contributed 27% of the total number of subjects in Trial 002, FDA requested that the confirmatory trial include at least 15 investigational sites (US and non-US), with no single site enrolling more than 15% of the total number of subjects. Also, at least 10% of the total randomized subjects would need to be from US and Canadian sites.[20]

By the time of the third review cycle for Makena, multiple clinical studies evaluating the consequences of "late preterm birth" (births between 34^0 to 36^6 weeks gestation) had emerged to show that late-preterm infants are less physiologically and metabolically mature than term infants and are thus at higher risk of morbidity and mortality than term infants.[21,22,23,24] This new evidence led the FDA to determine that PTB < 37 weeks was an acceptable surrogate endpoint that is reasonably likely to predict clinical benefit. This determination also led the FDA to reconsider data from Trial 002. For the endpoint of delivery at < 37 weeks, the results were deemed compelling (with a sizeable treatment difference between groups and a p value of 0.0004) and not driven by data obtained from the University of Alabama at Birmingham alone. FDA concluded that evidence in Trial 002 was sufficient to support Makena improving the proportion of PTB occurring at < 37 weeks under accelerated approval.[16] Furthermore, the Applicant initiated the confirmatory trial in 2009 and provided documentation supporting that this trial could be conducted and completed.

1.4. Hydroxyprogesterone and Progesterone Usage

1.4.1. Use During Pregnancy

FDA conducted a Sentinel query to assess the use of HPC or progesterone during the second or third trimester among pregnancies with live-birth deliveries and their potential reasons for use to characterize the context of real-world use of HPC, the drug substance in Makena. The query captured all pregnancies ending in live birth in the Sentinel Distributed Database, including

[20] Cross-Discipline Team Leader Review dated January 23, 2009 and Complete Response letter dated January 23, 2009.

[21] Engle WA, et al. Committee on Fetus and Newborn, American Academy of Pediatrics. Pediatrics 2007;120(6):1390-401.

[22] McIntire DD, et al. Neonatal mortality and morbidity rates in late preterm births compared with births at term. Obstet Gynecol 2008;111(1):35-41.

[23] Martin JA, et al. Born a bit too early: recent trends in late preterm birth. NCHS Data Brief 2009;Nov(4):1-8.

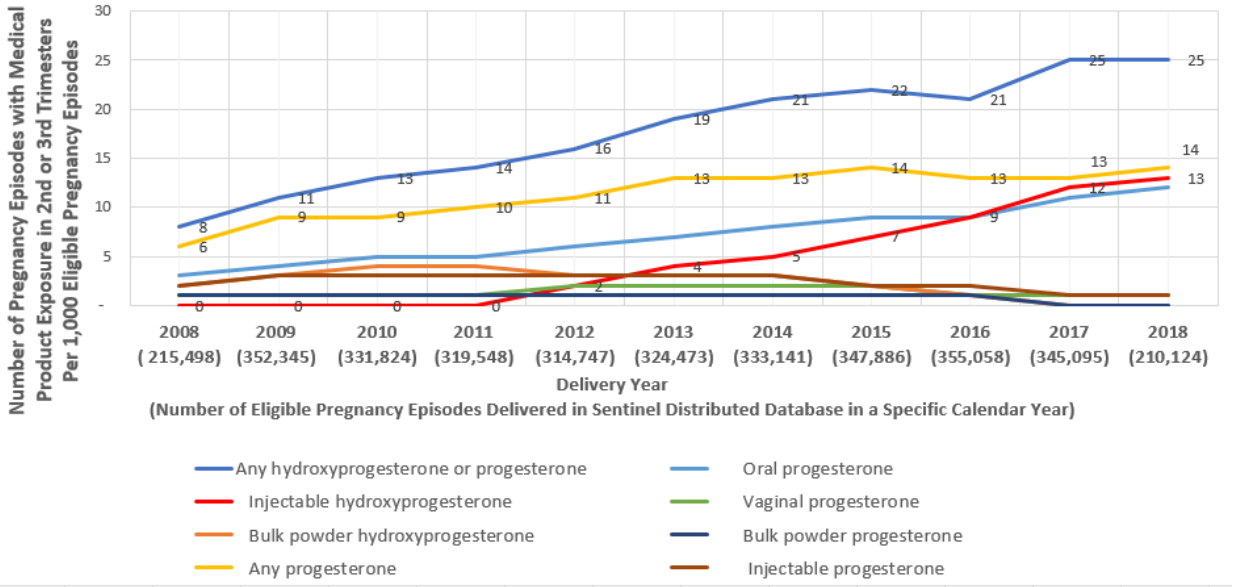
[24] Consortium on Safe Labor, Hibbard JU et al. Respiratory morbidity in late preterm birth. JAMA 2010;304(4):419-25.

singleton and multiple gestations. Progesterone use was included in this analysis because clinical guidelines recommend progesterone treatment for women at risk for preterm delivery.

Methods: This query was conducted in FDA's Sentinel Distributed Database (SDD) using electronic health care data from a distributed network of 15 data partners. The data were primarily comprised of patients with employer-based health care benefits and a small proportion of Medicaid recipients. The study population included women with a live-birth pregnancy (from the current pregnancy) between January 2008 and April 2019 (study period). The exposures of interest were HPC (injectable or bulk powder forms) and progesterone (injectable, oral, vaginal and bulk powder forms). Medical conditions related to potential reasons for HPC or progesterone use were identified by narrow and broad definitions using ICD-9 and ICD-10 diagnosis codes. Included under the narrow definition were diagnosis codes for: (1) history of preterm delivery recorded anytime until one day prior to the start of the current pregnancy, and (2) preterm labor or cervical shortening recorded during the current pregnancy. The broad definition expanded the narrow definition to add the diagnosis for (1) history of preterm labor or cervical shortening recorded anytime until one day prior to the start of the current pregnancy, and (2) preterm delivery recorded during the current pregnancy. Using the diagnostic codes, we could not determine whether the history of preterm delivery was spontaneous or indicated, or whether multiple gestations or other risk factors were present around the time of current pregnancy.

Results: We identified a total of 3,451,121 live-birth pregnancies (from 2,912,911 women) between 2008 and 2019 in FDA's SDD. Note that this number is not a total or annual number of live births in the U.S. Of these, 16,535 pregnancies (5 per 1,000 pregnancies) used injectable HPC during their second or third trimesters and 7,917 used bulk powder HPC (2 per 1,000 pregnancies). In addition, 40,144 (11 per 1,000 pregnancies) pregnancies were exposed to progesterone during the second or third trimesters. In total, approximately 18 per 1,000 pregnancies were exposed to HPC or progesterone during their second or third trimester. The number of exposed pregnancies in each year increased over the study period; the overall the number of exposed pregnancies is modest compared to total pregnancies. The use of HPC or progesterone remains low among pregnancies having a related medical condition, including history of preterm delivery (15%) (Table 3).

Figure 1: Hydroxyprogesterone or Progesterone Use in 2nd or 3rd Trimesters Among 3,449,739, Live-Birth Pregnancy Episodes With Live-Birth Deliveries in the Sentinel Distributed Database Between January 1, 2008, and December 31, 2018, by Delivery Year[1]



[1] Data from 2019 was incomplete and excluded from the figure

Table 3: Proportion of Total Pregnancy Episodes With Related Conditions and With Any Prevalent Hydroxyprogesterone or Progesterone Use During 2nd or 3rd Trimesters Among Women With Live-Birth Deliveries in Sentinel Distributed Database Between January 1, 2008, and April 30, 2019

Related Conditions	Total Number of Pregnancy Episodes with the Related Condition of Interest	Pregnancy Episodes (%) with the Related Conditions of Interest and Any Hydroxyprogesterone or Progesterone Use in the 2nd or 3rd Trimesters
	N	N (%)
Narrow Definition of Related Conditions		
History of preterm delivery[1]	82,255	12,416 (15%)
Preterm labor during the current pregnancy[2]	509,832	29,252 (6%)
Cervical shortening during the current pregnancy[2]	64,557	16,448 (26%)
Any of the narrowly defined conditions above	591,908	40,185 (7%)
Broad Definition of Related Conditions		
History of preterm labor or delivery[1] OR recorded personal history of preterm labor[2]	307,269	34,337 (11%)
Preterm labor or delivery during the current pregnancy[2]	657,719	34,809 (5%)
History of cervical shortening or cervical shortening during the current pregnancy[3]	73,899	17,857 (24%)
Any of the broadly defined conditions above[3]	860,043	51,152 (6%)

[1]Evaluated throughout available enrollment history until the day before pregnancy start date.
[2]Evaluated the day after pregnancy start date until 301 days after pregnancy start date.
[3]Evaluated throughout available enrollment history until 301 days after pregnancy start date.

Among pregnancies exposed to HPC or progesterone, 65% and 83% had at least one related medical condition by narrow and broad definitions, respectively (Table 4), most commonly preterm labor recorded during the current pregnancy. For the pregnancies exposed to injectable HPC, 73% and 98% had at least one narrowly or broadly defined medical condition, respectively.

Table 4: Proportion of Pregnancy Episodes with Related Conditions and Use of Hydroxyprogesterone or Progesterone During 2nd or 3rd Trimesters Among Women With Live-Birth Deliveries in Sentinel Distributed Database Between January 1, 2008, and April 30, 2019

	Any hydroxyprogesterone or progesterone		Hydroxyprogesterone				Progesterone							
			Injectable		Bulk powder		Any		Injectable		Oral		Vaginal	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Number of Pregnancy Episodes with Medical Product Exposure[1]	61,615		16,535		7,917		40,144		8,561		25,471		5,234	
Narrow Definition of Related Conditions														
History of preterm delivery[2]	12,416	20%	6,443	39%	2,568	32%	4,449	11%	1,646	19%	2,365	9%	318	6%
Preterm labor during the current pregnancy[3]	29,252	47%	8,137	49%	5,050	64%	17,969	45%	4,734	55%	10,337	41%	2,515	48%
Cervical shortening during the current pregnancy[3]	16,448	27%	3,228	20%	1,603	20%	12,650	32%	1,694	20%	8,404	33%	2,349	45%
Any of the narrowly defined conditions above[3]	40,185	65%	12,060	73%	6,240	79%	24,351	61%	5,717	67%	14,638	57%	3,499	67%
Broad Definition of Related Conditions														
History of preterm labor or delivery[2] OR recorded personal history of preterm labor during the current pregnancy[3]	34,337	56%	15,696	95%	6,387	81%	14,875	37%	5,381	63%	7,902	31%	1,285	25%
Preterm labor or delivery during the current pregnancy[3]	34,809	56%	8,861	54%	5,811	73%	22,256	55%	5,710	67%	12,875	51%	3,226	62%
History of cervical shortening or cervical shortening during the current pregnancy[4]	17,857	29%	3,982	24%	1,816	23%	13,199	33%	1,840	21%	8,745	34%	2,396	46%
Any of the broadly defined conditions above[4]	51,152	83%	16,240	98%	7,576	96%	30,268	75%	7,344	86%	18,109	71%	4,155	79%

[1] Numbers on the top row are not exclusive because a pregancy could use more than one medicaton of interest
[2] Evaluated throughout available enrollment history until the day before pregnancy start date.
[3] Evaluated the day after pregnancy start date until 301 days after pregnancy start date.
[4] Evaluated throughout available enrollment history until 301 days after pregnancy start date.

We note several study limitations. First, this analysis did not examine the timing of the related medical conditions relative to the use of HPC or progesterone. Therefore, we interpret the presence of the related medical conditions as possible reasons for use. It should be noted that this analysis captured all live-birth pregnancies in the Sentinel Distributed Database. However, we could not determine whether the recorded diagnosis for a history of preterm delivery was spontaneous or indicated, nor did we examine whether the current pregnancy was singleton or multiple gestation. Therefore, HPC exposed pregnancies may not entirely reflect the approved obstetrical indication of HPC. Second, given that women in the SDD were covered primarily by commercial insurance health plans, our findings may have limited generalizability to women without commercial health insurance. Third, we only examined HPC or progesterone use among pregnancies ending with live births. Lastly, the exposure could be under-estimated owing to the capture of pharmacy dispensing data and medication claims only (no capture of out of pocket payments). Some pharmacies create their own National Drug Codes (NDCs) for compounded HPC which would not have been captured in the analysis.

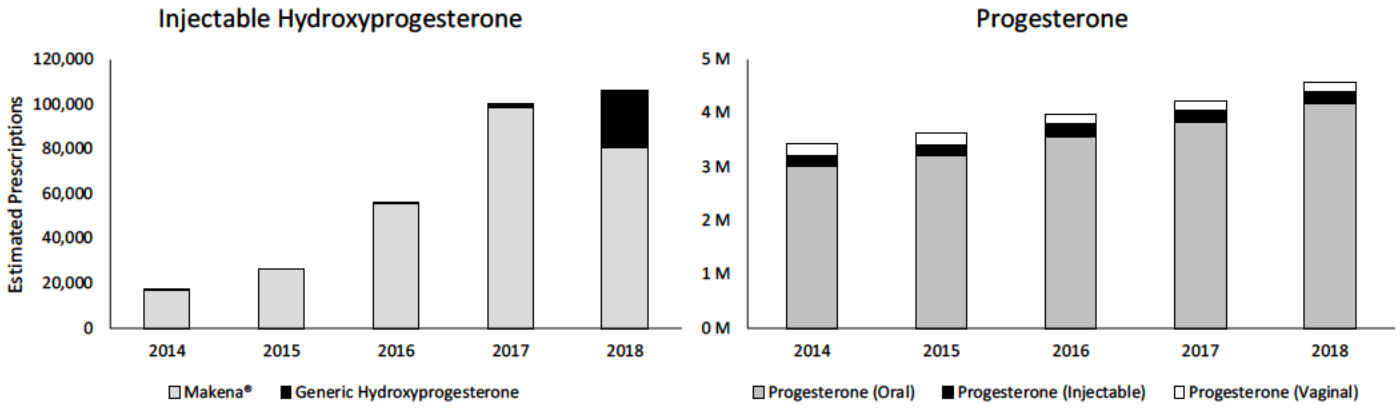
In summary, this analysis found modest use of HPC and progesterone during the second or third trimesters, even among pregnancies with a diagnostic code of a history of preterm delivery (15%). A high percentage (65% and 83% by narrow and broad definitions, respectively) of

pregnancies exposed to HPC or progesterone during their second or third trimester had at least one related medical condition recorded before or during the current pregnancy.

1.4.2. Estimated Use in U.S. Outpatient Settings

FDA analyzed use patterns of injectable HPC and oral, vaginal, or injectable dosage forms of progesterone. Prescriptions for bulk powder forms were excluded due to the inability to determine the final product form and the likelihood that these are underrepresented in the data. We used the Symphony Health PHAST™ Prescription monthly database to estimate the number of prescriptions for injectable HPC and oral, vaginal, or injectable progesterone products dispensed to patients of any age from U.S. outpatient retail or mail order/specialty pharmacies, stratified by molecule and form, annually from 2014 through 2018 (Figure 2). Total prescriptions dispensed for HPC or progesterone products (products with a non-proprietary name of 'hydroxyprogesterone' or 'progesterone') increased 35% from an estimated 3.5 million prescriptions in 2014 to 4.7 million prescriptions in 2018. During this time there was an increase in HPC dispensed prescriptions from an estimated 16,600 prescriptions in 2014 to 106,000 prescriptions in 2018. In 2018, 4.6 million prescriptions (98%) dispensed were for progesterone products.

Figure 2: Estimated Annual Number of Prescriptions Dispensed for Hydroxyprogesterone or Progesterone Products*, Stratified by Molecule and Form, From U.S. Retail or Mail Order/Specialty Pharmacies, Years 2014 to 2018



Prescriptions for bulk powder forms of hydroxyprogesterone and progesterone were not included.
* Products with a non-proprietary name of 'hydroxyprogesterone' or 'progesterone'
Source: Symphony Health PHAST™ Prescription Monthly. Years 2014-2018. Extracted July 2019. File: SH Progesterone and Hydroxyprogesterone Rx 07-29-2019.xlsx

The Symphony Health IDV® Integrated Dataverse was used to obtain the estimated number of 15- to 44-year-old patients who were dispensed prescriptions for injectable HPC and oral, vaginal, or injectable progesterone products from U.S. outpatient retail and mail order/specialty pharmacies, stratified by molecule and form, annually from 2014 through 2018. The total number of patients who were dispensed HPC or progesterone increased by 17% from an estimated 479,000 patients in 2014 to 560,000 patients in 2018 (Table 17 in the Appendix). In 2018, an estimated 42,000 patients (8%) were dispensed prescriptions for HPC, and an estimated 521,000 patients (93%) were dispensed prescriptions for progesterone products. The number of

patients who received a prescription for HPC increased from approximately 8,000 patients in 2014 to 25,500 patients in 2016 and 42,000 patients in 2018.

Table 18 in the Appendix provides the estimated number of drug use mentions of progesterone or HPC products among 15- to 44-year-old women, stratified by molecule and form, associated with a diagnosis as reported on U.S. office-based physician surveys from 2013 through 2018, aggregated. An estimated 50% of HPC use mentions were associated with a diagnosis of supervision of high-risk pregnancy (ICD-10 code O09), of which 78% were associated specifically with supervision of a pregnancy with a history of preterm labor (O09.21, data not shown) and 10% were associated specifically with supervision of elderly primigravida and multigravida (O09.5, data not shown). Twenty percent of HPC use mentions were associated with personal history of preterm labor (Z87.51, data not shown), 13% were associated with encounter for supervision of a normal pregnancy (Z34), and 10% were associated with preterm labor (in the current pregnancy, O60). Among progesterone products, an estimated 42% of progesterone injectable use mentions were associated with supervision of high-risk pregnancy and 41% were associated with female infertility (N97). An estimated 59% of progesterone vaginal use mentions were associated with female infertility.

Table 19 in the Appendix provides the estimated number of drug use mentions among women 15 to 44 years old associated with selected diagnoses as reported on U.S. office-based physician surveys from 2013 through 2018, aggregated. An estimated 42% of office visits with any drug use mentions that were associated with a diagnosis of history of preterm labor (O09.21 or Z87.51) mentioned Makena, and an additional 32% mentioned generic HPC products. Of office visits with drug use mentions that were associated with preterm labor in the current pregnancy, physicians mentioned Makena in 14% of visits. Of office visits associated with cervical shortening, physicians mentioned the use of progesterone products but no other products.

In summary, HPC use increased from 2014 to 2018 with the number of patients treated increasing over the same time period. However, HPC use represents a small proportion of the total use of progesterone in FDA's assessment. The primary use of HPC appeared related to obstetrical diagnoses whereas progesterone was used for both obstetrical and infertility related conditions.

2. Confirmatory Trial—Trial 003

2.1. Development of Trial 003

Please refer to Section 1.3 for a detailed discussion regarding the regulatory history of Makena. After the first non-approval of the NDA in 2006, FDA and the Applicant engaged in discussion regarding a clinical protocol to provide evidence verifying clinical benefit. In 2009, Trial 003 was initiated; the study design mirrored that of Trial 002, except that Trial 003 had coprimary endpoints of delivery prior to 35 weeks and the neonatal morbidity/mortality composite index. When Makena was approved under accelerated approval in 2011, the completion of Trial 003 became a requirement post-approval to verify and describe the clinical benefit of Makena.

Trial 003 was initiated in the United States to ensure at least 10% of subjects would be from the United States and Canada before expanding to Europe. However, after Makena's approval in

2011, enrolling U.S. subjects became increasingly difficult. Additional study sites were subsequently opened in Ukraine and Russia.

2.2. Trial Design

Trial 003 was a multicenter, randomized, double-blind, placebo-controlled clinical trial in women, aged 18 years or older, with a singleton pregnancy, and with a history of a previous singleton spontaneous preterm delivery.

2.2.1. Study Objectives

Primary objectives:

- Determine if treatment with Makena reduces the rate of preterm birth prior to 35^0 weeks of gestation.
- Determine if Makena reduces the rate of neonatal mortality or morbidity.

Secondary objectives:

- Exclude a doubling of the risk of fetal/early infant death, defined as spontaneous abortion/miscarriage (delivery from 16^0 through 19^6 weeks of gestation), early infant death (from minutes after birth until 28 days of life) occurring in livebirths prior to 24 weeks gestation, or stillbirth (antepartum or intrapartum death from 20 weeks gestation through term), in the Makena group compared to the placebo group.
- Determine if Makena reduces the rate of preterm birth prior to 32^0 and 37^0 weeks of gestation, respectively.
- Determine if Makena reduces the rate of stillbirth defined as all stillbirths/fetal deaths/in-utero fetal losses occurring from 20 weeks gestation until term.
- Determine if Makena reduces the rate of neonatal death (from minutes after birth until 28 days life) occurring in livebirths born at 24 weeks gestation or greater.

2.2.2. Trial Design and Conduct

Trial 003 was conducted in the United States, Canada, Russia, Ukraine, Hungary, Spain, Bulgaria, the Czech Republic, and Italy. Eligible subjects were randomized in a 2:1 ratio to receive either Makena or placebo and received weekly injections of study drug from randomization (16^0 through 20^6 weeks of gestation) until 36^6 weeks of gestation or delivery, whichever occurred first.

2.2.3. Eligibility Criteria

Major inclusion criteria:

1. Women aged 18 years or older.
2. Singleton gestation.
3. Estimated gestational age between 16^0 weeks and 20^6 weeks, inclusive, at the time of randomization.
4. Documented history of a previous singleton spontaneous preterm delivery. Spontaneous preterm birth was defined as delivery from 20^0 to 36^6 weeks of gestation following spontaneous preterm labor or preterm premature rupture of membranes (pPROM).

Major exclusion criteria:

1. Multifetal gestation.
2. Known major fetal anomaly or fetal demise;
3. Presence of a uterine anomaly (uterine didelphys or bicornuate uterus)
4. Maternal medical/obstetrical complications or had any significant medical disorder
5. Subjects who received a progestin during the current pregnancy AND met one of the following criteria:
    a. Progestin was administered in the 4 weeks preceding the first dose of study medication.
    b. Subjects received HPC
    c. Progestin was administered by a route other than oral or intra-vaginal.
6. Participation in an antenatal study in which the clinical status or intervention may have influenced gestational age at delivery.
7. Participation in this trial in a previous pregnancy.

2.2.4. Analysis Populations

The Applicant defined the following analysis populations:

- Intent-to-treat (ITT) population: all randomized subjects. Subjects were analyzed by the treatment group to which they were randomized, regardless of the blinded study medication (active or placebo) the subject received.
- Safety population: all subjects who received at least one dose of blinded study medication. Subjects were analyzed by the treatment that they received.
- Liveborn neonatal population: all babies of randomized women in the ITT Population who were liveborn and for whom morbidity/mortality data were available.

2.2.5. Efficacy Endpoints

There were two coprimary endpoints:

- Surrogate endpoint: PTB prior to 35^0 weeks of gestation
    - Scored as a 1 if any of the following events occurred: a delivery occurring from randomization up through 34^6 weeks of gestation, including a miscarriage occurring from 16^0 through 19^6 weeks of gestation, and an elective abortion.
    - Otherwise, scored as a 0.

- Clinical endpoint: Composite neonatal morbidity and mortality index
    - Scored as a 1 if the liveborn neonate had any of the following events occur at any time during the birth hospitalization up through discharge from the neonatal intensive care unit (NICU): neonatal death, grade 3 or 4 intraventricular hemorrhage (IVH), respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), necrotizing enterocolitis (NEC), or proven sepsis.
    - Otherwise, scored as a 0.

Key secondary endpoints:

- Neonatal death (from minutes after birth until 28 days of life) occurring in livebirths born at 24 weeks or older gestation
- Preterm birth prior to 32^0 weeks of gestation.

- Preterm birth prior to 37^0 weeks of gestation

Preterm birth endpoints were analyzed using the ITT population and neonatal endpoints were analyzed using the liveborn neonatal population.

The study was designed to detect a 30% reduction in PTB $<35^0$ weeks (from 30% to 21%) and 35% reduction (17% to 11%) in the neonatal composite index, based on the findings from Trial 002. An estimated sample size of 1707 provided at least 90% power to detect the hypothesized difference at alpha level 0.05, and approximately 83% power to rule out a doubling of risk of fetal/early infant death (upper bound of the 95% confidence interval of relative risk <2).

2.2.6. Statistical Analysis Methods

2.2.6.1. Primary Analyses

For each of the coprimary efficacy endpoints, the number and percentage of subjects for the event were presented by treatment groups. Statistical significance between Makena and placebo treatments for each endpoint was determined using a Cochran–Mantel–Haenszel test (CMH) stratified by gestational age at randomization (16^0 to 17^6 weeks and 18^0 to 20^6 weeks).

The interaction between treatment and gestational age at the time of randomization was assessed by a logistic regression model of preterm delivery prior to 35^0 weeks of gestation with terms for treatment, gestational age at randomization stratum, and treatment-by-gestational age at randomization stratum interaction. A similar analysis was performed for the neonatal composite index.

2.2.6.2. Exploratory Analyses

After Trial 003 failed to demonstrate efficacy with the coprimary endpoints, the Applicant conducted a series of exploratory subgroup analyses to understand the potential reasons for the negative findings in Trial 003. The Applicant analyzed the coprimary efficacy endpoints by subgroups defined in Table 5 for the overall study population in Trial 003 and its U.S. subgroup.

Table 5: Trial 003 Subgroup Categories

Subgroup	Categories
Geographic region	U.S., Non-U.S.
Gestational age at randomization	16^0-17^6 weeks, 18^0-20^6 weeks
Gestational age at qualifying delivery*	20^0-<28^0 weeks, 28^0-<32^0 weeks, 32^0-<35^0 weeks, 35^0-<37^0 Weeks
Gestational age at earliest prior PTBs	0-<20^0, 20^0-<28^0, 28^0-<32^0, 32^0-<35^0, 35^0-<37^0
Number of previous PTBs	1, 2, ≥3
Cervical length at randomization	<25 mm ≥25 mm
BMI before pregnancy (kg/m^2)	<18.5, 18.5 - <25, 25-<30, ≥30
Any substance use during pregnancy	Yes, No
Smoking	Yes, No
Alcohol	Yes, No
Illicit drugs	Yes, No
Race	Non-Hispanic black, non-Hispanic non-black
Ethnicity	Hispanic, non-Hispanic
Years of education	≤12, >12

* Qualifying delivery is the most recent preterm delivery.

Generally, FDA does not support unplanned exploratory subgroups analyses, especially when the overall result does not demonstrate efficacy. There are multiple reasons to not consider exploratory subgroup analyses to support establishing efficacy when treatment benefit in the overall population is not significant (FDA draft guidance on multiple endpoints in clinical trials,[25] E17 General Principles for Planning and Design of Multi-Regional Clinical Trials,[26] and E9 Statistical Principles for Clinical Trials[27]). The major statistical reason is inflation of type I error, that is, the heightened probability of incorrectly concluding treatment benefit. When such post-hoc subgroup analyses are used to search for evidence of benefit, there is a high probability that any observed favorable subgroup results are due to chance alone. Therefore, FDA considers exploratory analyses hypothesis-generating.

2.3. Trial Results

2.3.1. Subject Disposition

A total of 1708 subjects were randomized to either Makena (n=1130) or placebo (n=578). Almost all (99%) subjects completed the study and completed treatment (93%). Russia, Ukraine and the U.S. were the three highest enrolling countries, randomizing 621 (36%), 420 (25%) and 391 (23%) subjects, respectively, followed by Hungary, Spain, Bulgaria, Canada, the Czech Republic, and Italy, which each had less than 100 subjects (16% of all subjects).

[25] https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm536750.pdf

[26] https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM519603.pdf

[27] https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073137.pdf

Table 6: Trial 003 Subject Disposition

	Makena, N(%)	Placebo, N(%)
Subjects randomized (ITT population)	1130	578
Subjects who received at least one dose of study drug (safety population)	1128 (99.8)	578 (100)
Liveborn infant with morbidity data available (liveborn neonatal population)	1091 (96.5)	560 (96.9)
Subjects withdrawing from study	18 (1.6)	6 (1.0)
Subjects discontinuing study drug	80 (7.1)	43 (7.4)

Source: Applicant's study report

2.3.2. Demographics and Baseline Characteristics

The Makena and placebo groups were comparable across all demographic and baseline characteristics. The mean age was 30 years and pre-pregnancy BMI was 24.4 kg/m^2. Of the randomized subjects, 88% were white, 7% were black, and the rest included Native Hawaiian/Pacific Islanders, Asian and American Indian or Alaska native, mixed race and other. Almost all black subjects were from the United States. Approximately 10% of women were never married or divorced/widowed/separated, approximately 8% smoked, approximately 3% consumed alcohol, and 1.3% used illicit drugs.

The treatment groups were also well balanced with respect to obstetrical characteristics in the current and previous pregnancies. Slightly more subjects initiated study drug between 18^0 to 20^6 weeks of gestation (56% Makena, 58% placebo) than between 16^0 to17^6 weeks (44% Makena, 41% placebo). Overall, the median estimated gestational age at randomization was 18.1 weeks for the Makena group and 18.4 weeks for the placebo group.

2.3.3. Primary Efficacy Results

The neonatal composite index was scored as positive (value of 1) in 5.4% and 5.2% of liveborn infants in the Makena and placebo groups, respectively, with a difference of 0.2% (95% CI: -2.0%, 2.5%) as shown in Table 7. The rate of preterm births prior to 35^0 weeks gestation was 11.0% and 11.5% in the Makena and placebo groups, respectively, with a difference of -0.6% (95% CI: -3.8%, 2.6%). The treatment effect of Makena compared to placebo was not statistically significant for both coprimary endpoints.

The rates of preterm birth prior to 32 weeks gestation and prior to 37 weeks gestation were also not different between the Makena and placebo groups.

Table 7: Trial 003 Efficacy Results

Efficacy Endpoints	Makena (N=1130)	Placebo (N=578)	Difference (95% CI)*	P-value*
Neonatal composite index	5.4% (59/1091)	5.2% (29/560)	0.2% (-2.0, 2.5)	0.84
PTB <35^0 weeks (%)	11.0% (122/1113)	11.5% (66/574)	-0.6% (-3.8, 2.6)	0.72
PTB <32^0 weeks (%)	4.8% (54/1116)	5.2% (30/574)	-0.4% (-2.8, 1.7)	
PTB <37^0 weeks (%)	23.1% (257/1112)	21.9% (125/572)	1.3% (-3.0, 5.4)	

Abbreviations: N: number of randomized subjects, CI: confidence interval, PTB: preterm birth
*Difference, 95% CI and P-value were from CMH method stratified by gestational age at randomization
Source: FDA analysis

2.3.4. Exploratory Analyses Results

Applicant's subgroup analysis results: The Applicant's results for the subgroup analyses of the coprimary efficacy endpoints are presented in Table 21 and Table 22 in the Appendix.

FDA's subgroup analysis results:
FDA reviewed all results and conducted subgroup analyses by region and race because these subgroups are evaluated by FDA routinely. Also, they are important subgroups that differentiate the study populations between Trial 003 and Trial 002.

1. By geographic region (U.S. versus non-U.S.)

The Applicant asserts that the overall lower than expected rate of study outcomes substantially limited the ability of Trial 003 to assess the effects of Makena on these outcomes. The Applicant also believes that the lower rate of PTB in Trial 003 could be accounted for by significant geographic differences in PTB rates, where Russia and Ukraine enrolled more subjects but had much lower rates than the United States.

Generally, FDA does not support unplanned subgroup analyses but performed exploratory analysis by region (U.S. versus non-U.S.) to examine whether there were potentially important differences in treatment benefit between U.S. and non-U.S. patients in Trial 003.

For Trial 003, FDA calculated the rate difference between the Makena and placebo groups for each coprimary endpoint, and also the secondary endpoints of birth prior to 32 and 37 weeks gestation, using two methodologies, a stratified CMH method and shrinkage estimation through Bayesian modeling. Traditional subgroup analysis evaluates a particular subgroup category independently from other subgroup categories and relies only on the data from the subjects in that particular category, whereas the Bayesian shrinkage estimation analysis evaluates all subgroup categories jointly. In any trial, some subgroups will perform well, and others will perform poorly. The traditional subgroup analysis is likely to have an increase in the overall error of the estimates compared with the shrinkage analysis, which borrows strength across subgroups.

In the U.S. subgroup of Trial 003, both the neonatal composite index and preterm birth prior to 35 weeks endpoints showed no evidence of a treatment effect using stratified CMH and shrinkage estimation. Although the point estimates of -2.2%, based on the CMH analytic method, for the coprimary endpoints in the U.S. subgroup are in the direction of a beneficial treatment effect, the 95% confidence intervals around these point estimates include 0, indicating

no evidence of effect even in these exploratory subgroup analyses. Similarly, no evidence of a treatment effect was seen for the endpoints of delivery < 37 weeks or < 32 weeks. In addition, the interaction between treatment and region for each coprimary endpoint was assessed by a logistic regression model with treatment, region and treatment-by-region interaction; no significant interaction effect was noted. This Trial 003 subgroup analysis did not show that Makena had a favorable treatment effect compared to placebo for either coprimary endpoints in either the U.S. or non-U.S. region (see Table 8). The lack of evidence of an interaction between region and treatment and the lack of evidence of a treatment effect within the U.S. subgroup in Trial 003 does not provide support for regional differences explaining the differences in results between Trial 002 and 003.

Table 8: Trial 003 Results of Efficacy Endpoints by Region (U.S. vs. non-U.S.)

			Difference (95%CI) Makena vs. Placebo	
	Makena (N=1130)	**Placebo (N = 578)**	**Stratified CMH**	**Shrinkage Estimation**
Neonatal composite index	(N=1091)	(N=560)		
U.S.	7.1% (18/252)	9.5% (12/126)	-2.2% (-8.3, 3.9)	-0.2% (-4.9, 2.8)
Non-U.S.	4.9% (41/839)	3.9% (17/434)	1.0% (-1.4, 3.3)	0.6% (-1.6, 2.8)
Preterm birth <35^0 weeks gestation	(N=1113)	(N=574)		
U.S.	15.6% (40/256)	17.6% (23/131)	-2.2% (-10.1, 5.7)	-0.8% (-6.0, 3.5)
Non-U.S.	9.6% (82/857)	9.7% (43/443)	-0.2% (-3.6, 3.2)	0.4% (-3.6, 2.8)
Preterm birth <32^0 weeks gestation	(N=1116)	(N=574)		
U.S.	5.5% (14/256)	9.2% (12/131)	-3.9% (-9.6, 1.7)	-0.6% (-8.4, 3.8)
Non-U.S.	4.7% (40/860)	4.1% (18/443)	0.6% (-1.7, 2.9)	0.5% (-1.8, 2.8)
Preterm birth <37^0 weeks gestation	(N=1112)	(N=572)		
U.S.	33.2% (85/256)	28.2% (37/131)	4.7% (-5.0, 14.3)	1.8% (-3.6, 9.0)
Non-U.S.	20.1% (172/856)	20.0 % (88/441)	0.2% (-4.4, 4.8)	0.9% (-3.5, 5.2)

Source: FDA analysis

2. By race (black/African American vs. non-black/African American)

FDA conducted a subgroup analysis by race (black and non-black) for Trial 003. This race subgroup analysis did not provide evidence that Makena had a treatment effect on either coprimary efficacy endpoints in the black or non-black subgroups.

Table 9: Trial 003 Results of Coprimary Efficacy Endpoints by Race*

	Makena (N=1130)	Placebo (N=578)	Difference (95%CI)
Neonatal composite index			
Black/African American	8.7% (6/69)	7.5% (3/40)	0.8% (-9.9,11.5)
Non-black/African American	5.2% (53/1022)	5.0% (26/520)	0.2% (-2.1, 2.5)
PTB <35^0 weeks gestation			
Black/African American	23.6% (17/72)	19.5% (8/41)	3.0% (-12.5, 18.5)
Non-black/African American	10.1% (105/1041)	10.9% (58/533)	-0.8% (-4.1, 2.4)

*This is based on the entire Trial 003 study population
Source: FDA analysis

Considering the Applicant's and FDA's subgroup analyses results, Makena did not demonstrate any favorable effect (positive finding with nominal statistical significance) over placebo in the key efficacy endpoints in any of the evaluated subgroups.

2.4. Comparisons Between Trial 003 and Trial 002

FDA does not generally support cross-study comparisons to draw efficacy conclusions. Both Trials 003 and 002 were well-controlled and well-conducted, such that each should provide evidence of efficacy on its own merit. Nevertheless, we explored the potential for significant differences in key aspects between Trials 003 and 002 that might clarify their divergent results.

Study design:
Trials 002 and 003 were nearly identical in design. However, trial 002 was conducted entirely in the United States between 1999 to 2002 with preterm birth <37 weeks as the primary efficacy endpoint. Trial 003 was a multinational trial conducted between 2009 to 2018 with coprimary endpoints of a neonatal composite index and preterm birth <35 weeks and was approximately 3.5 times larger than Trial 002. Trial 003 was powered to detect the treatment difference in the coprimary endpoints based on the effect size observed in Trial 002.

Study populations and trial outcomes:
Trial 003 had the following notable differences compared to Trial 002:

Table 10: Comparisons of Selected Characteristics Between Trial 003 and Trial 002

	Trial 003 Overall (N=1708)	Trial 003 U.S. Subgroup (N=391)	Trial 002 (N=463)
Demographics			
Black race	7%	29%	59%
Single or without a partner	10%	31%	50%
Risk factors			
Use of substance* during pregnancy	10%	28%	26%**
Gestational age of qualifying delivery (weeks)	32	33	31
History of more than one previous PTB	15%	27%	28%/41%***
Rate PTB <35 weeks in placebo group+	12%	18%	30%
Rate PTB <37 weeks in placebo group+	22%	28%	55%

*Including tobacco, alcohol, illicit drugs
**Trial 002 collected information on substance use prior to the study pregnancy and not during the pregnancy; 26% is expected to be the higher end of the estimate because it assumes that all women who used substance prior to the pregnancy continued substance use after becoming pregnant.
***HPC – 28%; Placebo – 41%
+It is assumed that the rate in the placebo group approximates that of the contemporaneous intended population

The overall study population of Trial 003 appeared to be at lower risk for factors that might affect the risk of PTB. The 003-U.S. subgroup, however, was more similar to the Trial 002 study population (see Table 10). Yet, unlike Trial 002, there was no consistent evidence of benefit of Makena over placebo in the U.S. subgroup of Trial 003 (see Table 8). As noted above, no statistically significant interaction was seen between treatment and region in Trial 003.

In its briefing document, the Applicant presented post-hoc efficacy analyses exploring a potential relationship between efficacy and the proportion of subjects in a trial with more than one of 5 selective risk factors (history of > 1 prior PTB, black race, substance use in pregnancy, ≤ 12 years of education, unmarried with no partner). The Applicant concluded that Trial 002 had the "highest" risk population (based on the observation that this trial had the highest proportion of study subjects with more than one of these 5 factors), followed by the Trial 003-U.S. subgroup, and then the overall Trial 003 population as being the relatively lowest risk population. The Applicant's analysis showed a trend toward decreasing efficacy in subpopulations the Applicant considered as lower risk. As described earlier, subgroup analyses, especially when conducted post-hoc when the study findings are known, are exploratory and cannot be relied upon for inferences of efficacy.

In addition, it is challenging to identify specific patient subpopulations that may be more responsive to treatment based on the totality of the data. FDA conducted exploratory analyses of Trial 003 using logistic regression models for each coprimary efficacy endpoint with treatment, region, each of the aforementioned 5 risk factors, and its interaction with treatment. These analyses do not provide convincing evidence of efficacy over placebo in any subpopulation and there is no statistically significant interaction between Makena and any of these risk factors. Analogous analyses in the Trial 003-U.S. subgroup produced similar results. In summary, although these risk factors may have an impact on the overall PTB or neonatal composite index rate, there was no evidence in Trial 003 that they impact the treatment effect nor was there consistent convincing evidence of an effect within a specific subpopulation across the two trials. For example, while black women in the U.S. have a higher rate of PTB compared to non-black

women, there was no interaction between race (blacks vs. non-blacks) and treatment effect in Trial 002 or Trial 003, nor was there evidence of an effect in the U.S. subgroup in Trial 003. Similarly, women with > 1 prior PTB are considered at higher risk of having recurrent PTB. However, there was no consistent trend in treatment benefit in this population (see Table 22). In Trial 002, these women had a treatment benefit compared to placebo in reduced rate of delivery < 35 weeks (30% Makena vs. 44% placebo). This benefit was not observed in Trial 003, where women with > 1 PTB randomized to Makena had higher rates of birth < 35 weeks compared to placebo (Trial 003 overall: 26% Makena vs. 19% placebo; Trial 003 US subgroup: 25% Makena vs. 17% placebo). Importantly, Makena is approved in women with a singleton pregnancy and a prior sPTB, and evidence of efficacy must be based on that intended population.

In summary, Trial 003 did not demonstrate a treatment benefit of Makena on reducing the neonatal composite index or the rate of spontaneous preterm birth prior to 35 weeks gestation, nor was there evidence of a treatment benefit on the rate of spontaneous preterm birth prior to 37 weeks or 32 weeks gestation. The significant statistical limitations with exploratory subgroup analyses preclude reliable inference of efficacy based on findings from these analyses.

3. Other Evidence of Effects of Progesterone on Preterm Birth

There are published data on other progesterone formulations that have been investigated for the treatment of PTB. To explore the consistency of results, FDA evaluated pertinent published literature on the effect of progesterone on the risk of PTB from randomized, placebo-controlled trials and recent, larger meta-analyses. In its briefing document, the Applicant references several studies that evaluated 17-HPC.[28,29,30,31,32,33] However, most of these publications are not applicable to Makena's approved use because the studies assessed different clinical outcomes (early recurrent pregnancy losses or the prevention of preterm labor). There are additional publications that evaluated the effect of hydroxyprogesterone caproate intramuscular injections on pregnancy outcomes (with dosing regimens ranging from 500 mg weekly or twice weekly to

[28] Levine L. Habitual abortion. A controlled study of progestational therapy. West J Surg Obstet Gynecol. 1964;72:30-36.

[29] 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.

[30] Johnson JWC, et al. Efficacy of 17α-hydroxyprogesterone caproate in the prevention of premature labor. New Engl J Med. 1975;293:675-680.

[31] Yemini M, et al. Prevention of premature labor by 17 alpha-hydroxyprogesterone caproate. Am J Obstet Gynecol. 1985:151(5):574-577.

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

[33] Saghafi N, et al. Efficacy of 17α-hydroxyprogesterone caproate in the prevention of preterm delivery. J Obstet Gynaecol Res. 2011;37(10):1342-1345.

1000 mg weekly); however, they are not discussed further here because of the smaller sample size (80 subjects)[34] or the absence of a concurrent control group.[35,36,37,38]

3.1. Randomized, Placebo-Controlled Clinical Trials

The following six placebo-controlled trials evaluated the treatment effect of progesterone on preterm birth and included pregnant women with a history of a prior sPTB. Note that all these trials evaluated vaginal progesterone.

- The 2003 da Fonseca et al. publication reported findings from a single center trial in Brazil that randomized 142 women with a current singleton pregnancy and a history of previous PTB, cerclage, or uterine malformation in a 1:1 ratio to daily vaginal progesterone insert (100 mg) or placebo.[39] Study drug was applied from 24 to 34 weeks of gestation. The majority (>90%) of women enrolled had previous PTB (mean gestational age at delivery 33 weeks). The rate of PTB <37 weeks was 14% in the progesterone group compared to 29% with placebo (p=0.03).
- The 2007 O'Brien et al. publication reported findings from an international trial that randomized 659 women with a singleton pregnancy and a prior singleton sPTB (delivery between 20^0 and 35^0 weeks of gestation) in a 1:1 ratio to daily vaginal progesterone (8% gel, 90 mg) or placebo starting at 18 to 22^6 weeks until 37 weeks or delivery.[40] Both treatment groups had normal cervical length at randomization (3.7 cm). The primary endpoint, the rate of PTB ≤32 weeks, was not statistically different between the two study groups (10% progesterone vs. 11% placebo, odds ratio: 0.9). Similar results were seen for rate of PTB <37 weeks (42% progesterone vs. 41% placebo, odds ratio: 1.08) and ≤35 weeks (23% progesterone vs. 27% placebo., odds ratio: 0.9). No differences were seen in neonatal outcome (Apgar score, birth weight, NICU admission, respiratory distress syndrome, intraventricular hemorrhage, necrotizing enterocolitis, and death).

[34] Hauth JC, et al. The effect of 17 alpha- hydroxyprogesterone caproate on pregnancy outcome in an active-duty military population. Am J Obstet Gynecol. 1983;146(2):187-190.

[35] Katz Z, et al. Teratogenicity of progestogens given during the first trimester of pregnancy. Obstet Gynecol. 1985;65(6):775-780.

[36] Rozenberg P, Chauveaud A, Deruelle P, et al. Prevention of preterm delivery after successful tocolysis in preterm labor by 17 alpha-hydroxyprogesterone caproate: a randomized controlled trial. Am J Obstet Gynecol. 2012;206(3):206 e1-9.

[37] Senat MV, Porcher R, Winer N, et al. Prevention of preterm delivery by 17 alpha-hydroxyprogesterone caproate in asymptomatic twin pregnancies with a short cervix: a randomized controlled trial. Am J Obstet Gynecol. 2013;208(3):194 e1-8.

[38] Winer N, Bretelle F, Senat MV, et al. 17 alpha-hydroxyprogesterone caproate does not prolong pregnancy or reduce the rate of preterm birth in women at high risk for preterm delivery and a short cervix: a randomized controlled trial. Am J Obstet Gynecol. 2015;212(4):485 e481-485 e410.

[39] Da Fonseca EB, et al. 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 Feb;188(2):419-24

[40] O'Brien JM, 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

- The 2007 Fonseca et al. publication reported findings from an international trial that randomized, in a 1:1 ratio, 250 women with a singleton (N=226) or twin (N=24) pregnancy and a short cervix to daily 200 mg micronized progesterone capsule or placebo.[41] The qualifying risk factor was a cervical length ≤15 mm identified incidentally on routine anatomy ultrasound performed at 20 to 24 weeks of gestation, irrespective of history of PTB; the majority of women (>50%) were nulliparous, approximately a third had no prior PTBs, and 15% had a history of one or more PTB. The study medication was used from 24 to 33^6 weeks of gestation. The primary endpoint was spontaneous delivery <34 weeks. The rate of PTB <34 weeks was 19% in the progesterone group compared to 34% in the placebo group, and this difference was statistically significant (relative risk: 0.56; p=0.007). There was no between-group difference for birthweight, fetal/neonatal death, admission to the NICU or major adverse neonatal outcomes before discharge. Among women with a history of PTB (N=38), progesterone administration did not reduce the incidence of PTB before 34 weeks (95% confidence for relative risk included 1).
- In 2011, Hassan et al. reported results of an international (23 U.S. and 21 non-U.S. sites) trial that randomized 465 asymptomatic women with a singleton pregnancy and a shortened cervix (cervical length between 10 to 20 mm) to daily vaginal progesterone (8% gel, 90 mg) or placebo in a 1:1 ratio.[42] Enrollment was stratified by presence/absence of a history of PTB. Women received study drug from 20 to 23^6 weeks until 36^6 weeks or delivery. The primary endpoint was delivery <33 weeks of gestation. The progesterone group had a significantly lower rate of delivery <33 weeks of gestation compared with the placebo (9% vs. 16%, respectively, p=0.02). In women with a history of PTB (13% of the study population) <35 weeks gestation, vaginal progesterone gel administration was not associated with a reduction in the rate of delivery <33 weeks compared to placebo (relative risk: 0.77, 95% CI 0.29-2.06).
- Published in 2016, the OPPTINUM trial was conducted primarily in the United Kingdom and randomized 1228 women with a singleton pregnancy and at risk for PTB in a 1:1 ratio to daily vaginal progesterone (200 mg) or placebo from 22-24 weeks to 34 weeks of gestation.[43] Eligible women had the following risk factors: previous sPTB at ≤34 weeks gestation, a cervical length ≤25 mm, or a positive fetal fibronectin test combined with other clinical risk factors for preterm birth. Three primary outcomes were defined: fetal death or birth <34 weeks (obstetric), a composite of death, brain injury, or bronchopulmonary dysplasia (neonatal), and a standardized cognitive score at 2 years of age (childhood). After adjusting for multiplicity (i.e. overall type I error for multiple outcomes) progesterone was not found to have a significant benefit on the three primary outcomes. In the subgroup of women with a history of sPTB (N=903), there were no

[41] Fonseca EB, et al. Progesterone and the risk of preterm birth among women with a short cervix. N Engl J Med 2007;357:462-9.

[42] Hassan SS, et al. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial. Ultrasound Obstet Gynecol 2011; 38: 18–31.

[43] Norman JE, et al. Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial. Lancet 2016; 387: 2106–16.

significant differences in the rate of sPTB prior to 34 weeks gestation between the progesterone and placebo groups (odds ratio: 0.82, 95% confidence interval 0.58 to 1.16).

- The 2017 Crowther et al. publication reported findings of the PROGRESS trial, an international trial that randomized 787 women with a singleton or twin pregnancy and a history of sPTB <37 weeks gestation in a 1:1 ratio to vaginal progesterone pessary (100 mg) or placebo.[44] Women were asked to self-administer a vaginal pessary (equivalent to 100 mg vaginal progesterone as active substance) daily from 20 weeks gestation until 34 weeks or delivery. Progesterone treatment had no benefit on the primary outcome of neonatal respiratory distress syndrome (RDS) or other neonatal and maternal morbidities related to preterm birth. Progesterone treatment also had no effect on the incidence of PTB at <37 weeks gestation, a secondary outcome (37% in both treatment groups).

These randomized, placebo-controlled clinical trials enrolled women with varying risk factors for PTB, evaluated different vaginal progesterone doses and formulations, and assessed different outcome measures. Overall, the evidence from these publications does not suggest that vaginal progesterone is beneficial in reducing the risk of preterm birth in women with a history of PTB. Note that FDA has not approved vaginal progesterone for indications related to preterm birth.

3.2. Meta-Analyses

Two published meta-analyses of clinical trials studied the efficacy of progesterone on reducing the risk of PTB: Romero et al. (2018)[45] and Dodd et al. (2013)[46] (Table 11). This section summarizes the meta-analyses, discusses the limitations of each meta-analysis and the regulatory utility of these meta-analyses in supporting the efficacy of Makena. To be consistent with the coprimary endpoint used in Trial 003, we focus on PTB <35 weeks and neonatal composite index.[47]

[44] Crowther et al. 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 Sep 26;14(9):e1002390.

[45] Romero R, et al. Vaginal progesterone for preventing preterm birth and adverse perinatal outcomes in singleton gestations with a short cervix: a meta-analysis of individual patient data. Am J Obstet Gynecol 2018;218(2): 161-180.

[46] Dodd, Jodie M., et al. Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth. *Cochrane Database of Systematic Reviews*7 (2013).

[47] The components of neonatal composite index include neonatal death prior to discharge, grade 3/4 intraventricular hemorrhage, respiratory distress syndrome, bronchopulmonary dysplasia, necrotizing enterocolitis, and proven sepsis.

Table 11: Comparison of Study Designs

	Trial 003	Romero et al.	Dodd et al.
Number of subjects (Number of studies)	HPC (Makena): 1,130 Vehicle: 578 (1 RCT)	Progesterone: 498 Placebo: 476 (5 RCTs)	Progesterone: 1,029 Placebo: 869 (11 RCTs)
Study population	Women with singleton birth and history of spontaneous PTB	Women with singleton birth and short cervix	Women with singleton birth and history of spontaneous PTB
Dose	250 mg weekly	90-100 or 200 mg daily	<500 mg weekly or ≥500 mg weekly
Administration	Intramuscular	Intravaginal	Intramuscular, intravaginal, oral, intravenous
Number of subjects from the United States	HPC (Makena): 258 Placebo: 133	Progesterone: 115 Placebo: 117	No U.S. subjects

Source: Reviewer's table

Romero et al. (2018) assessed whether vaginal progesterone prevents PTB and improves perinatal outcomes in women with a singleton gestation and a mid-second trimester, sonographic short cervix (cervical length ≤25 mm). The authors defined a composite neonatal morbidity and mortality[48] outcome. The doses were either 90-100 mg/day or 200 mg/day by intravaginal administration. The authors performed a meta-analysis and estimated the pooled relative risk (RR) with an associated 95% confidence interval (CI). An additional post-hoc subgroup analysis was conducted using an interaction test to examine whether intervention effects differ between the country of enrollment (United States versus other countries). When the heterogeneity of treatment effect was substantial (I^2 >30%), the results were pooled using a random-effect model. Otherwise, a fixed-effect model was used.

The authors' meta-analysis included 5 studies (498 progesterone subjects versus 476 placebo subjects). The meta-analysis showed that vaginal progesterone significantly reduced the risk of PTB <35 weeks (RR [95% CI] = 0.72 [0.58–0.89]) and the risk of composite neonatal morbidity and mortality (RR [95% CI] = 0.59 [0.38–0.91]). A subgroup analysis compared the risk of PTB <33 weeks (PTB <35 weeks and composite neonatal morbidity and mortality not available) between women enrolled from the United States (RR [95% CI] = 0.73 [0.42–1.27]) and women from other countries (RR [95% CI] = 0.59 [0.43–0.80]). The interaction test for subgroup difference did not show significant difference (p = 0.51). Romero et al. included similar proportions of Caucasian subjects (37.2% vs. 39.7%, progesterone and placebo, respectively) and black subjects (36.3% vs. 37.0%, progesterone and placebo, respectively). The subgroup analysis for reduction of PTB among black subjects had a 95% confidence interval that crossed 1 (RR [95% CI] = 0.86 [0.58–1.26]), whereas that of Caucasian subjects had a 95% confidence interval that excluded 1 (RR [95% CI] = 0.45 [0.28–0.73]).

This meta-analysis included subjects with various dose levels (90-100 or 200 mg per day) and the analysis was mainly driven by 3 large studies. In addition, the meta-analysis was underpowered to evaluate interactions. Although both Trial 003 and Romero et al. included

[48] The only difference between neonatal composite index and composite neonatal morbidity and mortality is whether the intraventricular hemorrhages are restricted to grade 3/4 or all grades, respectively.

women with a singleton pregnancy, subjects of Trial 003 had a high prevalence of spontaneous PTB history (100%) with a low prevalence of short cervix (1.6%), while 30% of subjects in the Romero et al. meta-analysis had a history of sPTB history with a high prevalence of short cervix (100%). Romero et al. does not provide information for the approved dose of 250 mg per week administered by intramuscular injection. Because of the difference in study population, formulation, dose levels, and route of administration in Romero et al., the characteristics of the trials in this meta-analysis are not comparable to Trial 003 and the meta-analysis findings do not inform the efficacy of Makena.

Dodd et al. (2013) assessed the benefits and risks of progesterone for the prevention of PTB for women considered to be at increased risk of PTB. This article did not provide a composite neonatal outcome. However, components of the neonatal composite index, except bronchopulmonary dysplasia, were available. The authors performed a meta-analysis and estimated the pooled RR with an associated 95% CI. A random-effect model was employed when the heterogeneity of treatment effect was substantial (I^2 >30%). Otherwise, a fixed-effect model was used.

We focused on the results from the indicated population, women with a singleton pregnancy and history of spontaneous PTB. The authors dichotomized the weekly cumulative dose to either <500 mg or ≥500 mg per week, and the drug was administered through multiple routes: intramuscular, intravaginal, oral, and intravenous. The authors used a total of 11 clinical studies (1,029 progesterone subjects versus 869 placebo subjects) to conduct a meta-analysis in the indicated population. Not all 11 studies were used to analyze the outcomes. Because the result using an outcome of PTB <35 weeks of gestation was not available, we used the authors' outcome of PTB <34 weeks, which concluded that progesterone significantly reduced the risk of PTB (5 studies; RR [95% CI] = 0.31 [0.14–0.69]). The authors reported that neonatal death (6 studies; RR [95% CI] = 0.45 [0.27–0.76]) and necrotizing enterocolitis (3 studies; RR [95% CI] = 0.30 [0.10–0.89]) showed significant risk reduction.

The analysis using 5 studies to estimate the risk of PTB <34 weeks included subjects treated with multiple dose levels and routes of administration. Therefore, the treatment effect of the indicated dose (250 mg) and administration route is unclear. The I^2 from the five studies indicated substantial heterogeneity (I^2 = 56%), raising concerns of whether the trials were too different to be incorporate into the meta-analysis.

Compared to Trial 003, Dodd et al. neither studied the approved dose (250 mg weekly) nor used the intramuscular injection only for administration. Therefore, this meta-analysis is not directly comparable to Trial 003, providing limited inference from the pooled estimate of the treatment effect. None of the five pooled studies that estimated PTB<34 weeks were conducted in the United States; study sites were Iran, Turkey, Brazil, and India.

The two meta-analyses combined different patient populations, formulations, doses and routes of administration. Thus, these studies did not investigate Makena's indicated population, dose, and route of administration and are not comparable to Trial 003. In addition, we do not have access to the patient-level data, individual study protocols and study reports. Because of issues with the

relevancy and the unknown quality of these meta-analyses, the utility of these meta-analyses is limited in addressing the efficacy of Makena.

4. Safety

In Trial 002, total fetal/neonatal deaths included miscarriages (delivery from 16^0 up through 19^6 weeks, stillbirths ([antepartum or intrapartum death] from 20 weeks gestation through term) and neonatal deaths (death of a liveborn born from 20 weeks gestation through term). Of concern was the numerically higher rate of miscarriages and stillbirths in Trial 002. The number of these events were small, and no clear conclusions about the effect of HPC on this safety concern could be made. Trial 003 was powered to exclude a doubling of the risk of fetal/early infant deaths, the primary safety outcome. Fetal/early infant deaths were comprised of the following:

- Spontaneous abortion/miscarriage (delivery from 16^0 up through 19^6 weeks), and
- Stillbirth (antepartum or intrapartum death) from 20 weeks gestation through term, and
- Early fetal death (from minutes after birth until 28 days of life) occurring in livebirths born at < 24 weeks gestation

Fetal and early infant death data from Trial 002 and Trial 003 are juxtaposed in Table 12 and pooled results from both trials are shown in Table 13. Note that the "early fetal death," as defined in 003, was not analyzed as such in Trial 002. The results for "early fetal death" for Trial 002 in Table 12 and Table 13 were analyzed post-hoc for this efficacy supplement. As shown in Table 12, Trial 003 excluded a doubling of the risk of fetal/early infant deaths for Makena (upper bound of 95% was 1.81). When the data from Trial 002 and 003 were pooled, there was no difference in the overall incidence of fetal/early infant deaths with Makena compared to placebo in either trial. There appeared to be a trend toward an increase in stillbirths in both trials; however, the numbers are small, precluding reliable determination of risk. The pooled data from Trials 002 and 003 showed similar results.

Table 12: Fetal and Early Infant Deaths in Trial 002 and Trial 003 (Safety Population)

	Trial 002			Trial 003		
Safety Outcomes N[a], n[b] (%)	**Makena N=310**	**Placebo N=153**	**RR[c] (95% CI)**	**Makena N=1130**	**Placebo N=578**	**RR 95% CI**
Total fetal/early infant deaths[e]	15 (4.8%)	6 (3.9%)	1.22 (0.48, 3.1)	19 (1.7%)	11 (1.9%)	0.87 (0.42, 1.81)
Miscarriages (<20 weeks)	5 (2.4%)	0	N/A	4 (0.5%)	6 (1.3%)	0.32 (0.09, 1.14)
Stillbirths (≥20 weeks)	6 (2.0%)	2 (1.3%)	1.52 (0.31, 7.52)	12 (1.1%)	3 (0.5%)	2.07 (0.59, 7.29)
Early infant deaths	4 (1.3%)	4 (2.6%)	0.49 (0.13, 1.92)	3 (0.3%)	2 (0.4%)	0.73 (0.12, 4.48)

Abbreviations: RR = relative risk, calculated for 17-HPC relative to placebo; CI = confidence interval

[a]N = number of subjects in the Intent to Treat Population in the specified treatment group. The safety population consists of all subjects who received any amount of study medication.

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

[c]Relative risk of fetal/early infant death for Makena relative to placebo and is for the Cochran-Mantel-Haenszel test adjusted for gestational age at randomization

[e] Defined as spontaneous abortion/miscarriage, stillbirth, and early fetal death (from minutes after birth until 28 days of life) occurring in livebirths born at <24 weeks gestation

Source: Applicant's analysis (submitted September 25, 2019)

Table 13: Fetal and Early Infant Deaths in Trial 002 and Trial 003 Subjects Combined (Safety Population)

Safety outcomes Na, nb (%)	Trials 002 and 003 Combined		
	Makena N = 1438	Placebo N = 731	RR (95% CI)
Total fetal/neonatal deaths[e]	34 (2.4%)	17 (2.3%)	1.01 (0.57, 1.79)
Miscarriages (<20 weeks)	n = 1075 9 (0.8%)	n = 555 6 (1.1%)	0.73 (0.26, 2.04)
Stillbirths (≥20 weeks)	n = 1429 18 (1.3%)	n = 724 5 (0.7%)	1.86 (0.69, 4.99)
Early infant deaths	n = 1411 7 (0.5%)	n = 720 6 (0.8%)	0.58 (0.20, 1.73)

Source: Applicant's analysis (submitted September 25, 2019)

Birth at 24 weeks is traditionally considered to be the threshold for viability for a preterm neonate, and the Applicant counted only deaths in livebirths born < 24 weeks (early infant death) in the primary safety outcome. FDA, however, considers deaths occurring from minutes after birth until 28 days of life in livebirths born ≥ 20 weeks gestation (neonatal deaths) to be an important safety measurement. These results on fetal and neonatal deaths from Trial 002 and Trial 003 are juxtaposed in Table 14 and pooled results from both trials are shown in Table 15. Overall, these findings are consistent with those above.

Table 14: Fetal and Neonatal Deaths in Trial 002 and Trial 003 (Safety Population)

	Trial 002			Trial 003		
Safety Outcomes Na, nb (%)	Makena N=310	Placebo N=153	RR[c] (95% CI)	Makena N=1130	Placebo N=578	RR 95% CI
Total fetal/neonatal deaths[c]	19 (6.1%)	11 (7.2%)	0.83 (0.41, 1.70)	22 (2.0%)	13 (2.2%)	0.85 (0.43, 1.67)
Miscarriages (<20 weeks)	5 (2.4%)	0	N/A	4 (0.5%)	6 (1.3%)	0.32 (0.09, 1.14)
Stillbirths (≥20 weeks)	6 (2.0%)	2 (1.3%)	1.52 (0.31, 7.52)	12 (1.1%)	3 (0.5%)	2.07 (0.59, 7.29)
Neonatal deaths	8 (2.7%)	9 (6.0%)	0.44 (0.18, 1.12)	6 (0.5%)	4 (0.7%)	0.73 (0.21, 2.58)

[a]N = number of subjects in the Intent to Treat Population in the specified treatment group. The safety population consists of all subjects who received any amount of study medication.

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

[c] Defined as spontaneous abortion/miscarriage, stillbirth, and neonatal death (from minutes after birth until 28 days of life) occurring in livebirths born ≥ 20 weeks gestation

Source: Applicant's analysis (submitted September 27, 2019)

Table 15: Fetal and Neonatal Deaths in Trial 002 and Trial 003 Subjects Combined (Safety Population)

Safety outcomes N[a], n[b] (%)	Trials 002 and 003 Combined		
	Makena N = 1438	Placebo N = 731	RR (95% CI)
Total fetal/neonatal deaths[c]	41 (2.9%)	24 (3.3%)	0.85 (0.52, 1.40)
Miscarriages (<20 weeks)	n = 1075 9 (0.8%)	n = 555 6 (1.1%)	0.73 (0.26, 2.04)
Stillbirths (≥20 weeks)	n = 1429 18 (1.3%)	n = 724 5 (0.7%)	1.86 (0.69, 4.99)
Neonatal deaths	n = 1411 14 (1.0%)	n = 720 13 (1.8%)	0.54 (0.25, 1.31)

[a]N = number of subjects in the Intent to Treat Population in the specified treatment group. The safety population consists of all subjects who received any amount of study medication.

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

[c] Defined as spontaneous abortion/miscarriage, stillbirth, and neonatal death (from minutes after birth until 28 days of life) occurring in livebirths born ≥ 20 weeks gestation

Source: Applicant's analysis (submitted September 27, 2019)

In Trial 003, the same proportion of subjects in each treatment group (3%) experienced serious treatment-emergent adverse event (TEAE) or maternal pregnancy complications (MPC). The most frequently reported serious TEAE or MPC for subjects treated with Makena were premature separation of placenta (5 subjects, 0.4%), placental insufficiency (4 subjects, 0.4%), and pneumonia (3 subjects, 0.3%). The most frequently reported serious TEAE or MPC for subjects treated with placebo were cholestasis (3 subjects, 0.5%) and premature separation of placenta (2 subjects, 0.3%).

Table 16: Most Common (≥ 2 subjects Overall) Serious TEAE and MPC by Preferred Term in Trial 003 (Safety Population)

Preferred Term	Makena N = 1128 N (%)	Placebo N = 578 N (%)
Subjects with at least one serious TEAE/MPC	**34 (3%)**	**18 (3%)**
Cholestasis	0 (0)	3 (0.5)
Endometritis	1 (0.1)	1 (0.2)
Escherichia sepsis	2 (0.2)	0 (0)
Migraine	1 (0.1)	1 (0.2)
Placental insufficiency	4 (0.4)	1 (0.2)
Pneumonia	3 (0.3)	0 (0)
Premature separation of placenta	5 (0.4)	2 (0.3)
Pyelonephritis	2 (0.2)	1 (0.2)
Wound infection	2 (0.2)	0 (0)

Although the number of fetal and neonatal deaths are too low to draw definitive conclusions, the findings of this safety outcome appear to be similar between placebo and Makena. Otherwise, the safety profile of Makena remains unchanged.

5. Appendix

Table 17: Estimated Annual Number of 15- to 44-Year-Old Patients With Dispensed Prescriptions for Hydroxyprogesterone or Progesterone Products, Stratified by Molecule and Form, From U.S. Retail or Mail Order/Specialty Pharmacies 2014-2018

	2014		2015		2016		2017		2018	
	Patients (N)	**%**	**Patients (N)**	**%**	**Patients (N)**	**%**	**Patients (N)**	**%**	**Patients (N)**	**%**
Total Patients (Hydroxyprogesterone and Progesterone)*	478,567	100%	492,992	100%	513,900	100%	546,499	100%	559,985	100%
All Hydroxyprogesterone	8,039	2%	12,581	3%	25,477	5%	38,744	7%	42,320	8%
Makena®	8,035	100%	12,581	100%	25,126	99%	37,581	97%	31,684	75%
Generic Hydroxyprogesterone Caproate	0	0%	0	0%	117	<1%	769	2%	12,325	29%
All Progesterone Products	471,252	98%	481,858	98%	491,869	96%	510,955	93%	520,992	93%
Progesterone (Oral)	341,067	72%	358,172	74%	377,479	77%	403,335	79%	427,085	82%
Progesterone (Injectable)	94,578	20%	96,532	20%	100,647	20%	102,199	20%	113,736	22%
Progesterone (Vaginal)	117,579	25%	107,735	22%	96,986	20%	89,305	17%	77,378	15%

* Prescriptions for bulk powder forms of hydroxyprogesterone and progesterone were not included.
Source: Symphony Health IDV® Integrated Dataverse. Data years 2014-2018. Extracted August 2019. File: SH UPC Progesterone and Hydroxyprogesterone Pt 08-07-2019.xlsx.
Unique patient counts should not be added across time periods or drug categories due to the possibility of double counting those patients who received multiple products within the same calendar year or over multiple periods in the study. Generic hydroxyprogesterone caproate use in 2016 and 2017 were generic Delalutin products.

Table 18: Diagnoses Associated With the Estimated Number of Progesterone or Hydroxyprogesterone Use Mentions Among 15- to 44-Year-Old Women From U.S. Office-Based Physician Surveys, 2013 Through 2018, Aggregated

January 2013 - December 2018

	Uses (000)	95% CI (000)	% Share
Total Progesterone and Hydroxyprogesterone	**3,786**	**3,401-4,172**	**100%**
Hydroxyprogesterone Inj	**1,592**	**1,342-1,842**	**42%**
O09 Supervision of high-risk pregnancy	797	620-973	50%
Z87.51 Personal history of preterm labor	324	211-437	20%
Z34 Encounter for supervision of normal pregnancy	211	120-302	13%
O60 Preterm labor in current pregnancy	158	79-237	10%
O34 Maternal care for abnormality of pelvic organs	28	<0.5-61	2%
All Others	75	21-130	5%
Progesterone (all forms)	**2,194**	**1,901-2,488**	**58%**
Progesterone oral	**677**	**514-840**	**31%**
O20 Hemorrhage in early pregnancy	80	24-136	12%
N97 Female infertility	79	23-134	12%
Z34 Encounter for supervision of normal pregnancy	68	17-120	10%
N91 Absent, scanty and rare menstruation	68	16-119	10%
O26 Maternal care for pregnancy-related conditions	64	14-114	9%
All Others	318	206-430	47%
Progesterone injectable	**416**	**288-543**	**19%**
O09 Supervision of high-risk pregnancy	173	91-256	42%
N97 Female infertility	169	87-250	41%
O20 Hemorrhage in early pregnancy	41	1-81	10%
O60 Preterm labor in current pregnancy	17	<0.5-43	4%
O34 Maternal care for abnormality of pelvic organs	9	<0.5-28	2%
All Others	7	<0.5-23	2%
Progesterone vaginal	**1,054**	**851-1,258**	**48%**
N97 Female infertility	622	466-779	59%
O09 Supervision of high-risk pregnancy	125	55-195	12%
O20 Hemorrhage in early pregnancy	121	52-190	11%
O26 Maternal care for pregnancy-related conditions	105	41-170	10%
N96 Recurrent pregnancy loss	45	3-87	4%
All Others	36	<0.5-73	3%

Source: Syneos Health Research and Insights, TreatmentAnswers™ with Pain Panel. Data years 2013-2018. Extracted July 2019. File: Progesterone and Hydroxyprogesterone products by diagnosis 07-22-2019.xlsx. Diagnosis data are not directly linked to dispensed prescriptions but obtained from surveys of a sample of 3,200 office-based physicians reporting on patient activity one day a month. Drug use mentions below 100,000 may not represent reliable estimates of use and should be interpreted with caution because the sample size may be very small with corresponding large confidence intervals.

Table 19: Estimated Drug Use Mentions Among 15- to 44-Year-Old Women Associated With Selected Diagnoses From U.S. Office-Based Physician Surveys, 2013-2018, Aggregated

January 2013 through December 2018

	Uses (000)	95% CI Uses (000)	Share %
Current/history preterm labor or cervical shortening	**2,364**	**2,059-2,668**	**100%**
History of preterm labor (O09.21X, Z87.51)	1,277	1,054-1,501	54%
Makena	539	394-685	42%
17-Alpha Hydroxyprogesterone	290	184-397	23%
Hydroxyprogesterone	112	46-178	9%
Prenatal OTC	88	29-146	7%
Prenatal Rx	73	19-126	6%
All Others	175	92-258	14%
Preterm labor in current pregnancy (O60.XXX)	936	744-1,127	40%
Nifedipine	172	90-254	18%
Makena	135	62-207	14%
Procardia	132	60-203	14%
Terbutaline Inj	85	27-143	9%
Betamethasone Inj	75	21-129	8%
All Others	338	223-453	36%
Cervical shortening (O26.87X)	151	74-228	6%
Progesterone vaginal	73	20-127	48%
Prometrium	60	11-109	40%
Prochieve	11	<0.5-32	7%
Crinone	7	<0.5-23	5%

Source: Syneos Health Research and Insights, TreatmentAnswers™ with Pain Panel. Data years 2013-2018. Extracted July 2019. File: Progesterone and Hydroxyprogesterone products by diagnosis 07-22-2019.xlsx. Diagnosis data are not directly linked to dispensed prescriptions but obtained from surveys of a sample of 3,200 office-based physicians reporting on patient activity one day a month. Drug use mentions below 100,000 may not represent reliable estimates of use and should be interpreted with caution because the sample size may be very small with corresponding large confidence intervals.

Table 20: Comparison of Demographics and Baseline Characteristics: Studies 002 and 003

	Trial 003		Trial 003 U.S. subset		Trial 002	
Variable	**Makena (N=1130)**	**Placebo (N=578)**	**Makena (N=258)**	**Placebo (N=133)**	**Makena (N=310)**	**Placebo (N=153)**
Gestational age of qualifying delivery, weeks	31.3 ± 4.4	31.6 ± 4.2	32.5 ± 3.9	32.5 ± 3.9	30.6 ± 4.6	31.3 ± 4.2
Number of previous preterm deliveries						
1 previous PTB, N (%)	964 (85)	494 (86)	187 (72)	97 (73)	224 (72)	90 (59)
>1 previous PTB, N (%)	166 (15)	82 (14)	71 (28)	36 (27)	86 (28)	63 (41)
Number with cervical length <25 mm at randomization, N (%)	18 (2)	9 (2)	13 (5)	3 (2)	NA	NA
Age, years	30 ± 5	30 ± 5	28 ± 5	27 ± 5	26 ± 6	27 ± 5
Race, N (%)						
Black or African American/African Heritage	73 (6)	41 (7)	72 (28)	41 (31)	183 (59)	90 (59)
White	1004 (89)	504 (87)	170 (66)	84 (63)	79 (25)	34 (22)
Asian	23 (2)	22 (4)	4 (2)	2 (2)	2 (1)	1 (1)
Other	30 (3)	11 (2)	12 (5)	6 (5)	3 (1)	2 (1)
Ethnicity, N (%)						
Hispanic or Latino	101 (9)	54 (9)	31 (12)	23 (17)	43 (14)**	26 (17)**
Non-Hispanic or Latino	1029 (91)	524 (91)	227 (88)	110 (83)	267 (86)	127 (83)
Marital Status, N (%)						
Married or living with partner	1013 (90)	522 (90)	180 (70)	91 (68)	159 (51)	71 (46)
Never married	86 (8)	40 (7)	61 (24)	33(25)	119 (38)	64 (42)
Divorced, widowed or separated	31 (3)	16 (3)	17 (7)	9 (7)	32 (10)	18 (12)
BMI before pregnancy	24.3 ± 7.1	24.7 ± 8.7	27.4 ± 11.8	29.3 ± 15.3	26.9 ± 7.9	26.0 ± 7.0
Years of education	13 ± 2	13 ± 2	13 ± 2	13 ± 2	12 ± 2	12 ± 2
Any substance use during pregnancy, N (%)	105 (9)	51 (9)	69 (27)	40 (30)	85 (27)	36 (24)
Smoking	92 (8)	40 (7)	58 (22)	31 (23)	70 (23)	30 (20)
Alcohol	23 (2)	18 (3)	20 (8)	16 (12)	27 (9)	10 (7)
Illicit drugs	15 (1)	8 (1)	15 (6)	8 (6)	11 (4)	4 (3)

**Hispanic or Latino included in both race and ethnicity category for Study 002

Table 21: Summary of Neonatal Composite Index by Subgroups

Neonatal Composite Index, Subgroup	Trial 003		Trial 003 U.S. subset		Trial 002	
	Makena (N=1091)	Placebo (N=560)	Makena (n=252)	Placebo (n=126)	Makena (N=295)	Placebo (N=151)
GA at randomization (weeks)						
16^0-17^6	25/481 (5.2)	12/230 (5.2)	4/93 (4.3)	4/36 (11.1)	12/97 (12.4)	11/47 (23.4)
18^0-20^6	34/610 (5.6)	17/330 (5.2)	14/159 (8.8)	8/90 (8.9)	23/198 (11.6)	15/104 (14.4)
Overall	59/1091 (5.4)	29/560 (5.2)	18/252 (7.1)	12 /126 (9.5)	35/295 (11.9)	26/151 (17.2)
GA of qualifying delivery* (weeks)						
20^0 - $<28^0$	17/221 (7.7)	3/97 (3.1)	3/30 (10.0)	2/17 (11.8)	11/74 (14.9)	9/29 (31.0)
28^0 - $<32^0$	14/198 (7.1)	13/102 (12.7)	3/37 (8.1)	4/18 (22.2)	5/65 (7.7)	5/30 (16.7)
32^0 - $<35^0$	15/339 (4.4)	9/182 (4.9)	3/73 (4.1)	5/39 (12.8)	11/79 (13.9)	9/54 (16.7)
35^0 - $<37^0$	13/330 (3.9)	4/176 (2.3)	9/110 (8.2)	1/51 (2.0)	8/77 (10.4)	3/38 (7.9)
GA of earliest prior PTB** (weeks)						
0 - $<20^0$	24/445 (5.4)	11/228 (4.8)	5/75 (6.7)	3/35 (8.6)	6/46 (13.0)	1/16 (6.3)
20^0 - $<28^0$	13/153 (8.5)	2/71 (2.8)	4/27 (14.8)	1/18 (5.6)	10/47 (21.3)	9/23 (39.1)
28^0 - $<32^0$	9/112 (8.0)	7/59 (11.9)	2/29 (6.9)	3/13 (23.1)	4/39 (10.3)	4/20 (20.0)
32^0 - $<35^0$	7/198 (3.5)	6/99 (6.1)	2/59 (3.4)	4/29 (13.8)	8/55 (14.5)	6/34 (17.6)
35^0 - $<37^0$	6/183 (3.3)	3/102 (2.9)	5/62 (8.1)	1/31 (3.2)	5/40 (12.5)	2/26 (7.7)
Previous PTB, N (%)						
1	43/933 (4.6)	22/478 (4.6)	11/184 (6.0)	8/92 (8.7)	18/210 (8.6)	10/89 (11.2)
>1‡	16/158 (10.1)	7/80 (8.8)	7/78 (9.0)	4/34 (11.8)	17/85 (10.0)	16/62 (25.8)
2	14/125 (11.2)	5/66 (7.6)	6/52 (11.5)	4/28 (14.3)	12/55 (21.8)	8/45 (17.8)
≥3	2/33 (6.1)	2/14 (14.3)	1/16 (6.3)	0/6 (0.0)	5/30 (16.7)	8/17 (47.1)
Cervical length at randomization***, N (%)						
<25 mm	2/17 (11.8)	2/9 (22.2)	1/13 (7.7)	1/3 (33.3)	NA	NA
≥25 mm	44/890 (4.9)	23/444 (5.2)	11/110 (10.0)	10/63 (15.9)	NA	NA
BMI before pregnancy (kg/m^2)						
Underweight (<18.5)	4/80 (5.0)	3/37 (8.1)	0/11 (0)	0/2 (0)	4/25 (16.0)	2/10 (20.0)
Normal (18.5 - <25)	34/629 (5.4)	12/328 (3.7)	7/112 (6.3)	2/49 (4.1)	13/116 (11.2)	14/73 (19.2)
Overweight (25 - <30)	10/249 (4.0)	9/125 (7.2)	6/63 (9.5)	6/34 (17.6)	6/56 (10.7)	5/30 (16.7)
Obese (≥30)	11/133 (8.3)	5/69 (7.2)	5/66 (7.6)	4/41 (9.8)	10/86 (11.6)	5/34 (14.7)

Neonatal Composite Index, Subgroup	Trial 003		Trial 003 U.S. subset		Trial 002	
	Makena (N=1091)	Placebo (N=560)	Makena (n=252)	Placebo (n=126)	Makena (N=295)	Placebo (N=151)
Any substance use during pregnancy, N (%)						
Yes	8/101 (7.9)	5/49 (10.2)	5/67 (7.5)	4/38 (10.5)	12/82 (14.6)	6/35 (17.1)
No	51/990 (5.2)	24/511 (4.7)	13/185 (7.0)	8/88 (9.1)	23/213 (10.8)	20/116 (17.2)
Smoking						
Yes	8/89 (9.0)	4/39 (10.3)	5/57 (8.8)	3/29 (10.3)	10/67 (14.9)	6/29 (20.7)
No	51/1002 (5.1)	25/521 (4.8)	13/195 (6.7)	9/97 (9.3)	25/228 (11.0)	20/122 (16.4)
Alcohol						
Yes	0/23 (0)	4/17 (23.5)	0/19 (0)	3/15 (20.0)	3/26 (11.5)	0/10 (0)
No	59/1068 (5.5)	25/543 (4.6)	18/233 (7.7)	9/111 (8.1)	32/269 (11.9)	26/141 (18.4)
Illicit drugs						
Yes	1/14 (7.1)	1/7 (14.3)	1/13 (7.7)	1/7 (14.3)	2/10 (20.0)	0/4 (0)
No	58/1077 (5.4)	28/553 (5.1)	17/239 (7.1)	11/119 (9.2)	33/285 (11.6)	26/147 (17.7)
Race						
Non-Hispanic black	6/69 (8.7)	3/39 (7.7)	5/68 (7.4)	3/39 (7.7)	22/176 (12.5)	20/89 (22.5)
Non-Hispanic non-black	50/923 (5.4)	23/468 (4.9)	13/153 (8.5)	7/64 (10.9)	8/81 (9.9)	6/36 (16.7)
Ethnicity						
Hispanic	3/99 (3.0)	3/53 (5.7)	0/31 (0)	2/23 (8.7)	5/38 (13.2)	0/26 (0)
Non-Hispanic	56/992 (5.6)	26/507 (5.1)	18/221 (8.1)	10/103 (9.7)	30/257 (11.7)	26/125 (20.8)
Years of education						
≤12	28/458 (6.1)	18/249 (7.2)	9/116 (7.8)	9/69 (13.0)	29/213 (13.6)	18/101 (17.8)
>12	31/632 (4.9)	11/311 (3.5)	9/135 (6.7)	3/57 (5.3)	6/82 (7.3)	8/50 (16.0)

* If more than one prior delivery was sPTB, qualifying delivery was the most recent.
** The earliest PTB may be indicated or spontaneous.
***Cervical length measurement was not captured for all subjects in a treatment group.
GA = gestational age
NA = not available
Source: Applicant Analysis; ‡FDA Analysis.

Table 22: Summary of PTB <35^0 Weeks by Subgroups

	Trial 003		Trial 003 U.S. Subset		Trial 02	
Stratification Groups, n/N (%)	**Makena (N=1130)**	**Placebo (N=578)**	**Makena (N=258)**	**Placebo (N=133)**	**Makena (N=310)**	**Placebo (N=153)**
GA at randomization (weeks)						
16^0-17^6	61/493 (12.4)	31/238 (13.0)	16/96 (16.7)	9/40 (22.5)	22/103 (21.4)	21/47 (44.7)
18^0-20^6	61/620 (9.8)	35/336 (10.4)	24/160 (15.0)	14/91 (15.4)	41/203 (20.2)	26/106 (24.5)
Overall	122/1113 (11.0)	66/574 (11.5)	40/256 (15.6)	23/131 (17.6)	63/306 (20.6)	47/153 (30.7)
GA of qualifying delivery* (weeks)						
20^0 - <28^0	29/229 (12.7)	9/101 (8.9)	7/31 (22.6)	3/18 (16.7)	21/82 (25.6)	13/29 (44.8)
28^0 - <32^0	24/201 (11.9)	20/104 (19.2)	9/37 (24.3)	4/18 (22.2)	12/65 (18.5)	6/30 (20.0)
32^0 - <35^0	36/344 (10.5)	24/186 (12.9)	9/75 (12.0)	10/40 (25.0)	12/81 (14.8)	18/55 (32.7)
35^0 - <37^0	32/336 (9.5)	13/180 (7.2)	14/111 (12.6)	6/54 (11.1)	18/78 (23.1)	10/39 (25.6)
GA of earliest prior PTB** (weeks)						
0 - <20^0	53/459 (11.5)	26/234 (11.1)	13/78 (16.7)	5/36 (13.9)	9/46 (19.6)	3/16 (18.8)
20^0 - <28^0	21/156 (13.5)	7/73 (9.6)	7/27 (25.9)	3/19 (15.8)	21/55 (38.2)	11/23 (47.8)
28^0 - <32^0	15/113 (13.3)	12/60 (20.0)	8/30 (26.7)	3/13 (23.1)	7/39 (17.9)	5/20 (25.0)
32^0 - <35^0	18/201 (9.0)	12/100 (12.0)	5/59 (8.5)	6/29 (20.7)	9/56 (16.1)	13/35 (37.1)
35^0 - <37^0	15/184 (8.2)	9/106 (8.5)	7/62 (11.3)	6/34 (17.6)	10/40 (25.0)	5/26 (19.2)
Previous PTD, N (%)						
1	80/949 (8.4)	51/491 (10.4)	22/185 (11.9)	17/96 (17.7)	37/220 (16.8)	19/90 (21.1)
>1‡	42/164 (25.6)	15/81 (18.5)	18/71 (25.3)	6/35 (17.1)	26/86 (30.2)	28/63 (44.4)
2	29/127 (22.8)	10/67 (14.9)	13/52 (25.0)	4/29 (13.8)	18/56 (32.1)	17/46 (37.0)
≥3	13/37 (35.1)	5/14 (35.7)	5/19 (16.3)	2/6 (33.3)	8/30 (26.7)	11/17 (64.7)
Cervical length at randomization***, N (%)						
<25 mm	4/18 (22.2)	4/9 (44.4)	2/13 (15.4)	1/3 (33.3)	NA	NA
≥25 mm	92/907 (10.1)	45/455 (9.9)	21/112 (18.8)	13/66 (19.7)	NA	NA
BMI before pregnancy						
Underweight (<18.5)	13/83 (15.7)	4/38 (10.5)	0/11 (0)	0/3 (0)	5/25 (20.0)	6/10 (60.0)
Normal (18.5 - <25)	59/637 (9.3)	33/335 (9.9)	20/112 (17.9)	10/51 (19.6)	23/131 (17.6)	26/77 (33.8)
Overweight (25 - <30)	29/255 (11.4)	16/127 (12.6)	9/66 (13.6)	6/34 (17.6)	14/60 (23.3)	10/32 (31.3)
Obese (≥30)	21/138 (15.2)	13/74 (17.6)	11/67 (16.4)	7/43 (16.3)	21/90 (23.3)	5/34 (14.7)

Stratification Groups, n/N (%)	Trial 003		Trial 003 U.S. Subset		Trial 02	
	Makena (N=1130)	Placebo (N=578)	Makena (N=258)	Placebo (N=133)	Makena (N=310)	Placebo (N=153)
Any substance use during pregnancy, N (%)						
Yes	19/105 (18.1)	13/51 (25.5)	11/69 (15.9)	10/40 (25.0)	16/85 (18.8)	16/36 (44.4)
No	103/1008 (10.2)	53/523 (10.1)	29/187 (15.5)	13/91 (14.3)	47/221 (21.3)	31/117 (26.5)
Smoking						
Yes	18/92 (19.6)	11/40 (27.5)	10/58 (17.2)	8/30 (26.7)	13/70 (18.6)	15/30 (50.0)
No	104/1021 (10.2)	55/534 (10.3)	30/198 (15.2)	15/101 (14.9)	50/236 (21.2)	32/123 (26.0)
Alcohol						
Yes	1/23 (4.3)	5/18 (27.8)	1/19 (5.3)	4/16 (25.0)	5/27 (18.5)	2/10 (20.0)
No	121/1090 (11.1)	61/556 (11.0)	39/237 (16.5)	19/115 (16.5)	58/279 (20.8)	45/143 (31.5)
Illicit drugs	2/15 (13.3)	3/8 (37.5)	2/14 (14.3)	3/8 (37.5)	2/11 (18.2)	0/4 (0)
Yes						
No	120/1098 (10.9)	63/566 (11.1)	38/242 (15.7)	20/123(16.3)	61/295 (20.7)	47/149 (31.5)
Race						
Non-Hispanic black	17/72 (23.6)	8/40 (20.0)	16/71 (22.5)	8/40 (20.0)	39/183 (21.3)	32/90 (35.6)
Non-Hispanic non-black	92/940 (9.8)	50/480 (10.4)	19/154 (12.3)	10/68 (14.7)	28/127 (22.0)	15/63 (23.8)
Ethnicity						
Hispanic	13/101 (12.9)	8/54 (14.8)	5/31 (16.1)	5/23 (21.7)	10/41 (24.4)	4/26 (15.4)
Non-Hispanic	109/1012 (10.8)	58/520 (11.2)	35/225 (15.6)	18/108 (16.7)	53/265 (20.0)	43/127 (33.9)
Years of education						
≤12	64/474 (13.5)	40/256 (15.6)	24/120 (20.0)	18/74 (24.3)	49/223 (22.0)	32/103 (31.1)
>12	58/639 (9.1)	26/318 (8.2)	16/136 (11.8)	5/57 (8.8)	14/83 (16.9)	15/50 (30.0)

* If more than one prior delivery was sPTB, qualifying delivery was the most recent.
** The earliest PTB may be indicated or spontaneous.
***Cervical length measurement was not captured for all subjects in a treatment group.
GA = gestational age
NA = not available
Source: Applicant Analysis. ǂFDA Analysis.

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) Roster

---

Chairperson

Vivian Lewis, MD
Expertise: Obstetrics and Gynecology
Term: 7/1/2014 – 6/30/2020
Vice Provost for Faculty Development & Diversity
Professor, Obstetrics and Gynecology
University of Rochester
137 Wallis Hall
Rochester, New York 14627

Designated Federal Officer

Kalyani Bhatt, BS, MS
Division of Advisory Committee and Consultant Management
Food and Drug Administration
10903 New Hampshire Avenue
Silver Spring, Maryland 20993
(301) 796-9001
Fax: (301) 847-8533
Email: BRUDAC@fda.hhs.gov

Douglas C. Bauer, MD
Expertise: Bone Medicine, Epidemiology, Biostatistics
Term: 9/22/2015 – 6/30/2020
Professor of Medicine and Epidemiology & Biostatistics
University of California, San Francisco
1545 Divisadero Street
San Francisco, California 94115

Toby Chai, MD
Expertise: Urology
Term: 7/1/2019 – 6/30/2023
Vice Chair of Research
Co-Director of Female Pelvic Medicine and Reconstructive Surgery Program
Department of Urology
Yale School of Medicine
P.O. Box 208058
New Haven, Connecticut 06520

James Q. Clemens MD, FACS, MSCI
Expertise: Urology
Term: 7/1/2018 – 6/30/2022
Professor of Urology
The University of Michigan Medical Center
3875 Taubman Center
1500 East Medical Center Drive, SPC 5330
Ann Arbor, Michigan 48109

Matthew T. Drake, MD, PhD
Expertise: Endocrinology, Diabetes, Metabolism, Nutrition
Term: 9/22/2015 – 6/30/2021
Associate Professor of Medicine
Chair, Metabolic Bone Disease Core Group
Division of Endocrinology
Mayo Clinic College of Medicine
200 First Street SW
Rochester, Minnesota 55905

Beatrice J. Edwards, MD, MPH, FACP
Expertise: Geriatric Medicine
Term: 9/30/2016 – 6/30/2020
Associate Professor
Department of General Internal Medicine
Division of Internal Medicine
University of Texas MD Anderson Cancer Center
1400 Pressler Street, Unit 1465
Houston, Texas 77030

Margery Gass, MD
Expertise: Obstetrics and Gynecology
Term: 7/28/2017– 6/30/2021
Consultant
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue North
Seattle, Washington 98109

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) Roster (cont.)

****Gerard G. Nahum, MD, FACOG**
Expertise: General Medicine
Term: 3/31/2016 – 10/31/2019
Vice President of Global Development, General Medicine
Women's Healthcare, Long-Acting Contraception, Medical Devices, and Special Projects
Bayer HealthCare Pharmaceuticals, Inc.
100 Bayer Boulevard
Parsippany, New Jersey 07054

Christian P. Pavlovich, MD
Expertise: Urology and Oncology
Term: 7/28/2017 – 6/30/2021
Director of Urologic Oncology and Professor of Urology and Oncology
James Buchanan Brady Urological Institute
Department of Urology
John Hopkins Bayview Medical Center, A-345
Suite 3200, 301 Building, 4940 Eastern Avenue
Baltimore, Maryland 21224

Gloria Richard-Davis, MD, MBA, NCMP, FACOG
Expertise: Obstetrics and Gynecology
Term: 8/29/2019 – 6/30/2023
Division Director, Reproductive Endocrinology and Infertility
University of Arkansas Medical Sciences
Department of Obstetrics and Gynecology
4301 W. Markham Street
Little Rock, Arkansas 72205

Pamela Shaw, PhD
Expertise: Biostatistics
Term: 7/28/2017 – 6/30/2021
Professor, Department of Biostatistics and Epidemiology
University of Pennsylvania School of Medicine
423 Guardian Drive, Room 606
Philadelphia, Pennsylvania 19104

*** Consumer Representative (vacant)**
**** Industry Representative (non-voting)**

Updated: September 23, 2019

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) Meeting
FDA White Oak Campus, Building 31 Conference Center, the Great Room (Rm. 1503)
10903 New Hampshire Avenue, Silver Spring, Maryland
October 29, 2019

AGENDA

The committee will discuss supplemental new drug application (sNDA 021945/S-023) for MAKENA (hydroxyprogesterone caproate injection, 250 milligrams per milliliter) manufactured by AMAG Pharmaceuticals. In 2011, MAKENA received approval under the accelerated approval pathway (21 CFR part 314, subpart H, and section 506(c) of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 356(c)) for reducing the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. MAKENA was shown in the preapproval clinical trial to reduce the proportion of women who delivered at less than 37 weeks gestation, a surrogate endpoint that FDA determined was reasonably likely to predict a clinical benefit of preterm birth prevention, such as improved neonatal mortality and morbidity. As required under 21 CFR 314.510, the Applicant conducted a postapproval confirmatory clinical trial to verify and describe clinical benefit. AMAG Pharmaceuticals has disclosed that this completed confirmatory trial did not demonstrate a statistically significant difference between the treatment and placebo arms for the co-primary endpoints of reducing the risk of recurrent preterm birth or improving neonatal mortality and morbidity. The committee will consider the trial's findings and the sNDA in the context of AMAG Pharmaceuticals' confirmatory study obligation.

8:15 a.m.	Call to Order and Introduction of Committee	**Vivian Lewis, MD** Chairperson, BRUDAC
8:25 a.m.	Conflict of Interest Statement	**Kalyani Bhatt, BS, MS** Designated Federal Officer, BRUDAC
8:30 a.m.	FDA Opening Remarks	**Christine Nguyen, MD** Deputy Director for Safety Division of Bone, Reproductive and Urologic Products (DBRUP) Office of Drug Evaluation III (ODE III) Office of New Drugs (OND), CDER, FDA
8:45 a.m.	**APPLICANT PRESENTATIONS**	**AMAG Pharmaceuticals, Inc.**
	Introduction	**Julie Krop, MD** Chief Medical Officer Executive Vice President, Development & Regulatory Affairs AMAG Pharmaceuticals, Inc.
	Clinical Background and Unmet Need	**Michelle Owens, MD** Professor and Medical Director School of Medicine Department of Obstetrics and Gynecology The University of Mississippi Medical Center

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

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

AGENDA (cont.)

APPLICANT PRESENTATIONS (CONT.)

	Meis Study Design and Results	**Baha Sibai, MD** Professor Department of Obstetrics, Gynecology, and Reproductive Sciences Investigator, MFMU University of Texas Health Science Center of Houston MFMU[1] Network
	PROLONG: Efficacy and Safety	**Laura Williams, MD, MPH** Sr. Vice President, Clinical Development & Biostatistics AMAG Pharmaceuticals, Inc.
	Prevention of Preterm Birth: Clinical Perspective	**Sean Blackwell, MD** Professor and Chair Department of Obstetrics, Gynecology, and Reproductive Sciences Principal Investigator, MFMU University of Texas Health Science Center of Houston MFMU[1] Network
	Conclusion	**Julie Krop, MD**
10:00 a.m.	Clarifying Questions to Applicant	
10:25 a.m.	**BREAK**	
10:35 a.m.	**FDA PRESENTATIONS**	
	Clinical Overview	**Barbara Wesley, MD, MPH** Medical Officer DBRUP, ODEIII, OND, CDER, FDA
	Efficacy in Confirmatory Trial 003	**Jia Guo, PhD** Statistical Reviewer Division of Biometrics 3 (DB3) Office of Biostatistics (OB) Office of Translational Sciences (OTS), CDER, FDA
	Hydroxyprogesterone Caproate (HPC) Utilization in the United States	**Huei-Ting Tsai, PhD** Epidemiologist Division of Epidemiology II (DEPI-II) Office of Pharmacovigilance and Epidemiology (OPE) Office of Surveillance and Epidemiology (OSE) CDER, FDA

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

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

AGENDA (cont.)

FDA PRESENTATIONS (CONT.)

	Summary Remarks	**Christina Chang, MD, MPH** Clinical Team Leader DBRUP, ODEIII, OND, CDER, FDA

11:40 a.m. Clarifying Questions to FDA

12:00 p.m. **LUNCH**

1:00 p.m. **OPEN PUBLIC HEARING**

2:00 p.m. Clarifying Questions to Applicant or FDA

2:20 p.m. **BREAK**

2:30 p.m. Questions to the Committee/Committee Discussion and Voting

5:00 p.m. **ADJOURNMENT**

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) Meeting
FDA White Oak Campus, Building 31 Conference Center, the Great Room (Rm. 1503)
10903 New Hampshire Avenue, Silver Spring, Maryland
October 29, 2019

MEETING ROSTER

DESIGNATED FEDERAL OFFICER (Non-Voting)

Kalyani Bhatt, BS, MS
Division of Advisory Committee and Consultant Management
Office of Executive Programs, CDER, FDA

BONE, REPRODUCTIVE AND UROLOGIC DRUGS ADVISORY COMMITTEE MEMBERS (Voting)

Douglas C. Bauer, MD
Professor of Medicine and Epidemiology & Biostatistics
University of California, San Francisco
San Francisco, California

Matthew T. Drake, MD, PhD
Associate Professor of Medicine
Chair, Metabolic Bone Disease Core Group
Division of Endocrinology
Mayo Clinic College of Medicine
Rochester, Minnesota

Vivian Lewis, MD
(Chairperson)
Vice Provost for Faculty Development & Diversity
Professor, Obstetrics and Gynecology
University of Rochester
Rochester, New York

Pamela Shaw, PhD
Associate Professor
Department of Biostatistics and Epidemiology
University of Pennsylvania School of Medicine
Philadelphia, Pennsylvania

TEMPORARY MEMBERS (Voting)

Jonathan M. Davis, MD
Vice-Chair of Pediatrics
Chief of Newborn Medicine
The Floating Hospital for Children at Tufts Medical Center
Professor of Pediatrics
Tufts University School of Medicine
Boston, Massachusetts

Ahizechukwu Eke, MD, MPH
Assistant Professor
Division of Maternal Fetal Medicine
Department of Gynecology & Obstetrics
Johns Hopkins University School of Medicine
Baltimore, Maryland

Annie Ellis
(Patient Representative)
White Plains, New York

Daniel Gillen, PhD
Professor and Chair, Statistics
University of California, Irvine
Irvine, California

MEETING ROSTER (cont.)

TEMPORARY MEMBERS (Voting) (cont.)

Kimberly Hickey, MD
Colonel, Medical Corps, US Army
Chief, Maternal Fetal Medicine
Walter Reed National Military Medical Center
Deputy Director, National Capital Consortium
Uniformed Services University of the Health Sciences
Bethesda, Maryland

Sally Hunsberger, PhD
Mathematical Statistician
Division of Clinical Research
National Institute of Allergy and Infectious Disease
National Institute of Health
Rockville, Maryland

Michael K. Lindsay, MD, MPH
Luella Klein Associate Professor
Chief, Gynecology and Obstetrics Service Grady Health Systems
Director, Division of Maternal Fetal Medicine
Emory University
Atlanta, Georgia

Michele Orza, ScD
(Acting Consumer Representative)
Chief of Staff
Patient-Centered Outcomes Research Institute (PCORI)
Washington, District of Columbia

Uma M. Reddy, MD, MPH
Professor, Department of Obstetrics, Gynecology and Reproductive Sciences
Division Chief, Maternal Fetal Medicine
Section Chief, Maternal Fetal Medicine of Yale New Haven Hospital
Program Director, Maternal-Fetal Medicine Fellowship
Department of Obstetrics, Gynecology and Reproductive Sciences
Yale School of Medicine
New Haven, Connecticut

Brian Smith MD, MPH, MHS
Samuel L. Katz Professor of Pediatrics
Division of Neonatal-Perinatal Medicine
Duke University Medical Center
Durham, North Carolina

Kelly Wade, MD, PhD, MSCE
Attending Neonatologist
Children's Hospital of Philadelphia (CHOP)
Associate Professor of Clinical Pediatrics
University of Pennsylvania
CHOP Newborn Care
Philadelphia, Pennsylvania

Deborah A. Wing, MD, MBA
Senior Client Partner
Los Angeles, California
Formerly, Professor of Obstetrics-Gynecology
Division of Maternal Fetal Medicine
University of California, Irvine
Orange, California

MEETING ROSTER (cont.)

ACTING INDUSTRY REPRESENTATIVE TO THE COMMITTEE (Non-Voting)

Venkateswar Jarugula, PhD
(Acting Industry Representative)
Executive Director
Translation Medicine
Novartis Institutes for Biomedical Research
East Hanover, New Jersey

FDA PARTICIPANTS (Non-Voting)

Christine Nguyen, MD
Deputy Director for Safety
Division of Bone, Reproductive and Urologic Products (DBRUP)
Office of Drug Evaluation III (ODE III)
Office of New Drugs (OND), CDER, FDA

Barbara Wesley, MD, MPH
Medical Officer
DBRUP, ODEIII, OND, CDER, FDA

Christina Chang, MD, MPH
Clinical Team Leader
DBRUP, ODEIII, OND, CDER, FDA

Jia Guo, PhD
Statistical Reviewer
Division of Biometrics 3 (DB3)
Office of Biostatistics (OB)
Office of Translational Sciences (OTS), CDER, FDA

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) Meeting
FDA White Oak Campus, Building 31 Conference Center, the Great Room (Rm. 1503)
10903 New Hampshire Avenue, Silver Spring, Maryland
October 29, 2019

QUESTIONS

---

1. **DISCUSSION:** Discuss the effectiveness of Makena on recurrent preterm birth and neonatal morbidity and mortality.

2. **DISCUSSION:** If a new confirmatory trial were to be conducted, discuss the study design, including control, dose(s) of study medication, efficacy endpoints and the feasibility of completing such a trial.

3. **DISCUSSION:** Discuss the potential consequences of withdrawing Makena on patients and clinical practice.

4. **VOTE:** Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal outcomes?

   Provide rationale for your vote.

5. **VOTE**: Based on the findings from Trial 002 and Trial 003, is there substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth?

   Provide rationale for your vote.

6. **VOTE:** FDA approval, including accelerated approval, of a drug requires substantial evidence of effectiveness, which is generally interpreted as clinically and statistically significant findings from two adequate and well-controlled trials, and sometimes from a single adequate and well-controlled trial. For drugs approved under the accelerated approval pathway based on a surrogate endpoint, the Applicant is required to conduct adequate and well-controlled postapproval trial(s) to verify clinical benefit. If the Applicant fails to conduct such postapproval trial(s) or if such trial(s) do not verify clinical benefit, FDA may, following an opportunity for a hearing, withdraw approval.

   Should FDA:

   A. Pursue withdrawal of approval for Makena
   B. Leave Makena on the market under accelerated approval and require a new confirmatory trial
   C. Leave Makena on the market without requiring a new confirmatory trial

   Provide rationale for your vote and discuss the following:

   - Vote (A) (withdraw approval) may be appropriate if you believe the totality of evidence does not support Makena's effectiveness for its intended use.

     - Discuss the consequences of Makena removal (if not previously discussed in Discussion point 3)

QUESTIONS (cont.)

- Vote (B) (require a new confirmatory trial) may be appropriate if you believe the totality of evidence supports Makena's effectiveness in reducing the risk of recurrent preterm birth, but that there is no substantial evidence of effectiveness on neonatal outcomes AND you believe that a new confirmatory trial is necessary and feasible.

    - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth, based on the surrogate endpoint of gestational age at delivery.

    - Also discuss key study elements, including study population, control, dose(s), and efficacy endpoints of the new confirmatory trial (if not previously discussed in Discussion point 2) and approaches to ensure successful completion of such a trial.

- Vote (C) (leave Makena on the market without a new confirmatory trial) may be appropriate if you believe Makena is effective for reducing the risk of recurrent preterm birth and that it is not necessary to verify Makena's clinical benefit in neonates.

    - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth and why it is not necessary to verify Makena's clinical benefits in neonates.

Food and Drug Administration
Center for Drug Evaluation and Research

Summary Minutes of the of the
Bone, Reproductive and Urologic Drugs Advisory Committee Meeting
October 29, 2019

Location: FDA White Oak Campus, Building 31 Conference Center, the Great Room (Rm. 1503), 10903 New Hampshire Avenue, Silver Spring, Maryland.

Topic: The committee discussed supplemental new drug application (sNDA 021945/S 023) for MAKENA (hydroxyprogesterone caproate injection, 250 milligrams per milliliter) manufactured by AMAG Pharmaceuticals. In 2011, MAKENA received approval under the accelerated approval pathway (21 CFR part 314, subpart H, and section 506(c) of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 356(c)) for reducing the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. MAKENA was shown in the preapproval clinical trial (Trial 002) to reduce the proportion of women who delivered at less than 37 weeks gestation, a surrogate endpoint that FDA determined was reasonably likely to predict a clinical benefit of preterm birth prevention, such as improved neonatal mortality and morbidity. As required under 21 CFR 314.510, the Applicant conducted a post approval confirmatory clinical trial (Trial 003) to verify and describe clinical benefit. AMAG Pharmaceuticals has disclosed that this completed confirmatory trial did not demonstrate a statistically significant difference between the treatment and placebo arms for the co-primary endpoints of reducing the risk of recurrent preterm birth at less than 35 weeks gestation or improving neonatal mortality and morbidity. The committee considered the trial's findings and the sNDA in the context of AMAG Pharmaceuticals' confirmatory study obligation.

These summary minutes for the Ocotber 29, 2019, meeting of the Bone, Reproductive and Urologic Drugs Advisory Committee of the Food and Drug Administration were approved on November 22, 2019.

I certify that I attended the November 22, 2019, meeting of the Bone, Reproductive and Urologic Drugs Advisory Committee of the Food and Drug Administration and that these minutes accurately reflect what transpired.

/S/	**/S/**
Kalyani Bhatt, BS, MS *Designated Federal Officer, BRUDAC*	Vivian Lewis, MD *Chairperson, BRUDAC*

Summary Minutes of the
Bone, Reproductive and Urologic Drugs Advisory Committee Meeting
October 29, 2019

The Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) of the Food and Drug Administration, Center for Drug Evaluation and Research met on October 29, 2019, at the FDA White Oak Campus, Building 31 Conference Center, the Great Room (Rm. 1503), 10903 New Hampshire Avenue, Silver Spring, Maryland. Prior to the meeting, the members and temporary voting members were provided the briefing materials from the FDA and AMAG Pharmaceuticals. The meeting was called to order by Vivian Lewis, MD (Chairperson). The conflict of interest statement was read into the record by Kalyani Bhatt, BS, MS (Designated Federal Officer). There were approximately 175 people in attendance. There were sixteen (16) Open Public Hearing (OPH) presentations.

A verbatim transcript will be available, in most instances, approximately ten to twelve weeks following the meeting date.

Agenda: The committee discussed supplemental new drug application (sNDA 021945/S-023) for MAKENA (hydroxyprogesterone caproate injection, 250 milligrams per milliliter) manufactured by AMAG Pharmaceuticals. In 2011, MAKENA received approval under the accelerated approval pathway (21 CFR part 314, subpart H, and section 506(c) of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 356(c)) for reducing the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. MAKENA was shown in the preapproval clinical trial (Trial 002) to reduce the proportion of women who delivered at less than 37 weeks gestation, a surrogate endpoint that FDA determined was reasonably likely to predict a clinical benefit of preterm birth prevention, such as improved neonatal mortality and morbidity. As required under 21 CFR 314.510, the Applicant conducted a post approval confirmatory clinical trial (Trial 003) to verify and describe clinical benefit. AMAG Pharmaceuticals has disclosed that this completed confirmatory trial did not demonstrate a statistically significant difference between the treatment and placebo arms for the co-primary endpoints of reducing the risk of recurrent preterm birth at less than 35 weeks gestation or improving neonatal mortality and morbidity. The committee considered the trial's findings and the sNDA in the context of AMAG Pharmaceuticals' confirmatory study obligation.

Attendance:
Bone, Reproductive and Urologic Drugs Advisory Committee Members Present (Voting): Douglas C. Bauer, MD; Matthew T. Drake, MD, PhD; Vivian Lewis, MD (Chairperson); Pamela Shaw, PhD

Bone, Reproductive and Urologic Drugs Advisory Committee Members Not Present (Voting): Toby Chai, MD; James Q. Clemens, MD, FACS, MSCI; ; Beatrice Edwards, MD, MPH, FACP; Margery Gass, MD; Christian P. Pavlovich, MD; Gloria Richard Davis, MD, MBA, NCMP, FACOG

Bone, Reproductive and Urologic Drugs Advisory Committee Member Not Present (Non-Voting): Gerard G. Nahum, MD, FACOG (Industry Representative)

Temporary Members (Voting): Jonathan M. Davis, MD; Ahizechukwu Eke, MD, MPH; Annie Ellis (Patient Representative); Daniel Gillen, PhD; Kimberly Hickey, MD; Sally Hunsberger, PhD; Michael K. Lindsay, MD, MPH; Michele Orza, ScD (Acting Consumer Representative); Uma M. Reddy, MD, MPH; Brian Smith MD, MPH, MHS; Kelly Wade, MD, PhD, MSCE; Deborah A. Wing, MD, MBA

Acting Industry Representative to the Committee (Non-voting): Venkateswar Jarugula, PhD (Acting Industry Representative)

FDA Participants (Non-Voting): Christine Nguyen, MD; Barbara Wesley, MD, MPH; Christina Chang, MD, MPH; Jia Guo, PhD

Open Public Hearing Speakers: Meena M. Aladdin, PhD (Public Citizen); Adam C. Urato, MD (MetroWest Medical Center); Stephanie Fox-Rawlings, PhD (National Center for Health Research); Washington Clark Hill, MD, FACOG (Florida Department of Health, Sarasota County); John R. Barton, MD, MS (Baptist Health Lexington); Danielle Boyce (statement read by Robin Osman); Mary Norton, MD (Society for Maternal-Fetal Medicine); Anabel Jimenez-Gomez (statement read by Amelia Chiaverini); Kelle Moley, MD (March of Dimes); Allison Johnson; Glory M. Joseph (statement read by Allison Johnson); Marc Jackson, MD, MBA (The American College of Obstetricians and Gynecologists); Amelia Chiaverini; Michael Randell, MD, MBA; Steven Caritis, MD (University of Pittsburgh School of Medicine); Elizabeth Thom, PhD (George Washington University)

The agenda was as follows:

Call to Order and Introduction of Committee	**Vivian Lewis, MD** Chairperson, BRUDAC
Conflict of Interest Statement	**Kalyani Bhatt, BS, MS** Designated Federal Officer, BRUDAC
FDA Opening Remarks	**Christine Nguyen, MD** Deputy Director for Safety Division of Bone, Reproductive and Urologic Products (DBRUP) Office of Drug Evaluation III (ODE III) Office of New Drugs (OND), CDER, FDA
APPLICANT PRESENTATIONS	**AMAG Pharmaceuticals, Inc.**
Introduction	**Julie Krop, MD** Chief Medical Officer Executive Vice President, Development & Regulatory Affairs

	AMAG Pharmaceuticals, Inc.
Clinical Background and Unmet Need	**Michelle Owens, MD** Professor and Medical Director School of Medicine Department of Obstetrics and Gynecology The University of Mississippi Medical Center
Meis Study Design and Results	**Baha Sibai, MD** Professor Department of Obstetrics, Gynecology, and Reproductive Sciences Investigator, MFMU University of Texas Health Science Center of Houston MFMU[1] Network
PROLONG: Efficacy and Safety	**Laura Williams, MD, MPH** Sr. Vice President, Clinical Development & Biostatistics AMAG Pharmaceuticals, Inc.
Prevention of Preterm Birth: Clinical Perspective	**Sean Blackwell, MD** Professor and Chair Department of Obstetrics, Gynecology, and Reproductive Sciences Principal Investigator, MFMU University of Texas Health Science Center of Houston MFMU[1] Network
Conclusion	**Julie Krop, MD**
Clarifying Questions to Applicant	

BREAK

FDA PRESENTATIONS

Clinical Overview	**Barbara Wesley, MD, MPH** Medical Officer DBRUP, ODEIII, OND, CDER, FDA
Efficacy in Confirmatory Trial 003	**Jia Guo, PhD** Statistical Reviewer Division of Biometrics 3 (DB3) Office of Biostatistics (OB) Office of Translational Sciences (OTS), CDER, FDA

Hydroxyprogesterone Caproate (HPC) Utilization in the United States	**Huei-Ting Tsai, PhD** Epidemiologist Division of Epidemiology II (DEPI-II) Office of Pharmacovigilance and Epidemiology (OPE) Office of Surveillance and Epidemiology (OSE) CDER, FDA
Summary Remarks	**Christina Chang, MD, MPH** Clinical Team Leader DBRUP, ODEIII, OND, CDER, FDA

Clarifying Questions to FDA

LUNCH

OPEN PUBLIC HEARING

Clarifying Questions to Applicant or FDA

BREAK

Questions to the Committee/Committee Discussion and Voting

ADJOURNMENT

Questions to the Committee:

1. **DISCUSSION:** Discuss the effectiveness of Makena on recurrent preterm birth and neonatal morbidity and mortality.

   Committee Discussion: *There was general consensus among committee members that neither Trial 002 nor Trial 003 showed a treatment benefit of Makena on neonatal morbidity or mortality. The committee members further agreed that the data regarding preterm birth rates were conflicting, but there was a range of opinion as to which of the two trials better informed the efficacy of Makena for this outcome. Certain committee members opined that Trial 003 was large enough to show that there were no effect modifiers that could explain the differences in efficacy findings between 002 and 003. Further, the members could not identify a subgroup of patients where the efficacy results were consistent between Trials 002 and 003. Several members of the committee questioned the high rate of preterm birth in the placebo arm in Trial 002. Several commented on the smaller size of the US cohort in Trial 003 (23% of the total), making it difficult to interpret findings. Others were encouraged by the trend of positive treatment effect in the US subgroup in Trial 003, although the findings were not statistically significant. See the transcript for details of the committee discussion.*

2. **DISCUSSION:** If a new confirmatory trial were to be conducted, discuss the study design, including control, dose(s) of study medication, efficacy endpoints and the feasibility of completing such a trial.

   Committee Discussion: *The committee members agreed that, given the years to complete Trial 003, the number of sites used, and professional societies' guidelines, a new placebo-controlled trial would be extremely challenging and likely not feasible. Several committee members commented that pharmacokinetic studies should be performed to assess dosing, timing of drug administration and drug metabolism. Committee members also noted that studies should include an "enriched" population, such as pregnant women who are obese, with family histories of preterm birth, with substance abuse history, and recurrent preterm birth. Some committee members also recommended inclusion of other populations that might benefit, such as patients of different ages and racial groups. Some members recommended a study to look at "responders" vs "non-responders" and perhaps study pharmacogenetics. Other study design alternatives noted by committee members included comparing Makena to vaginal progesterone, a dose escalation study, a dose-response study, or creating a registry of women who used Makena. Some members noted that only a randomized control trial, and not observational studies, could provide the data needed. See the transcript for details of the committee discussion.*

3. **DISCUSSION**: Discuss the potential consequences of withdrawing Makena on patients and clinical practice.

   Committee Discussion: *Several members noted that Makena withdrawal from the US market would lead to resumption of use of compounded (hydroxyprogesterone caproate) HPC and use of other progesterone products. Some expressed concerns over unknown risks of compounded HPC from a safety perspective and quality perspective. Committee members also noted that the greatest burden could be felt by the most vulnerable groups (e.g., lower socioeconomic groups). Committee members also commented on the emotional burden for patients, and their providers, who are desperate for a treatment. On the other hand, some members commented on the potential positive consequences of Makena's withdrawal. These included the opportunity to bring the discussion of Makena's efficacy back to equipoise to allow the conduct of an adequate and well-controlled trial to inform Makena's efficacy in a defined population. See the transcript for details of the committee discussion.*

4. **VOTE:** Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal outcomes?

   Provide rationale for your vote.

   Vote Result: Yes: 0 No: 16 Abstain: 0

   Committee Discussion: *The committee unamiously agreed that the findings from Trial 003 do not verify the clinical benefit of Makena on neonatal outcomes. The committee*

members noted that there were no other data that supported the clinical benefit on the neonate. A neonatogist commented that significantly adverse neonatal outcomes in infants born after 32 – 34 weeks gestation are relatively rare. To detect treatment effect of Makena on these outcomes would likely require a trial larger than Trial 003. See the transcript for details of the committee discussion.

5. **VOTE**: Based on the findings from Trial 002 and Trial 003, is there substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth?

   Vote Result: Yes: 3 No: 13 Abstain: 0

   Committee Discussion: *The majority of the committee members agreed that, based on the findings from Trial 002 and Trial 003, there is not substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth. The committee members who voted "No" based their vote on the statutory and scientific definition of "substantial evidence of effectiveness," because Trial 003 did not substantiate the positive findings on preterm birth seen in Trial 002. These members also noted there was no treatment effect seen in any of the Trial 003 subgroups analyzed, and that there was no evidence of an interaction between the treatment effect of Makena and risk factors for preterm birth to explain the differences in the efficacy findings between Trials 003 and 002. Because no subgroup could be identified to have benefitted from Makena in both Trials 002 and 003, the appropriate patient population could not be determined. Those who voted "Yes" stated that the findings from Trial 002 were compelling and the positive trend seen in the U.S. subgroup in Trial 003 was encouraging. Although there was no evidence of effectiveness of Makena in Trial 003, they opined that the study's population, a majority of whom were from Russia and Ukraine, was not relevant to the U.S. and that the population's low-risk of pre-term birth may have obscured the evidence of effectiveness in U.S. women. See the transcript for details of the committee discussion.*

6. **VOTE:** FDA approval, including accelerated approval, of a drug requires substantial evidence of effectiveness, which is generally interpreted as clinically and statistically significant findings from two adequate and well-controlled trials, and sometimes from a single adequate and well-controlled trial. For drugs approved under the accelerated approval pathway based on a surrogate endpoint, the Applicant is required to conduct adequate and well-controlled post approval trial(s) to verify clinical benefit. If the Applicant fails to conduct such post approval trial(s) or if such trial(s) do not verify clinical benefit, FDA may, following an opportunity for a hearing, withdraw approval.

   Should FDA:

   A. Pursue withdrawal of approval for Makena
   B. Leave Makena on the market under accelerated approval and require a new confirmatory trial
   C. Leave Makena on the market without requiring a new confirmatory trial

   Provide rationale for your vote and discuss the following:

- Vote (A) (withdraw approval) may be appropriate if you believe the totality of evidence does not support Makena's effectiveness for its intended use.

    - Discuss the consequences of Makena removal (if not previously discussed in Discussion point 3)

- Vote (B) (require a new confirmatory trial) may be appropriate if you believe the totality of evidence supports Makena's effectiveness in reducing the risk of recurrent preterm birth, but that there is no substantial evidence of effectiveness on neonatal outcomes AND you believe that a new confirmatory trial is necessary and feasible.

    - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth, based on the surrogate endpoint of gestational age at delivery.

    - Also discuss key study elements, including study population, control, dose(s), and efficacy endpoints of the new confirmatory trial (if not previously discussed in Discussion point 2) and approaches to ensure successful completion of such a trial.

- Vote (C) (leave Makena on the market without a new confirmatory trial) may be appropriate if you believe Makena is effective for reducing the risk of recurrent preterm birth and that it is not necessary to verify Makena's clinical benefit in neonates.

    - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth and why it is not necessary to verify Makena's clinical benefits in neonates.

Vote Result: **A: 9** **B: 7** **C: 0**

Committee Discussion: *The committee members who voted "A" noted that the totality of evidence did not provide substantial evidence of effectiveness of Makena in the reducing the risk of recurrent preterm birth. Furthermore, there is no evidence from Trials 002 and 003 that Makena benefits the neonate, which is the goal of treatment. These members stated that the only way to definitely determine whether Makena is effective would be to conduct a well-designed, prospective, randomized clinical trial. They expressed that the withdrawal of Makena would facilitate the conduct of such a trial in the US and that professional societies should take a leadership role in communicating the importance of gathering this information. Some of these committee members, however, expressed concerns over Makena's withdrawal, because of potential clinical and societal repurcussions.*

The committee members who voted "B" acknowledged the efficacy data for reducing the risk of recurrent preterm birth are conflicting and not particularly persuasive. They also

recognized the need for more data, especially to identify subpopulations that might benefit from Makena. However, these members did not believe another randomized, controlled trial would be feasible under any circumstance, including after withdrawal of Makena's approval. They were concerned that prescribers and patients would insist on receiving treatment, regardless of the evidence of efficacy, and would resort to compounded products or other progesterone products with even less evidence. Some members indicated that withdrawal of Makena would be warranted only if the drug was unsafe.

None of the committee members voted "C."

See the transcript for details of the committee discussion.

The meeting adjourned at approximately 4:30 p.m.

17 α-Hydroxyprogesterone Caproate (Makena®) for Women with Singleton Pregnancy and Prior Spontaneous Preterm Birth

FDA Advisory Committee Meeting

Division of Bone, Reproductive and Urologic Products

AMAG Pharmaceuticals, Inc.

October 29, 2019

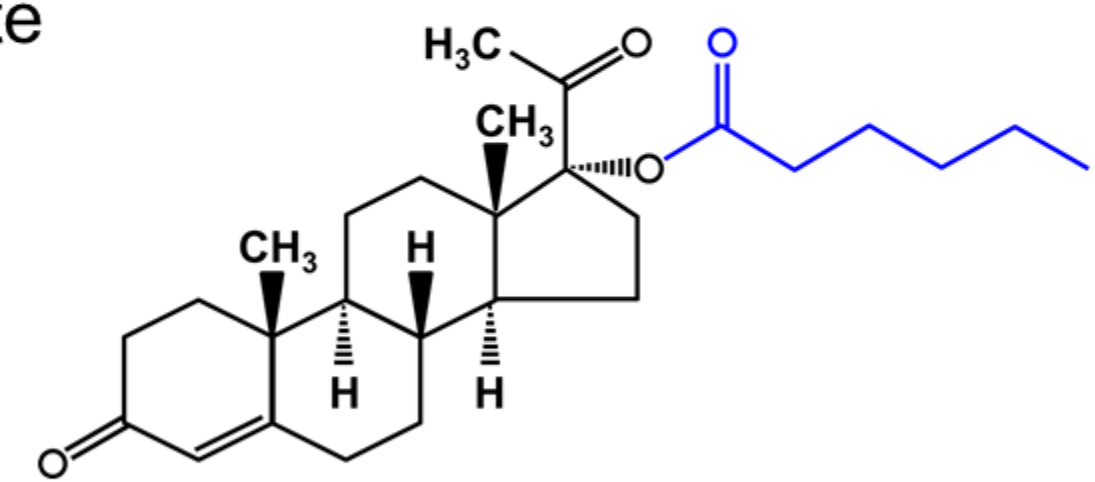
Introduction

Julie Krop, MD
Chief Medical Officer
EVP Clinical Development and Regulatory Affairs
AMAG Pharmaceuticals, Inc.

Makena and Generic 17P Formulations: Only FDA-Approved Therapy to Reduce Recurrent Preterm Birth

- Synthetic progestin
- Active pharmaceutical ingredient: 17 α-hydroxyprogesterone caproate
  - Not same as progesterone
- Proposed MOA
  - Decreases inflammation
  - Inhibits uterine activity
- Not metabolized into androgens, estrogen, or corticosteroids

17P



17P is an Orphan Drug

- Indicated for women with singleton pregnancy and prior spontaneous preterm birth
- Subset of all preterm birth
  - Affects ~ 3% (130,000) of all pregnancies
- Orphan Drug designation received

17P's Prolonged Half-Life Allows Once-Weekly Administration

- 17P treatment
  - Begins between 16^0 and 20^6 weeks pregnancy
  - Continued until 37 weeks or delivery
- Previously only available through pharmacy compounding

17P Approved Under Subpart H Accelerated Approval Pathway in 2011

- Subpart H applies to therapies that
  - Treat serious or life-threatening conditions with unmet need
  - Demonstrate efficacy on surrogate endpoint reasonably likely to predict clinical benefit
- Preterm birth (PTB) < 37 weeks accepted surrogate endpoint
  - Multiple studies established preterm infants at high risk of morbidity and mortality
- Required confirmatory trial of clinically relevant endpoints

FDA 21CFR314

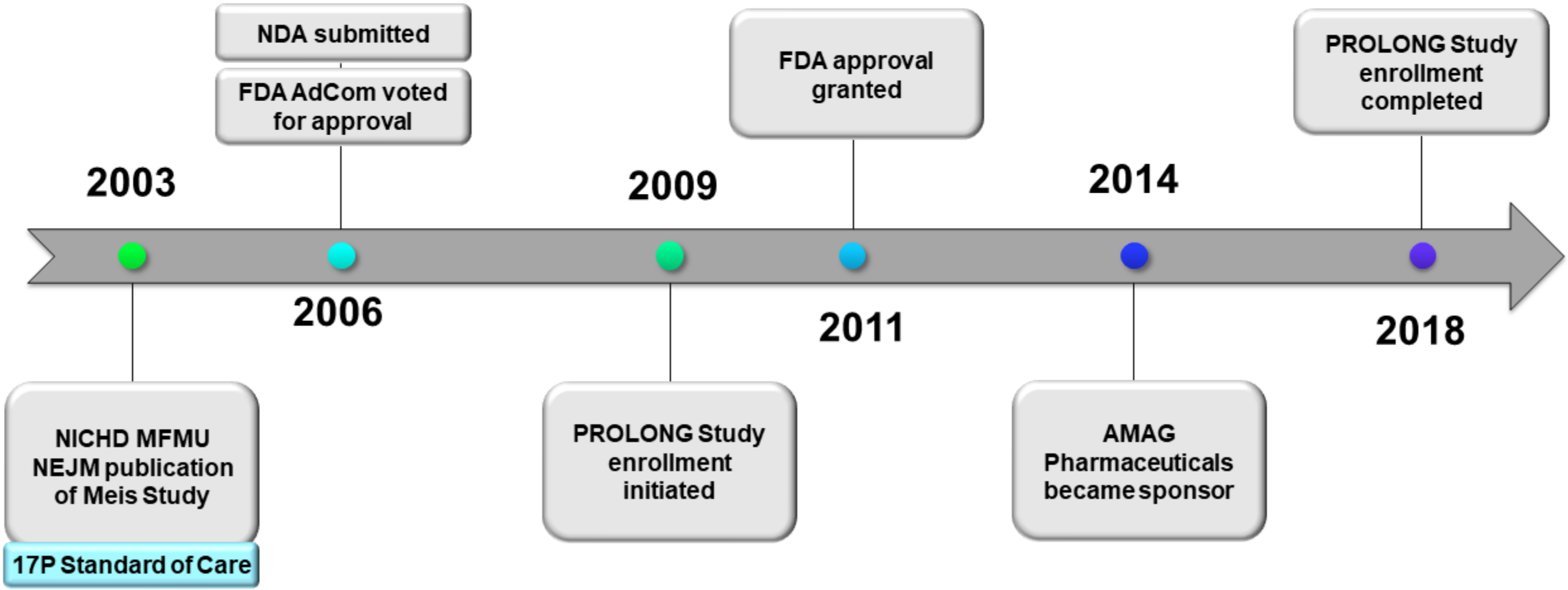
17P Approved Based on Compelling Results of Study 002 (Meis)

- Meis study conducted through NICHD MFMU
  - All US population
- Established substantial evidence of 17P efficacy
  - Highly statistically significant reduction in PTB rate vs. placebo < 37 weeks (p=0.0003)
  - Also reduced PTB < 35 weeks and < 32 weeks
    - Associated with highest incidence of neonatal complications

Meis, NEJM, 2003

NICHD = National Institute of Child Health and Human Development Health; MFMU = Maternal Fetal Medicine Units

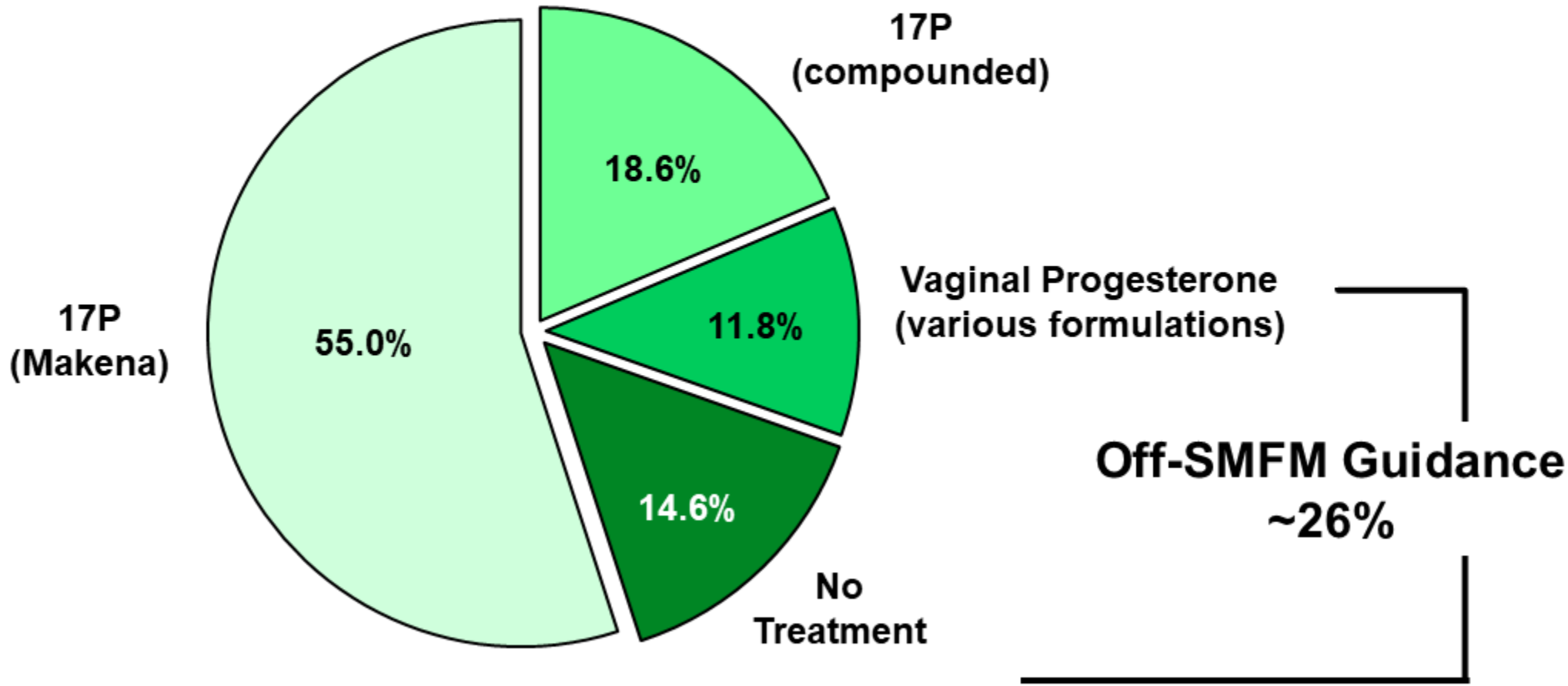
Key Events in 17P Approval Pathway



Preterm Birth is Major US Public Health Concern

- Leading cause of infant morbidity and mortality
- Can lead to serious long-term health complications
- Recurrent PTB represents only a small proportion of all PTBs

~75% of Indicated Patients Treated with 17P in 2017



Gallagher AJP Reports, 2018

Generalizability of PROLONG Efficacy Data to US is Challenging

- Key differences in study populations and background rates
- Meis trial enrolled in US inner city academic medical centers
  - ~30% background rate of PTB < 35 weeks
- PROLONG enrolled population with low PTB rate
  - ~11% background rate of PTB < 35 weeks
- Strong efficacy from Meis and other clinical trials along with favorable safety profile

Agenda

Clinical Background / Unmet Need	**Michelle Y. Owens, MD** Professor and Medical Director School of Medicine Department of Obstetrics and Gynecology The University of Mississippi Medical Center
Meis Study Design and Results	**Baha M. Sibai, MD** Professor Department of Obstetrics, Gynecology and Reproductive Sciences McGovern Medical School-UTHealth Principal Investigator, MFMU
PROLONG Efficacy and Safety	**Laura A. Williams, MD, MPH** Sr Vice President, Clinical Development AMAG
Clinical Perspective / Benefit / Risk	**Sean C. Blackwell, MD** Professor and Chair Department of Obstetrics, Gynecology, and Reproductive Sciences McGovern Medical School-UTHealth Principal Investigator, MFMU
AMAG Actions Following PROLONG	**Julie Krop, MD** CMO, EVP Clinical Development and Regulatory Affairs, AMAG

Additional Expert Consultants

Hugh Miller, MD	Principal Investigator, PROLONG Founder, Watching Over Mothers & Babies (WOMB)
Anita Das, PhD	Statistician AD Stat Consulting
Eugene Poggio, PhD	Statistician Biostatistical Consulting

Clinical Background and Need

Michelle Y. Owens, MD
Professor and Medical Director
School of Medicine Department of Obstetrics and Gynecology
The University of Mississippi Medical Center

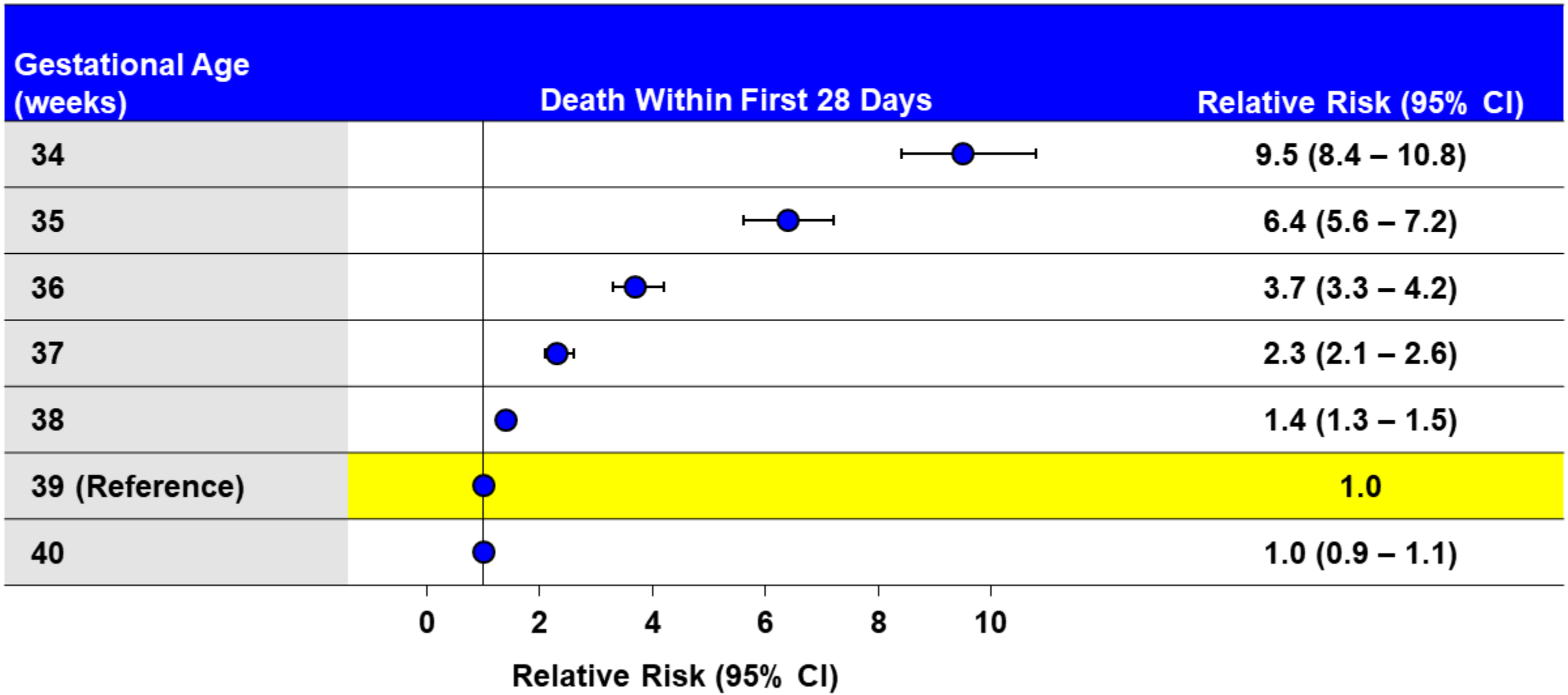
Preterm Birth: Significant Problem in US

- 1 in 10 babies born prematurely in US
- Disadvantaged women – socioeconomically, educationally, limited healthcare access
- PTB puts infant at substantial risk
- Critical to have access to FDA-approved 17P for subset of women with prior PTB

What is at Stake: The Health of Infants

Photo: https://nyuwinthrop.org/newsroom/cornerstone/winter-spring-2018/nyu-winthrop-neonatal-breakthroughs/

Neonatal and Infant Mortality Significantly Higher for Babies Born at 34 – 36 Weeks Gestation



Reddy, et al., Pediatrics 2009

Preterm Birth and Complications #1 Cause of Death of Babies in US

Short-Term Complications

- Respiratory distress syndrome (RDS)
- Bronchopulmonary dysplasia (BPD)
- Intraventricular hemorrhage (IVH)
- Periventricular leukomalacia (PVL)
- Necrotizing enterocolitis (NEC)
- Apnea
- Jaundice
- Anemia
- Infections

Long-Term Consequences

- Chronic respiratory problems
- Rehospitalization
- Metabolic disorders
- Neurodevelopmental problems
  - Cerebral palsy
  - Cognitive deficits
  - Hearing and vision impairment
  - Learning disorders

Spong CY. Obstet Gynecol 2003, Stoll, et al, Pediatrics, 2010, Gyamfi-Bannerman, NEJM, 2016

Babies Born at Lower Gestational Ages Have Higher Rates of Neonatal Morbidity and Mortality

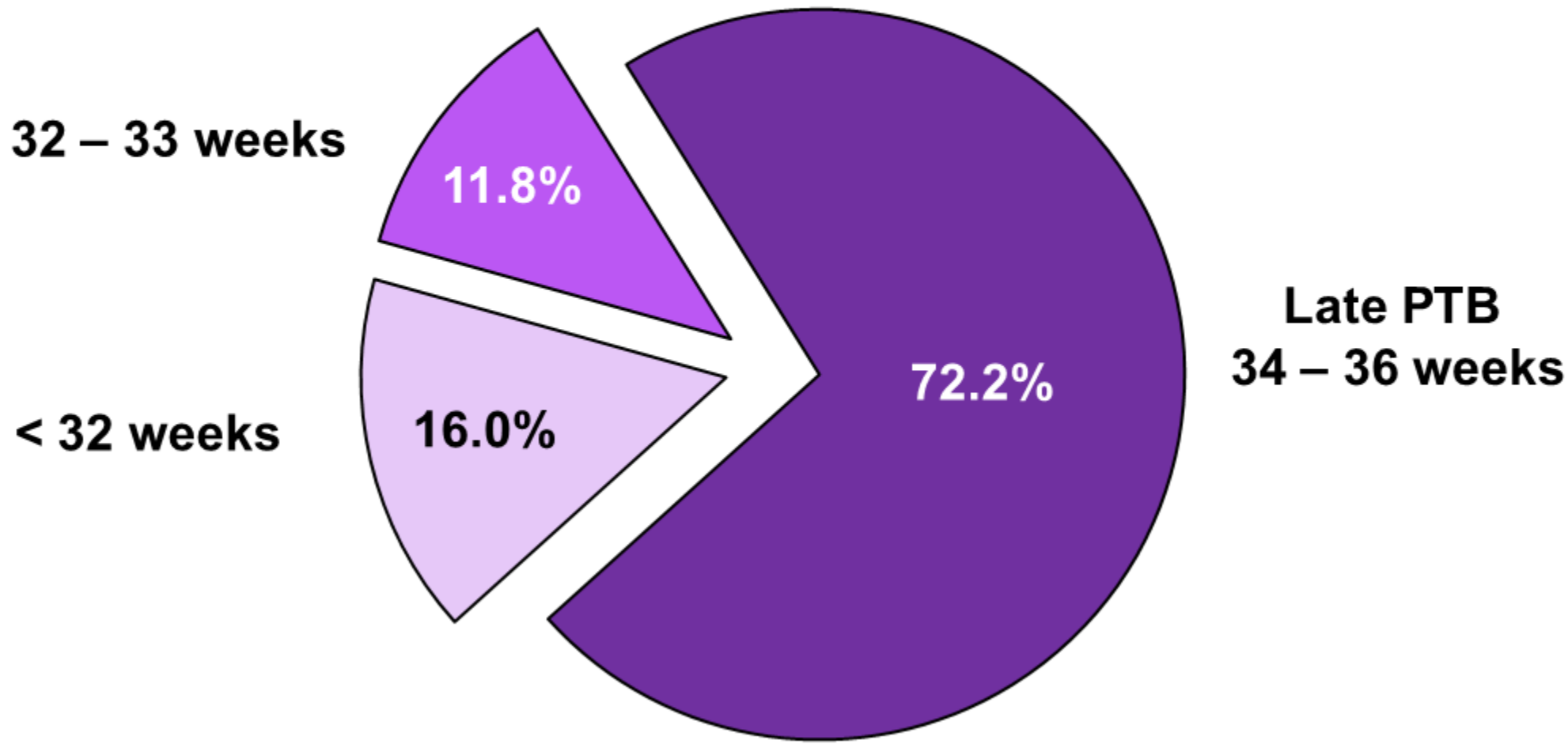
Delivery Gestational Age (Weeks)	Death n (%)	Major Morbidity n (%)	Death or Major Morbidity n (%)
< 32	117 (3)	448 (11)	565 (14)
< 35	119 (2)	560 (9)	679 (11)
36	0 (0)	55 (2)	55 (2)

Major morbidities

- Persistent pulmonary hypertension
- IVH grade 3 / 4
- Seizures
- Hypoxic-ischemic encephalopathy
- NEC stage II / III
- Bronchopulmonary dysplasia

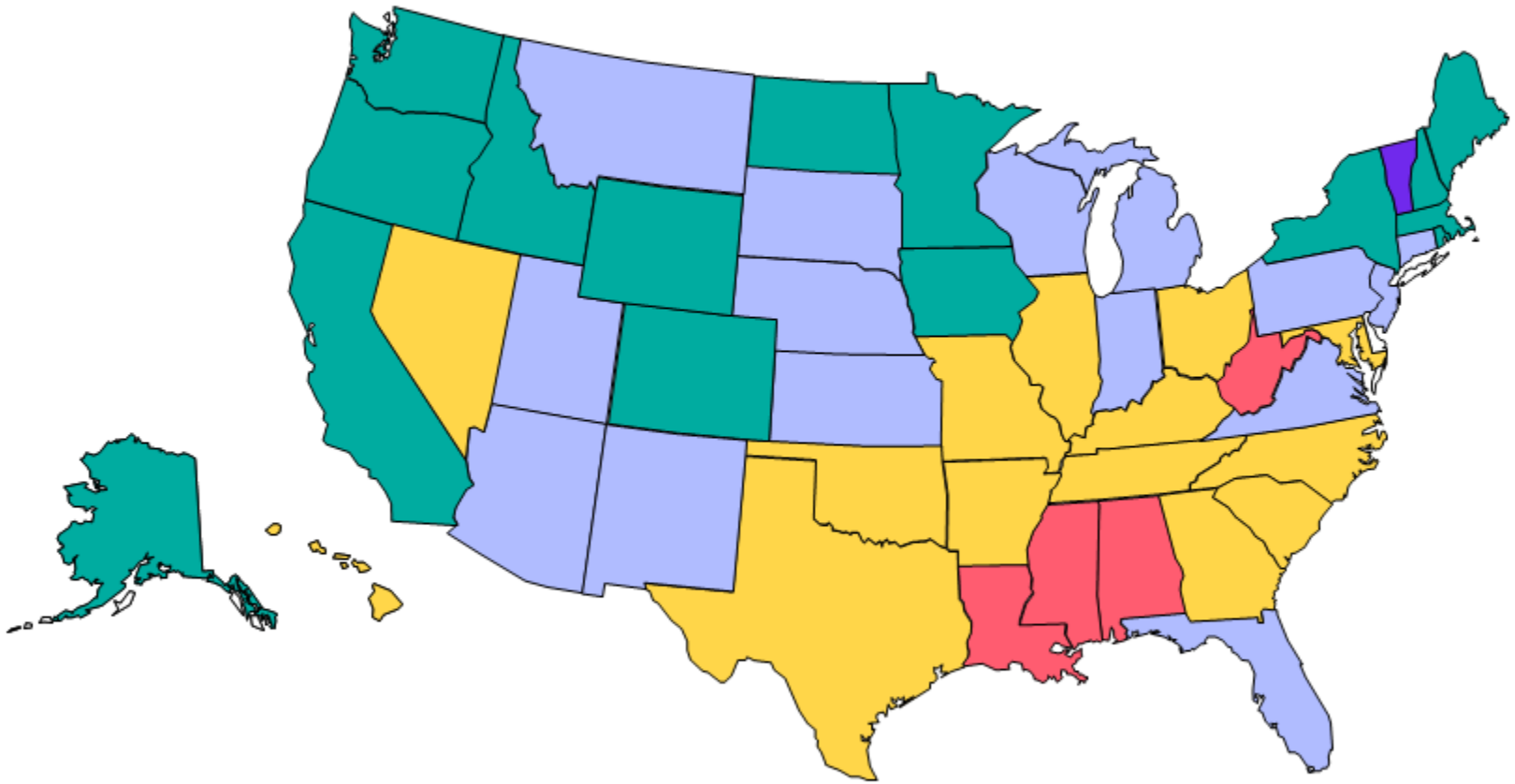
Manuck, AJOG 2016

Preterm Birth by Gestational Age



National Vital Statistics Reports “Births: Final Data for 2017”

US Ranks 131st of 184 Countries for Preterm Birth



Preterm Birth Rate

- Grade A: ≤ 8.1%
- Grade B: 8.2 – 9.2%
- Grade C: 9.3 – 10.3%
- Grade D: 10.4 – 11.4%
- Grade F: ≥ 11.5%

March of Dimes, 2018 Premature Birth Report Card
National Vital Statistics Reports "Births: Preliminary Data for 2015"

Risk Factors for Singleton Preterm Birth

Maternal Characteristics

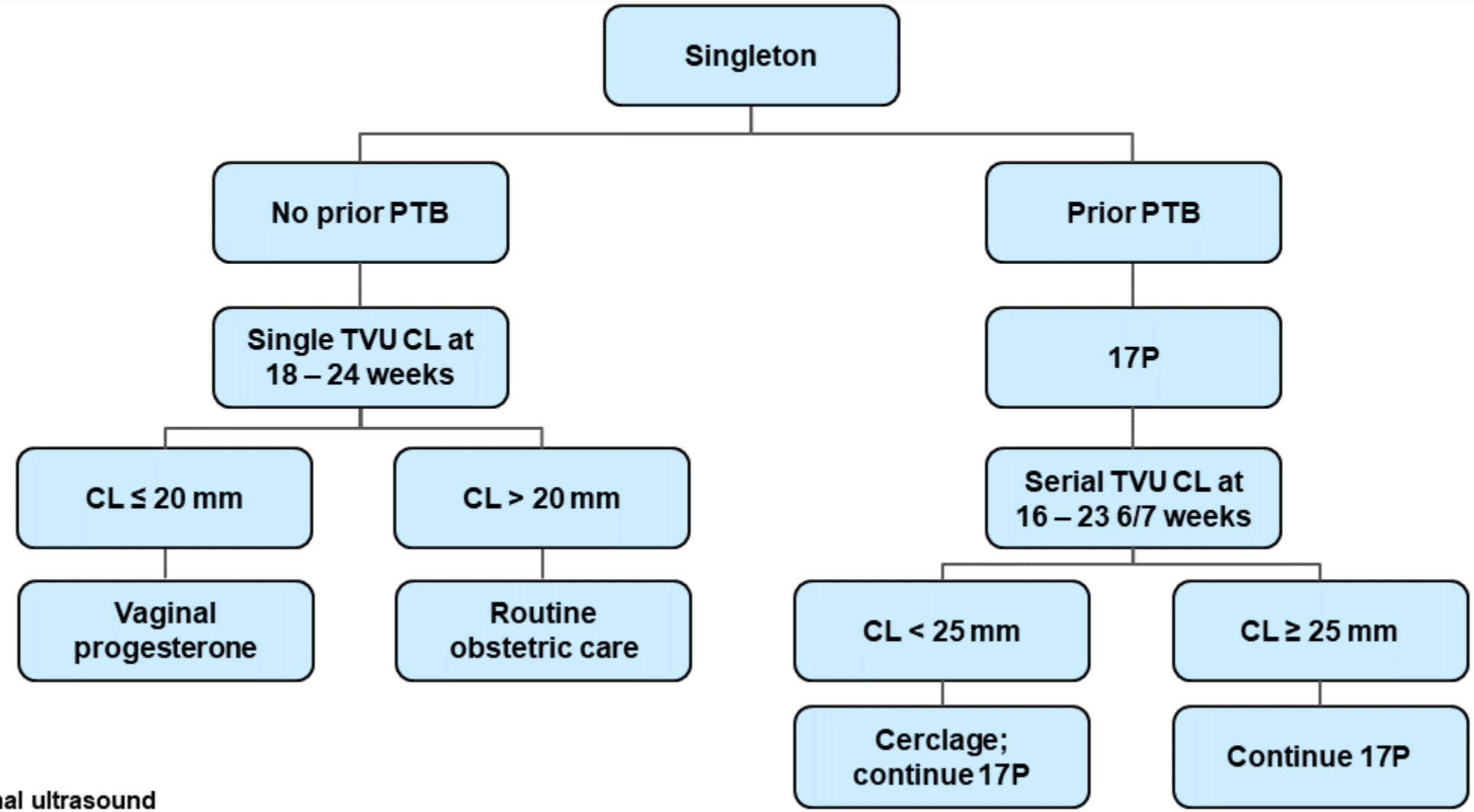
- **History of SPTB < 37 weeks**
- Short cervix
- African American
- Genitourinary infections
- Short intervals between pregnancies
- Advanced maternal age
- Low pre-pregnancy BMI

Social Determinants of Health

- Low socioeconomic status (i.e., education, income, marital status, nutrition)
- Stress (e.g., domestic violence, housing instability)
- Nicotine, alcohol, or drug use

Goldenberg Lancet 2008
ACOG Committee Opinion #729 2018

SMFM 2012 Clinical Guidelines



TVU = transvaginal ultrasound

SMFM Publications Committee, with assistance of Vincenzo Berghella AJOG 2012

Preterm Birth is Major US Public Health Concern, Disproportionately Affecting Lower SES Groups

- Infants spend weeks or months in NICU
- Must reduce preterm birth rate and prevent complications
- Physicians and patients need continued access to 17P

Meis Study Design and Results

Baha Sibai, MD

Professor, Department of Obstetrics, Gynecology and Reproductive Sciences

McGovern Medical School-UTHealth

Investigator, MFMU Network

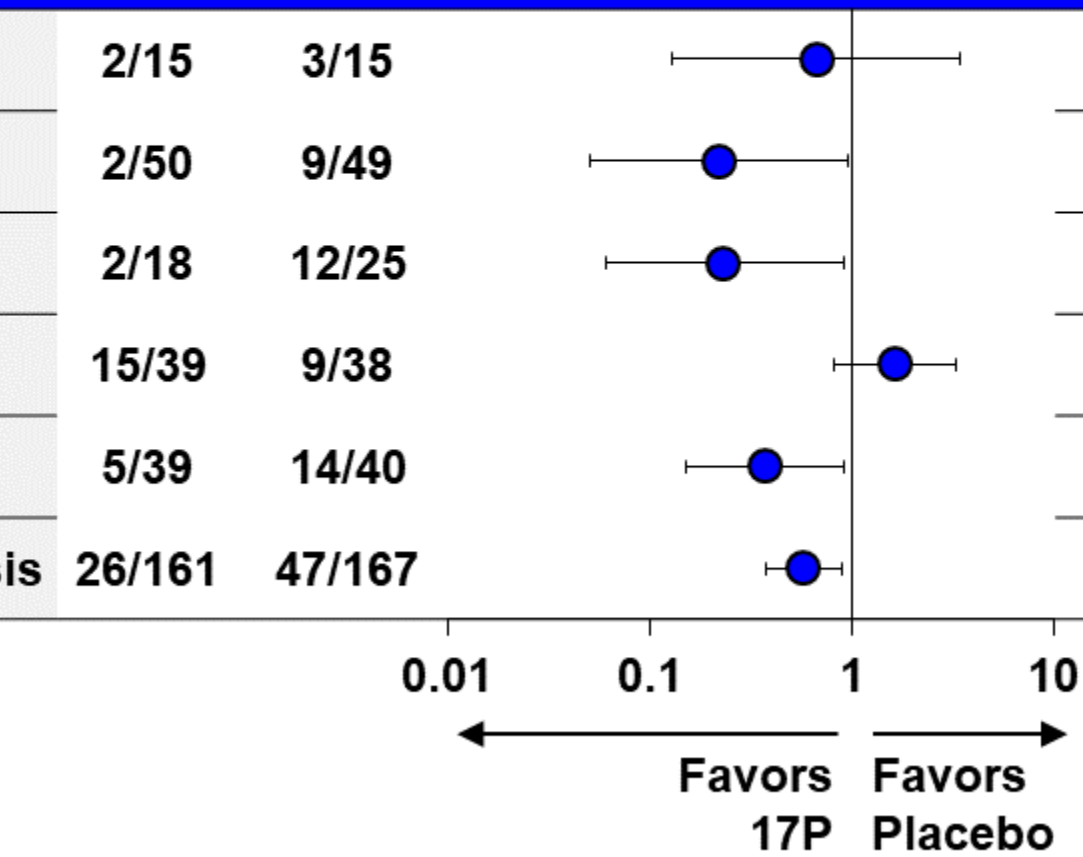
MFMU Network Established to Promote Rigorous Clinical Trials in Pregnancy

- Primary aim to reduce rate of preterm birth
- Rigorous process to select centers and studies
  - Network centers with ≥ 40% high-risk obstetric population
  - Diverse patient populations

MFMU established in 1986 by NICHD

Meta-Analysis of 17P Demonstrated 42% Reduction in Recurrent PTB

Study	17P n/N	Placebo n/N	17P vs. Placebo Relative Risk (95% CI)
LeVine (1964)	2/15	3/15	0.67 (0.13 – 3.44)
Papiernick-Berkhauer (1970)	2/50	9/49	0.22 (0.05 – 0.96)
Johnson (1975)	2/18	12/25	0.23 (0.06 – 0.91)
Hartikainen-Sorri (1980)	15/39	9/38	1.62 (0.81 – 3.25)
Yemini (1985)	5/39	14/40	0.37 (0.15 – 0.92)
Meta-Analysis	26/161	47/167	0.58 (0.38 – 0.90)



Adapted from Keirse, BJOG 1990

Meis Study Designed to Evaluate 17P in Women with History of SPTB

- Women ≥ 18 years
- Singleton pregnancy
- Hx of previous singleton SPTB
- Enrolled 16^0 - 20^6 week of pregnancy

Randomization 2:1

17P
250 mg IM q7 days

Weekly injections of study drug until 36^6 week of pregnancy OR delivery

Placebo (vehicle)
IM q7 days

Meis, et al. NEJM. 2003

Meis Study High-Risk Population for Preterm Birth

Demographics and baseline characteristics	17P (N=310)	Vehicle (N=153)
Age (years), mean ± SD	26.0 ± 5.6	26.5 ± 5.4
> 1 Previous PTB	27.7%	41.2%
Black or African American	59.0%	58.8%
White	25.5%	22.2%
Non-Hispanic or Latino	86.1%	83.0%
Married or living with partner	51.3%	46.4%
BMI before pregnancy (kg/m^2), mean ± SD	26.9 ± 7.9	26.0 ± 7.0
Educational level (years), mean ± SD	11.7 ± 2.3	11.9 ± 2.3
Gestational age of qualifying delivery (weeks), mean ± SD	30.6 ± 4.6	31.3 ± 4.2
Any substance use* during pregnancy	27.4%	23.5%

2 – 4% of patients were Asian, 2 – 3% were Other (Native Hawaiian/Pacific Islander, American Indian or Alaska native, mixed race and other)
*Smoking, alcohol or illicit drugs

Meis et al. NEJM 2003

Meis Study Primary Outcome: Preterm Delivery < 37 Weeks

- < 37 weeks gestation current definition of prematurity
- Sample size N = 500 women
  - Based on expected recurrent PTB rate of 37% in placebo group
  - Expected 1/3 reduction of recurrence with 17P

Meis, et al. NEJM 2003

Meis Study: High Rate of Completion and Compliance

	17P (N=310)	Vehicle (N=153)
Completed study, %	98.7%	100%
Number of injections (mean)	14.1	13.7
Full compliance (< 10 days between doses)	91.5%	91.5%

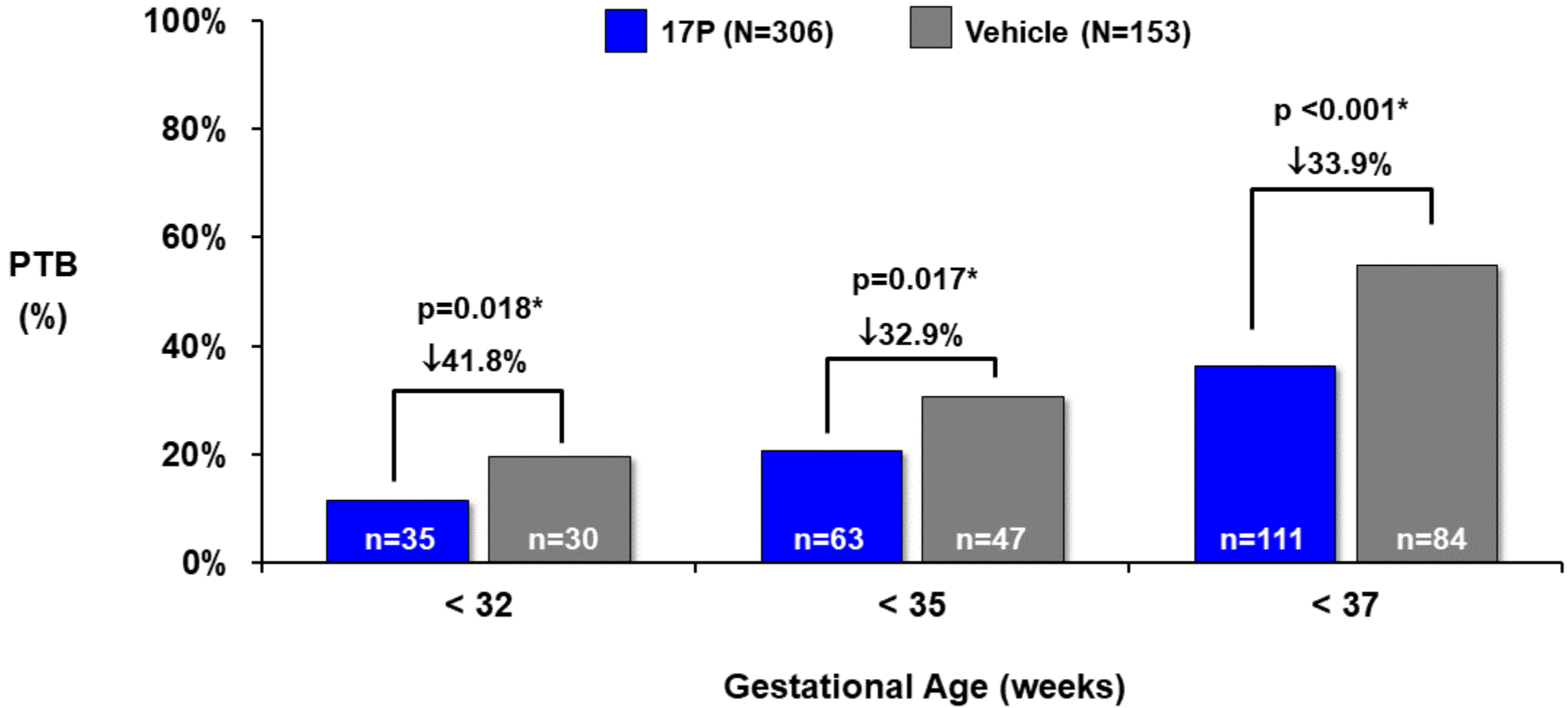
Meis, et al. NEJM 2003

Meis Study Stopped Early Due to Clear Evidence of 17P Benefit

- Study conducted from 1999 – 2002
- Planned interim analysis with pre-specified stopping criterion for efficacy (p=0.015)
  - Study halted at second interim analysis
  - Data available for 93% of planned sample (463/500)

Meis, et al. NEJM 2003

Meis Study Demonstrated Significant Reduction of PTB with 17P Compared to Vehicle

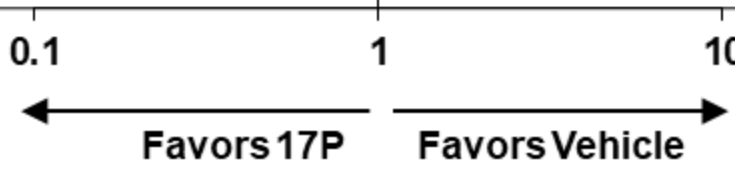


*p-values unadjusted for imbalance in prior PTBs

Meis, et al. NEJM 2003

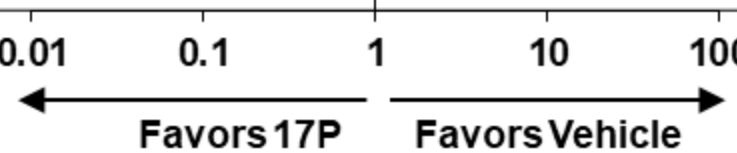
Meis Study: Consistent Reduction in PTB < 37 Weeks with 17P Across Subgroups

Subgroup	17P (n/N)	Vehicle (n/N)	Relative Risk (95% CI)
All patients	111/310	84/153	0.66 (0.54, 0.81)
> 1 Prior PTB	41/86	44/63	0.68 (0.52, 0.90)
Only 1 prior PTB	70/220	40/90	0.72 (0.53, 0.97)
Black	64/181	47/90	0.68 (0.51, 0.90)
Non-Black	47/125	37/63	0.64 (0.47, 0.87)
Unmarried	50/150	43/82	0.64 (0.47, 0.86)
Married	61/156	41/71	0.68 (0.51, 0.90)
Smoke or substance use	28/85	23/36	0.52 (0.35, 0.76)
No smoke or substance use	83/221	61/117	0.72 (0.57, 0.92)
Education ≤ 12 years	80/223	55/103	0.67 (0.52, 0.86)
Education > 12 years	31/83	29/50	0.64 (0.45, 0.93)



17P Reduced Neonatal Complications vs. Placebo

Subgroup	17P (n/N)	Vehicle (n/N)	Relative Risk (95% CI)
Necrotizing enterocolitis	0/305	4/152	0 (p<0.05)
IVH Any Grade	4/305	8/153	0.25 (0.08, 0.82)
Retinopathy	5/305	5/152	0.50 (0.15, 1.70)
Bronchopulmonary dysplasia	4/305	5/152	0.40 (0.11, 1.46)
Ventilatory support	26/303	22/151	0.59 (0.35, 1.00)
Respiratory distress syndrome	29/305	23/152	0.63 (0.38, 1.05)
Transient tachypnea	11/305	11/152	0.50 (0.22, 1.12)
Supplemental oxygen	45/303	36/151	0.62 (0.42, 0.92)
Proven sepsis	9/305	4/152	1.12 (0.35, 3.58)
Patent ductus arteriosus	7/305	8/151	0.43 (0.16, 1.17)



Meis, et al., NEJM 2003

Reduced Neonatal Intensive Care Unit (NICU) Admissions and Days With 17P Compared to Vehicle

	17P	Vehicle	17P vs Vehicle
Admitted to NICU, n/N (%)	82/295 (27.8%)	55/151 (36.4%)	Relative Risk = 0.76 95% CI (0.58, 1.01)
Number of NICU days, mean ± SD	23.9 ± 32.4	29.2 ± 37.6	Δ = -5.3 95% CI (-17.5, 6.9)

Perinatal Death in Meis Study

Complication	17P (N=306)[1] n (%)	Vehicle (N=153) n (%)
Total deaths	19 (6.2)	11 (7.2)
Neonatal deaths	8 (2.6)	9 (5.9)
Miscarriages < 20 weeks gestation [2]	5 (2.4)	0
Stillbirth	6 (2.0)	2 (1.3)

1. 4 patients in the 17P group were lost to follow-up and perinatal death status could not be determined
2. Percentage adjusted for the number at risk (17P n=209, Vehicle n=107) enrolled at <20 weeks gestation.

Follow-Up Observational Study of Meis Trial Babies Confirmed Long-Term Safety of 17P

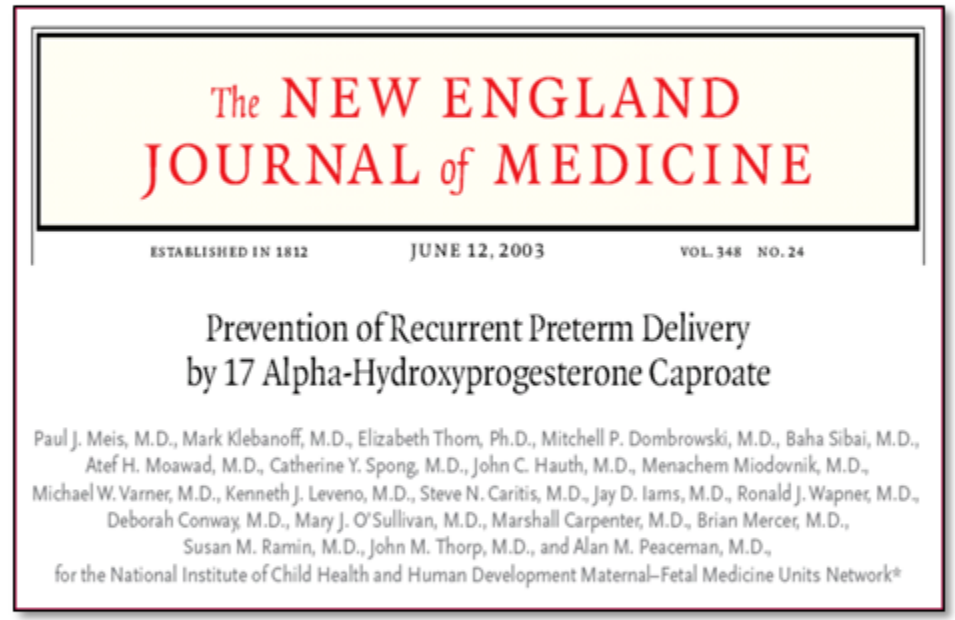
Ages and Stages Questionnaire (ASQ) Scored below cutoff on	17P (N=193)	Vehicle (N=82)	p-value
At least one area	27%	28%	0.9
Communication	11%	11%	0.9
Gross motor	3%	4%	0.7
Fine motor	21%	18%	0.6
Problem solving	10%	11%	0.9
Personal-social	4%	1%	0.4
Preschool Activities Inventory (PAI)			
Mean score for boys	67	67	0.3
Mean score for girls	33	32	0.5

Northen et al., Am J Obstet Gynecol, 2007

Median age (48 months, 41.8 to 55.0, 25th to 75th percentile)

Meis Results Considered Significant Advance in Obstetrics

- Relative risk
  - 0.66 (95% CI, 0.54 to 0.81)
- Absolute difference in preterm
  - 18.6%
- Number Need to Treat
  - 5.4 women to prevent 1 PTB



Meis, NEJM 2003

Meis Study Established Substantial Evidence of 17P Efficacy and Formed Foundation of PTB Prevention

- Clinicians have relied on 17P since 2003
- Only FDA-approved treatment to reduce risk of recurrent PTB since 2011
- Patients and clinicians need 17P as available option to prevent recurrent PTB

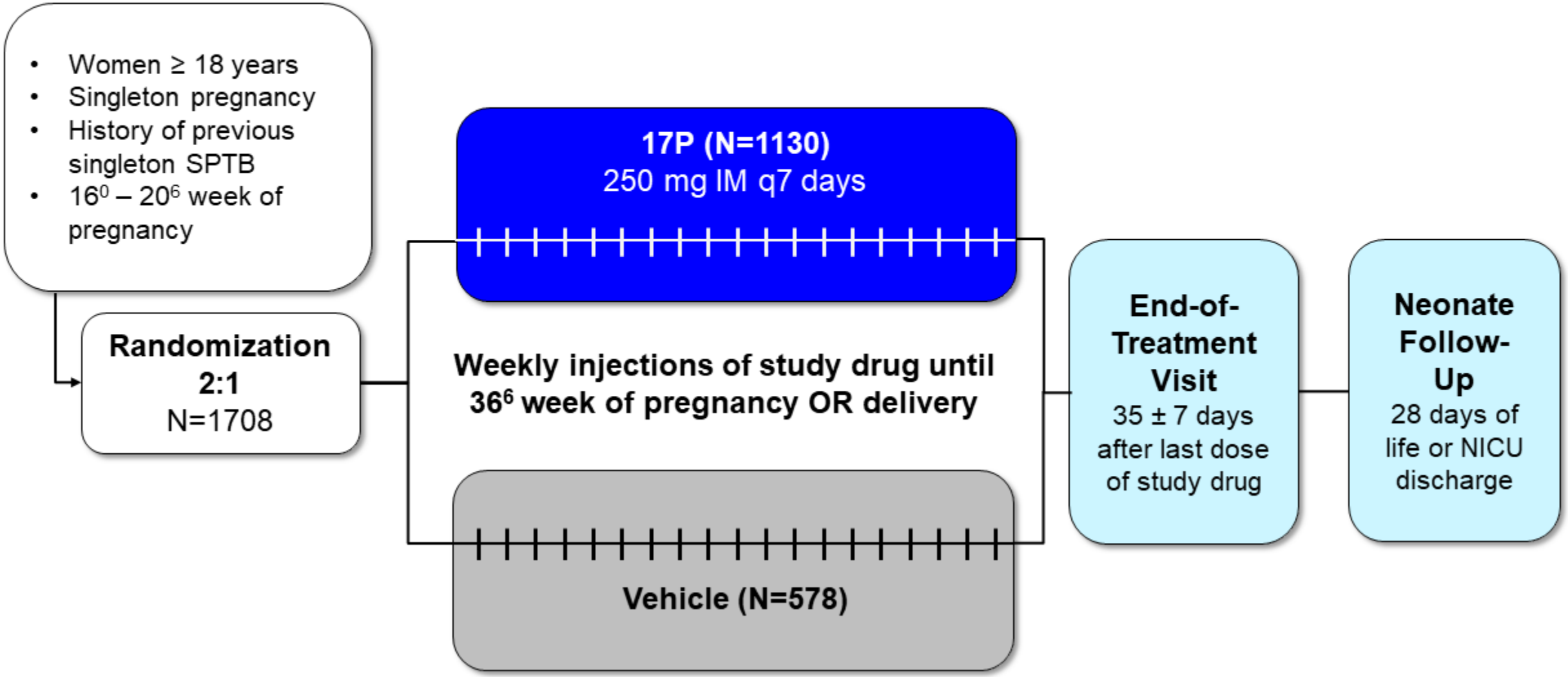
PROLONG: Efficacy and Safety

Laura A. Williams, MD, MPH
Senior Vice President, Clinical Development & Biostatistics
AMAG Pharmaceuticals, Inc.

PROLONG Designed to Mirror Meis Trial

- PROLONG did not meet its co-primary outcomes
  - Lower background PTB rates in PROLONG compared to Meis

PROLONG: Double-Blind, Vehicle-Controlled, Multi-Center, Randomized Study



PROLONG: Co-Primary Outcomes

- Reduction of PTB < 35 weeks gestation
- Reduction in composite neonatal morbidity and mortality index
  - Respiratory distress syndrome
  - Bronchopulmonary dysplasia
  - Grade 3 or 4 intraventricular hemorrhage
  - Necrotizing enterocolitis
  - Proven sepsis
  - Neonatal death

PROLONG: Key Secondary Efficacy and Primary Safety Outcomes

Secondary Outcomes

- Reduction of PTB by gestational age at delivery

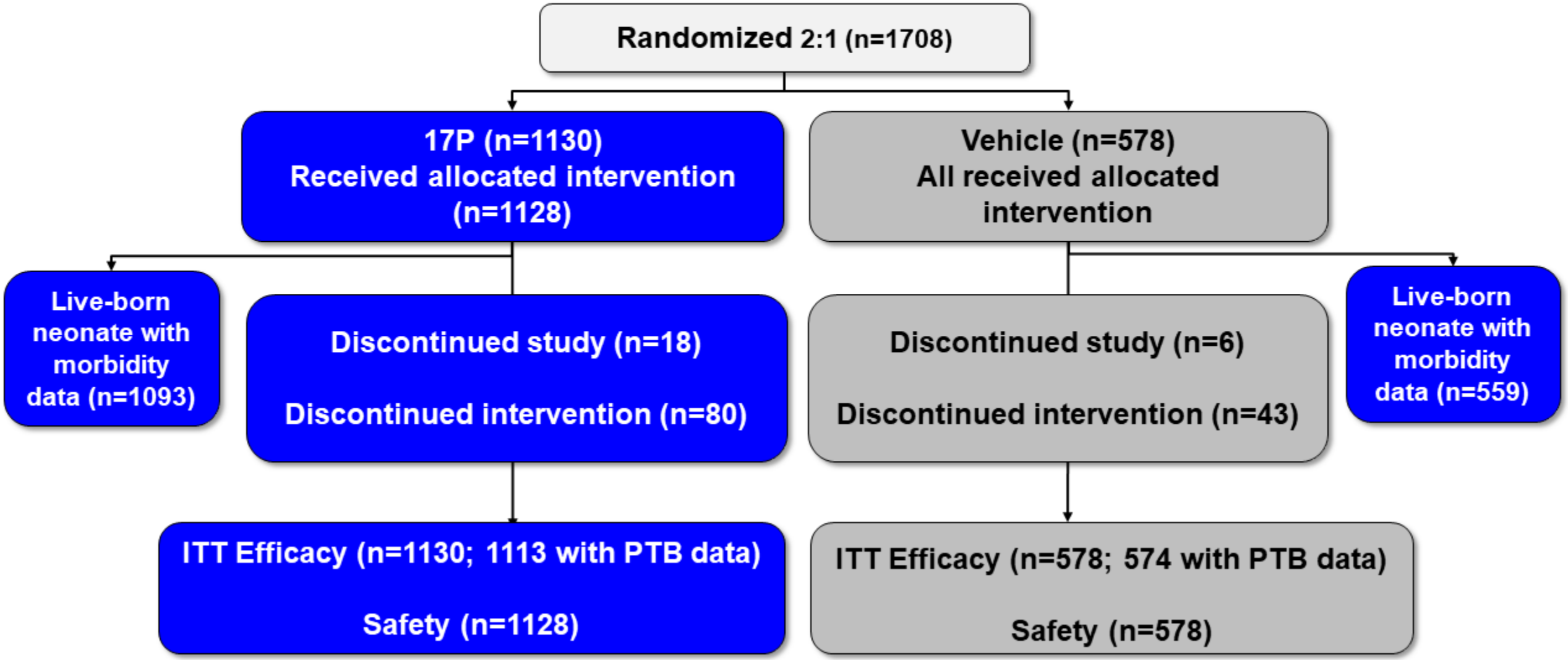
Primary Safety Outcome

- Exclude doubling in risk of fetal or early infant death

PROLONG: Sample Size and Powering Based on Conservative Estimates of Meis Results

- Sample size of 1707 provide
  - 98% power to detect 30% reduction in PTB < 35 weeks
  - 90% power to detect 35% reduction in neonatal composite index
  - 83% power to rule out doubling in risk of fetal/early infant death

PROLONG Patient Disposition: ~ 99% of Patients Completed Study



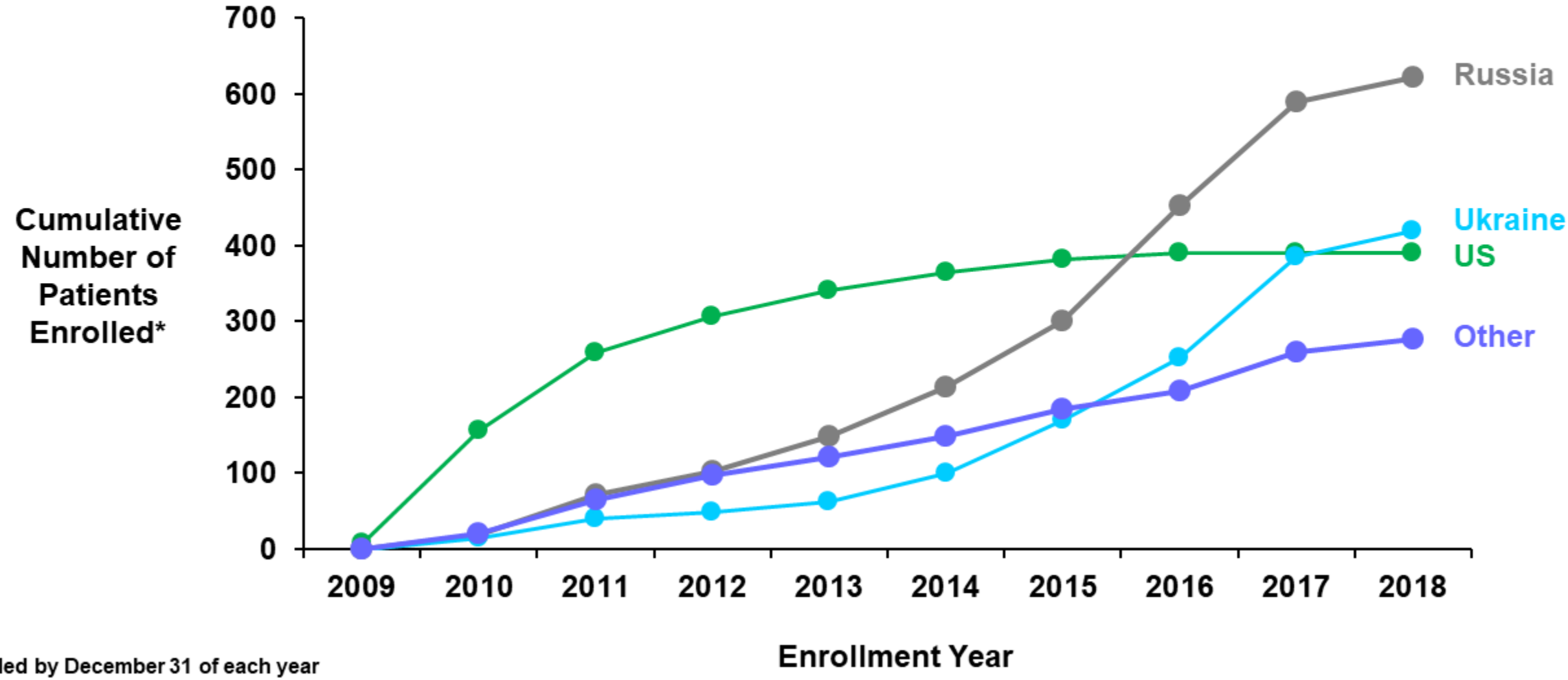
PROLONG: Enrollment by Geographic Region ~ 75% of Patients Enrolled Ex-US

Country	Number of Patients (N=1708) n (%)
United States	391 (22.9)
Ex-US	1,317 (77.1)
Russia	621 (36.4)
Ukraine	420 (24.6)
Hungary	91 (5.3)
Spain	85 (5.0)
Bulgaria	50 (2.9)
Canada	31 (1.8)
Czech Republic	14 (0.8)
Italy	5 (0.3)

61% of patients from Russia/Ukraine

16% (n=276 total) from other Ex-US Countries

PROLONG: Enrollment (Year End)



*Enrolled by December 31 of each year
Other: Bulgaria, Canada, Czech Republic, Hungary, Italy, Spain

PROLONG: Demographics and Baseline Characteristics Similar Across Treatment Groups

Demographics & Baseline Characteristics	17P (N=1130)	Vehicle (N=578)
Age (years), mean ± SD	30.0 ± 5.17	29.9 ± 5.22
Race / ethnicity		
White	88.8%	87.2%
Black, African American / African heritage	6.5%	7.1%
Non-Hispanic or Latino	91.1%	90.7%
Married or living with partner	89.6%	90.3%
BMI before pregnancy (kg/m^2), mean ± SD	24.3 ± 7.1	24.7 ± 8.7
Educational level (years), mean ± SD	13 ± 2.4	13 ± 2.4
Transvaginal cervical length < 25 mm at ≤ 20 weeks	1.2%	1.9%
Any substance use* during pregnancy	9.3%	8.8%

2 – 4% of patients were Asian, 2 – 3% were Other (Native Hawaiian/Pacific Islander, American Indian or Alaska native, mixed race and other) *Smoking, alcohol or illicit drugs

PROLONG: Prior Pregnancy History Similar Across Treatment Groups

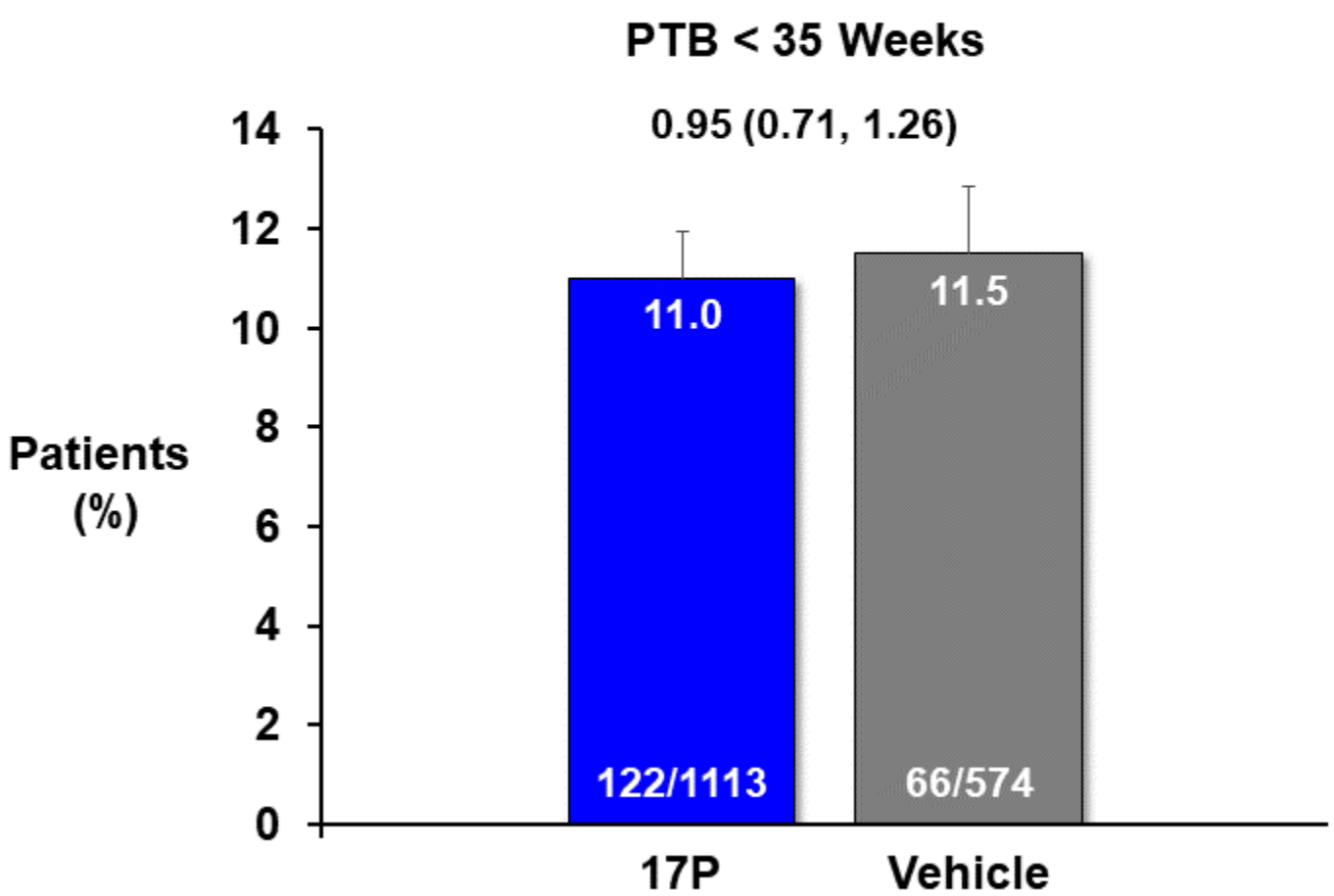
Pregnancy Characteristics	17P (N=1130)	Vehicle (N=578)
Prior SPTB – median (min, max)	1.0 (1, 7)	1.0 (0*, 5)
> 1 previous SPTB n (%)	148 (13.1)	70 (12.1)
Gestational age of prior qualifying delivery (weeks) mean ± SD median (min, max)	 31.3 ± 4.35 32 (20, 36)	 31.6 ± 4.16 33 (20,36)

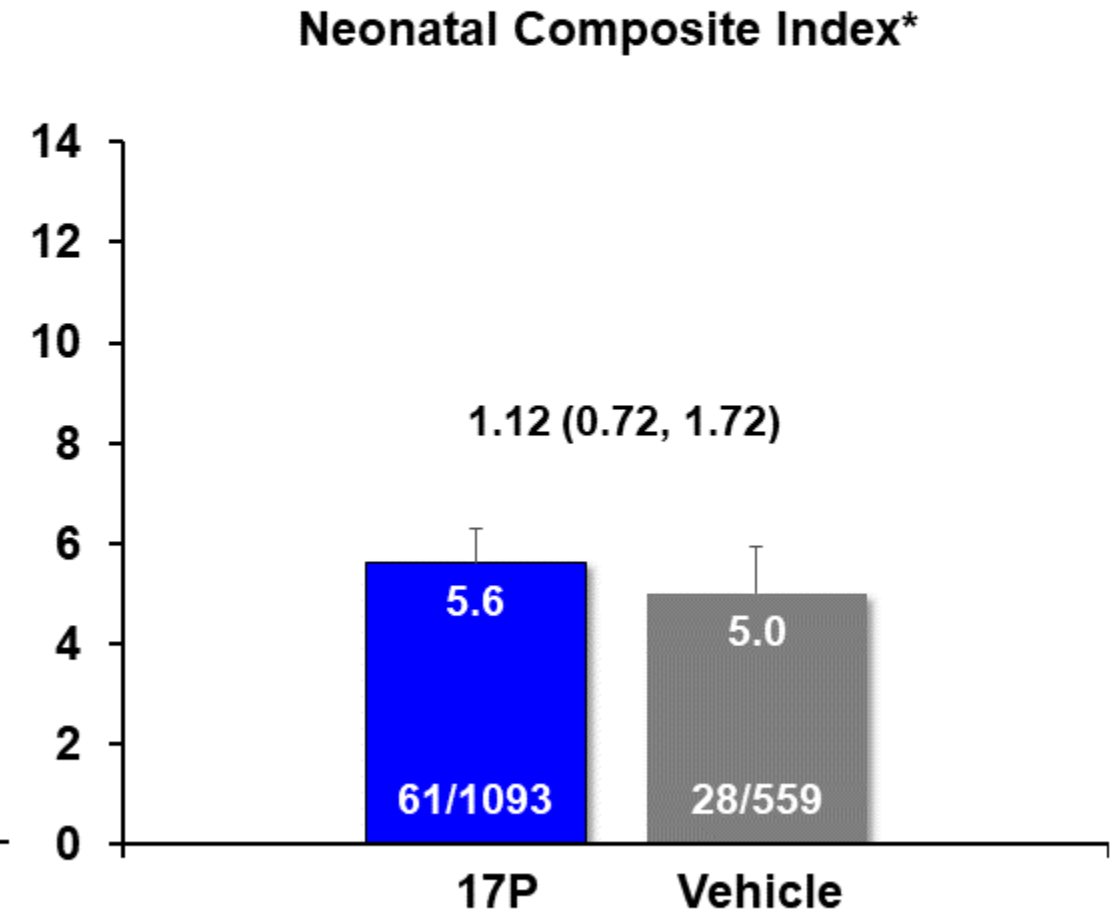
*1 patient in vehicle arm did not have SPTD (protocol deviation)

PROLONG: Comparable Study Drug Compliance Across Treatment Groups

Study Drug Exposure	17P (N=1128)	Vehicle (N=578)
Injections received		
Mean (SD)	17.6 (3.65)	17.5 (3.81)
Median	18	18
Min, Max	(1, 22)	(1, 22)
Patients with full compliance	91.4%	92.4%

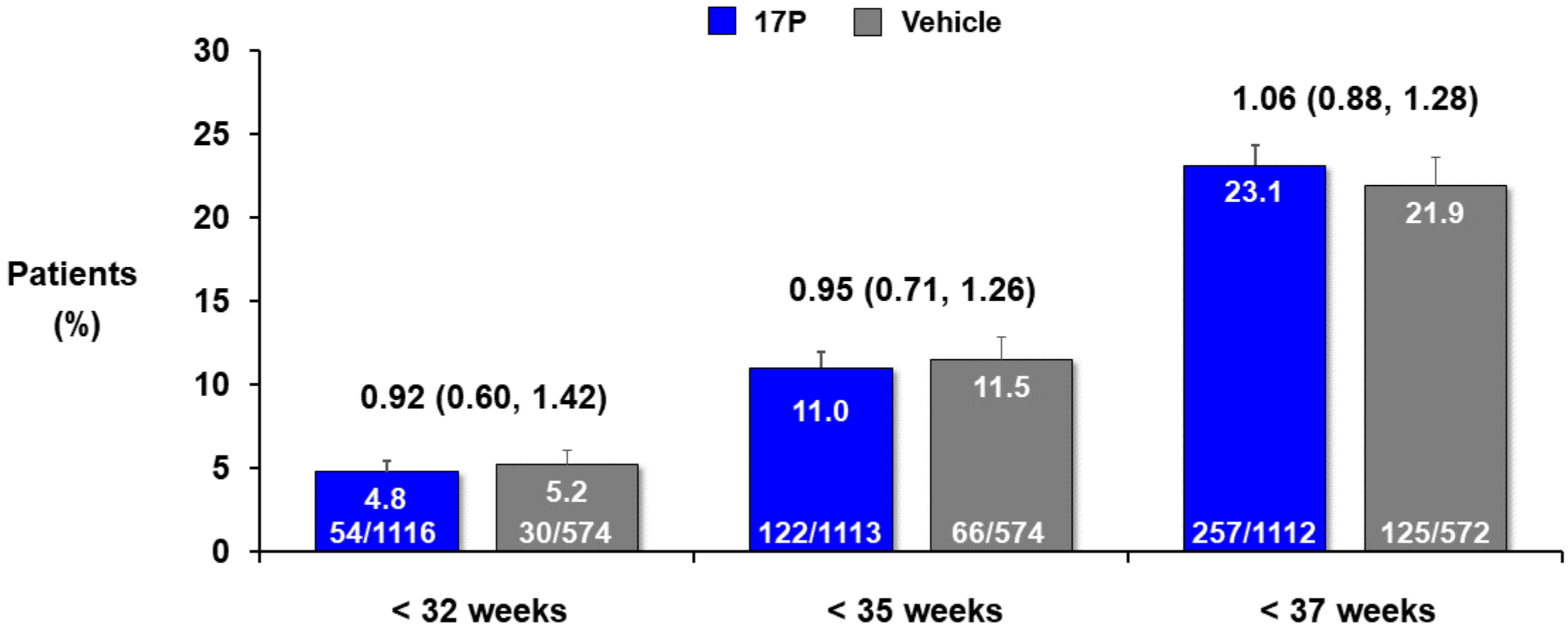
PROLONG: Co-Primary Endpoint Results
PTB < 35 Weeks and Neonatal Composite Index





*Composite included: Death, RDS, BPD, Grade 3 or 4 IVH, NEC and proven sepsis

PROLONG: Key Secondary Endpoint Results (PTB by Gestational Age at Delivery)



PROLONG: Co-Primary Efficacy Outcome Event Rates Higher in US Compared to Ex-US

Preterm Birth < 35 weeks

	17P (N=1130)	Vehicle (N=578)
US, n/N (%)	40/256 (15.6)	23/131 (17.6)
Ex-US, n/N (%)	82/857 (9.6)	43/443 (9.7)

Neonatal Composite Index

	17P (N=1093)	Vehicle (N=559)
US, n/N (%)	18/252 (7.1)	11/125 (8.8)
Ex-US, n/N (%)	43/841 (5.1)	17/434 (3.9)

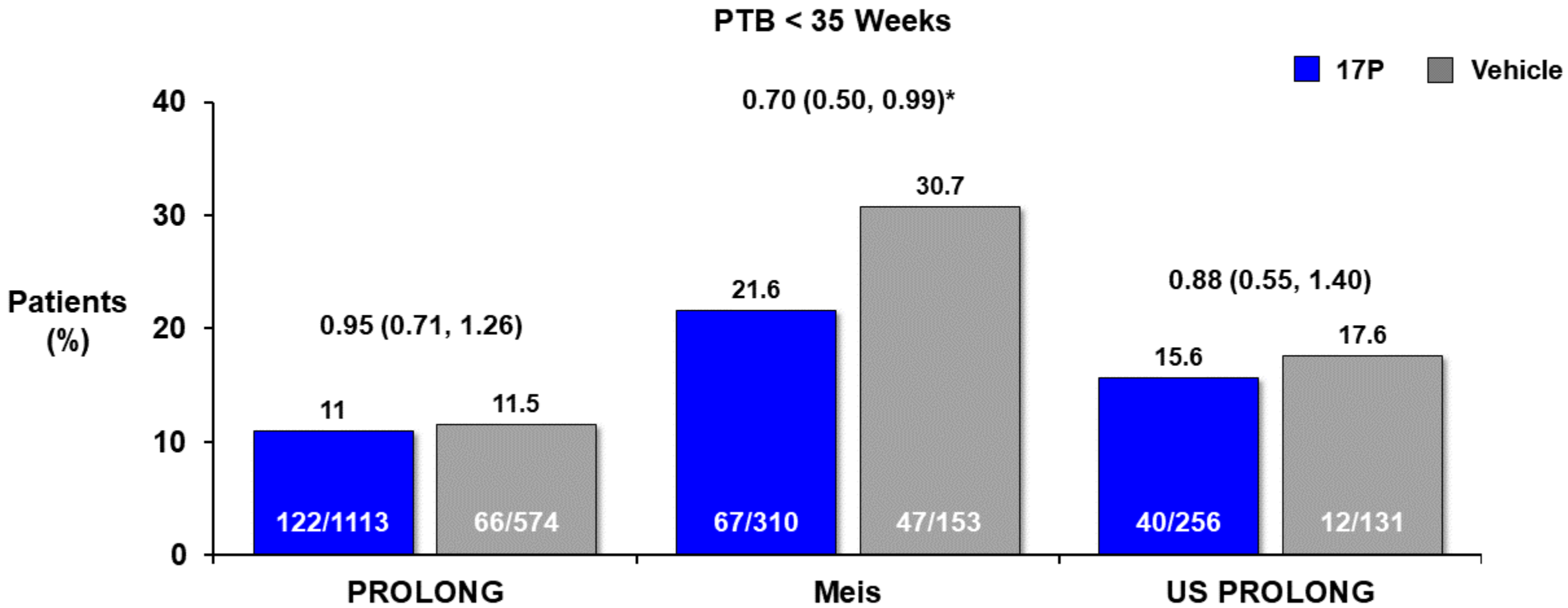
Differences Between PROLONG and Meis Study Populations Less Notable in US PROLONG Subset

Demographics/Baseline Characteristics	PROLONG (N=1708) %	Meis (N=463) %	US PROLONG (N=391) %
Age (years), mean ± SD	30.0 ± 5.2	26.2 ± 5.6	27.6 ± 5.1
> 1 previous SPTB	14.5	28.9	27.4
Black / African American	6.7	59.0	28.9
Hispanic or Latino	9.1	14.9	13.8
Unmarried with no partner	10.1	50.3	30.7
Educational status (≤ 12 years)	43.7	71.3	50.5
Any substance use during pregnancy	9.3	26.1	28.4
Smoking	7.8	21.6	22.8
Alcohol	2.5	8.0	9.2
Illicit drugs	1.4	3.2	5.9

Multifactorial Causes of PTB Make it Challenging to Identify Markers of Response

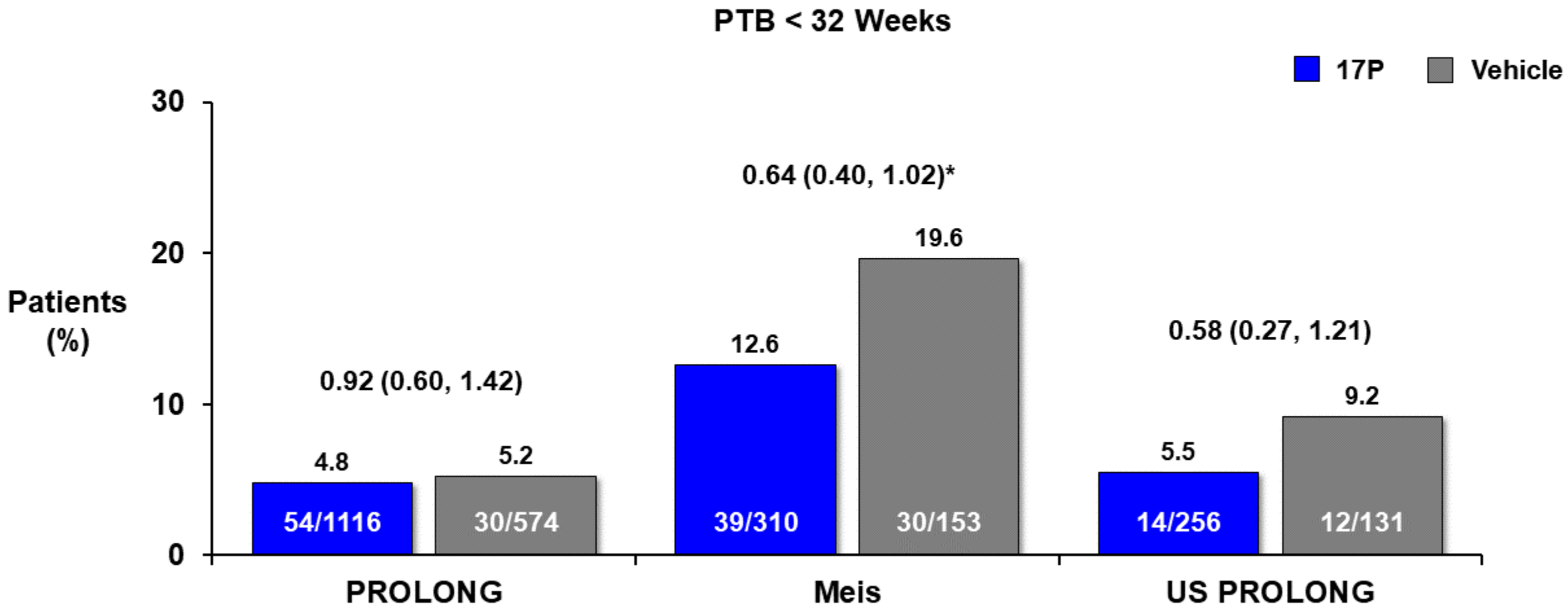
- Additional post hoc analyses conducted
- US PROLONG subset more similar demographics and background characteristics to Meis

Trends for Reductions in PTB Rates at < 35 Weeks in US PROLONG Align (Directionally) with Meis



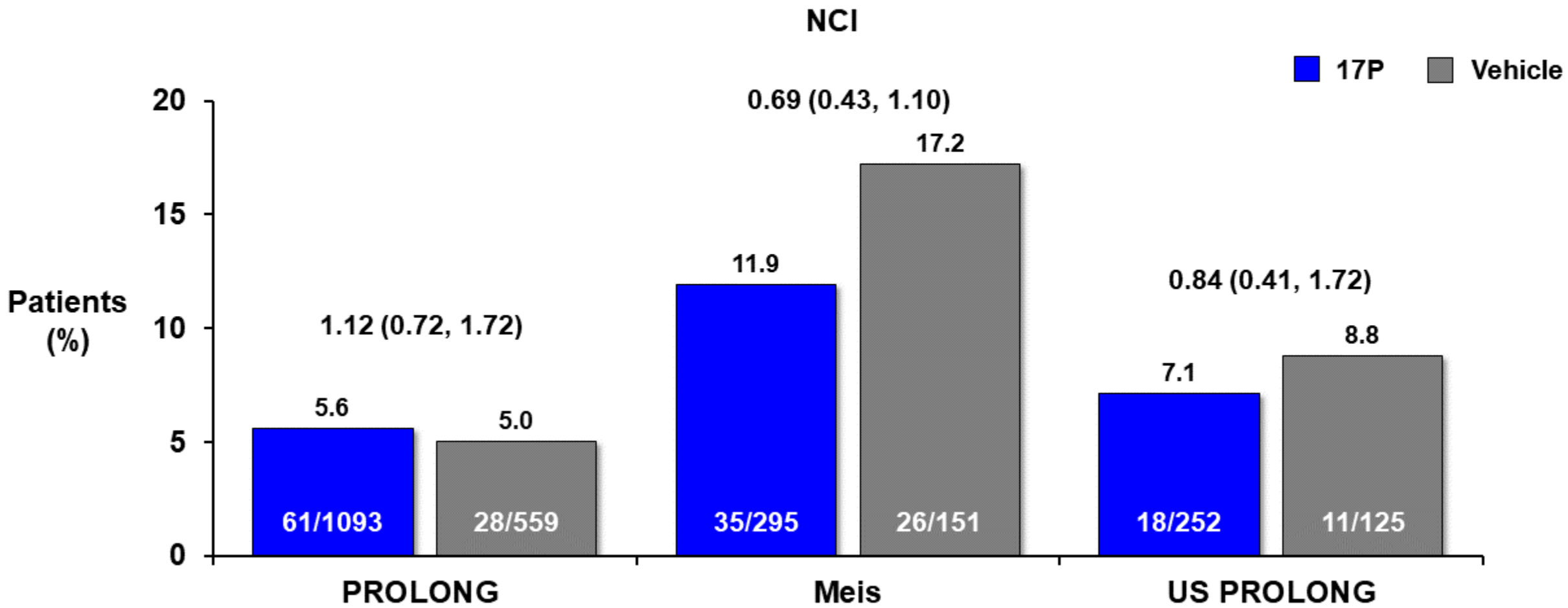
*The CI is a 96.5% CI to adjust for the interim analyses.
Relative risks (RR) and confidence intervals (CI) for the PROLONG study are adjusted for gestational age at randomization stratum.

Trends for Reductions in PTB Rates at < 32 Weeks in US PROLONG Align (Directionally) with Meis



*The CI is a 96.5% CI to adjust for the interim analyses.
Relative risks (RR) and confidence intervals (CI) for the PROLONG study are adjusted for gestational age at randomization stratum.

Trends for Reductions in Neonatal Composite Index* Rates in US PROLONG Align (Directionally) with Meis



*Composite included: Death, RDS, BPD, Grade 3 or 4 IVH, NEC and proven sepsis

PROLONG Efficacy Summary

- Study did not meet co-primary endpoints
  - Findings do not refute robust efficacy seen in Meis
  - Lower background PTB rates in PROLONG compared to Meis
- Trends for benefit favoring 17P seen in smaller subset study population (US PROLONG)

PROLONG Safety

PROLONG: Primary Safety Outcomes

- **Exclude doubling in risk of fetal or early infant death in 17P group vs. vehicle, defined as**
  - Spontaneous abortion/miscarriage (delivery 16^0 – 19^6)
  - Stillbirth at ≥ 20 weeks
  - Early infant death at ≤ 24 weeks (occurring minutes after birth until 28 days of life)

- **Overall perinatal death most relevant outcome**

PROLONG: Overall Rates of Perinatal Death Low and Similar Across Treatment Groups

Fetal or Early Infant Deaths		17P (N=1128)	Vehicle (N=578)	RR (95% CI)[1]
Non-Liveborn	Miscarriages (< 20 weeks)	4/866 (0.5)	7/448 (1.3)	0.28 (0.08, 0.94)
	Stillbirths (≥ 20 weeks)	12/1124 (1.1)	3/571 (0.5)	2.07 (0.59, 7.29)
Liveborn	Early Infant Deaths (≥ 20 to ≤ 24 weeks)	3/1112 (0.3)	1/568 (0.2)	1.48 (0.14, 15.24)
Total Fetal/Early Infant Deaths		19/1128 (1.7)	11/578 (1.9)	0.87 (0.42, 1.81)

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

PROLONG: Overall Incidence of Treatment Emergent Adverse Events (TEAEs) Comparable Between 17P and Vehicle

Summary of TEAEs	17P (N=1128) n (%)	Vehicle (N=578) n (%)
Any AEs	646 (57.3)	334 (57.8)
Any maternal pregnancy complication (MPC)	115 (10.2)	64 (11.1)
Any AEs leading to study drug withdrawal	11 (1.0)	5 (0.9)
Any SAEs	34 (3.0)	18 (3.1)
Maternal deaths	0	0

PROLONG: No Clinically Meaningful Differences in AEs or Maternal Pregnancy Complications (MPCs)* Between Treatment Groups

AEs and MPCs (≥ 3%) Preferred Term	17P (N=1128) / n (%)	Vehicle (N=578) / n (%)
Patients with ≥1 TEAE or MPC	653 (57.9)	336 (58.1)
Anemia	104 (9.2)	56 (9.7)
Headache	68 (6.0)	28 (4.8)
Nausea	55 (4.9)	26 (4.5)
Back pain	50 (4.4)	20 (3.5)
After birth pain	48 (4.3)	24 (4.2)
Urinary tract infection	44 (3.9)	23 (4.6)
Abdominal pain	40 (3.5)	27 (4.7)
Vomiting	42 (3.7)	19 (3.3)
Injection site pruritis	42 (3.7)	23 (4.0)
Vaginal infection	41 (3.6)	21 (3.6)
Nasopharyngitis	39 (3.5)	27 (4.7)
Constipation	38 (3.4)	17 (2.9)
Dyspepsia	37 (3.3)	25 (4.3)
Insomnia	36 (3.2)	13 (2.2)
Injection site pain	36 (3.2)	24 (4.2)
Vaginitis bacterial	35 (3.1)	21 (3.6)
Gestational diabetes	35 (3.1)	22 (3.8)
Cervical incompetence*	34 (3.0)	16 (2.8)

*MPC = maternal pregnancy complication

PROLONG: TEAEs and MPCs* Leading to Premature Discontinuation of Study Medication

TEAE/MPC Leading to Discontinuation Preferred Term	17P (N=1128)	Vehicle (N=578)
Patients with ≥ 1 TEAE/MPC leading to discontinuation	11 (1.0)	5 (0.9)
Hypothyroidism*	1 (0.1)	0
Nausea/vomiting[1]	1 (0.1)	0
Injection site AEs (erythema, nodule, pruritus, rash, reaction)	4 (0.4)	1 (0.2)
Cholestasis*	0	2 (0.3)
Headache	0	1 (0.2)
Fetal growth restriction*	1 (0.1)	0
Preeclampsia*	0	1 (0.2)
Mood altered[1]	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

*MPC [1]AE occurred in same patient

PROLONG: Most Commonly Reported Serious Adverse Events (SAEs) and MPCs*

SAEs and MPCs (≥ 2 patients) Preferred Term	17P (N=1128) n (%)	Vehicle (N=578) n (%)
Patients ≥ 1 SAE/MPC	34 (3.0)	18 (3.1)
Cholestasis*	0	3 (0.5)
Endometritis	1 (0.1)	1 (0.2)
Escherichia coli sepsis	2 (0.2)	0
Migraine	1 (0.1)	1 (0.2)
Placental insufficiency*	4 (0.4)	1 (0.2)
Pneumonia	3 (0.3)	0
Premature separation of placenta*	5 (0.4)	2 (0.3)
Pyelonephritis	2 (0.2)	1 (0.2)
Wound infection	2 (0.2)	0

2 patients each had 1 SAE considered possibly related to study drug: 1 in 17P group hospitalized for mild nephrolithiasis; 1 in vehicle group with severe cholestasis

Post-Marketing Surveillance: Safety Consistent with Clinical Trial Data

- Cumulative exposure of 294,781 patients since approval
- Post-marketing data consistent with safety data obtained from Meis and PROLONG
  - No new or unexpected safety findings
- Most frequent AE reports consistent with registration studies
  - Injection site pain and/or other injections site localized reactions (e.g., pruritus, nodule, swelling)

Post-Marketing Surveillance: Makena SAEs Around Perinatal Mortality

SAE: Death		Estimated Post-marketing Reporting Rate*
Non-Liveborn	Abortion spontaneous	0.1%
	Stillbirth	0.1%
Liveborn	Death Neonatal	0.003%
Total Perinatal Deaths		0.2%

*Reporting Rate is computed based on cumulative patient exposure of 294,781 as of Aug 2019

PROLONG Reaffirmed Safety of 17P Demonstrated in Meis Study

- No new or unexpected safety findings
- No clinically meaningful difference in safety profile between treatment arms
- Consistent, favorable maternal and fetal safety comparable to vehicle
- Consistent findings in post-marketing surveillance data

Prevention of Preterm Births: Clinical Perspective

Sean C. Blackwell, MD

Professor and Chair, Department of Obstetrics, Gynecology, and Reproductive Sciences

McGovern Medical School-UTHealth

3 Key Clinical Questions

1. Why did PROLONG efficacy results differ from Meis results?

2. Is it feasible to do another confirmatory trial?

3. What do we do from here?

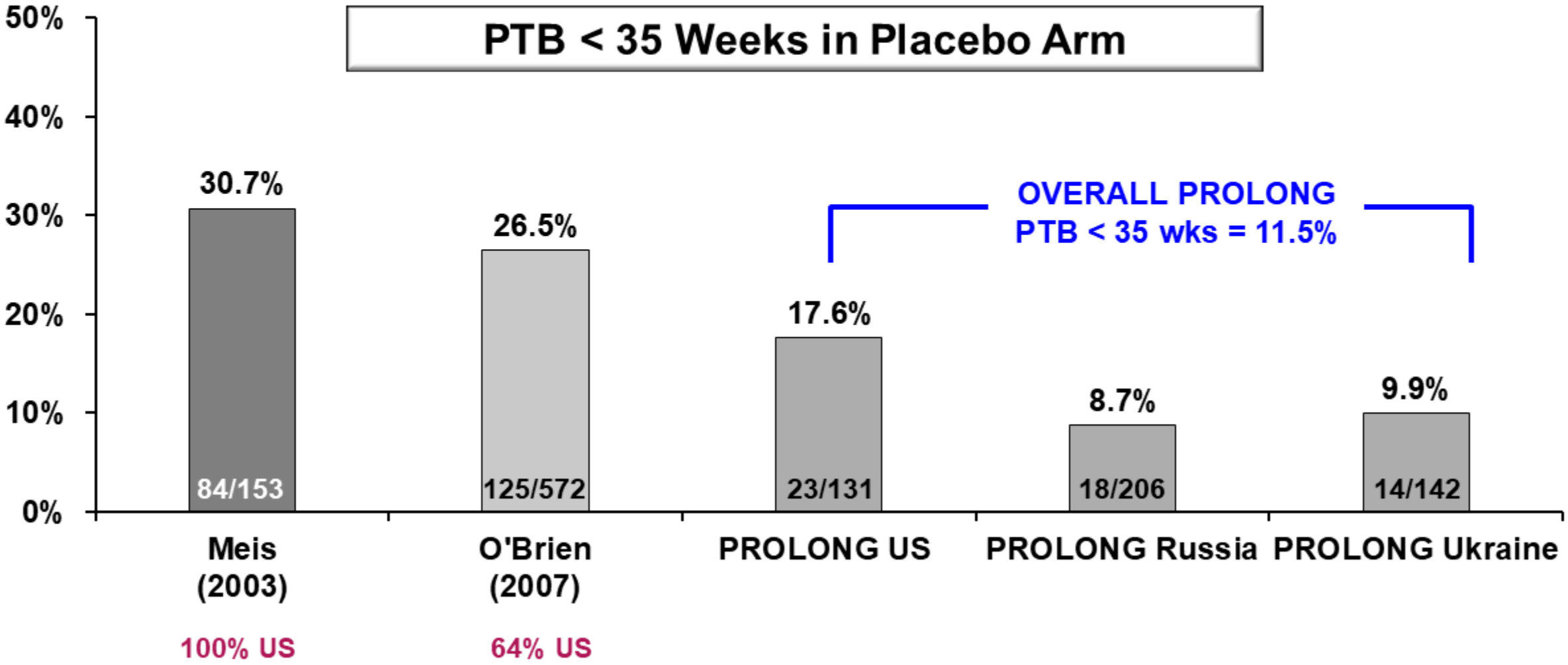
Question #1

Why did PROLONG efficacy results differ from Meis?

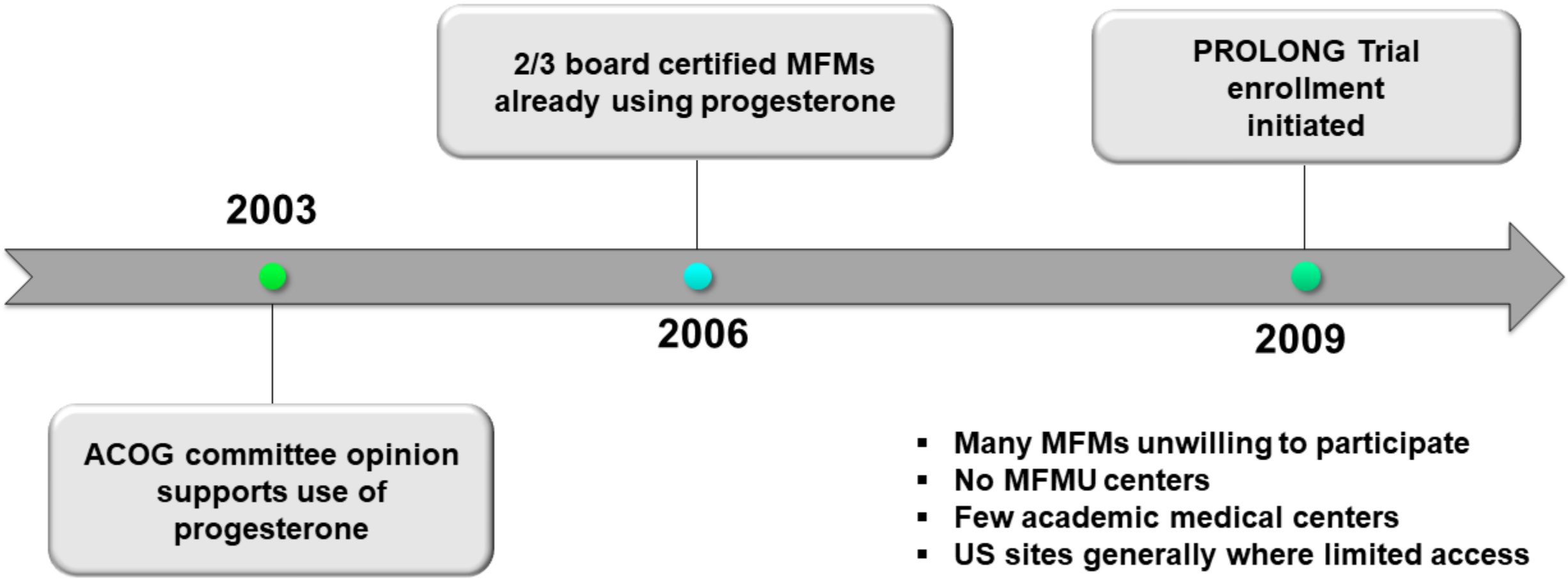
Differences in Clinical Characteristics Between Meis And PROLONG Study Populations

	Meis	PROLONG
Recruitment	100% US Academic Medical Centers	75% ex-US 60% Russia & Ukraine
Race	60% Black	7% Black
Surrogates of SES		
Unmarried with no partner	50%	10%
Smoking	22%	8%
≤ 12 years of education	71%	44%
> 1 prior PTB	27%	14%
PTB in placebo groups		
< 32 wks	19.6%	5.2%
< 35 wks	30.7%	11.5%
< 37 wks	54.9%	21.9%

Placebo Arm PTB Rates Across Different Clinical Trial Populations



PROLONG US: Recruitment Challenges



- Many MFMs unwilling to participate
- No MFMU centers
- Few academic medical centers
- US sites generally where limited access to 17P

Ness, AJOG, 2006

PROLONG US: Enrollment Challenges

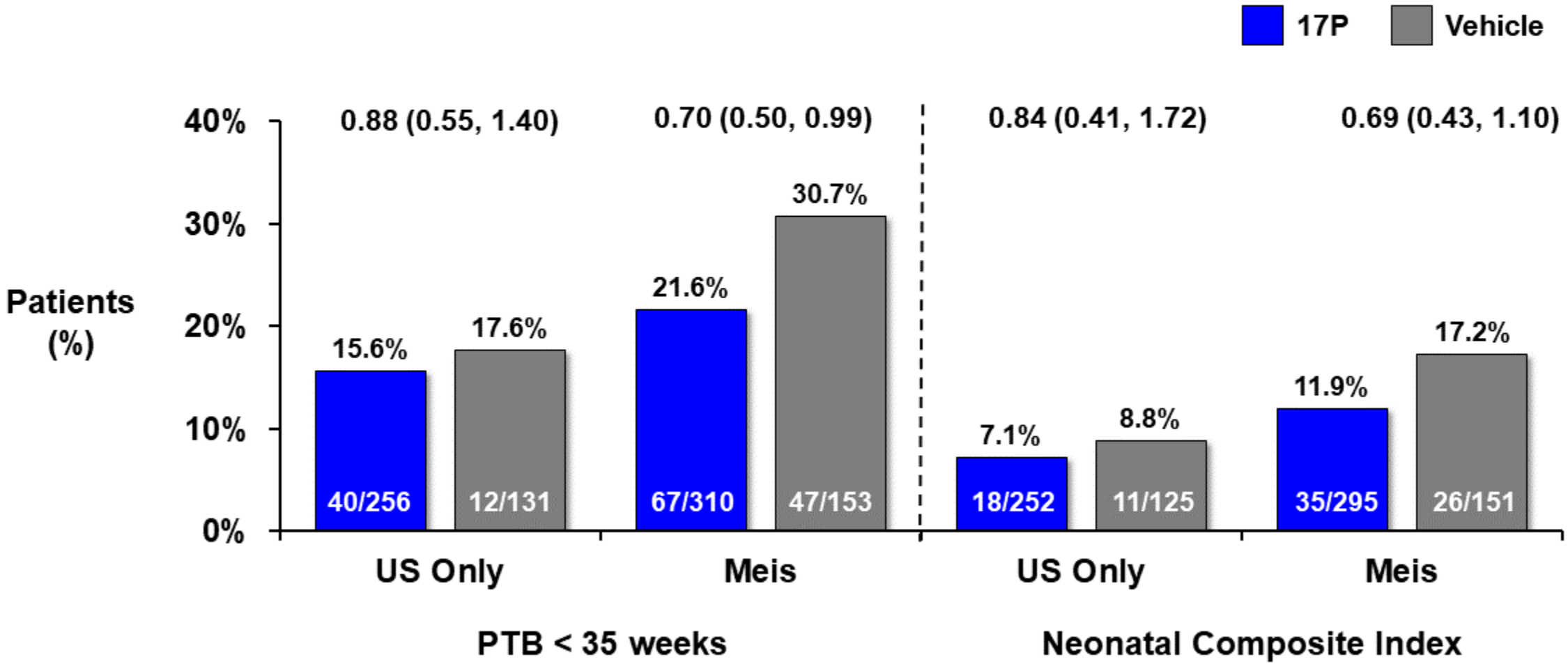
- Patients potentially steered from RCT to get open-label therapy (compounded 17P, vaginal progesterone, other)

- PROLONG
  - Low rate PTB > 1
  - Very low rate short cervix (< 2%)

PROLONG: Low Event Rates

- Sample size and expected event rate based on Meis trial
- 50% lower rates in PROLONG than Meis
- To design a trial today based on these low rates
  - 90% power would require
    - 3,600 women for PTB < 35 weeks
    - 6,000 for neonatal composite morbidity
- Population differences and low event rates make PROLONG results inconclusive

PROLONG: Treatment Effect Trends in US Only



96.5% CI to adjust for the interim analyses and number of prior preterm birth for PTB<35
95%CI adjusted for number of prior preterm birth for NCI

Question #2

Is it feasible to do another confirmatory trial?

Another Confirmatory Trial is Not Feasible

- US physicians won't accept placebo controlled RCT

- Where could we do another placebo controlled RCT?
  - Difficulty finding combination of
    - High-risk women
    - No access to 17P
    - Research infrastructure to conduct major trial

Feasibility of Another Confirmatory Trial: Trial of Two Therapies?

- No evidence-based therapies vs. 17P
- Vaginal progesterone: no benefit in 3 large RCTs
- Cervical cerclage and pessary also no proven benefit

	N	Endpoint	Vaginal progesterone	Placebo	95% CI
O'Brien	659	PTB ≤ 32 weeks	10%	11.3%	OR 0.9 (0.52 to 1.56)
OPPTIMUM*	1,053	PTB < 34 weeks or fetal death	15.9%	18.8%	OR 0.82 (0.58 to 1.16)
PROGRESS*	787	PTB < 37 weeks	36.5%	37.2%	aRR 0.97 (0.81 to 1.17)

*Included 12 women with twin pregnancies

1.O'Brien et al., 2007. 2. Norman et al., The Lancet. 2016. 3. Crowther et al., 2017

Question #3

What should we do from here?

SMFM Statement (October 25, 2019)

"Based on the evidence of effectiveness in the Meis study, which is the trial with the largest number of US patients, and given the lack of demonstrated safety concerns, SMFM believes that it is reasonable for providers to use 17-OHPC in women with a profile more representative of the very high-risk population reported in the Meis trial."

ACOG Practice Advisory (October 25, 2019)

"ACOG is not changing our clinical recommendations at this time and continues to recommend offering hydroxyprogesterone caproate as outlined in Practice Bulletin # 130, Prediction and Prevention of Preterm Birth."

What Will Happen if FDA-Approved 17P is Not Available?

- Many clinicians will use compounded 17P
  - Lack of GMP
  - Potential variance/sterility issues
  - No labeling
- Most physicians will not tolerate NO TREATMENT
  - Other off-label therapies (non-evidence based)
  - Many will choose cerclage (surgical therapy)

What Will I Do?

- Meis and PROLONG not contradictory
  - Meis showed efficacy in population similar to my patients
  - PROLONG reaffirms safety
  - Overall positive benefit/risk ratio

- Essential to be able to offer FDA-approved 17P

AMAG Actions Following PROLONG

Julie Krop, MD
Chief Medical Officer
EVP Clinical Development and Regulatory Affairs
AMAG Pharmaceuticals, Inc

Totality of Evidence Support Continued 17P Access

- Meis study demonstrated substantial evidence of efficacy
  - Basis of medical societies recommending 17P as standard of care
- PROLONG results inconclusive given differences in patient populations

What Have We Learned from PROLONG?

- Impossible to conduct confirmatory trial once 17P was recommended by medical societies as standard of care
  - Lead to bias towards lower risk population
- PROLONG confirmed favorable safety profile
  - Supported by 8 years of post-marketing surveillance

Does Meis Trial Alone Meet Criteria for Single Trial as Basis for Approval?

- FDA guidance for single trial approval
  - Second trial not feasible or ethical
  - Statistically persuasive findings
  - Clinically relevant endpoint
  - Robust, consistent results across multiple subgroups
- PTB at <37, < 35 and < 32 weeks increases risk to neonate
  - Should no longer require a confirmatory trial
- Orphan disease with NO alternative treatment options

FDA Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products, 1998

17P is an Important Treatment Option for Pregnant Women With History of Preterm Birth

- Physicians and patients can make informed decisions together
- PROLONG results recently published in American Journal of Perinatology
- Label update with PROLONG safety and efficacy data

Considerations for Another Confirmatory Study

- Randomized placebo-controlled trial
  - Not feasible given current clinical practice guidelines
- Observational study
  - Challenging to control for known and unknown PTB risk factors

Positive Benefit-Risk Profile of 17P Supports Continued Access for Physicians and Patients

- Preterm birth remains major US public health concern
- Critical to keep 17P available to patients who need it most

17 α-Hydroxyprogesterone Caproate (Makena®) for Women with Singleton Pregnancy and Prior Singleton Spontaneous Birth

FDA Advisory Committee Meeting

Division of Bone, Reproductive and Urologic Products

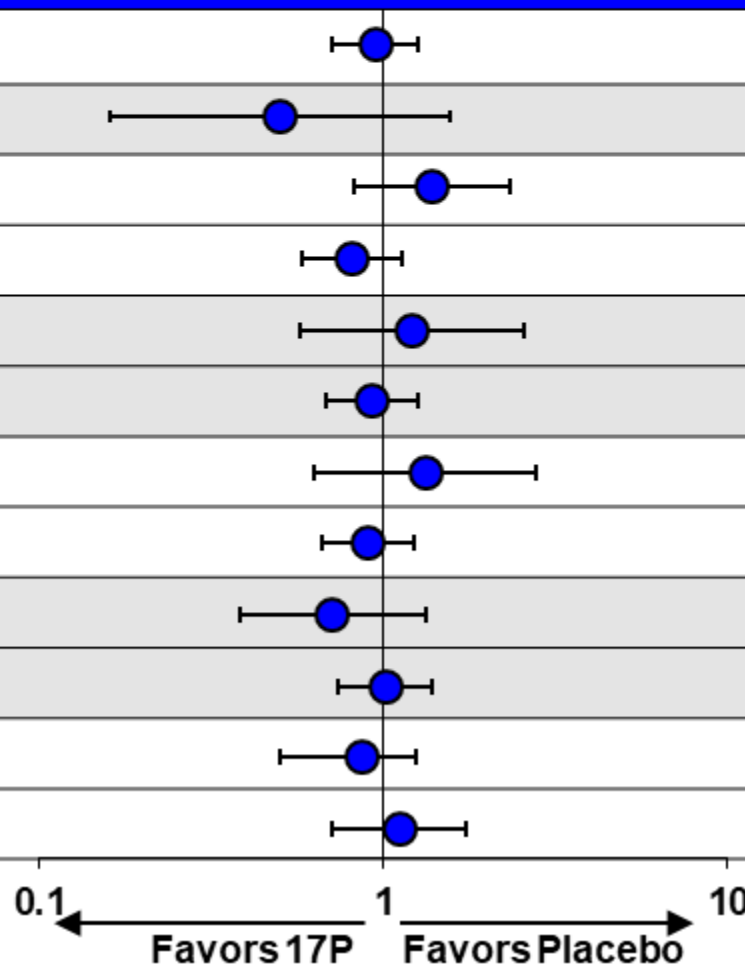
AMAG Pharmaceuticals, Inc.

October 29, 2019

Back-up Slides Shown

PROLONG Study: PTB < 35 Weeks with 17P Across Multiple Subgroups

Subgroup	17P (n/N)	Vehicle (n/N)	Relative Risk (95% CI)
All patients	122/1113	66/574	0.95 (0.71, 1.26)
Cervical Length < 25 mm	4/18	4/9	0.50 (0.16, 1.55)
> 1 Prior SPTB	42/164	15/81	1.38 (0.82, 2.34)
Only 1 Prior SPTB	80/949	51/491	0.81 (0.58, 1.13)
Black	17/72	8/41	1.21 (0.57, 2.56)
Non-Black	105/1041	58/533	0.93 (0.68, 1.26)
Unmarried	22/117	8/56	1.32 (0.63, 2.77)
Married	100/996	58/518	0.90 (0.66, 1.22)
Smoke or substance use	19/105	13/51	0.71 (0.38, 1.32)
No smoke or substance use	103/1008	53/523	1.01 (0.74, 1.38)
Education ≤ 12 years	64/474	40/256	0.86 (0.50, 1.24)
Education > 12 years	58/638	26/318	1.11 (0.71, 1.73)



Stillbirth: PROLONG and Meis MFM Review

Stillbirth affects 1 in 160 pregnancies each year in general population
Several underlying fetal/maternal causes[1]

- **PROLONG**:
  - 17P: 12/1128 (1.1%)
    - 11 underlying factors; 1 unknown
  - Vehicle: 3/578 (0.5%)
    - All had underlying factors
- **Meis**:
  - 17P: 6/306 (2.0%)
    - 5 underlying factors; 1 unknown
  - Vehicle: 2/153 (1.3%)
    - 1 underlying factor; 1 unknown

1. Silver JAMA 2011

Makena (hydroxyprogesterone caproate injection)
New Drug Application 021945/Supplement 023

Opening Remarks

Bone, Reproductive and Urologic Drugs Advisory Committee Meeting

October 29, 2019

Christine P. Nguyen, M.D.
Deputy Director for Safety
Division of Bone, Reproductive and Urologic Products
Office of New Drugs, Center for Drug Evaluation and Research
Food and Drug Administration

Clinical Background

- Neonatal mortality and morbidity from preterm birth (PTB) is a significant public health concern
- No therapies approved to reduce the risk of neonatal mortality and morbidity from prematurity
- Progestogens (intravaginal or intramuscular) used to reduce the risk of PTB
  - Only Makena (hydroxyprogesterone caproate injection) approved for reducing the risk of recurrent PTB

Regulatory History

- Makena approved in 2011 under accelerated approval to reduce the risk of PTB in women with a singleton pregnancy and a prior spontaneous PTB
- Approval: a single trial conducted 1999-2002 in the U.S., based on surrogate endpoint of gestational age (GA) of delivery <37 weeks
- As required under accelerated approval regulations, the Applicant conducted a postapproval confirmatory trial to verify clinical benefit for the neonate

Confirmatory Trial - 003

- International, randomized, double-blind, placebo-controlled trial in 1708 pregnant women
  - Russia, Ukraine, and U.S. enrolled 36%, 25%, and 23% subjects
- Design, eligibility criteria similar to Trial 002, except for primary endpoints
  - Trial 002: GA at delivery <37 weeks
  - Trial 003: GA at delivery <35 weeks, neonatal morbidity/mortality index
- Conducted 2009-2018

Trial 003 Results: No Treatment Effect

Efficacy Endpoints* (% of patients)	Makena (N=1130)	Placebo (N=578)	Difference (95% CI)	P-value
Coprimary: Neonatal composite index (%)	5.4	5.2	0.2 (-2.0, 2.5)	0.84
Coprimary: PTB <35^0 weeks (%)	11.0	11.5	-0.6 (-3.8, 2.6)	0.72
PTB <32^0 weeks (%)	4.8	5.2	-0.4 (-2.8, 1.7)	
PTB <37^0 weeks (%)	23.1	21.9	1.3 (-3.0, 5.4)	

*FDA's Analysis

Trial 003 Exploratory Subgroup Analyses

- No statistically significant treatment difference or interaction between treatment effect and these factors:
  - Region (U.S. vs. non-U.S.)
  - Race (Black vs. Non-Black)
  - Elements that may increase PTB risk:
    - 1 vs. >1 prior PTB, substance use in pregnancy, ≤12 years of education, single/no partner

  ❖ These factors may be prognostic, but they do not appear to be effect modifiers

- There was no consistent, convincing evidence of a treatment effect within any particular subpopulation across Trials 002 and 003.

Totality of Evidence: Trial 002 and Trial 003

- Trial 002 - efficacy on gestational age of delivery (*surrogate endpoint*)
  - Conducted 1999-2002 in the U.S.
  - Issues regarding generalizability: ~60% self-identified black, all from academic centers, 27% from a single center, high recurrent preterm birth rate <37 weeks in placebo arm (55%)
- Trial 003 – no efficacy on neonatal outcomes (*clinical endpoint*) or gestational age at delivery (*surrogate endpoint*)
  - Conducted 2009-2018, powered to detect treatment effect in Trial 002
  - International (23% from the U.S.), lower risk population, lower recurrent preterm birth rate in placebo arm than in Trial 002

Totality of Evidence

Endpoint	Efficacy on Endpoint	Approval Efficacy Requirement Issues
Surrogate endpoint: GA at delivery	Yes (Trial 002) No (Trial 003) ❖ **Conflicting efficacy findings**	**Issue 1:** Substantial Evidence of Effectiveness
Clinical endpoint: Neonatal composite index	No (Trial 003) ❖ **No verification of clinical benefit**	**Issue 2:** Accelerated Approval

Issue 1: Substantial Evidence of Effectiveness

- Statutory standard of establishing efficacy for FDA drug approval*, including accelerated approval
  - Traditionally, significant findings from ≥ 2 adequate and well-controlled trials, each convincing on its own (independent substantiation) on the efficacy endpoint(s), reduces risk false positive from chance or bias

- When appropriate, a single adequate, well-controlled trial with persuasive findings may be accepted as substantial evidence

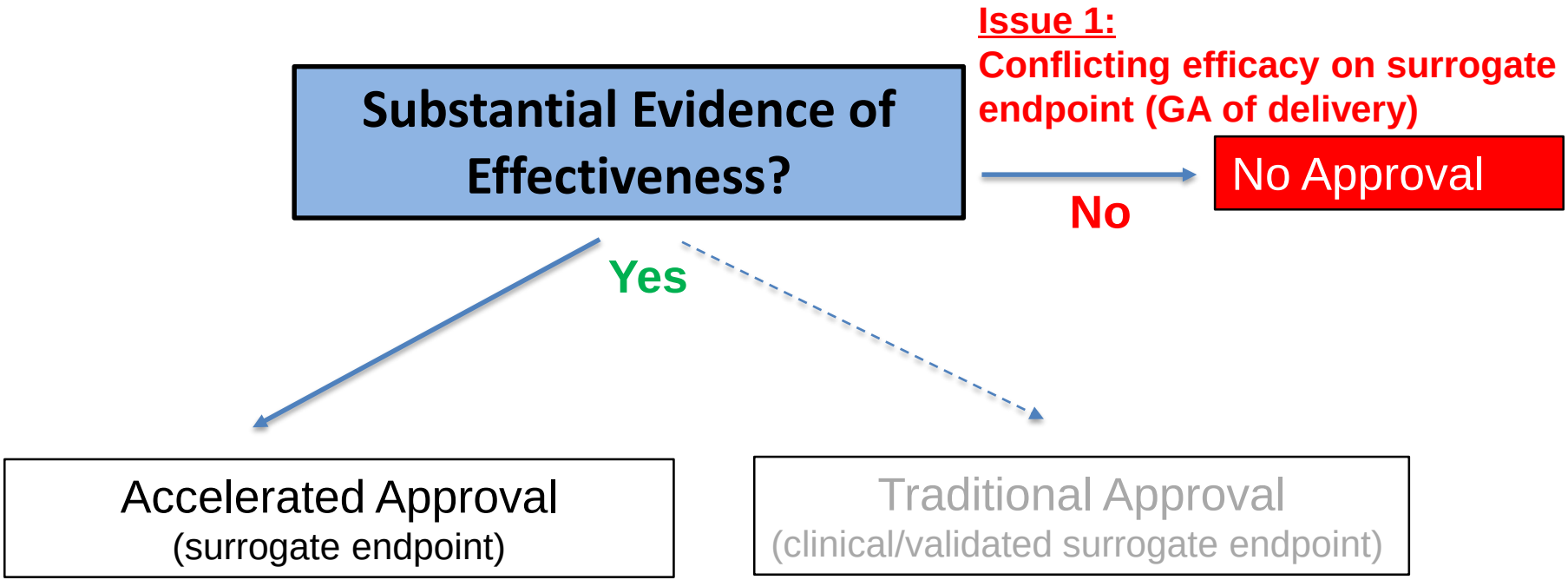
*Substantial evidence defined in section 505(d) of the Act as "evidence consisting of adequate and well-controlled investigations.."

Issue 1: Substantial Evidence of Effectiveness

- 2011 accelerated approval of Makena based on a single trial
- If there were additional adequate and well-controlled trials in 2011, FDA would have considered those data when deciding about substantial evidence of effectiveness
- Now there are 2 adequate and well-controlled trials (Trials 002 and 003)

Issue 1: Trial 003 did not substantiate Makena's treatment effect on GA of delivery: Is there still substantial evidence of the drug's effect on reducing the risk of preterm birth?

Issue 1: Substantial Evidence of Effectiveness



Issue 2: Accelerated Approval

- Traditional approval: based on *clinical endpoint* (directly measures how patients feel, function, or survive) or *validated surrogate endpoint*
- Accelerated approval: based on a *surrogate endpoint* reasonably likely to predict clinical benefit
  - Expedited drug development pathway
  - Reserved for certain drugs treating serious/life-threatening conditions with unmet medical need
  - Must meet same statutory effectiveness standards as those for traditional approval

Issue 2: Accelerated Approval

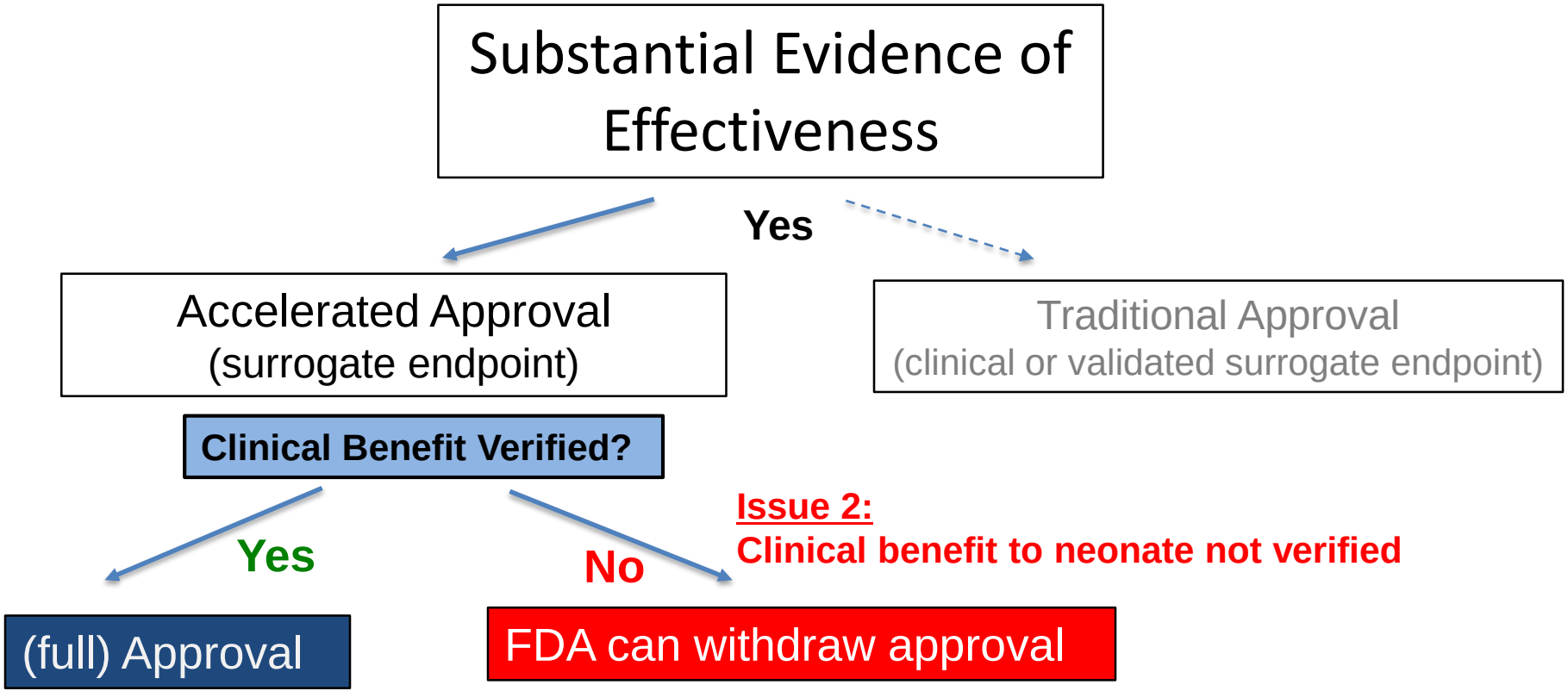
- Makena accelerated approval based on treatment effect on *surrogate endpoint (GA of delivery)*
  - GA of delivery is not a direct measure of how neonates feel, function, or survive
  - Spontaneous PTB poorly understood syndrome with potential for multiple pathophysiologic pathways
  - Prolonging GA of delivery may not consistently translate into improved neonatal outcomes

Issue 2: Accelerated Approval

- More uncertainty at the time of approval that the treatment effect on surrogate endpoint (GA at delivery) will translate into clinical benefit (neonatal outcomes)
  - Therefore, must undergo a postapproval confirmatory trial to verify clinical benefit
- FDA can withdraw approval of the drug or indication if the Applicant does not conduct the required trial(s) with due diligence or the trial(s) fail to verify clinical benefit

Issue 2: Trial 003 did not verify Makena's clinical benefit to the neonate

Issue 2: Accelerated Approval



Discussion and Voting Questions

Discussion Question 1

- Discuss the effectiveness of Makena on recurrent preterm birth and neonatal morbidity and mortality.

Discussion Question 2

- If a new confirmatory trial were to be conducted, discuss the study design, including control, dose(s) of study medication, efficacy endpoints and the feasibility of completing such a trial.

Discussion Question 3

- Discuss the potential consequences of withdrawing Makena on patients and clinical practice.

Voting Question 4

- Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal outcomes?
  - Provide rationale for your vote.

Voting Question 5

- Based on the findings from Trial 002 and Trial 003, is there substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth?
  - Provide rationale for your vote.

Voting Question 6

FDA approval, including accelerated approval, of a drug requires *substantial evidence of effectiveness (Issue 1).*

For drugs approved under the accelerated approval pathway based on a surrogate endpoint, the Applicant is required to conduct confirmatory trial(s) to *verify clinical benefit (Issue 2)*. If the Applicant fails to conduct such a trial(s) or if such trial(s) does not verify clinical benefit, FDA may, following an opportunity for a hearing, withdraw approval.

Voting Question 6 Continued

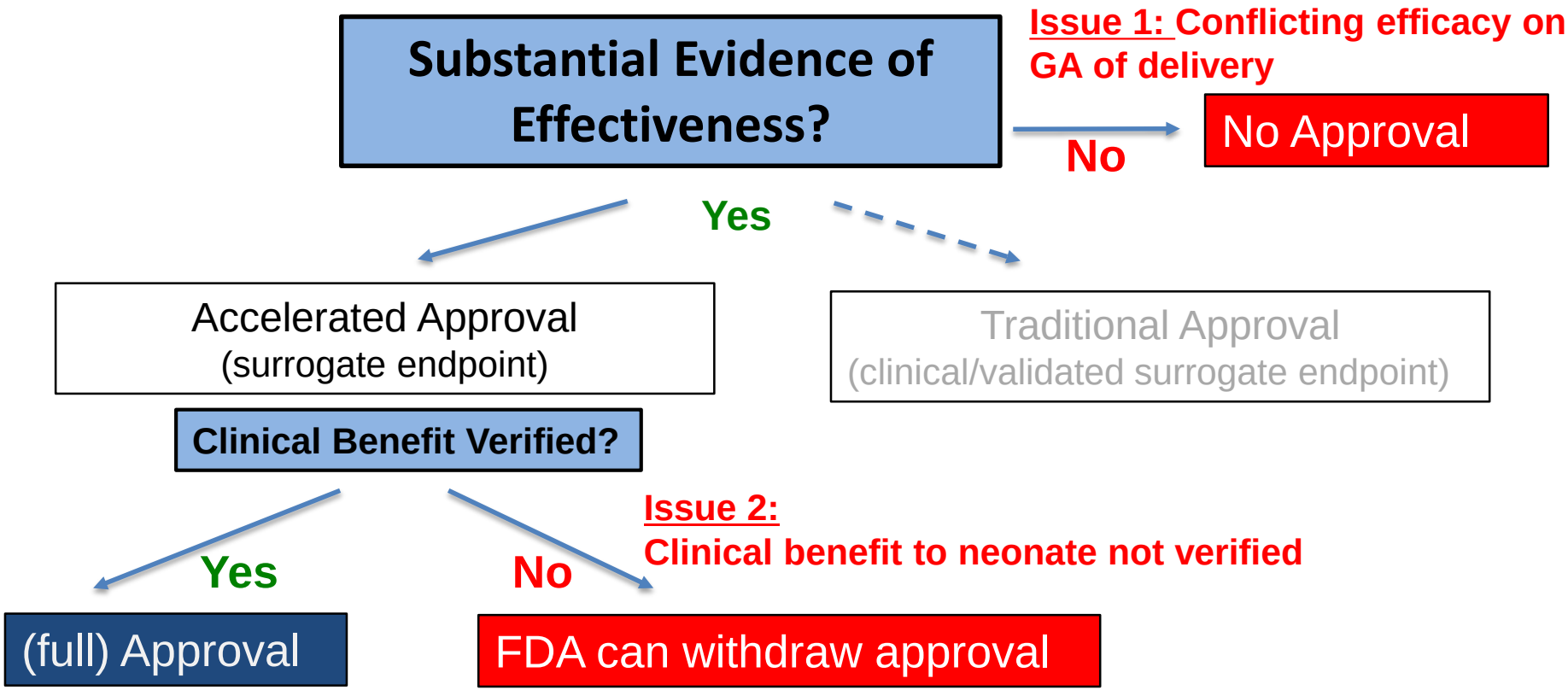
- Should FDA:

  (A) Pursue withdrawal of approval for Makena

  (B) Leave Makena on the market under accelerated approval and require a new confirmatory trial

  (C) Leave Makena on the market without requiring a new confirmatory trial

Approval: Efficacy Requirement Issues



Voting Question 6 Continued

- **Vote A** (withdraw approval) may be appropriate if you believe the totality of evidence does not support Makena's effectiveness for its intended use.
  - Discuss the consequences of Makena removal

Voting Question 6 Continued

- **Vote B** (require a new confirmatory trial) may be appropriate if you believe the totality of evidence supports Makena's effectiveness in reducing the risk of recurrent PTB, but that there is no substantial evidence of effectiveness on neonatal outcomes AND you believe that a new confirmatory trial is necessary and feasible.

  - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent PTB, based on the surrogate endpoint of gestational age at delivery.

  - Also discuss key study elements, including study population, control, dose(s), and efficacy endpoints of the new confirmatory trial (if not previously discussed in Discussion point 2) and approaches to ensure successful completion of such a trial.

Voting Question 6 Continued

- **Vote C** (leave Makena on the market without a new confirmatory trial) may be appropriate if you believe Makena is effective for reducing the risk of recurrent PTB and that it is not necessary to verify Makena's clinical benefit to neonates.
  - Discuss how the existing data provide substantial evidence of effectiveness of Makena in reducing the risk of recurrent PTB and why it is not necessary to verify Makena's clinical benefit to neonates.


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Makena (hydroxyprogesterone caproate injection)
New Drug Application 021945/Supplement 023

Clinical Overview

Bone, Reproductive and Urologic Drugs Advisory Committee Meeting

October 29, 2019

Barbara Wesley, M.D., M.P.H.
Medical Officer
Division of Bone, Reproductive and Urologic Products
Office of New Drugs, Center for Drug Evaluation and Research
Food and Drug Administration

Outline

- Trial 002 and its history (1999-2011)
  - Findings, areas of controversy
- 2006 Advisory Committee
- FDA Actions (2006, 2009, 2011)
- Accelerated approval postmarketing requirement - Confirmatory Trial 003

Background of Trial 002

- 1999-2002: Funded by National Institute of Child Health and Human Development NICHD; conducted by Maternal-Fetal Medicine Units Network (MFMU).
- 2003: Positive findings of hydroxyprogesterone caproate (HPC) reducing the risk of preterm birth <37 weeks published in the New England Journal of Medicine*
- 2006: Submission of new drug application (NDA) for HPC 250 mg/mL

*Meis PJ, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. N Engl J Med. 2003;348(24):2379-85.

Makena

Indication

- To reduce the risk of preterm birth in women with a singleton pregnancy and a history of spontaneous preterm birth

Dosage & Administration

- 250 mg once a week beginning between 16^0 weeks and 20^6 weeks gestation to week 37 of gestation or birth

Trial 002 Design

Study Medications

- HPC in castor oil
- Placebo

Primary Efficacy Endpoint

- Birth $<37^0$ weeks

Additional Efficacy Endpoints (post hoc)

- $<35^0$ weeks and $<32^0$ weeks
- Composite index of neonatal morbidity
  - Death, respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), Grade 3 or 4 intraventricular hemorrhage (IVH), proven sepsis, necrotizing enterocolitis (NEC)

Trial 002: Preterm Births <37^0 Weeks Gestation

Primary Efficacy Endpoint P= 0.001

HPC N = 310	Placebo N = 153	% Difference [Adjusted 95% Confidence Interval]
Number (%) Preterm Births		
115 (37%)	84 (55%)	-18% [-28%, -7%]

- PTB rate of **55%** in placebo arm considerably greater than rate in other MFMU Network studies (~36%)
- PTB rate of **37%** in HPC arm similar to PTB rate in placebo arms in other MFMU Network study

PTB Rate in Placebo Arm by Race in Trial 002

Race	Placebo - n/N (%)
Black	47/90 **(52%)**
Non-black	37/63 **(59%)**

Percent of Preterm Births at Various Gestational Age Thresholds (Trial 002)

Age at Delivery (Weeks)	HPC N=310	Placebo N=153	% Difference [Adjusted 95% Confidence Interval]
	Percent Delivered		
$<37^0$	37	55	-18.0% [-28%, -7.4%]
$<35^0$	21	31	-9.4% [-19.0%, -0.4%]
$<32^0$	12	20	-7.7% [-16.1%, -0.3%]

Makena prescribing information, Drugs@FDA

Confidence intervals **adjusted for the interim analyses and the final analysis**. To preserve overall Type I error rate of 0.05, p-value boundary of 0.035 used for the adjustment (equivalent to a 96.5% confidence interval).

Composite Neonatal Morbidity (Trial 002)

Morbidity	HPC N=295 n (%)	Placebo N=151 n (%)
Death (live births only)	8 (2.6)	9 (5.9)
Respiratory distress syndrome	29 (9.9)	23 (15.3)
Bronchopulmonary dysplasia	4 (1.4)	5 (3.3)
Gr. 3/4 intraventricular hemorrhage	2 (0.7)	0 (0.0)
Proven sepsis	9 (3.1)	4 (2.6)
Necrotizing enterocolitis	0 (0.0)	4 (2.7)
Composite Index of Morbidity*	**35 (12%)**	**26 (17%)**

*** No. subjects with one or more of the listed morbidities**

Summary of Effectiveness Issues

- Applicant sought approval for HPC based on
  - Findings from a single clinical trial
  - A surrogate endpoint for infant mortality/morbidity (preterm birth <37 weeks)

- Concern about generalizability to general U.S. population
  - Notably high preterm birth rate in placebo arm (55%)
  - Approximately 60% Black or African American
  - Enrollment from academic centers only; 27% from one academic center

2006 Advisory Committee Meeting

Which gestational age at birth is an adequate surrogate? (21 members voting)

- PTB <37 weeks – yes = 5
- PTB <35 weeks – yes = 13
- PTB <32 weeks – yes = 20

2006 FDA Action: Not Approved

- Major deficiency: New trial to provide substantial evidence of efficacy - direct benefit on neonatal morbidity and mortality or the surrogate PTB <35 and <32 weeks of gestation

- Address the concern regarding early pregnancy loss

Between 2009 and 2011 FDA Actions: Effect of Late-Preterm Birth

- **Late-Preterm Infants** – defined as infants born between 34 $^{0/7}$ and 36 $^{6/7}$ weeks of gestation: “are often mistakenly believed to be as physiologically and metabolically as mature as term infants”
- Higher rates of infant mortality and morbidity than term infants.

Engle, WA, et al. “Late-preterm” infants: a population at risk. Pediatrics 2007;120:1390-1401.
ACOG Obstetrics Practice Committee Opinion, Number 404, April 2008

2011 FDA Action: Accelerated Approval

- Recent data on effect of FDA to reconsider gestational age at delivery
- FDA concluded that delivering at <37 weeks of gestation was an adequate surrogate endpoint
- Findings of Trial 002 now deemed sufficient to support accelerated approval
- Trial 003 was ongoing and Applicant demonstrated that it could be successfully completed

Applicant's Obligation

As a condition of accelerated approval, the Applicant was required to complete the confirmatory clinical trial of Makena (Trial 003) to verify the clinical benefit to neonates from the reduction in the risk of PTB.


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Makena (hydroxyprogesterone caproate injection)
New Drug Application 021945/Supplement 023

Efficacy in Confirmatory Trial 003

Bone, Reproductive and Urologic Drugs Advisory Committee Meeting

October 29, 2019

Jia Guo, Ph.D.
Statistical Reviewer
Division of Biometrics 3
Office of Biostatistics, Center for Drug Evaluation and Research
Food and Drug Administration

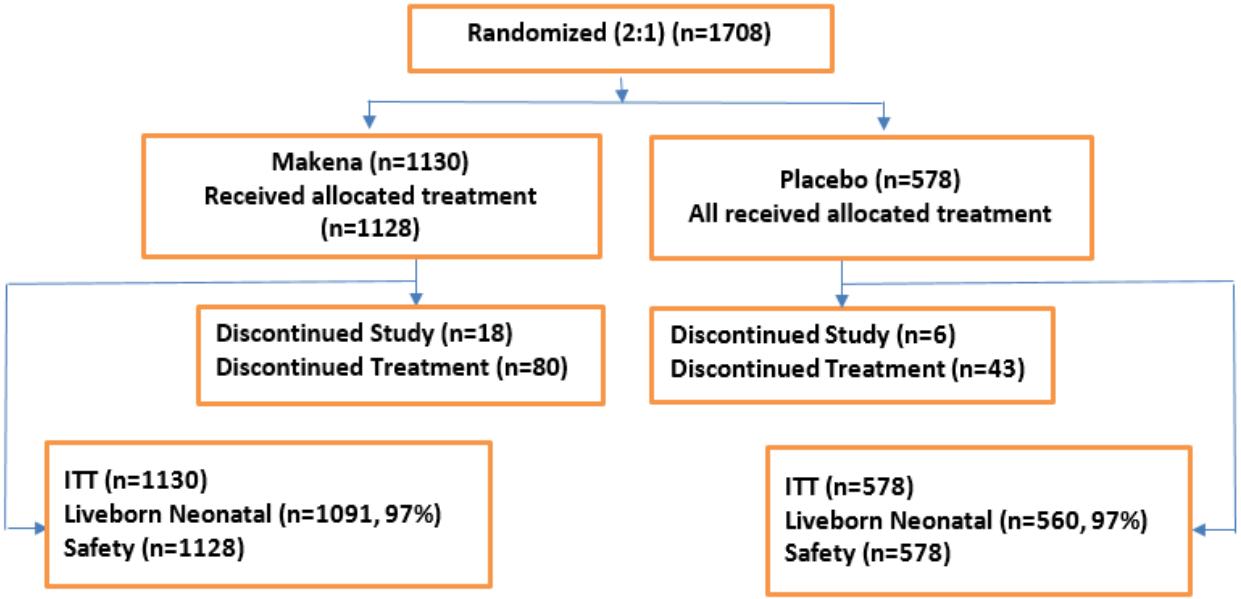
Outline

- Overview of Trial 003
  - Trial Design
  - Subject Disposition
  - Demographics and Baseline Characteristics
  - Efficacy Results
- FDA's Exploratory Analyses
- Concluding Remarks

Trial 003 Study Design

- **Study Design**
  - Multicenter, randomized, double-blind, placebo-controlled
  - Makena or placebo (2:1) stratified by study site and gestational age at randomization (16^0-17^6 weeks, 18^0-20^6 weeks)
- **Power**
  - 90% to detect a 35% reduction (from 17% to 11%) in the rate of the neonatal composite index
  - 98% to detect a 30% reduction (from 30% to 21%) in the rate of preterm birth <35^0 weeks of gestation
- **Key Inclusion Criteria**
  - Aged ≥18 years
  - With a previous singleton spontaneous preterm delivery
  - Gestational age between 16^0 to 20^6 weeks
- **Key Exclusion Criteria**
  - Had significant medical disorder
  - Multifetal gestation
  - Known major fetal anomaly or fetal demise

Trial 003 Subject Disposition



- Intent-to-treat (ITT) population: all randomized subjects
- Liveborn neonatal population: all neonates of randomized subjects who were liveborn and had morbidity/mortality data available

Trial 003 Demographics and Baseline Characteristics

Makena and placebo groups were comparable across all demographics and baseline characteristics.

Variable	Makena (N=1130) n (%)	Placebo (N=578) n (%)	All (N=1708) n (%)
Race			
White	**1004 (89)**	**504 (87)**	**1508 (88)**
Black	**73 (6)**	**41 (7)**	**124 (7)**
Other	**53 (5)**	**33 (6)**	**86 (5)**
Single or without a partner	**117(10)**	**56 (10)**	**173 (10)**
≤12 years	**488 (43)**	**259 (45)**	**747 (44)**
Any substance use during pregnancy	**106 (9)**	**52 (9)**	**158 (9)**
>1 previous SPTB	**166 (15)**	**82 (14)**	**248 (15)**
Region, United States	**258 (23)**	**133 (23)**	**391 (23)**

SPTB = spontaneous preterm birth

Trial 003 Efficacy Endpoints

- **Coprimary Endpoints**
  - Preterm birth (PTB) prior to 35^0 weeks of gestation (Yes/No)
  - Neonatal composite morbidity and mortality index: Yes, if the liveborn neonate had any of
    - RDS
    - BPD
    - Grade 3 or 4 IVH
    - NEC
    - Proven Sepsis
    - Death
- **Secondary Endpoints**
  - PTB prior to 32^0 Weeks
  - PTB prior to 37^0 Weeks

Trial 003 Efficacy Results

Efficacy Endpoint	Makena (N=1130)	Placebo (N=578)	Difference* (95% CI)	P value*
Neonatal Composite Index (%)	**5.4**	**5.2**	**0.2 (-2.0, 2.5)**	**0.84**
PTB <35^0 weeks (%)	**11.0**	**11.5**	**-0.6 (-3.8, 2.6)**	**0.72**
PTB <32^0 weeks (%)	4.8	5.2	-0.4 (-2.8, 1.7)	
PTB <37^0 weeks (%)	23.1	21.9	1.3 (-3.0, 5.4)	

N: number of randomized subjects
* CMH method stratified by gestational age at randomization
FDA analysis

No statistically significant benefit of Makena (vs. placebo) was demonstrated in either coprimary and secondary efficacy endpoints.

FDA's Position

- Generally FDA does not support subgroup analyses for inference of efficacy when the primary analysis result does not demonstrate efficacy (FDA 1998, FDA 2017b)
  - Inflation of type I error
  - FDA considers such analyses for hypothesis-generating

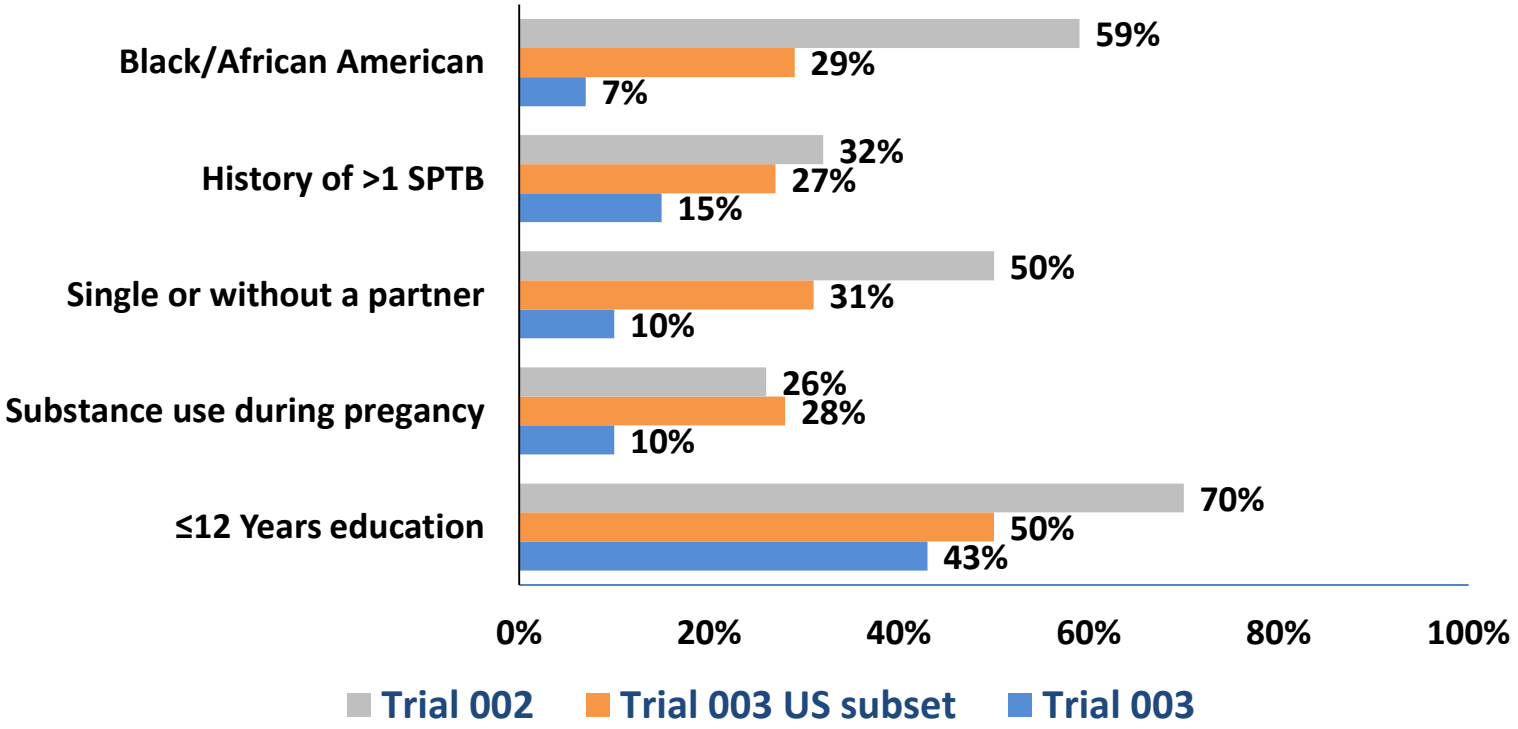
Guidance for Industry *E9 Statistical Principles for Clinical Trials* (September 1998) https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9-statistical-principles-clinical-trials

Draft Guidance for Industry *Multiple Endpoints in Clinical Trials* (January 2017) https://www.fda.gov/media/102657/download

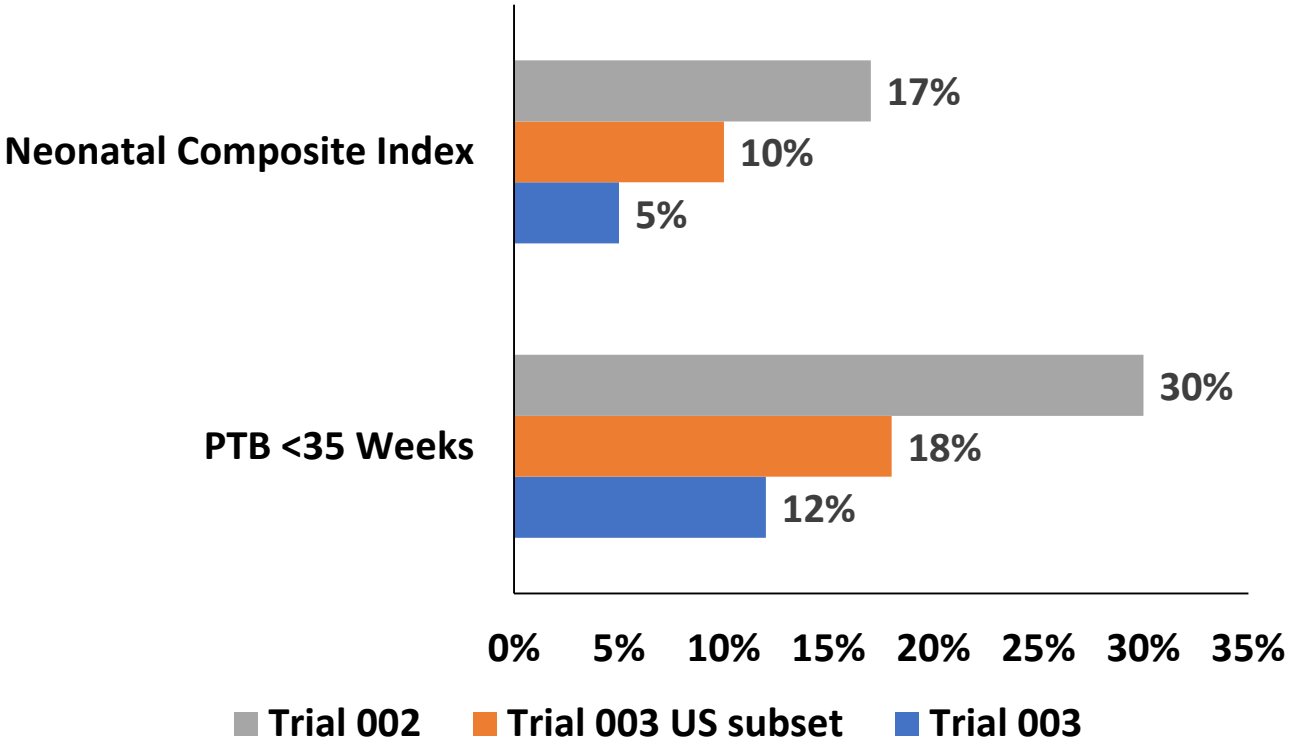
FDA Exploratory Analyses

- FDA reviewed the Applicant's post hoc subgroup analyses results to explore if differences in key aspects of Trials 003 and 002 might clarify the divergent results
  - Comparison between Trial 002 and Trial 003
  - Subgroup analyses

Comparison Between Trials 003 and 002 – Study Population

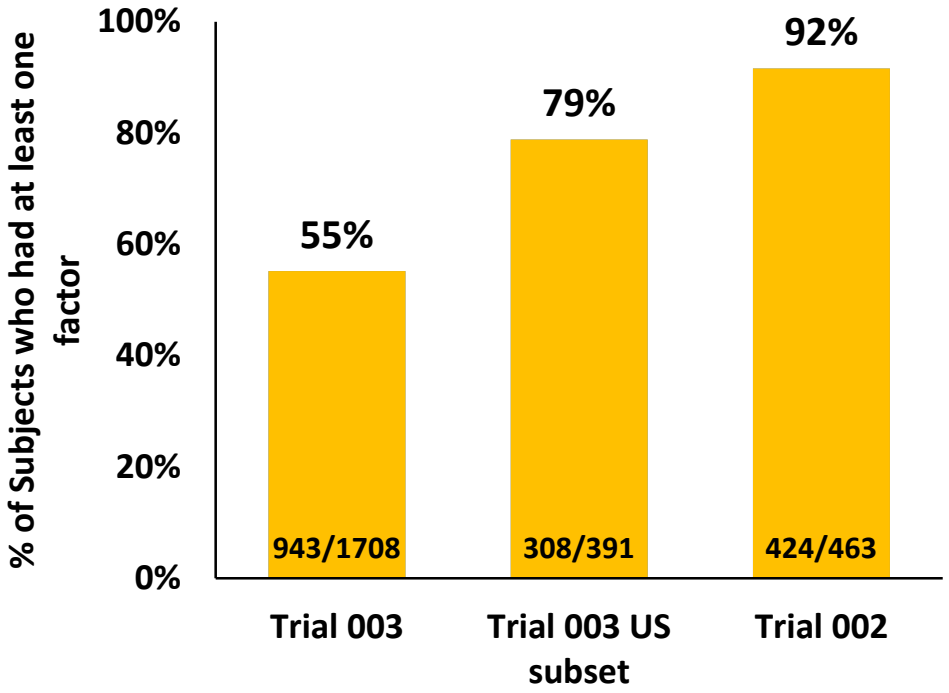


Comparison Between Trials 003 and 002 – Placebo Group



Comparison Between Trials 003 and 002 – "Composite" Risk at Baseline

- **"Composite" Risk Profile:**
  - Black
  - History of >1 prior SPTB
  - Single or without a partner
  - Substance use during pregnancy
  - ≤12 years of education



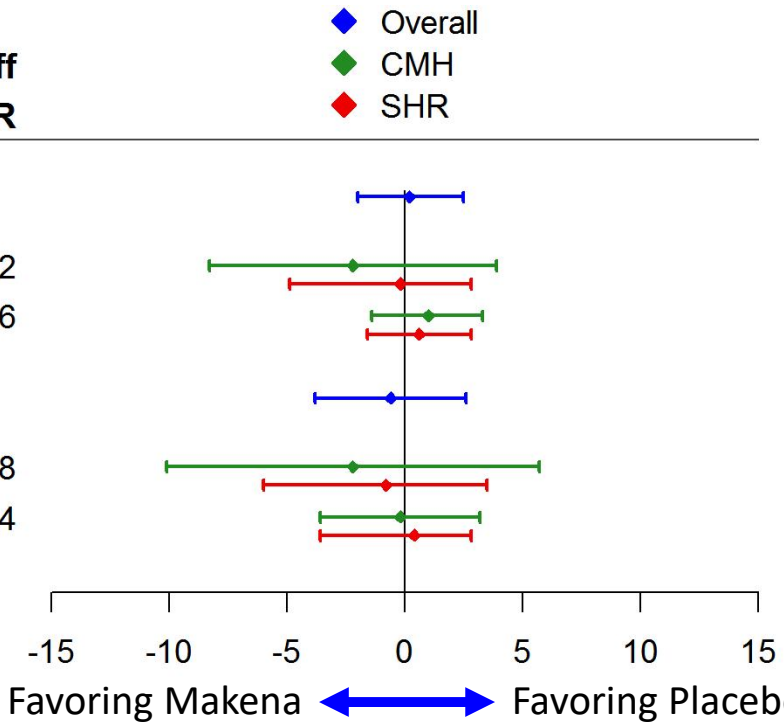
FDA Subgroup Analyses

- By single factor (stratified Cochran–Mantel–Haenszel (CMH) and shrinkage estimation)
  - Region (U.S., non-U.S.)
  - Race (Black, non-black)
  - History of SPTB (1 previous SPTB, >1 previous SPTB)
- By “composite” risk at baseline (no factor, ≥1 factor, ≥2 factors)

FDA Subgroup Analysis – by Region (003)

- No evidence of treatment effect on coprimary endpoints in either regional subgroup

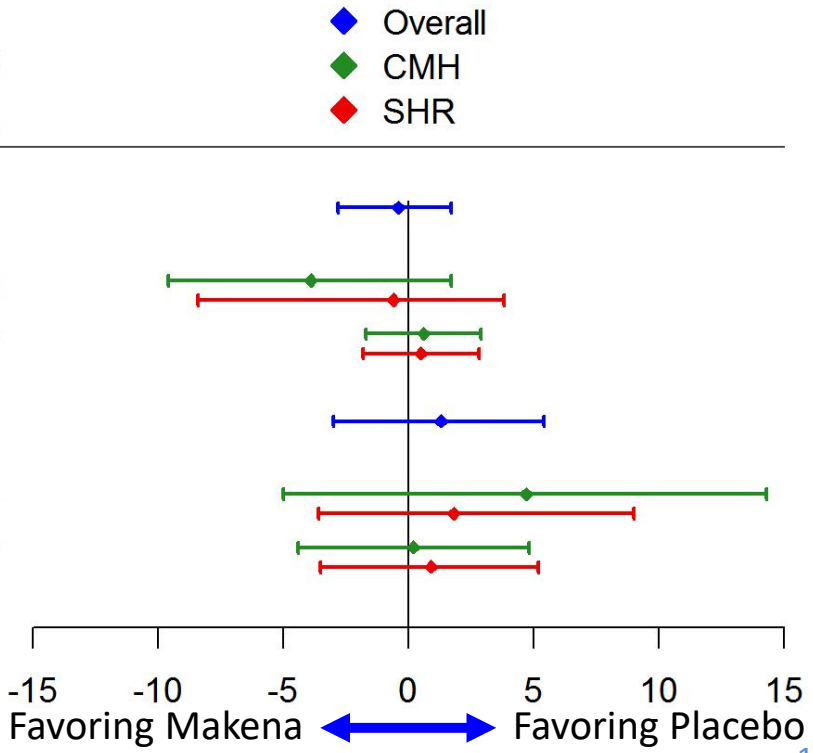
Endpoint Subgroup	Makena	Placebo	Diff CMH	Diff SHR
Neonatal Index (%)	5.4	5.2	0.2	
US (252, 126)	7.1	9.5	-2.2	-0.2
Non_US (839, 434)	4.9	3.9	1.0	0.6
PTB<35 Weeks (%)	11.0	11.5	-0.6	
US (256, 131)	15.6	17.6	-2.2	-0.8
Non_US (857, 443)	9.6	9.7	-0.2	0.4



FDA Subgroup Analysis – by Region (003)

- No evidence of treatment effect on secondary efficacy endpoints in either regional subgroup

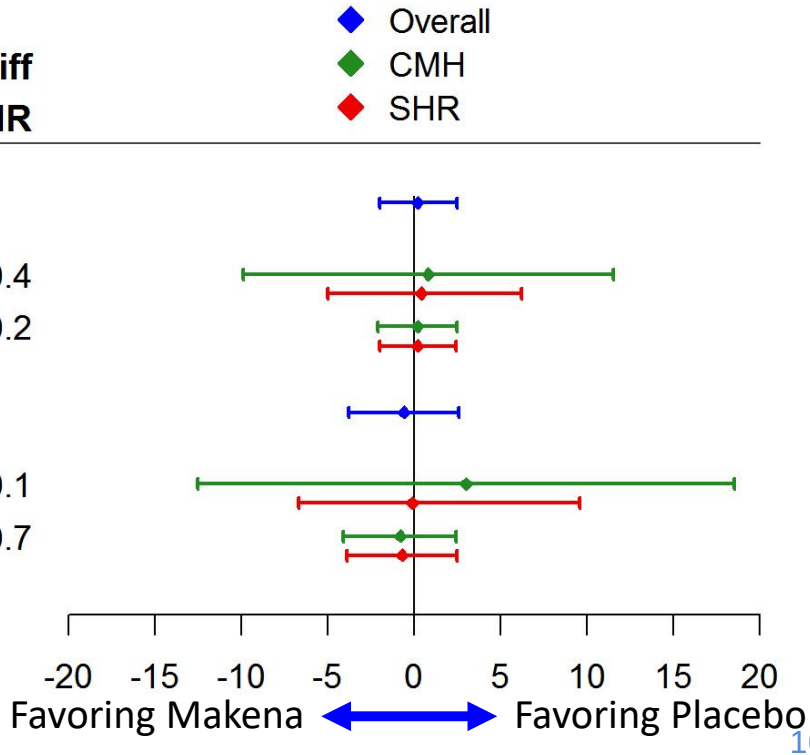
Endpoint Subgroup	Makena	Placebo	Diff CMH	Diff SHR
PTB<32 Weeks (%)	4.8	5.2	-0.4	
US (256, 131)	5.5	9.2	-3.9	-0.6
Non_US (860, 443)	4.7	4.1	0.6	0.5
PTB<37 Weeks (%)	23.1	21.9	1.3	
US (256, 131)	33.2	28.2	4.7	1.8
Non_US (856, 441)	20.1	20.0	0.2	0.9



FDA Subgroup Analysis – by Race (003)

- No evidence of treatment effect on coprimary endpoints in Black or non-Black subgroups

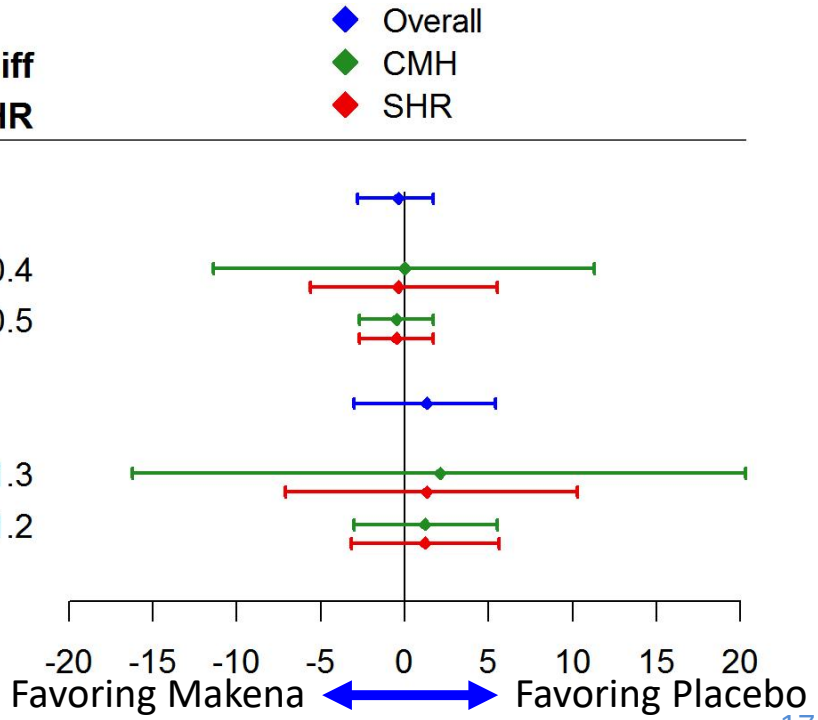
Endpoint Subgroup	Makena	Placebo	Diff CMH	Diff SHR
Neonatal Index (%)	5.4	5.2	0.2	
Black (69, 40)	8.7	7.5	0.8	0.4
Non-Black (1022, 520)	5.2	5.0	0.2	0.2
PTB<35 Weeks (%)	11.0	11.5	-0.6	
Black (72, 41)	23.6	19.5	3.0	-0.1
Non-Black (1041, 533)	10.1	10.9	-0.8	-0.7



FDA Subgroup Analysis – by Race (003)

- No evidence of treatment effect on secondary endpoints in Black or non-Black subgroups

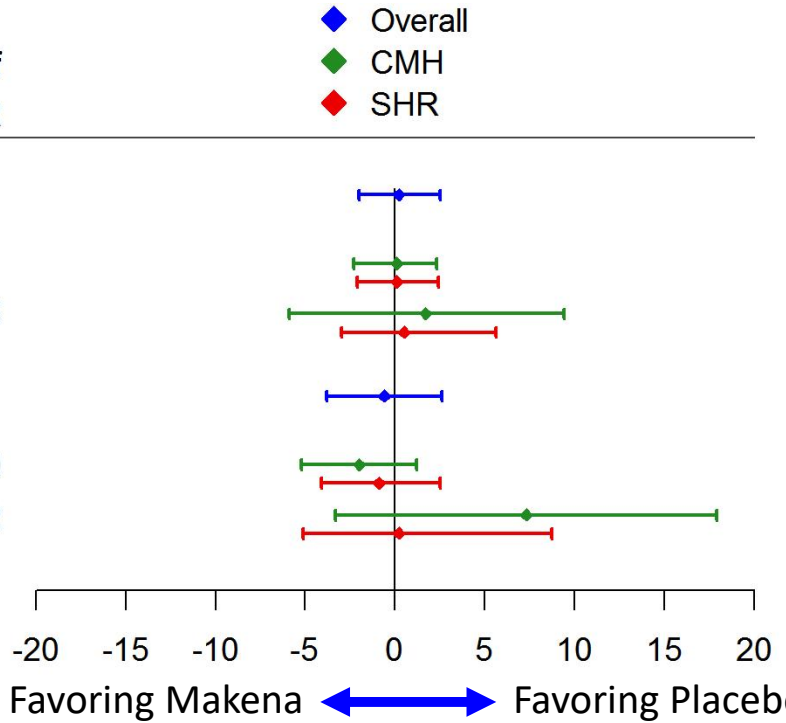
Endpoint Subgroup	Makena	Placebo	Diff CMH	Diff SHR
PTB<32 Weeks (%)	4.8	5.2	-0.4	
Black (72, 41)	11.1	9.8	0	-0.4
Non-Black (1044, 533)	4.4	4.9	-0.5	-0.5
PTB<37 Weeks (%)	23.1	21.9	1.3	
Black (72, 41)	37.4	34.2	2.1	1.3
Non-Black (1041, 533)	22.1	20.9	1.2	1.2



FDA Subgroup Analysis – by History of SPTB (003)

- No evidence of treatment effect on coprimary endpoints in either subgroup defined by history of SPTB

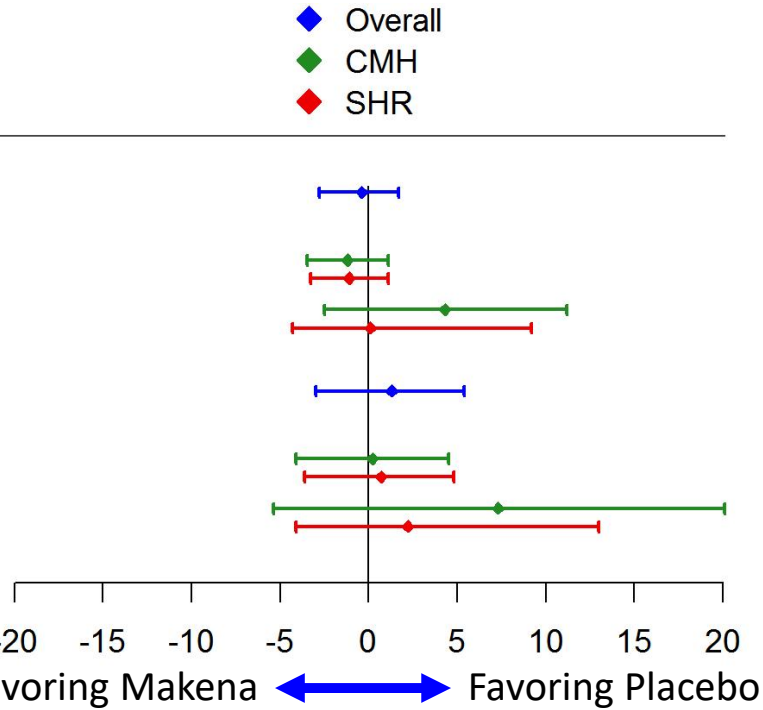
Endpoint Subgroup	Makena	Placebo	Diff CMH	Diff SHR
Neonatal Index (%)	5.4	5.2	0.2	
1 (933, 478)	4.6	4.6	0	0.1
>1 (158, 80)	10.1	8.8	1.7	0.5
PTB<35 Weeks (%)	11.0	11.5	-0.6	
1 (949, 491)	8.4	10.4	-2.0	-0.9
>1 (164, 81)	25.6	18.5	7.3	0.2



FDA Subgroup Analysis – by History of SPTB (003)

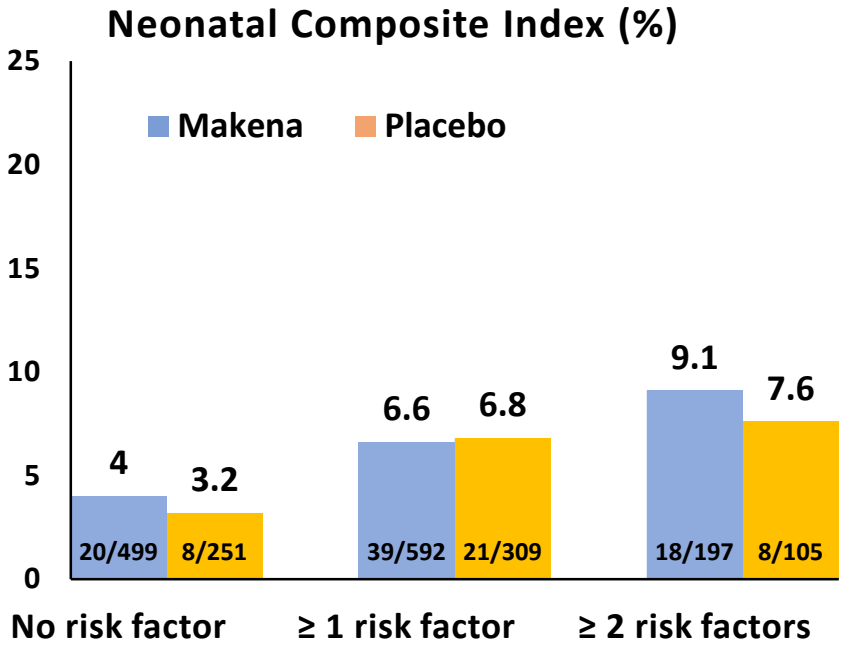
- No evidence of treatment effect on the secondary efficacy endpoints in either subgroup with history of SPTB

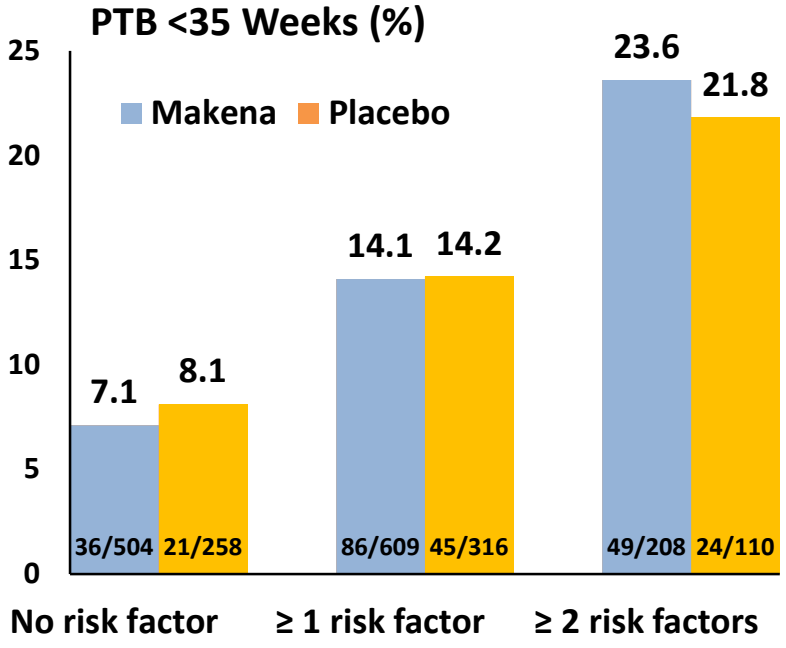
Endpoint Subgroup	Makena	Placebo	Diff CMH	Diff SHR
PTB<32 Weeks (%)	4.8	5.2	-0.4	
1 (951, 491)	3.9	5.1	-1.2	-1.1
>1 (165, 81)	10.3	6.2	4.3	0.1
PTB<37 Weeks (%)	23.1	21.9	1.3	
1 (948, 489)	19.8	19.6	0.2	0.7
>1 (164, 81)	42.1	35.8	7.3	2.2



FDA Analysis by "Composite" Risk Level (003)

- No evidence of treatment effect in any risk groups defined using the 5 selected factors.





Concluding Remarks

- **Primary Analysis**
  - Makena did not demonstrate statistically significant treatment benefit vs. placebo on either gestational age at delivery or the neonatal composite index in Trial 003
- **Exploratory Analyses**
  - No evidence that Makena had a treatment effect on the efficacy endpoints vs. placebo in the subgroups
  - Although baseline risk factors can impact the overall probability of a PTB or the neonatal composite index, there is no evidence that they are effect modifiers to Makena's treatment effect


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Makena (hydroxyprogesterone caproate injection)
New Drug Application 021945

Hydroxyprogesterone caproate (HPC) Utilization in the United States

Bone, Reproductive and Urologic Drugs Advisory Committee Meeting

October 29, 2019

Huei-Ting Tsai, Ph.D.
Epidemiologist
Division of Epidemiology II
Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research
Food and Drug Administration

Outline

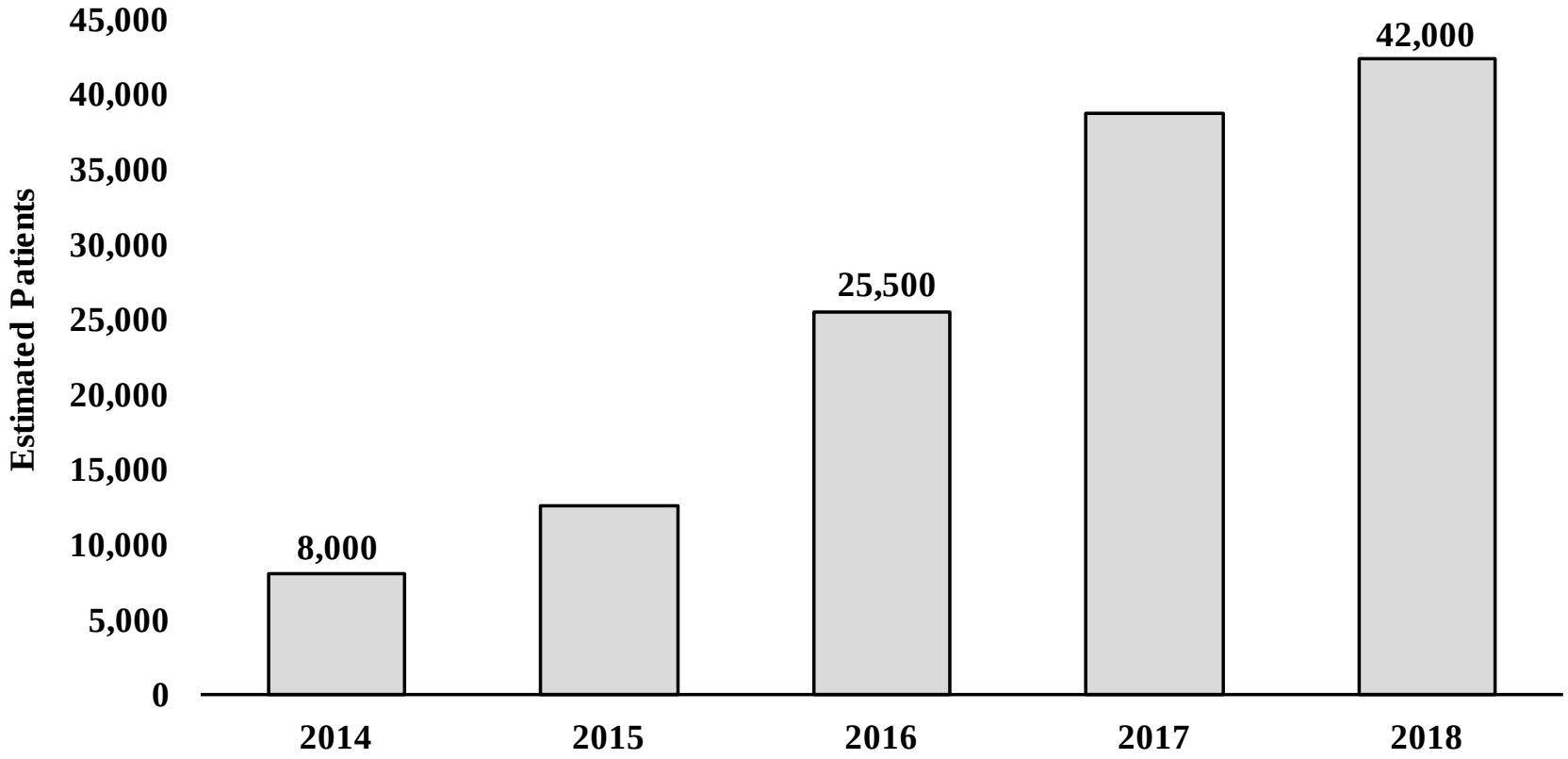
We evaluated 1) HPC utilization and 2) possible reasons for HPC use in each of two separate analyses below:

1. In U.S outpatient settings
   - Patients, pregnant and non-pregnant
   - National estimates

2. During 2nd or 3rd trimesters in live-birth pregnancies
   - In Sentinel Distributed Database
   - Not national estimates

HPC Utilization in U.S. Outpatient Settings

Increased Number of Patients With HPC Prescriptions (2014-2018)



Estimated annual number of 15- to 44-year-old patients with dispensed prescriptions for injectable hydroxyprogesterone, from U.S. retail and mail order/specialty pharmacies, 2014 through 2018

Source: Symphony Health IDV® Integrated Dataverse. Data years 2014-2018. Extracted August 2019. Unique patient counts should not be added across time periods due to the possibility of double counting those patients who received multiple products within the same calendar year or over multiple periods in the study. Prescriptions for bulk powder forms of hydroxyprogesterone were not included.

Physician Survey for Diagnoses Associated With Injectable HPC Use Among 15- to 44-Year-Old Women

- Injectable HPC
  - Supervision of high risk pregnancy (50%)
    - Of which 78% for supervision of pregnancy with history of preterm labor
  - History of preterm labor (20%)
  - Supervision of normal pregnancy (13%)
  - Preterm labor in current pregnancy (10%)

- Progesterone Products
  - Supervision of high risk pregnancy (14%); female infertility (40%)

Source: Syneos Health Research and Insights, TreatmentAnswers™ with Pain Panel. Data years 2013-2018. Extracted July 2019. Diagnosis data are not directly linked to dispensed prescriptions but obtained from surveys of a sample of 3,200 office-based physicians reporting on patient activity one day a month.

Limitations and Summary

- Limitations
  - Patient estimates obtained for retail and mail-order pharmacy settings, not hospital or clinics
  - Diagnoses related to HPC use were obtained from physician survey data
    - Do not directly link to dispensed prescriptions
    - Do not necessarily result in dispensed prescriptions
- Summary
  - Outpatient injectable HPC use increased from 2014 to 2018; use was low
  - HPC use was largely associated with history of preterm labor diagnosis

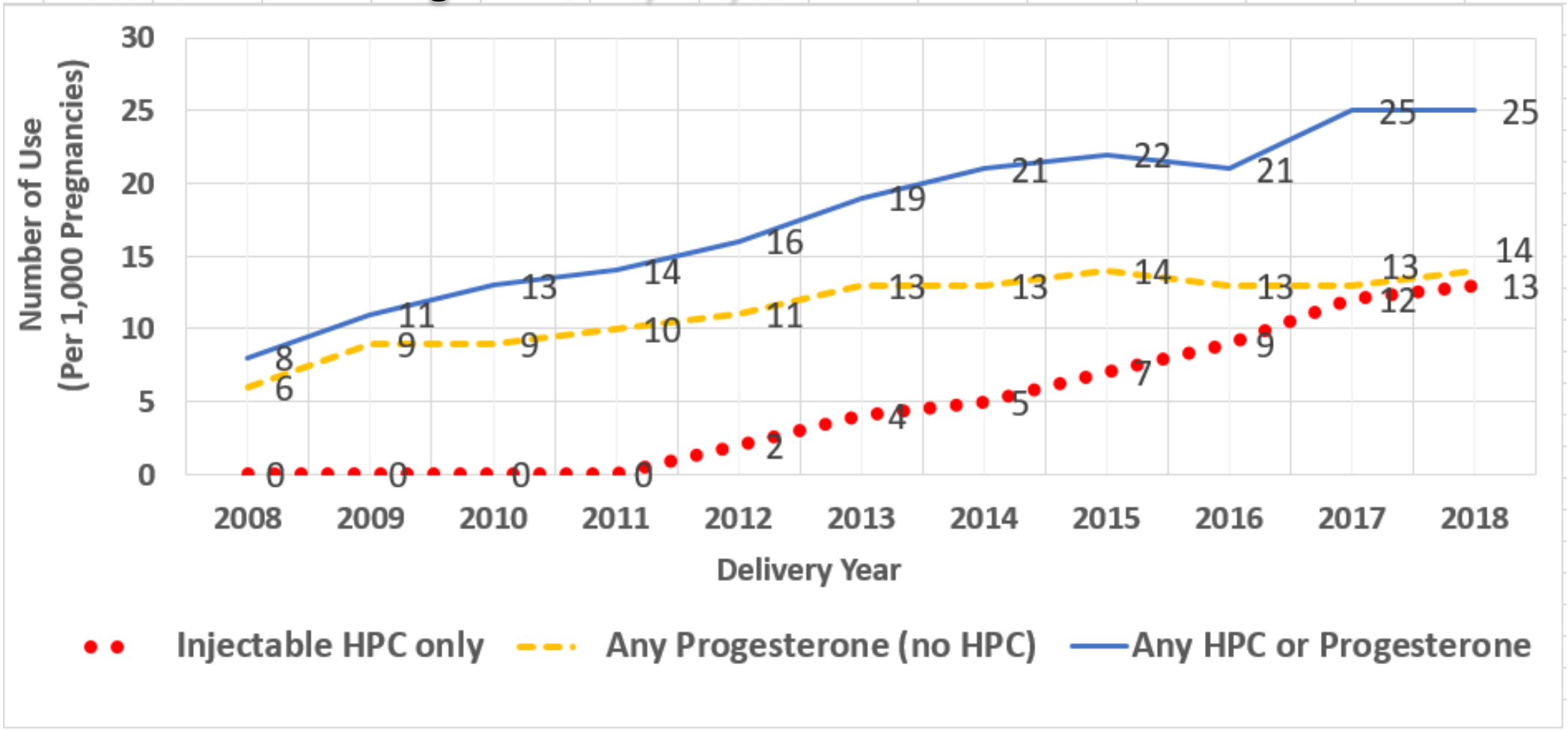
Utilization During 2nd or 3rd Trimesters in Pregnancy in Sentinel Distributed Database

Methods: Utilization in 2nd or 3rd Trimesters of Pregnancy

- Database: Sentinel Distributed Database
- Population: Live-birth pregnancies delivered Jan 2008-Apr 2019
- Medications of interest: HPC or progesterone
- Related obstetrical conditions (possible reasons for use):
  - Narrow definition:
    - Preterm delivery in a prior pregnancy
    - Preterm labor in a current pregnancy
    - Cervical shortening in a current pregnancy
  - Broad definition:
    - Same three obstetrical conditions above recorded in a prior or current pregnancy

Temporal Trend on Number of Pregnancies With HPC Use Per 1,000 Pregnancies

- Total Live-Birth Pregnancies: 3,451,121



[1] Data from 2019 was incomplete and excluded from the figure

Injectable HPC Users: Most Had a Related Obstetrical Diagnosis Code

Related Obstetrical Conditions	Injectable HPC (N=16,535)	Progesterone (N= 40,144)	Any HPC or Progesterone (N= 61,615)
Narrow Definition			
1. Preterm delivery in a prior pregancy	39%	11%	20%
2. Preterm labor in a current pregnancy	49%	45%	47%
3. Cervical shortening in a current pregnancy	20%	32%	27%
Any of the conditions above	73%	61%	65%
Broad Definition			
1. Preterm labor or delivery in a prior pregnancy	95%	37%	56%
2. Preterm labor or delivery in a current pregnancy	54%	55%	56%
3. Cervical shortening in a past or current pregnancy	24%	33%	29%
Any of the conditions above	98%	75%	83%

Limitations and Summary of Sentinel Analysis

- Limitations
  - May not be generalizable to women without a commercial health plan
  - Unspecified timing between related obstetrical conditions and injectable HPC use
  - Inability to capture out of pocket payment
- Summary
  - Overall modest use of injectable HPC during 2nd or 3rd trimesters among pregnancies with a live birth
  - A high percentage (at least 73%) of pregnancies using injectable HPC had a related obstetrical condition recorded before or during the current pregnancy.


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Makena (hydroxyprogesterone caproate injection)
New Drug Application 021945/Supplement 023

Summary Remarks

Bone, Reproductive and Urologic Drugs Advisory Committee Meeting

October 29, 2019

Christina Chang, M.D., M.P.H.
Clinical Team Leader
Division of Bone, Reproductive and Urologic Products
Office of New Drugs, Center for Drug Evaluation and Research
Food and Drug Administration

Background

- Neonatal morbidity and mortality from preterm birth (PTB) is a significant public health concern
- No drugs are approved to reduce the risk of neonatal mortality and morbidity due to prematurity
- Progestogens have been used to reduce the risk of preterm birth*

*American College of Obstetricians and Gynecologists (ACOG) Practice Bulletin (2012, reaffirmed 2018) and Society for Maternal-Fetal Medicine Statement (March 2017)

NDA 021945 Makena

- Received accelerated approval 2011 based on a single clinical trial

- Indication
  - To reduce the risk of preterm birth in pregnant women with a singleton pregnancy who have a history of spontaneous preterm birth

- Dosage & Administration
  - Administered at a dose of 250 mg once a week beginning between 16^0 weeks and 20^6 weeks gestation to week 37 of gestation or birth

Pre-Approval Data (Trial 002)

- Completed in 2002
- Double blind, randomized, placebo-controlled
- 463 U.S. women randomized to receive either HPC (n=310) or placebo (n= 153)
- Efficacy evaluated using a surrogate endpoint
  - Delivery at <37 weeks gestation
  - "Reasonably likely to predict a clinical benefit" in reducing adverse clinical outcomes, such as infant mortality/morbidity
- Makena reduced proportion of women who delivered prior to 37 weeks by 18% (37% Makena vs. 55% placebo)
- Possible safety signal of fetal loss

Design: Confirmatory Trial (Trial 003)

- Completed in 2018
- Double-blind, randomized, placebo-controlled, international trial
- Virtually identical design as Trial 002 except:
  - Gestational age surrogate endpoint
  - Adding clinical outcome
- Efficacy evaluated with two coprimary endpoints:
  - Delivery prior to 35 weeks gestation
  - Neonatal morbidity/mortality composite index*

*The neonatal morbidity/mortality composite index includes neonatal death, Grade 3 or 4 intraventricular hemorrhage, respiratory distress syndrome, bronchopulmonary dysplasia, necrotizing enterocolitis, and proven sepsis.

Results: Confirmatory Trial (Trial 003)

- Total number of subjects randomized = 1708
  - Makena (n=1130) vs. placebo (n=578)
  - Total U.S. subjects randomized (n=391, 23%)

- No statistically significant treatment effect for either coprimary endpoints:
  - Proportion of women delivering <35 weeks (11% Makena vs. 12% placebo-vehicle, p=0.72)
  - Neonatal composite index (5.4% Makena vs. 5.2% placebo-vehicle, p = 0.84)

- Proportions of women delivering <32 weeks and <37 weeks were also not different between the Makena and placebo groups.

Results: Confirmatory Trial (Trial 003)

- No relevant differences in the treatment effect when analyzed by region (U.S. vs. non-U.S.) or subgroups (e.g., race, previous # of spontaneous PTB)

- In the U.S. subgroup:
  - Makena did not improve the neonatal outcome
  - Makena did not reduce the risk of delivery <35 weeks (16% Makena vs. 18% placebo)

- Safety findings:
  - Number of fetal/neonatal deaths were low but were similar between groups
  - The study met the prespecified endpoint of excluding a doubling of the risk of fetal/early infant deaths for Makena

Effectiveness Standard for Drug Approval

- All approved drugs, including those approved under accelerated approval, must meet the statutory standard of "substantial evidence" of effectiveness

  Evidence consisting of adequate and well-controlled investigations, including clinical investigations… to evaluate the effectiveness of the drug involved...*

*21 U.S.C. § 355(d), FD&C Act Section 505(d)

Trial 002 vs. Trial 003

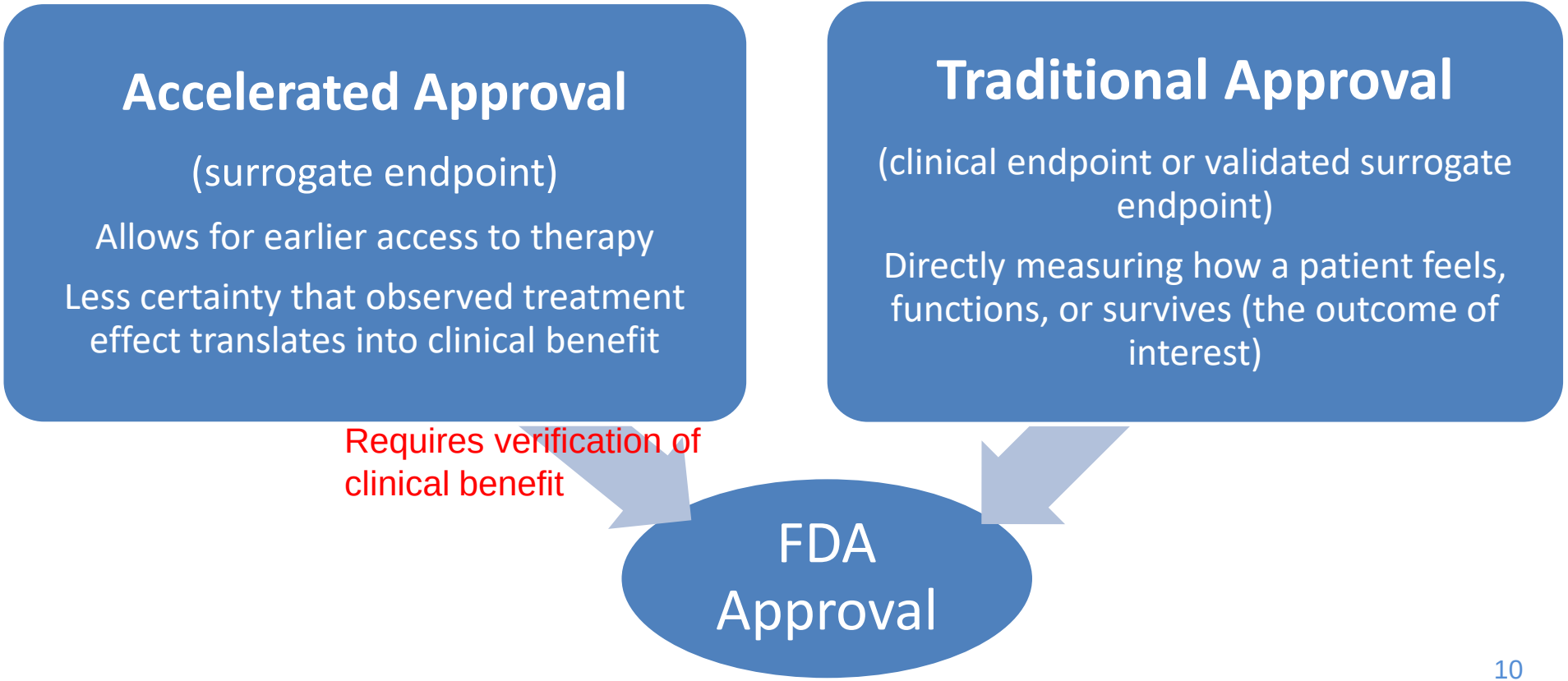
Trial 002

- Assessed efficacy based on gestational age at delivery (surrogate)
- U.S. academic centers only
- ~60% blacks
- Unusually high PTB rate (55%) in placebo group
- Makena reduced proportion of PTB <37 weeks by 18%

Trial 003

- Assessed efficacy based on neonatal outcomes (clinical benefit) and gestational age at delivery (surrogate)
- International trial (but 23% from United States)
- Makena had no treatment effect for proportion of delivery <35 weeks, <32, or <37 weeks
- No difference in neonatal outcomes

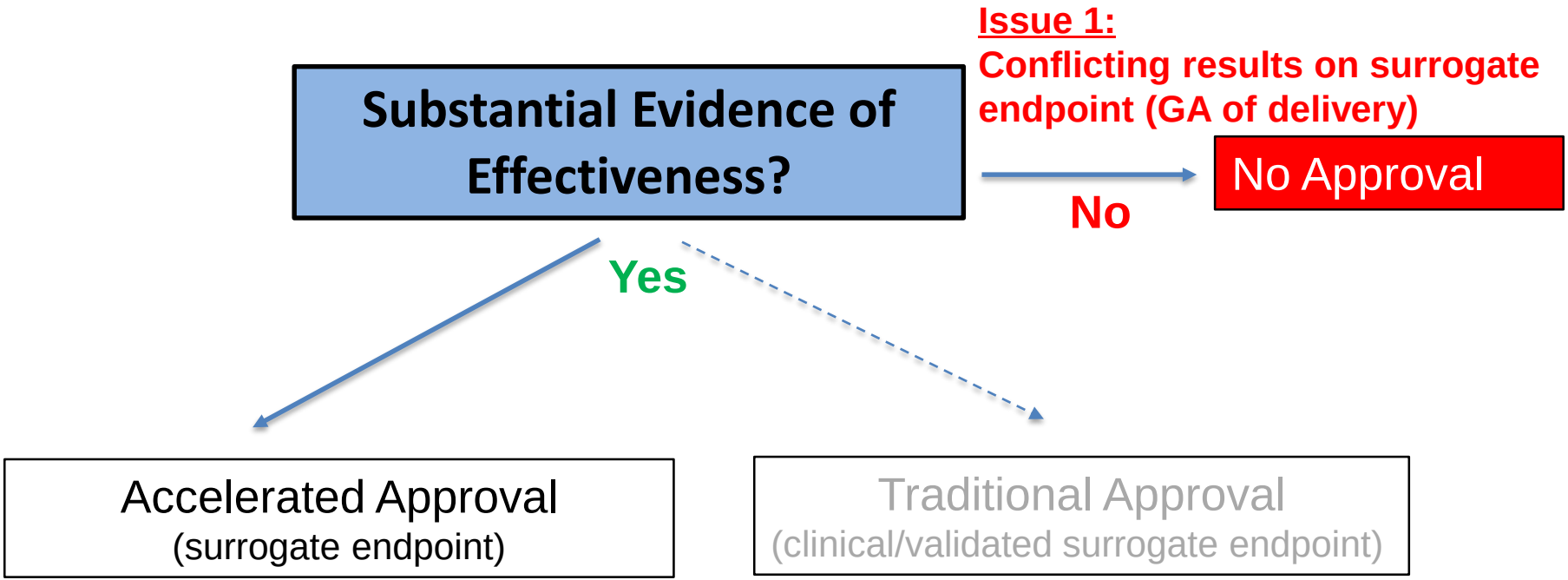
Substantial Evidence of Effectiveness



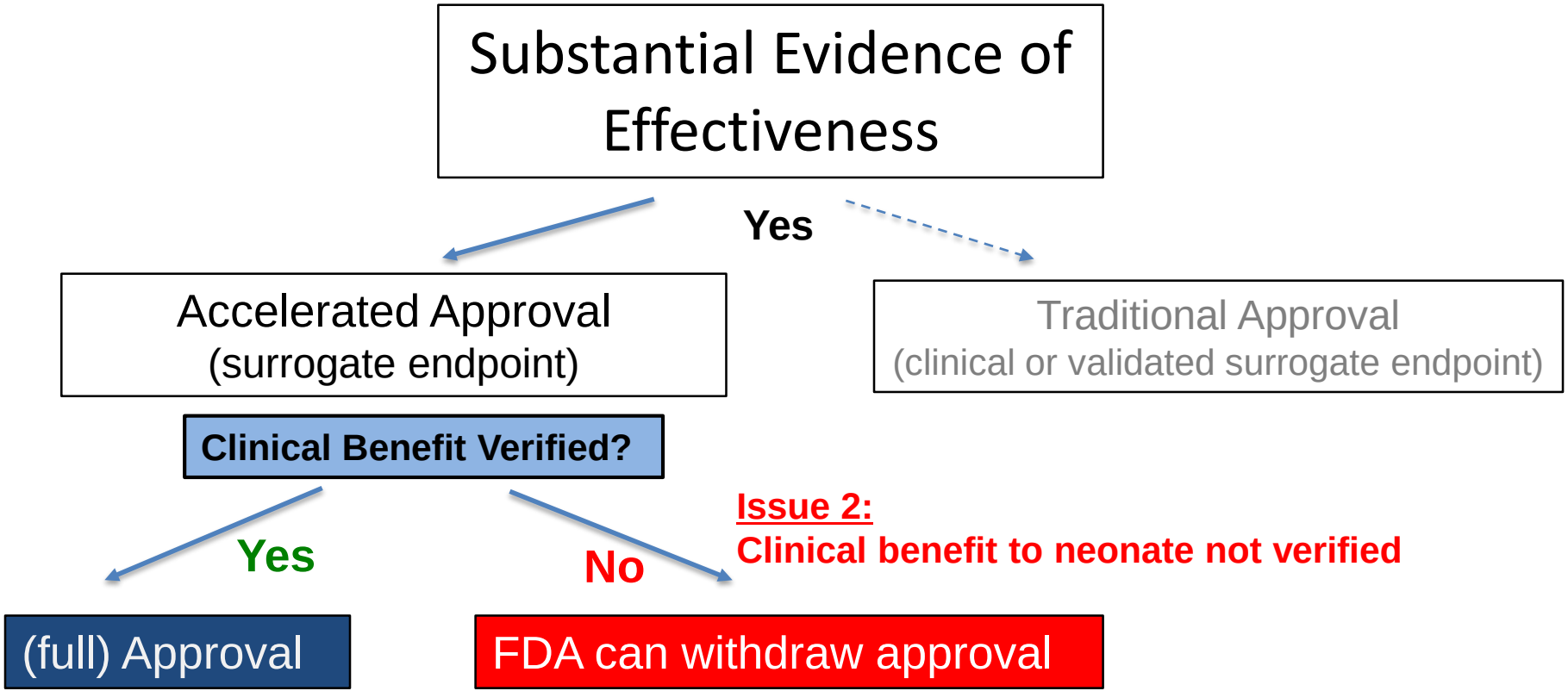
Why the Discrepant Results?

- Trial 002 (with the surrogate endpoint only) falsely positive?
- Trial 003 falsely negative?
- Discrepant results between Trials 002 and 003 due to unknown factors?

Issue 1: Substantial Evidence of Effectiveness



Issue 2: Accelerated Approval




FDA
U.S. FOOD & DRUG
ADMINISTRATION

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

BONE, REPRODUCTIVE. AND UROLOGIC DRUGS ADVISORY COMMITTEE (BRUDAC)

Tuesday, October 29, 2019

8:15 a.m. to 4:26 p.m.

FDA White Oak Campus

White Oak Conference Center

Building 31, The Great Room

10903 New Hampshire Avenue

Silver Spring, Maryland

Meeting Roster

DESIGNATED FEDERAL OFFICER (Non-Voting)

Kalyani Bhatt, BS, MS

Division of Advisory Committee and Consultant Management

Office of Executive Programs, CDER, FDA

BONE, REPRODUCTIVE AND UROLOGIC DRUGS ADVISORY COMMITTEE MEMBERS (Voting)

Douglas C. Bauer, MD

Professor of Medicine and Epidemiology & Biostatistics

University of California, San Francisco

San Francisco, California

Matthew T. Drake, MD, PhD

Associate Professor of Medicine

Chair, Metabolic Bone Disease Core Group

Division of Endocrinology

Mayo Clinic College of Medicine

Rochester, Minnesota

Vivian Lewis, MD

(Chairperson)

Vice Provost for Faculty Development & Diversity

Professor, Obstetrics and Gynecology

University of Rochester

Rochester, New York

Pamela Shaw, PhD

Associate Professor

Department of Biostatistics and Epidemiology

University of Pennsylvania School of Medicine

Philadelphia, Pennsylvania

TEMPORARY MEMBERS (Voting)

Jonathan M. Davis, MD

Vice-Chair of Pediatrics

Chief of Newborn Medicine

The Floating Hospital for Children at Tufts

Medical Center

Professor of Pediatrics

Tufts University School of Medicine

Boston, Massachusetts

Ahizechukwu Eke, MD, MPH

Assistant Professor

Division of Maternal Fetal Medicine

Department of Gynecology & Obstetrics

Johns Hopkins University School of Medicine

Baltimore, Maryland

Annie Ellis

(Patient Representative)

White Plains, New York

Daniel Gillen, PhD

Professor and Chair, Statistics

University of California, Irvine

Irvine, California

Kimberly Hickey, MD

Colonel, Medical Corps, US Army

Chief, Maternal Fetal Medicine

Walter Reed National Military Medical Center

Deputy Director, National Capital Consortium

Uniformed Services University of the Health Sciences

Bethesda, Maryland

Sally Hunsberger, PhD

Mathematical Statistician

Division of Clinical Research

National Institute of Allergy and Infectious Disease

National Institute of Health

Rockville, Maryland

Michael K. Lindsay, MD, MPH

Luella Klein Associate Professor

Chief, Gynecology and Obstetrics Service Grady Health Systems

Director, Division of Maternal Fetal Medicine

Emory University

Atlanta, Georgia

Michele Orza, ScD

(Acting Consumer Representative)

Chief of Staff

Patient-Centered Outcomes Research Institute (PCORI)

Washington, District of Columbia

Uma M. Reddy, MD, MPH

Professor, Department of Obstetrics,
Gynecology and Reproductive Sciences
Division Chief, Maternal Fetal Medicine
Section Chief, Maternal Fetal Medicine of
Yale New Haven Hospital
Program Director, Maternal-Fetal Medicine
Fellowship
Department of Obstetrics, Gynecology and
Reproductive Sciences
Yale School of Medicine
New Haven, Connecticut

Brian Smith MD, MPH, MHS

Samuel L. Katz Professor of Pediatrics
Division of Neonatal-Perinatal Medicine
Duke University Medical Center
Durham, North Carolina

Kelly Wade, MD, PhD, MSCE
Attending Neonatologist
Children's Hospital of Philadelphia (CHOP)
Associate Professor of Clinical Pediatrics
University of Pennsylvania
CHOP Newborn Care
Philadelphia, Pennsylvania

Deborah A. Wing, MD, MBA
Senior Client Partner
Los Angeles, California
Formerly, Professor of Obstetrics-Gynecology
Division of Maternal Fetal Medicine
University of California, Irvine
Orange, California

Venkateswar Jarugula, PhD
(Acting Industry Representative)
Executive Director
Translation Medicine
Novartis Institutes for Biomedical Research
East Hanover, New Jersey

FDA PARTICIPANTS (Non-Voting)

Christine Nguyen, MD

Deputy Director for Safety

Division of Bone, Reproductive and Urologic Products (DBRUP)

Office of Drug Evaluation III (ODE III)

Office of New Drugs (OND), CDER, FDA

Barbara Wesley, MD, MPH

Medical Officer

DBRUP, ODEIII, OND, CDER, FDA

Christina Chang, MD, MPH

Clinical Team Leader

DBRUP, ODEIII, OND, CDER, FDA

Jia Guo, PhD

Statistical Reviewer

Division of Biometrics 3 (DB3)

Office of Biostatistics (OB)

Office of Translational Sciences (OTS), CDER, FDA

C O N T E N T S

C O N T E N T S (continued)

P R O C E E D I N G S

(8:15 a.m.)

Call to Order

Introduction of Committee

DR. LEWIS: Good morning. I would first like to remind everyone to please silence your cell phones and any other devices if you haven't already done so. I would also like to identify the FDA press contact, Amanda Turney. She's standing there in the back. We're going to get started with the meeting.

My name is Vivian Lewis, and I'm the chair of the Bone, Reproductive, and Urologic Drugs Advisory Committee, and I'll be chairing this meeting. I will now call upon today's Bone, Reproductive, and Urologic Drugs Advisory Committee members to introduce themselves. The meeting's now call to order. We'll start with the FDA on my left, and we'll go around the table for everyone to say their name.

DR. NGUYEN: Thank you, Dr. Lewis. Good morning. I'm Christine Nguyen, and I am the deputy director for safety in the Division of Bone, Reproductive, and Urologic Products; otherwise known as

DBRUP.

DR. CHANG: Good morning, everyone. My name is Christina Chang. I am a clinical team leader in the division.

DR. WESLEY: Good morning. I'm Barbara Wesley. I'm the primary medical reviewer and have been since the beginning of this drug.

DR. GUO: Good morning. My name is Jia Guo. I'm the statistical reviewer from the Office of Biostatistics.

DR. EKE: Good morning, everyone. My name is Ahizechukwu Eke. I am a maternal fetal medicine physician at Johns Hopkins.

DR. HICKEY: Good morning. I'm Kimberly Hickey. I'm one of the maternal fetal medicine physicians at Walter Reed.

DR. LINDSAY: Good morning. I'm Michael Lindsay. I'm a maternal fetal medicine specialist at Emory University.

DR. REDDY: Hi. I'm Uma Reddy, maternal fetal medicine division director at Yale.

DR. WING: Good morning. I'm Deborah Wing. I

am the senior client partner at Korn Ferry. I'm a former professor of OB/GYN and division director of maternal fetal medicine at the University of California Irvine.

DR. DRAKE: Good morning. My name is Matthew Drake. I'm an adult endocrinologist at the Mayo Clinic in Rochester, Minnesota.

MS. BHATT: Good morning. I'm Kalyani Bhatt. I'm the designated federal officer for this advisory committee.

DR. BAUER: Good morning. My name is Doug Bauer. I'm from the departments of medicine, epidemiology, and biostatistics from UCSF in San Francisco.

DR. SHAW: Good morning. I'm Pam Shaw. I'm at the Department of Biostatistics, Epidemiology, and Informatics at University of Pennsylvania.

MS. ELLIS: Good morning. I'm Annie Ellis, and I'm a patient representative.

DR. ORZA: Good morning. I'm Michele Orza. I'm the chief of staff at the Patient-Centered Outcomes Research Institute, and I'm the acting consumer

representative today.

DR. GILLEN: Good morning. Daniel Gillen, professor and chair of statistics at UC Irvine.

DR. HUNSBERGER: Good morning. I'm Sally Hunsberger at the biostatistics research branch at NIAID, at NIH.

DR. SMITH: Good morning. I'm Brian Smith. I'm a neonatologist at Duke.

DR. WADE: Good morning. I'm Kelly Wade. I'm a neonatologist for Children's Hospital of Philadelphia and the chair of the Pediatric Advisory Committee.

DR. DAVIS: Good morning. I'm Jon Davis, chief of neonatology at Tufts Medical Center in Boston and chair of the Neonatal Advisory Committee at FDA.

DR. LEWIS: Thank you. We'll have one other panel member, and that will be Dr. Jarugula. He's stuck in traffic. He'll introduce himself once he gets here.

For topics such as those being discussed at today's meeting, there are often a variety of opinions, some of which are strongly held. Our goal is that today's meeting will be a fair and open forum for

discussion of the issues and that individuals can express those views without interruption. Thus, as a gentle reminder, individuals will be allowed to speak into the record only if recognized by the chair. We look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Government in the Sunshine Act, we ask that the advisory committee members take care that their conversations about the topic at hand take place in the open forum of the meeting.

We are aware that members of the media are anxious to speak with the FDA about these proceedings, however, FDA will refrain from discussing the details of this meeting with the media until its conclusion. Also, the committee is reminded to refrain from discussing the meeting topic during breaks or during lunch. Thank you.

I'd now like to pass it to Kalyani Bhatt, who will read the Conflict of Interest Statement.

Conflict of Interest Statement

MS. BHATT: The Food and Drug Administration is convening today's meeting of the Bone, Reproductive,

and Urologic Drugs Advisory Committee under the authority of the Federal Advisory Committee Act, FACA, of 1972. With the exception of the industry representative, all members and temporary voting members of the committee are special government employees or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations.

The following information on the status of this committee's compliance with federal ethics and conflict of interest laws, covered by but not limited to those found at 18 U.S.C. Section 208, is being provided to participants in today's meeting and to the public. FDA has determined that members and temporary voting members of this committee are in compliance with federal -- [inaudible - audio gap].

(Pause.)

MS. BHATT: -- statistically significant difference between the treatment and placebo arms for the co-primary endpoints of reducing the risk of recurrent preterm birth or improving neonatal mortality and morbidity. The committee will consider the trial's

findings and the supplement NDA in the context of AMAG Pharmaceutical's confirmatory study application.

This is a particular matters meeting during which specific matters related to AMAG and the supplemental NDA will be discussed. Based on the agenda for today's meeting and all financial interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting.

To ensure transparency, we encourage all standing committee members and temporary voting members to disclose any public statements that they have made concerning the product at issue. With respect to FDA's invited industry representative, we'd like to disclose that Dr. Jarugula is participating in this meeting as a nonvoting industry representative, acting on behalf of regulated industry. Dr. Jarugula's role at this meeting is to represent industry in general and not any particular company. Dr. Jarugula is employed by Novartis Institutes for Biomedical Research.

We'd like to remind members and temporary voting members that if the discussions involve any

other products or firms not already on the agenda for which an FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement, and their exclusion will be noted for the record. FDA encourages all participants to advise the committee of any financial relationship that they may have with the firm at issue. Thank you.

DR. LEWIS: Thank you.

Before we go to the FDA opening remarks, I'd like the one last panel member who just got here to please introduce himself.

DR. JARUGULA: Good morning, everybody. Sorry. I got stuck in heavy traffic. I didn't anticipate this heavy D.C. traffic. My name is Venkat Jarugula. I'm representing the industry here. I am from Novartis Pharmaceuticals. Thank you.

DR. LEWIS: Thank you. We will now proceed with the FDA opening remarks from Dr. Nguyen.

FDA Opening Remarks - Christine Nguyen

DR. NGUYEN: Good morning, everyone. I want to thank each one of you for sacrificing a beautiful

holiday to be here with us. We are convening this advisory committee meeting to discuss the evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth and improving neonatal outcomes. In my introductory remarks, I will be covering the key issues that you will hear about and discuss throughout the day.

We appreciate that neonatal mortality and morbidity from preterm birth is a significant public health concern. Currently, there are no therapies approved to reduce the risk of these adverse neonatal outcomes from prematurity. Progestogens, which include progesterone and progestins, have been used in clinical practice over the years to reduce the risk of preterm birth. However, only Makena has been approved to reduce the risk of recurrent preterm birth.

In 2011, we approved Makena under accelerated approval to reduce the risk of preterm birth in women with a singleton pregnancy and a prior spontaneous singleton preterm birth. This approval was based on a single trial conducted between 1999 and 2002 in approximately 460 women in the U.S., and this trial

showed persuasive efficacy findings on the surrogate endpoint of gestational age of delivery of less than 37 weeks.

I will refer to this trial as Trial 002. As required under accelerated approval regulations, the applicant conducted a post-approval confirmatory trial to verify the clinical benefit for the neonates, and I'll be expanding on these key concepts that are underlined later in my presentation.

The confirmatory trial was an international, randomized, double-blind, placebo trial that enrolled approximately 1700 pregnant women. The top three enrolling countries were Russia, Ukraine, and the U.S., with the U.S. enrolling 23 percent of total subjects. I would note that the number enrolled in Trial 003 from the U.S., which was about 390, was not substantially less than the number that was enrolled in Trial 002, which is 460.

The design eligibility criteria were similar to Trial 002, except for the primary endpoints. Trial 002's primary efficacy endpoint was gestational age of delivery less than 37 weeks, and for child Trial 003,

it was gestational age of delivery less 35 weeks and the clinical endpoint of neonatal morbidity and mortality Index. This trial was conducted between 2009 and 2018.

As you can see here, there are no treatment effects between Makena and placebo for the co-primary endpoints, and there also no treatment effects for the two key secondary endpoints, which were preterm birth of less than 32 weeks and less than 37 weeks. I remind you that the endpoint of preterm birth of less than 37 weeks was the primary efficacy endpoint for Trial 002.

Because of the contradictory results for the gestational age of delivery endpoint, we conducted multiple exploratory subgroup analyses for factors that were dissimilar between the two trials. The subgroup analyses included that for region, race, and certain elements that the applicant identified that may increase the risk of preterm birth. These included the number of previous preterm birth, substance use in pregnancy, number of years of formal education, and partner status.

There were no statistically significant

treatment difference for any of these subgroup analyses. In addition, there was no statistically significant interaction between treatment effect and these factors, meaning that these factors may be prognostic for preterm birth, but they do not appear to be effect modifiers; meaning that if a woman has these factors, she may be at increased of having preterm birth, but these factors do not render her having more favorable response to Makena.

Also, there are no consistent convincing evidence of a treatment effect within any particular subpopulation across the two trials.

This is the totality of the evidence in front of us today. Trial 002 shows efficacy on gestational age of delivery, which is a surrogate endpoint. However, this trial was conducted almost 20 years ago, but it was conducted in the United States. There were issues regarding generalizability to the general U.S. population that I've listed in my slide.

Trial 003, on the other hand, did not show any efficacy on neonatal outcomes or gestational age at delivery. It was conducted more recently, and it was

adequately powered to the treatment effect that was observed in Trial 002. However, it was an international trial, but I'll remind you, approximately 1 in 4 women enrolled in 003 was from the U.S., and it evaluated a low-risk population who showed a low recurrent preterm birthrate in placebo arm than 002.

The efficacy in Makena was evaluated by two different types of endpoints. The first endpoint is a surrogate endpoint of gestational age of delivery. Both Trials 002 and 003 evaluate this endpoint. While 002 show efficacy, 003 did not. So we concluded there's conflicting efficacy findings for this endpoint, and this raises the first issue regarding the approval requirement of substantial evidence of effectiveness.

The second type of endpoint evaluated was a clinical endpoint of neonatal composite index. This endpoint was only appropriately evaluated in 003, and as you can see, Trial 003 did not show a treatment effect in this endpoint, so we conclude that there's not been verification of the clinical benefit of Makena to the neonates, so this raises the second approval issue concerning accelerated approval.

Going back to issue 1, substantial evidence of effectiveness, this is the statutory standard for establishing efficacy for FDA drug approval, including accelerated approval. Traditionally, we look for significant findings from at least two adequate and well-controlled trials, each convincing on its own to provide independent substantiation on the efficacy endpoint. This approach also reduces the risk of false positive from chance or bias, which may remain undetected from a single trial.

The concept of independent substantiation is the scientific principle that underlies the legal standard of substantial evidence of effectiveness. That said, when appropriate, a single adequate and well-controlled trial with persuasive findings may be accepted as substantial evidence, and this is what happened for Makena in 2011 when we approved it based on Trial 002.

Note that if there were additional adequate and well-controlled trials at the time of approval, we would have considered those data when deciding about substantial evidence. In 2019, we now have two

adequate and well-controlled trials, and the first issue is that Trial 003 did not substantiate Makena's treatment effect on gestational age of delivery. So is there still substantial evidence of a drug's effect on reducing the risk of recurrent preterm birth?

Here in this diagram, I wanted to lay out where this first issue lies. To gain approval, any approval, a drug must demonstrate substantial evidence of effectiveness. Whether or not it receives accelerated approval or traditional approval depends on the efficacy endpoint that was evaluated. For accelerated approval, it will be the surrogate endpoint, which is what happened for Makena. If there lacks substantial evidence of effectiveness, then there will be no approval.

At this point, we have contradictory efficacy findings on the gestational age of delivery. So that puts in question whether or not there is still substantial evidence of a drug's effectiveness for that endpoint.

The second issue relates to accelerated approval. As I've shown in this earlier slide,

traditional approval is granted when there is substantial evidence of the drug's effect on a clinical endpoint, and that is one that directly measures how patients feel, function, or survive, or a validated surrogate endpoint, which is one that is known to predict clinical benefit.

We grant accelerated approval when there's a drug's effect on the surrogate endpoint, which is one that reasonably likely predicts clinical benefit. Accelerated approval is an expedited drug development pathway, and we reserve it only for certain drugs treating serious or life-threatening conditions with unmet medical need. As I mentioned, it must meet the same statutory effectiveness standards, that is substantial evidence of effectiveness, as those for traditional approval.

I will take a second here to explain why gestational age of delivery is not a clinical endpoint, and we do not consider at this time a validated surrogate endpoint. Gestational delivery is not a clinical endpoint because it doesn't directly measure how neonates feel, function, or survive. When we're

talking about treatment for prematurity, it is the improved outcomes to a neonate that is most meaningful.

It's not considered a validated surrogate endpoint because spontaneous preterm birth is a poorly understood syndrome with potential for multiple pathophysiologic pathways. So prolonging gestation may not consistently translate into improved neonatal outcomes.

Let's take a hypothetical example of a woman going to preterm labor at 35 weeks due to some subclinical, undiagnosed, low inflammatory process. We now iatrogenically prolong that pregnancy for another week, and the baby is delivered at 36 weeks. However, the fetus has been exposed for an additional week in a relatively unhealthy in utero environment, so it's unclear whether or not that fetus, when born, will have improved neonatal outcomes.

As you can see, there's more uncertainty, at the time of accelerated approval, that the treatment effect on the surrogate endpoint will translate into clinical benefit. Therefore, the drug must undergo a post-approval confirmatory trial to verify its clinical

benefit.

FDA can withdraw approval of the drug or the indication if the applicant does not conduct such required trial, or if the trial fails to verify the clinical benefit. That's the second issue that we face, which is that Trial 003 did not verify Makena's clinical benefit to the neonates.

Back to this diagram, let's assume we don't have a problem with substantial evidence of effectiveness. Makena now still sits under accelerated approval. Its clinical benefit must still be verified. If the clinical benefit is not verified, FDA can withdraw approval.

I'll wrap up my presentation by walking you through 3 three discussion questions and 3 voting questions, or 6 questions total that you'll be seeing later on today. The first discussion question, discuss the effectiveness of Makena on recurrent preterm birth and neonatal morbidity and mortality.

Discussion question 2. If a new confirmatory trial were to be conducted, discuss the study design, including control, dose(s) of study medication,

efficacy endpoints, and importantly, the feasibility of completing such a trial.

Discussion question 3. Discuss the potential consequences of withdrawing Makena on patients and clinical practice.

Voting question 4. Do the findings from Trial 003 verify clinical benefit of Makena on neonatal outcomes? Provide your rationale.

Voting questions 5. Based on the findings from Trial 002 and 003, is there substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth based on the surrogate endpoint of gestational age of delivery? Provide your rationale.

Voting question 6 requires a preamble. FDA approval, including accelerated approval of a drug, requires that there is a demonstration of substantial evidence of effectiveness of the drug on the efficacy endpoint. This is the first approval issue that I discussed earlier.

For drugs approved under accelerated approval, the applicant is required to conduct a confirmatory

trial to verify the clinical benefit. That is the second approval issue that I discussed earlier. If the applicant fails to conduct such a trial, or if such a trial does not verify the clinical benefit, FDA may, following an opportunity for a hearing, withdraw approval.

There are three voting options for this question. Should FDA, A, pursue withdrawal of approval from Makena; B, leave Makena on the market under accelerated approval and require a new confirmatory trial; or C, leave Makena on the market without requiring a new trial?

Back to this diagram, I wanted to remind you, again, the approval steps and how one could take these two issues into consideration within the context of the three voting options. As I mentioned, at the very top, to gain approval, a drug must demonstrate substantial evidence of effectiveness; and if it doesn't, then there will be no approval.

So that's where our first issue lies. There are contradictory efficacy findings on gestational age of delivery. Assuming that substantial evidence of

effectiveness is not an issue, Makena is still sitting in the accelerated approval box, which means that its clinical benefit must be verified. And if the clinical benefit has not been verified, FDA can withdraw approval.

I remind you that either issue in and of itself can impact approval so that you not have to have problems with both issues to impact approval. Let's go back to option A, which is to remove the approval of Makena. That will be appropriate if you find that issue 1, or issue 2, or both, is such that Makena's approval should be removed.

Option B, which is, to leave Makena on the market under accelerated approval -- so again, it will be sitting in the accelerated approval box but require a new confirmatory trial -- would be appropriate if you believe that issue 1 has been adequately resolved so that accelerated approval is still appropriate, but that there is no substantial evidence of effectiveness on the neonatal outcomes and that a new trial is necessary and feasible.

Option C, which is to leave Makena on the

market without a new trial, would be appropriate if you believe issue 1 has been adequately resolved and that the clinical benefit of Makena to the neonate does not need to be verified, so that issue 2 is moot.

I'll walk you through this. Vote A, may be appropriate if you believe that the totality of the evidence does not support Makena is effective for its intended use. If you vote A, please discuss the consequences of Makena's removal.

B, which is to leave Makena on the market under accelerated approval but to require a new confirmatory trial, may be appropriate if you believe that the totality of the evidence supports Makena's effectiveness in reducing the risk of recurrent preterm birth, but that there is no substantial evidence on neonatal outcomes; and you believe that a new confirmatory trial is necessary and feasible.

Let me just comment on this new confirmatory trial being necessary. This will be appropriate if you find that Trial 003, which is a large, adequate and well-controlled trial, is significantly flawed in some way such that its results are not usable or could be

discounted.

If you vote B, please discuss how the existing data provides substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm birth, and also discuss the key study elements of this new trial and approaches to ensure its successful completion.

Lastly, vote C, which is the leave Makena on the market without doing anything else, without requiring a new trial, may be appropriate if you believe Makena is affective for reducing the risk of recurrent preterm birth and that is not necessary to verify Makena's clinical benefit to neonates. If you vote C, discuss how the existing data provide substantial evidence of Makena in reducing the risk of recurrent preterm birth and why it is not necessary to verify its clinical benefit to neonates.

Thank you for your attention, and I now turn the meeting back to Dr. Lewis.

DR. LEWIS: Thank you.

Both the Food and Drug Administration and the public believe in a transparent process for information

gathering and decision making. To ensure such transparency of the advisory committee meeting, FDA believes that it is important to understand the context of every individual's presentation.

For this reason, FDA encourages all participants, including the sponsor's non-employee presenters, to advise the committee of any financial relationships that they have with the firm at issue, such as consulting fees, travel expenses, honoraria, and interests in the sponsor, including equity interests in those based upon the outcome of the meeting.

Likewise, FDA encourages you at the beginning of your presentation to advise the committee if you do not have any such financial relationship. If you choose not to address the issue of financial relationships at the beginning of your presentation, it will not preclude you from speaking.

We will now have presentations from AMAG Pharmaceuticals.

Applicant Presentation - Julie Krop

DR. KROP: Good morning, Dr. Lewis, members of

the committee, FDA colleagues. My name is Julie Krop, and I'm the chief medical officer at AMAG Pharmaceuticals. Thank you for this opportunity to share the results from the PROLONG study and review them in the context of prior clinical trials evaluating 17P.

17P, including our product Makena and recently approved generic formulation, is the only FDA-approved therapy to reduce the risk of recurrent preterm birth. 17P is a synthetic progestin. It contains the active pharmaceutical ingredient 17 alpha hydroxyprogesterone caproate. It is not the same as progesterone or vaginal progesterone.

While its exact mechanism of action is unknown, it is thought to support gestation by decreasing inflammation and inhibiting uterine muscular activity. It's important to note that unlike progesterone, 17P is not metabolized into androgens, estrogens, or corticosteroids. For the rest of the presentation. to be clear, we'll refer to the product we're talking about today as 17P since the discussion is about the entire class, including both Makena and

the recently approved generics.

17P is approved to treat women with a singleton pregnancy who've had a prior singleton spontaneous preterm birth. This population represents a subset of all pregnant women, affecting about 3 percent. That's 130,000 pregnancies every year, and that is why Makena qualifies as a orphan drug.

17P has a prolonged half-life and is administered weekly. Treatment is initiated between 16 and 20 weeks of pregnancy and continues until 37 weeks or delivery, whichever comes first. Prior to the FDA approval of Makena, 17P was available only through pharmacy compounding, which is not held to good manufacturing standards, and that creates the potential for safety and efficacy concerns.

FDA approved 17P under the Subpart H accelerated pathway in 2011. Subpart H approvals are reserved for therapies that treat serious or life-threatening conditions with an important unmet medical need, where efficacy is demonstrated on a surrogate endpoint that is considered reasonably likely to predict clinical benefit.

As FDA pointed out in its briefing book, by the time of 17P's approval, multiple clinical studies evaluating the consequences of late preterm birth had established that preterm infants are less physiologically and metabolically mature than term infants, and therefore at a higher risk of morbidity and mortality. Based on these studies, FDA accepted preterm birth less than 37 weeks as a surrogate endpoint that was reasonably likely to predict clinical benefit.

A condition of accelerated approval was to conduct a confirmatory trial with clinically relevant endpoints. 17P received approval based on the compelling results of study 002, which from this point on we'll refer to as the Meis study. This landmark study was conducted by the National Institute of Child Health and Human Development's maternal fetal medicine units. It was enrolled entirely within the United States.

The Meis study established substantial evidence of efficacy, demonstrating that 17P significantly reduced the rate of preterm birth

compared to placebo. The highly statistically significant results demonstrated the superiority of 17P compared to placebo at the primary endpoint of less than 37 weeks, but also at less than 35 weeks and less than 32 weeks, which have the highest incidence of neonatal complications.

I'd like to highlight some key events in 17P's approval pathway, starting in 2003 when the Meis trial results were published in the New England Journal of Medicine. The Meis results were hailed as a significant advance in obstetrics and ultimately led medical societies to recommend its use to prevent recurrent preterm birth.

After the completion of the study, Adeza Biomedical was granted full access to the data to pursue FDA approval for 17P and submitted an NDA in 2006. Later that year, an FDA advisory committee concluded that the Meis data provided substantial evidence of 17P's safety and efficacy. Most panelists agreed that an effect on early preterm birth at less than 35 weeks and particularly at less than 32 weeks were clinically meaningful, and could therefore serve

as adequate surrogates for reducing neonatal morbidity and mortality. The advisory committee recommended a confirmatory study to verify and describe 17P's clinical benefit.

With increasing adoption of 17P as the standard of care, clinical experts and investigators raised concerns about the feasibility of conducting a placebo-controlled trial in the U.S. In November of 2009, the first patient was enrolled in study 003, from this point on we'll refer to as the PROLONG study.

In 2011, 17P was approved with two required post-approval studies, the confirmatory efficacy and safety study and the associated incident follow-up study, which is still ongoing. Not surprisingly, given the rarity of the condition and the fact that 17P became quickly adopted as the standard of care, recruitment for the PROLONG study was challenging.

Enrolling the requisite 1700 patients required going to sites outside of the United States. In 2014, AMAG became the sponsor, inheriting the study with approximately 50 percent of the patients enrolled. In total, recruitment took 9 years. Enrollment was

finally completed in 2018.

Preterm birth is a major public health concern in the United States, particularly in the most vulnerable patients. It is one of the leading causes of infant morbidity and mortality and can lead to serious long-term health consequences. It's important to remember that recurrent preterm birth represents only a small proportion of all preterm births. While the impact on the total preterm birth rate is minimal, the impact on these women is substantial.

Today, based on the Meis data, clinicians rely on 17P. In fact, based on the sample of nearly a thousand patient charts published in 2018, about 75 percent of patients with a prior spontaneous preterm birth were treated with 17P. 17P is the only FDA-approved therapy to reduce recurrence of preterm birth, supported since 2008 by the American College of Obstetricians and Gynecologists and the Society for Maternal Fetal Medicine, as the standard of care to prevent recurrent preterm birth.

Today, we face a unique challenge. How do we make sense of the PROLONG study in the context of the

prior positive Meis study, which demonstrated consistent and statistically significant efficacy across multiple clinically important endpoints. In the presentations that follow, we'll highlight key differences in study population and background rates of preterm birth that we believe account for the inability of the PROLONG study to demonstrate significant reductions in preterm birth.

The Meis study enrolled patients exclusively in the United States at inner city academic medical centers with high rates of preterm birth. The background or placebo rate of preterm birth at less than 35 weeks was high, around 30 percent. In contrast, the PROLONG study enrolled patients with much lower rates of preterm birth, particularly in Russia and Ukraine.

Background rates of preterm birth at less than 35 weeks were approximately 11 percent, far lower than the rates seen in the Meis study, highlighting the difference in the patient populations, which likely contributed to the different results between the two studies. That said, the strong consistent efficacy

demonstrated in the Meis study, along with previous supporting clinical trial data, and most important, a favorable and reassuring safety profile, all support the continued availability of 17P.

Now let's review the agenda. Next, Dr. Michelle Owens will discuss the clinical background and continued need for 17P; then Dr. Baha Sibai will present the clinical design and the key results from the Meis study. Dr. Laura Williams will present the PROLONG study efficacy and safety data, followed by Dr. Sean Blackwell, who will provide his clinical perspective on the PROLONG data and the overall benefit-risk of 17P.

Finally, I will conclude by summarizing AMAG's action following PROLONG and then moderate the question and answer session. We also have additional experts with us today to help answer your questions. All external experts or their institutions have been compensated for their time and travel with the exception of Dr. Blackwell, who has been reimbursed only for travel.

Thank you, and I will now turn the

presentation over to Dr. Owens.

Applicant Presentation - Michelle Owens

DR. OWENS: Good morning, everyone. I'm Michelle Owens, a maternal fetal medicine physician and professor at the University of Mississippi. I appreciate the opportunity to discuss preterm birth, a significant problem in the United States. One in 10 babies, nearly 400,000, are born prematurely in the United States each year. The rate is even higher for a subset of pregnant women who are disadvantaged socioeconomically, educationally, or by limited access to health care and healthy lifestyle choices. It puts their unborn children at substantial risk, both in the short term and long term.

Fortunately, we have an FDA-approved therapy, 17P, to prevent this in that small subset of women with a prior spontaneous preterm birth, and it's critical that doctors and pregnant women have continued access to it. The stakes are high. We're talking about the health of infants in the short term and throughout their life. I see babies like this one far too often. They can spend weeks or months in the neonatal

intensive care unit.

These babies are often on ventilators because their lungs are immature. They're at high risk for infections. They're also more likely to suffer brain damage or a brain bleed. And even if they get to leave the NICU, many of them don't get a chance to see their first birthday. And for those who do survive, they often face a lifetime of complications.

Let's use 39 weeks as the reference point for the risk of infant mortality with a relative risk of 1. Babies born at 34 weeks are nearly 10 times more likely to die than those who go full term, and babies who make it to 36 weeks are nearly 4 times more likely to die.

Preterm birth and its complications are the number one cause of death of babies in the United States. I've mentioned just a few of the short term risks, and even when we deal with those, the risks don't just go away by getting these infants out of the NICU. While the long-term complications are rare, they are profound and can affect these infants throughout their lives. These babies are at increased risk of learning difficulties, hearing and vision impairments,

and chronic respiratory problems, including asthma.

Babies born at lower gestational ages have higher rates of neonatal morbidity and mortality. An analysis from Manuck, published in the American Journal of Obstetrics and Gynecology in 2016, including more than 100,000 women and their babies, demonstrated a higher rate of death and major morbidities in babies born earlier than 32 and 35 weeks. Approximately 14 percent, that's 1 in 7 babies, born at less than 32 weeks either die or have a major morbidity. At less than 35 weeks, it's 1 in 10 babies.

For context. Let's discuss some background on preterm birth. One in six of all preterm birth occur earlier than 32 weeks gestation, a critical timepoint because of the high prevalence of serious neonatal complications. Our goal is to prolong pregnancy so that we can decrease the chance of these serious complications.

Across the United States, preterm birth rates vary substantially by geography. The March of Dimes assigned the grades of A to F to individual states based on preterm birth rates. The highest rates are

found predominantly in the southeast. My state, Mississippi, has consistently received an F despite our best efforts, though recently we have seen improvements in preterm birth rates.

In addition to where a woman lives, there are many other risk factors for singleton preterm birth, including a multitude of social determinants that, quite frankly, are often overlooked in research. But I can tell you as a clinician practicing in a poor state, these make a difference in overall health, particularly as it pertains to pregnancy. Lower socioeconomic status, higher psychosocial stress, and less access to healthcare all contribute to prematurity.

17P is an effective and integral part of how I help women at risk avoid a subsequent preterm birth. Like most OB/GYNs, I follow the guidelines set forth by SMFM in 2012. For women with no prior history of preterm birth and a short cervix, SMFM recommends vaginal progesterone. For the subset of women with a prior spontaneous preterm birth, SMFM recommends 17P.

Now, it's important to note that this is not a treatment for preterm birth, but the one tool we have

to prevent it. We don't always know which specific patients will benefit, similar to a flu shot or other preventive therapies. In patients with both a prior preterm birth and a short cervix, we continue 17P and place a cervical suture known as cerclage.

In summary, preterm birth remains a major public health concern, particularly in this country. Too many infants are spending weeks or months in the NICU, and too many women with a history of preterm delivery have to watch their babies fight for life. They are afraid to live through that again. As a maternal fetal medicine specialist, my vision is that every child receives the best possible start in life by reducing the preterm birth rate and preventing its complication.

For the small subset of women with a prior preterm birth, 17P provides more than just preventive therapy. It actually provides hope for mothers who are traumatized by the experience of preterm birth, and taking it away would deprive the patients who need it most. Thank you, and I'll now turn the presentation over to Dr. Sibai.

Applicant Presentation - Baha Sibai

DR. SIBAI: Thank you, Dr. Owens.

Good morning. My name is Baha Sibai. I am a maternal fetal medicine physician and professor at UT Health in Houston. I have been in practice for more than 40 years, and I was one of the study investigators. I am here today to describe and summarize the study design and the results that led to 17P's approval, but before jumping into study details, let me explain the premise of studying 17P for recurrent preterm birth.

In 1986, the National Institute of Child Health and Human Development established the Maternal Fetal Medicine Units Network, known as the MFMU. The network's primary aim is to reduce preterm birth by conducting rigorous clinical trials. I was one of the original investigators with the MFMU. I continue to be active in numerous studies.

The MFMU has a rigorous process for selecting both network centers and determining which randomized trials to conduct, given the limited resources. Network centers are selected, in part, based upon the

adequate obstetric populations being at least 40 percent high risk. Additionally, the network has a diverse patient population available for conducting research. The hospitals that are part of the MFMU serve patients at the highest risk due to their social circumstances, and they are often considered safety net hospitals.

Let's review some of the earlier studies of preterm birth. There have been a number of meta-analyses of progestogen. In 1990, Keirse restricted the meta-analysis to only 17P, as this was the most well studied progestational agent. Although these five studies are small and not definitive on their own, they come together. There is a statistically significant relative risk of 0.58, which translates to a 42 percent reduction in recurrent preterm birth with 17P compared to a placebo. Of note, the only study that did not favor 17P was in twin pregnancies for which 17P is not recommended.

This meta-analysis served as the basis for evaluating 17P in a large multicenter trial, which was a research proposal championed by Dr. Paul Meis for the

Maternal Fetal Medicine Network. The Meis study involved women with a history of singleton spontaneous preterm births at less than 37 weeks. Women were randomized in a 2 to 1 ratio to 17P or a matching vehicle placebo.

Women began receiving weekly intramuscular injections between 16 weeks and 20 weeks and 6 days. The Meis population was very high risk for recurrent preterm births given the populations served by centers and the Maternal Fetal Medicine Units Network. There was an imbalance in the proportion of women with more than one previous preterm birth, with 28 percent in the 17P group and 41 percent in the vehicle group. However, this was subsequently and appropriately adjusted for in the statistical analysis.

The other demographics and baseline characteristics were well balanced between treatment groups. The majority were black. The gestational age of the qualifying delivery was about 31 week and approximately 25 percent used substances such as smoking, alcohol, illicit drugs during pregnancy.

The primary outcome was preterm delivery at

less than 37 weeks. We estimated that the sample size of 500 women was needed, expecting a recurrence rate of 37 percent in the placebo group and a reduction of recurrent preterm births with 17P by one third. The Meis study had a very high rate of completion and treatment compliance. The main number of injections was about 40 in both groups. Compliance was defined as not missing 10 days or more between doses. More than 90 percent were compliant in each group.

We began the study in 1999, and it was stopped early due to 17P's clear benefit. In 2002, at a second planned interim analysis, the prespecified stopping criteria for efficacy had been met. The MFMU and the Data Safety Monitoring Board determined that if 17P demonstrated efficacy with a p-value of 0.015, recruitment would be halted. This decision was made so that once 17P's efficacy was established, women at risk for recurrent preterm birth would not receive a placebo.

Outcome data were available for 463 out of the total 500 patients. This represented 93 percent of the planned study population. The data you see here are

from our New England Journal of medicine publication. We found a significant reduction in preterm birth rates with 17P compared to vehicle at 37 weeks, at 35 weeks, and at 32 weeks. These women who are at very high risk for preterm birth, 17P significantly reduced recurrent preterm birth compared to vehicle.

When we certified the results by these factors for preterm birth, we saw consistent reduction across all subgroups. Importantly, regardless of the number of prior preterm births, the relative risks were similar. However, these are just some of the no-risk factors for preterm birth. There are many more unknown factors as described by Dr. Owens, but across the board, these results demonstrate the robust and consistent efficacy of 17P.

Turning now to neonatal complications, the reductions I just showed you in preterm birth rates translated to direct clinical benefit for the neonates. Although the Meis trial was not adequately powered to evaluate neonatal complications, there were consistent reductions with 17P. With the exception of neonatal sepsis, all point estimates of relative risk favors 17P

with some significance.

These neonatal complications, particularly some of those listed at the top, have important clinical implications for long-term outcomes. We clearly see the benefits of 17P by looking at neonatal intensive care unit admissions. Mothers receiving 17P were less likely to have their infant admitted to an ICU; and if their infant was admitted the mean days in the NICU were shortened.

Let's look closer at perinatal death. The overall perinatal deaths were similar between groups. The rate of neonatal deaths with 17P was half that of the vehicle. There was a small and non-significant increase in the rate of miscarriage and stillbirth in the 17P group. This was evaluated further in the PROLONG study, which you will hear about shortly from Dr. Williams.

When we give medications in pregnancy, long-term safety of the babies and healthy development is always a concern. The MFMU conducted a follow-up of babies enrolled in the Meis study and confirmed the long-term safety of 17P exposure in utero. Nearly

80 percent of eligible children completed development assessment, including the Ages and Stages Questionnaire shown here. That includes five domains.

The median age at follow-up was 4 years. There were no differences between 17P and vehicle. Caretakers also administered the preschool activities inventory, which showed no gender-specific differences. Also, this follow-up study reassured long-term safety and development of babies exposed to 17P.

When we published our findings in the New England Journal of Medicine in 2003, the results were considered a significant advance in obstetrics. Overall, 17P reduced preterm birth by about one-third, which was highly statistically and clinically significant, with a absolute difference in preterm delivery of nearly 19 percent.

Numbers needed to treat are often used to convey efficacy of medications. A number needed to treat of hundred is typically considered an appropriate threshold for a clinical value. Remarkably, based on these data, we need to treat with 17P only 5 to 6 women who have had a prior singleton spontaneous preterm to

prevent one recurrent preterm birth.

In summary, the Meis study established substantial evidence of 17P's efficacy and formed the foundation of today's standard of care for high-risk pregnant patients where a history of spontaneous preterm delivery. Since 2003, clinicians have relied on 17P. I have seen 17P reduce recurrent preterm birth in my patients with a history of spontaneous preterm birth, and I continue to routinely prescribe it for these patients.

Without FDA-approved 17P, there will be no acceptable alternative to prevent recurrent preterm birth in this patient population. Moreover, our obstetric community has extensive clinical experience with 17P and supports its use in this subset of patients who are at high risk for preterm birth. Thank you. I now would ask Dr. Williams to come.

Applicant Presentation - Laura Williams

DR. WILLIAMS: Good morning, and thank you Dr. Sibai.

I'm Laura Williams, senior vice president at AMAG and head of clinical development and

biostatistics. Today I'll be reviewing the efficacy and safety results from the PROLONG study.

PROLONG was designed to mirror the Meis trial, and as you've heard, it did not meet its co-primary endpoints. Despite similar entry criteria, background preterm birth rate in the placebo group were much lower in PROLONG compared to Meis, which likely played a significant role.

Let me first take you through the PROLONG study design. PROLONG was a double-blind, vehicle-controlled, multicenter, randomized study in women with a singleton pregnancy and a history of a previous singleton spontaneous preterm birth. The key objective was to further demonstrate the safety and efficacy of 17P in this study population. Eligible women could be randomized between 16 weeks 0 days and 20 weeks 6 days of pregnancy.

In total, 1708 were randomized in a 2 to 1 ratio to receive either 17P or vehicle, respectively. Women received weekly intramuscular injections of study drug until 36 weeks 6 days of pregnancy or delivery, whichever occurred first.

In addition to routine follow-up for the mom following study completion, a prospective, non-interventional, infant follow-up study, similar to what was done in Meis, is also being conducted for PROLONG. This study remains blinded to complete the follow-up with database lock anticipated in late 2020.

The co-primary outcomes for PROLONG were preterm birth at less than 35 weeks gestation and a neonatal composite index that highlights the significant morbidity and mortality often associated with preterm birth, which Dr. Owens previously highlighted. The index included respiratory distress syndrome, bronchopulmonary dysplasia, grade 3 or 3 intraventricular hemorrhage, necrotizing enterocolitis, sepsis, or death.

Key secondary outcomes were the reduction in preterm birth by gestational age at delivery. The primary safety outcome was to exclude a doubling in the risk of perinatal deaths. This was included to address concerns from the original review. The sample size and powers assumptions for the PROLONG study were based on results from the Meis trial.

Based on preterm birth rates in the vehicle group in Meis, a sample size of 1707 patients provided 98 percent power to detect a 30 percent reduction in preterm birth at less than 35 weeks gestation and a 90 percent power to detect a 35 percent reduction in the neonatal composite index. Assuming a 4 percent fetal or early infant death rate in both treatment arms, the sample size provided 83 percent power to exclude a doubling in risk of perinatal death.

Let's look at the patient disposition. Impressively, 99 percent of patients completed the study; 1113 in the 17P arm and 574 in the vehicle arm had data for the preterm birth endpoint and were included in the intent-to-treat or ITT population to evaluate efficacy. The most common reasons for treatment discontinuation were withdrawal of consent or lost to follow-up. All patients who received at least one dose of study drug were included in the safety evaluation.

Now, let's take a look at enrollment by geographic region. As you heard earlier, since 17P was recommended in treatment guidelines and had rapid

uptake in clinical practice, enrollment in the U.S. was extremely challenging. The first patient was enrolled in November of 2009, and as expected, enrollment in the U.S. became increasingly difficult. For that reason, approximately 75 percent of patients in PROLONG were enrolled outside of the U.S. Notably, 61 percent were from Russia and Ukraine.

Let's take a closer look at enrollment over time. The study enrolled from 2009 to 2018, and nearly all U.S. patients enrolled by 2014. In the last four years of the study, only 49 additional U.S. patients were enrolled. With enrollment rates plateauing in the U.S. it was clear that in order to complete the study, ex-U.S. sites would be needed. And beginning in 2014, enrollment increased in Russia and Ukraine, allowing for study completion.

Turning now to demographics and baseline characteristics, demographics and other baseline characteristics thought to be associated with preterm birth were similar across treatment groups. The mean age was 30, most women were white, non-Hispanic or Latino, and married or living with a partner during

this study. The mean prepregnancy BMI was around 24 with a small percentage of patients having a short cervix, that is less than 25 millimeters at the less than or equal to 20 weeks gestational age.

Less than 10 percent in both treatment arms reported any substance used during pregnancy at baseline. Prior pregnancy history was also similar across treatment groups. A prior spontaneous preterm birth was an entry criteria such that the median was 1. Only 12 to 13 percent of women had more than one prior spontaneous preterm birth, and the mean and median age of the prior qualifying delivery was around 32 and 33 weeks, respectively.

Let's move now to study drug compliance. The number of study drug injections were comparable across treatment groups, injections were administered at the investigator site, and more than 90 percent of patients were fully compliant with their scheduled appointment to receive weekly injections.

Now let's review the study results. Here we show the preterm birth endpoint on the left and the neonatal composite index on the right. The relative

risk with 95 percent confidence intervals are provided above the bar graphs for each endpoint. As you can see, the results were not statistically significant between treatment groups for either endpoint. Preterm birth rates at less than 35 weeks were around 11 percent and neonatal composite index rates were around 5 percent.

In addition to the preterm birth rates at less than 35 weeks, there were similar results for preterm birth rate at less than 32 and less than 37 weeks gestation. Recognizing that most patients were enrolled outside the U.S., we also looked at efficacy by geographic region, which was a prespecified analysis, and we found no statistically significant difference between treatment groups by region. However, the preterm birth rates were notably higher in the U.S. compared to ex-U.S.

In fact, they were one and a half to 2 times higher, at nearly 18 percent in the U.S. compared to almost 10 percent ex-U.S. The neonatal composite index rate was around 9 percent in the U.S. compared to only 4 percent ex-U.S.

Given the lower background preterm birth rates seen here in PROLONG compared to Meis, we conducted various exploratory analyses in an effort to better understand the efficacy results from the two registrational studies, Meis and PROLONG. We first examined baseline characteristics between these two study populations, and differences in PROLONG compared to Meis were noteworthy.

Patients in PROLONG were nearly 4 years older. They were 50 percent less likely to have had more than one prior spontaneous preterm birth. Only 7 percent were black and 9 percent were Hispanic. Only 10 percent were unmarried and only 9 percent reported substance use during pregnancy. But interestingly, and perhaps not entirely unexpected, those differences were far less prominent when looking at the U.S. PROLONG population, which was clearly more similar to Meis. That said, it's also important to reiterate differences in background preterm birth rates in the placebo group in Meis at 31 percent versus U.S. PROLONG at nearly 18 percent.

As FDA has noted, the cause of preterm birth,

or causes of preterm birth, are multifactorial, and the uncertainty around the relative contribution of any given risks makes finding markers of response very challenging. We thought a lot about how best to interrogate the data to provide additional insights and have conducted various additional analyses, some of which were post hoc, exploratory, and hypothesis generating.

Although the U.S. PROLONG subset population was not identical to Meis, given the more similar demographics and background characteristics, we were compelled to look at the subset population in much more detail. And here you see the aforementioned results for preterm birth rates at less than 35 weeks for PROLONG on the far left, Meis in the middle, and U.S. PROLONG to the far right.

In the U.S. PROLONG subset population, there are trends and relative risk reductions indicating benefit favoring 17P, and the relative risk of 0.88 is directionally aligned to that seen in Meis at 0.70. We also saw similar findings for preterm birth rate at less than 32 weeks, with relative risk reductions in

preterm birth at less than 32 weeks, again, indicating benefit favoring 17P, and the relative risk of 0.58 is even lower than that seen in Meis at 0.64.

Importantly, those trends in reductions in preterm birth rates also translated to relative risk reductions in the neonatal composite index in the U.S. PROLONG subset, similar to what was seen in Meis. So while analyses of efficacy by geographic region were prespecified, we fully acknowledged that these analyses are exploratory and in no way change the overall efficacy findings. However, these trends that favor 17P in a smaller subset U.S. population that was not powered to show these differences are promising and directionally aligned with results from Meis.

So how do we summarize these efficacy data? PROLONG did not meet its primary efficacy outcomes, but these findings do not refute the efficacy results seen in the Meis trial. Key differences in background rates of preterm birth across different study populations are the most plausible reason, and as you evaluate subset populations like U.S. PROLONG, which had higher background preterm birth rates than PROLONG overall,

there were trends for benefit favoring 17P in a much smaller subset population that was not powered to demonstrate efficacy. Nevertheless, these findings are promising as they directionally align to those from the Meis trial.

Now then, let's take a look at the safety data. The key safety outcome was to exclude a doubling in risk of perinatal death in the 17P group compared to vehicle. If the upper bound of the confidence interval is less than or equal to 2, a doubling in risk of perinatal or neonatal death would be excluded. Fetal and early infant death, or neonatal death, was defined as a spontaneous abortion or miscarriage occurring from 16 weeks to 20 weeks gestation, a stillbirth occurring at greater or equal to 20 weeks gestation, or an early infant death, which is a liveborn death at less than or equal to 24 weeks gestation with death occurring from minutes after birth until 28 days of life.

With anticipated low rates for this outcome, sample size considerations to exclude a lower risk level were taken into account for this orphan population when the FDA defined and added this specific

endpoint. However, I think we all agree that the most important outcome is the overall rate of all perinatal deaths.

As shown here, the prespecified primary safety outcome, total fetal or early infant deaths had low and similar rates across both treatment groups. Rates of miscarriage were numerically lower in the 17P group compared to vehicle, while rates of stillbirth were numerically higher. Most importantly, the rates of all perinatal deaths were low and similar across treatment groups.

Overall, the incidence of adverse events and maternal pregnancy complications were comparable between treatment groups. Rates of adverse events leading to study drug withdrawal and serious adverse events were also low and similar, and there were no maternal deaths occurring during the study.

This table shows adverse events and maternal pregnancy complications occurring in at least 3 percent of patients in the 17P arm. Maternal pregnancy complications are denoted by an asterisk. As shown, the rates were low and comparable between the two

treatment groups. Only 15 patients in the entire study discontinued study medication due to an adverse event or a maternal pregnancy complication, again with low and similar rates across treatment groups.

This table captures serious adverse events in maternal pregnancy complications that occurred in two or more patients, and, again, the rates were low and comparable across treatment groups. As is usually done with similar design registration studies, a pooled safety data analysis combining Meis and PROLONG was also conducted as a post hoc analysis. Additional details of those pooled safety data are included in the briefing package, but they are similar to what I've shown for PROLONG.

Finally, we will review postmarketing safety findings. Among the estimated cumulative U.S. Makena exposure of nearly 300,000 patients, safety data obtained from postmarketing surveillance remains very consistent with both Meis and PROLONG. The most frequent adverse event reports were consistent with the registration studies with injection site reactions leading the list. The overall postmarketing safety

data in general and around perinatal deaths in particular had very low reporting rates and are, again, also consistent with what was seen in the registration studies.

So how do we summarize the safety data? PROLONG reaffirmed the safety of 17P that was demonstrated in the Meis study. We saw no new or unexpected findings and no clinically meaningful difference in safety between treatment arms. Overall, across both studies and in clinical practice, 17P has consistently demonstrated favorable maternal and fetal safety.

Thank you. I'll now turn the presentation over to Dr. Blackwell.

Applicant Presentation - Sean Blackwell

DR. BLACKWELL: Thank you, Dr. Williams.

Good morning. I'm grateful for the opportunity to provide my perspectives on the role of 17P in this high-risk patient population. I was the lead author of the PROLONG publication, and I have thought a lot about why the findings were different from the Meis trial. I am also a maternal fetal

medicine physician and departmental chair at McGovern Medical School at the University of Texas in Houston. I lead a physician team, which includes 25 maternal fetal medicine physicians, 50 obstetricians, 12 maternal fetal medicine fellows, and 48 OB/GYN residents across 10 hospitals.

One of my jobs is to make sure that physicians are providing the best care for our patients, and as a high risk pregnancy specialist, this definitely includes trying to prevent recurrent preterm birth. So these discussions and decisions about 17P are not theoretical or abstract. They will affect what we do every day.

The goal of my presentation is to address three key questions? Why did the PROLONG efficacy results differ from the Meis trial; is it feasible to conduct another confirmatory trial; and what should we do from here; and how should I guide my team of physicians in the care of their patients?

To the first question, why did PROLONG efficacy results differ from the Meis trial? You have heard from Dr. Sibai as he described the Meis trial and

Dr. Williams explain PROLONG. It was perplexing at first. How could two studies with the same enrollment criteria in the same treatment protocol, that both performed with high methodologic rigor, have such different results?

The bottom line is that these two clinical trials ended up studying two very different groups of women. The Meis trial studied women from university based academic medical centers in the United States. This population included a very high percentage of African American women and women with lower socioeconomic status. These women enrolled in Meis had a very high background rate of preterm birth and were motivated to participate based on their obstetrical history.

PROLONG recruitment was 75 percent outside the United States, and the two countries with the largest recruitment were Ukraine and Russia. There were only 7 percent of women in PROLONG who were black, and their socioeconomic status in PROLONG appeared to be greater, on average, than women enrolled in the Meis trial. The percentage of women with greater than one prior preterm

birth was half that of the Meis trial. These facts are manifest in the comparison of the rates of preterm birth in the placebo arm of these two trials. e can see marked differences in the preterm birth rates at 32 weeks, 35 weeks, and 37 weeks.

This slide illustrates these differences between three trials using preterm birth less than 35 weeks as a proxy for baseline risk of preterm birth. and I've chosen preterm birth less than 35 weeks since it was a co-primary outcome for the PROLONG trial. This slide not only highlights the differences in the baseline risk between me and PROLONG but also the differences between women recruited in the U.S. versus outside the U S for a PROLONG.

I have also included the O'Brien trial for additional context. This was an international, placebo-controlled trial of vaginal progesterone, which was also studied in women with a prior spontaneous preterm birth, and the vast majority of women were recruited from the United States. The importance of this slide is to emphasize the differences in the recurrent preterm birth rate in the U.S. versus non-

U.S. sites across various study populations.

Recruitment challenges in the United States were a second major factor for why PROLONG had such a lower risk patient population. The first patient recruited for PROLONG was in 2009, but in 2003, less than 5 months after publication of Meis, ACOG published a committee opinion supporting the use of progesterone for women with a prior spontaneous preterm birth.

In 2006, a survey published in the American Journal of Obstetrics and Gynecology indicated that two-thirds of board certified maternal fetal medicine physicians were already using progesterone for women with a prior spontaneous preterm birth. By the time prolonged started its recruitment in 2009, most maternal fetal medicine physicians in the United States were already using this treatment, and therefore most likely not willing to participate in a placebo-controlled trial.

As an example, no center in the Maternal Fetal Medicine Units Network and very few university academic medical centers in the United States were recruitment sites for PROLONG. Neither Dr. Sibai nor I, while at

different institutions, felt it proper to refer our patients to PROLONG. In our minds, a placebo-controlled trial was only appropriate where 17P was not accessible.

These challenges resulted in enrollment bias in PROLONG favoring a lower risk patient population. Due to this bias, women at greater risk for preterm birth, such as those with a short cervix or more severe obstetrical history, were potentially steered away from participating in PROLONG in favor of some other open-label therapy. PROLONG had one-half the number of women with greater than one prior preterm births than Meis, and less than 2 percent of women in PROLONG had a short cervix, a percentage much lower than one would expect from prior trials.

The sample size estimates for PROLONG were based on the Meis trial, yet the rates in PROLONG were 50 percent lower than Meis. If we were to design a new trial today based on these lower event rates, 3,600 women would be required for a 90 percent power for preterm birth less than 35 weeks and 6,000 women would be needed for the neonatal composite index. Based on

these population differences and low event rates in PROLONG compared to Meis, the results are inconclusive regarding efficacy.

In PROLONG, there was a preplanned subgroup analysis of 17P treatment effect by U.S. versus the non-U.S. population. These analyses by their nature are exploratory and hypothesis generating and not meant to be conclusive. In the U.S.-only subgroup, there are trends for benefit for both co-primary outcomes with relative risks 0.88 and 0.84, respectively. Although less robust, these are in a similar direction as Meis and would be clinically significant.

The second question, is it feasible to do another confirmatory trial? As a maternal fetal medicine physician who conducts clinical trials, my ears perk up when someone proposes we do another one. However, in this case, the answer is no. I do not think another interventional trial or a confirmatory trial is feasible. I do not believe physicians or patients will accept a placebo in this patient population, even with the lack of benefit noted in the PROLONG trial. At worst, the trial would be futile,

and at best, the same enrollment bias would occur.

This is certainly true in the United States, but I also believe would occur outside the United States in any developed country. In order to conduct this trial, we would have to identify a population of women at sufficiently high risk who also have no access to 17P and be in a setting where there is research infrastructure to conduct a major trial. All this seems improbable.

Now, another option would be a comparison of two therapies, thus no one would receive a placebo. The problem is that there are no other evidence-based therapies that would be a good alternative to 17P. Vaginal progesterone has been studied in women with a prior spontaneous preterm birth. Three recent large placebo-controlled trials -- O'Brien, Norman, and Crowther -- included 2000 women with a high baseline risk of preterm birth. All reported no benefit for this population. Other potential therapies such as cervical cerclage or cervical pessary have also not shown benefit for women with a prior spontaneous preterm birth.

Finally, what should we do from here, given the robust findings from the Meis trial, and then a larger trial, PROLONG, that is inconclusive? Following the publication of PROLONG trial, both SMFM and ACOG have given updated guidance to physicians regarding the role of 17P. I am the past president and prior chair of the SMFM Publications Committee, but due to my involvement with PROLONG, I was not involved in the new SMFM guidelines statement.

SMFM states that based on the evidence of effectiveness in the Meis study, which is the trial with the largest number of U.S. patients, and given the lack of demonstrated safety concerns, SMFM believes that it is reasonable for providers to use 17P in women with a profile more representative of the very high-risk population reported in the Meis trial.

ACOG has not changed their clinical recommendation at this time and continues to recommend offering 17P as outlined in their practice bulletin. We also have to consider what will happen if an FDA-approved 17P would no longer be available. It is my belief that many experts and clinicians will still

consider the risks and benefits of 17P in a positive balance that supports its use. If there is not a 17P FDA-approved version available, many will turn to a compounded 17P. Others will advise off-label, unproven medical therapies or choose a surgical option with cervical cerclage, which has not been proven to work and has a greater risk for patient harm.

Finally, last question, what will I do? How do I recommend we take care of our patients? First, I believe that the Meis and PROLONG studies do not contradict each other. Meis shows robust treatment effects for a high-risk U.S. population similar to my patients. PROLONG did not confirm treatment efficacy in a much lower risk population and was inconclusive due to its sample size. PROLONG does provide reassuring data regarding safety, miscarriage, pregnancy loss, and gestational diabetes.

Overall, the benefit to risk ratio is positive considering the totality of efficacy data and the low safety risk profile. That is why I will continue to offer and recommend 17P to my patients. It's my belief, after counseling many women with a prior

preterm birth, especially those who deliver at a very early gestational age, or those whose child suffered from complications related to preterm birth, we'll choose 17P therapy based on the available data.

In order for my team of physicians to provide the best care for our patients, it's essential that we have the ability to offer an FDA-approved 17P, especially to those at the highest risk. Thank you.

Applicant Presentation - Julie Krop

DR. KROP: Thank you, Dr. Blackwell.

I'd like to conclude our presentation by summarizing what you heard today and sharing the actions AMAG is taking following the PROLONG study. We have just reviewed the totality of the evidence that supports continued access to 17P. The Meis study demonstrated robust and substantial evidence of efficacy and was the basis of ACOG and SMFM's recommendation of 17P.

Last week, after reviewing the PROLONG publication, ACOG and SMFM announced their continued support of 17P. Because the placebo birthright in the placebo arm of the PROLONG study was much lower than

rates typically seen in the United States, the results are inconclusive and difficult to apply to the U.S. population. Despite these differences, it neither refutes nor invalidates the findings of the Meis study.

So what have we learned over the 10 years it took to complete the PROLONG study? We've learned that since 17P was recommended by medical societies as the standard of care, it was not possible to conduct a placebo-controlled trial to confirm the Meis results. Once efficacy was established, U.S. physicians would not withhold an efficacious treatment from their patients. Bias was introduced. This bias skewed enrollment towards a low-risk patient population. Despite this bias, the U.S. subset still demonstrated trends favoring 17P for the co-primary endpoint. However, the U.S. subset was not powered to evaluate efficacy.

The PROLONG study did confirm 17P's favorable safety profile. We also have eight years of postmarketing surveillance, which firmly supports its safety in this population. While we successfully conducted and completed the confirmatory trial, the

results are inconclusive. This leaves us with a question. If the Meis study was being reviewed here today, would Meis alone have met the criteria for full approval?

According to FDA's guidance on establishing evidence of effectiveness, approval may be supported by a single trial if a second trial is not feasible or ethical. To qualify, that single trial should demonstrate statistically persuasive findings on a clinically relevant endpoint, as well as robust, consistent results across multiple subgroups in the study. If so, the results of a single trial are frequently sufficient to support approval in the context of a rare or orphan condition.

Today, almost a decade after 17P's approval, there is now compelling evidence delivery at less than 37 weeks, but especially at less than 35 weeks and less than 32 weeks, are associated with significant increases in neonatal morbidity and mortality. This newer data strongly suggests preterm birth endpoints evaluated in the Meis study should no longer be considered surrogate endpoints that require a

confirmatory study.

It's important to note that this population of women with a prior preterm birth still qualify today as an orphan condition with no available treatment options. Given what we know today, we believe 17P's reduction in preterm birth rates at less than 32, less than 35, and less than 37 weeks in the Meis study, coupled with its consistent statistically significant efficacy across multiple endpoints and subgroups, and 17P's overall reassuring safety profile, strongly support its continued availability.

It is vital that we put the PROLONG study into the proper context so we make the right decisions for these high-risk patients. It's critical to remember that 17P is not a treatment for preterm birth; it's a treatment aimed at reducing risks. Like other preventive measures, we do not expect to see a benefit in a low-risk patient population. We trust physicians and their patients to weigh the potential benefits and risks of treatment together.

To better inform these decisions, the PROLONG results have recently been published in the American

Journal of Perinatology. In addition, we propose working closely with FDA to update all relevant sections of the label with the PROLONG study data in order to provide clinicians with a comprehensive understanding of all available safety and efficacy data.

A question you face today is whether or not another confirmatory trial needs to be done. We have grappled extensively with this question and if any study could serve as a confirmatory study of the Meis study. As you've heard from Dr. Blackwell, another randomized, placebo-controlled trial is simply not feasible. Worse, it might even be considered unethical given the current clinical practice guidelines that recommend 17P's use in this high-risk subset of preterm birth.

We've also carefully considered alternative study designs such as an observational study. The challenge, how do account for the myriad of known and unknown risk factors for preterm birth that would be difficult or impossible to control for in a non-randomized trial. That said, we look forward to

hearing your thoughts today. We are committed to working with the FDA to look for other potential studies that might better inform providers on the appropriate use of 17P.

The totality of the data we share today and nearly a decade of routine clinical use, support 17P's positive benefit-risk profile and the importance of continuing to make it available to physicians and their patients. Preterm birth remains a major public health concern, particularly in the most underserved and most vulnerable patients. These patients have the highest preterm birth rates, and they are the very patient population who benefited the most in the Meis study.

We look forward to today's discussion and partnering closely with the FDA on next steps. Most important, as we complete this work, it is critical that we do not take this medication away from the patients who need it the most. Thank you.

Before we take your questions, I wanted to mention that the lead statistician for the Meis and the PROLONG study, Dr. Anita Das, is unable to be here due to an emergency. Dr. Das lives in the area impacted by

the current wildfires in California, and her neighborhood is under mandatory evacuation. She left to be with her family, but she will be joining us by phone today, so we're happy to take your questions.

Clarifying Questions to Applicant

DR. LEWIS: Thank you.

Are there any clarifying questions for AMAG Pharmaceuticals? Please remember to state your name for the record before you speak, and please identify which presenter your question is for, or if it is a general question for all presenters. We'll start with Dr. Davis.

DR. DAVIS: Thank you very much for the presentation. There's a lot of work and effort that goes into that. I was curious about a few things. One is if your group could clarify how you chose the sites and in what order. Clearly, I think we all recognize there are tremendous regional disparities globally with things such as preterm birth, so I was curious how you ended up in Russia and the Ukraine with the majority of your patients, and then the European sites look like they came later and had a much smaller percentage.

That's my first question, and once you answer that, I'll follow up with one more short

DR. KROP: Yes. The sites were selected in the United States based on specific criteria to make sure that they have the adequate neonatal care, level 3/level 4 NICUs, and appropriate experience doing research. It was quite challenging because the majority of centers that qualify for that were already part of the network and would not participate.

We had 42 sites in the United States attempt to enroll, and when it became clear, because of the entrenched guidelines, it became impossible to recruit at those centers, we had other centers in Europe as well as Ukraine and Russia. But we saw that those recruitments were going much better than the United States, and we continued to add sites there in order to complete the study. It's very difficult in an orphan population to get, as you can imagine, 1700 patients. Those were the sites that were the highest recruiters. We had sites also in Italy. We had sites in Spain. Unfortunately, they were not strong recruiters.

DR. DAVIS: Just one more brief question. It

involves this neonatal morbidity index. This is by far the healthiest group of babies I've ever seen in my lifetime, and using it as an outcome measure, when you have a 98 percent survival and you have more deaths than any intraventricular hemorrhage, something didn't make a lot of sense to me.

At least to me, it suggested that these were mostly older, very healthy babies. The ones we are really concerned about were the ones delivering less than 30 weeks, or 28 weeks I guess was some of the data, and that didn't seem to have much of an influence by progesterone.

DR. KROP: Again, I think we did have a much healthier patient population. Our event rates in the neonatal index were much lower than we anticipated. Unfortunately, that made it very difficult to show benefit, I think, compared to the Meis trial, where there were much higher incidences of adverse affects in the infants, a much higher background rate of preterm birth and higher number of risk factors.

DR. LEWIS: Thank you. Dr. Bauer?

DR. BAUER: Thank you. I have a question for

Dr. Sibai about the Meis trial. Again, through much of the presentation, it's been discussed how this was really a landmark study, and it certainly was. But it's interesting. I really was struck by the unexpectedly high event rate in the placebo group, almost 55 percent. In fact, that is much, much higher than even the meta-analysis numbers that you showed, where it looks like it was about 28 percent above the other trials.

I'm wondering if you can discuss that because it looked like, based on the power estimates, that actually they expected the event rate in the placebo group to be closer to 36 percent, I believe, and it was 55; and in fact the event rate in the active treatment group was close to the placebo group, or expected in the placebo group. I don't know if you can mention that.

Also, if you could also just then comment what particular risk factor profile you think accounted for that really astronomically high event rate.

DR. SIBAI: Thank you for your question. The rate that we estimated the sample size was, we expected

the rate to be 37 percent. However, given the nature of the network and the patients in the network, and considering the fact when the trial was performed, there was no other drug available, it required a woman to receive 20 intramuscular injections. So it became obvious, people who agreed to enroll in the trial pre-selected themselves to be at highest risk. If you look at that population, very high-risk women had more than one prior preterm birth. In addition, we had a high percentage of women who their qualifying prior preterm birth was at very risk.

Given all of this information, the risk factors for recurrent preterm birth, not only having a prior spontaneous preterm birth, it depends on the gestational age, when you had the prior preterm birth, as well as the number of prior preterm births. Because we had this very high rate in the placebo, we expected it to be 37 percent based on a study we did, an observational study with collected data, prospectively, to know what will be the baseline, so we ended up having a much higher rate.

However, this was wasn't surprising because

the network did another study, which was a randomized trial of women who were assigned to Omega 3 versus a placebo to prevent recurrent preterm birth. All of these women received 17P, and still we had a very high rate of recurrent -- Omega 3 didn't work, but the rate was still the same.

More importantly, when we did a study after the availability of 17P, the compounded form, earlier we looked at data collected by one of the home health agencies that enrolled more than 5400 women in 40 states in the United States, all of these women received 17P, and the rate of recurrent preterm birth, at less than 37 weeks and at 35 weeks, was similar. So it seems as if the patient populations receiving the 17P are really at a very high risk of preterm birth. It wasn't only unique to the network.

DR. KROP: And I would add, I think these patients are still quite prevalent. I would ask Dr. Owens also to comment in terms of her experience at her center.

DR. OWENS: Michelle Owens, Jackson, Mississippi. My patient population is probably more

similar to the Meis population that was studied. I do practice in a state that has led the country for years with the highest rates of preterm birth. We have significantly higher rates of not only preterm birth, but also, subsequent to that, infant mortality.

My patients reflect very similar demographics. They are socioeconomically disadvantaged, in many cases, educationally disadvantaged, and we have a high percentage of African American patients as well. Many of the patients where I live in my state, while I am in a metropolitan area, the largest city in my state, many of my patients will travel 3 or 4 hours from many more rural areas in order to receive their care.

I've been using 17P for women with a history of spontaneous preterm birth, and I have actually seen the benefits. The greatest complaint that we have come to expect from the women, who have had a preterm birth and then turn around and subsequently come in for care, is that they end up being more pregnant than they've ever been, and typically much more uncomfortable because they're carrying their pregnancies to longer gestations,

This particular day is really important because I feel like we know that we have some seemingly confusing information in a lower risk population, but we do have really compelling data that tells us that this works exceptionally well in a very unique subset of women, and it's so integral that they continue to have access to this medication.

DR. KROP: It's also important to remember that about 50 percent of our sales are to Medicaid patients, which is representative of the population. I think about 43 percent of pregnant women are on Medicaid, so it is a high-risk patient population.

DR. LEWIS: Thank you. I have a quick question, and I'm not sure who would best answer it. That is, what have been the trends in U.S. preterm delivery rates, by race, I guess.

DR. KROP: I'll answer the last part of that question. The rates of preterm birth in United States have been about 10 percent, and they've been fairly steady over the last several years. You have to remember this as a very small subset of patients that this affects, so therefore, we wouldn't really expect

to see a difference in the preterm birth rate. In fact, there was a survey done based on the Meis -- not a survey, an analysis done based on the Meis trial, where if you assume all 10,000 births that would be affected, it would only improve -- I think it would only decrease the overall preterm birth rate by like 0.3 percent, so it would be very difficult to detect, based on that.

DR. LEWIS: Thank you. Dr. Gillen?

DR. GILLEN: Thank you. I'm trying to put the general logic together in my mind here. The preface here is that the two studies disagree. Meis and PROLONG disagree because they have different patient populations. The implication would be that there is a different point estimate in effective treatment in those two populations due to effect modification by subgroups.

If we can start with -- and there is a question coming here, but I need to set it up. If we can start with slide C-034, which is the Meis study, which very beautifully -- and I think the sponsor presented this in 2006 -- shows consistency of results

across all subpopulations, and quite strikingly in that consistency of results. I'm starting with, are there any subpopulations that were found in the Meis study for which there was a differential effect; in other words, for which we would expect effect modification if we had oversampled those individuals?

That's the first. Then if we go to slide C-056, I think there's a very strong preface here that says that it's a U.S. issue, that we've oversampled individuals outside of the United States. And if we focus on those individuals within the United States, we can see that we now have a similar patient demographic to that that was observed in Meis.

Then if we go to slide C-058, and here will be my question, alas, when we stratify on the U.S. population in PROLONG, first of all, isn't that point estimate of 0.88 with a confidence interval ranging from 0.55 to 1.40 exactly consistent with what is seen as the point estimate and confidence interval that's seen in the overall PROLONG population? We've seem to have treat it differently, and I think that the words were, "It's in the right direction, so with adequate

power, it would have been significant." That presumes that 0.88 is the true estimate. That's not what it is. The confidence interval ranges from 0.55 to 1.40 there.

So my question is, was there any effect modification that was tested and observed in PROLONG with respect to the U.S. population, or with respect to any other subpopulation inside of PROLONG, where you can simply say, yes, there is a differential effect of this therapy in this subgroup?

DR. KROP: We conducted a number of post hoc group analyses looking at race, ethnicity, many of the traditional factors that you would think of, composites, level of background. I think we have a forest plot of the various subgroups that we looked at in PROLONG that we can bring up in a second.

I think you have to keep in mind, the PROLONG U.S. subset is substantially underpowered. It was not powered, obviously, to look at those endpoint. And when we went back retrospectively and tried to calculate the power we would have had in the U.S. subset, it was less than 20 percent, so that's a challenge.

I think with the subgroup analysis up here, you can see there really isn't anything, based on what we can understand of traditional risk factors, but one has to remember that there are a whole hosts and a myriad of other risk factors, as FDA points out, that we don't fully understand. When you enroll a very different patient population with different social characteristics, it's hard to understand what those impacts would be.

As Dr. Owens stated, in her practice, there are huge impacts of social determinants of health in terms of disadvantage that are impossible to incorporate into a clinical study. They're just different patient populations. I Ukraine and Russia, there are preventive services that are far more significant than we have here in the United States. Women are counseled before they ever become pregnant. There's a universal health care system; I mean, just a host of different factors.

DR. GILLEN: I appreciate that, but what I am as a committee member am struggling with is -- and this is Dr. Owens' words, "This works well in a selected

population," but who was that population? Who are we talking about? In other words, we can't have it both ways. We can say, "Oh no, no, no, the population was what we had seen in Meis, but it was the wrong population in PROLONG." But we can't find that subpopulation in PROLONG to justify what was seen in Meis.

So I'm asking, what is that selective population that you're asking me to consider here?

DR. KROP: I'm going to call up Dr. Sibai in a minute, but I think it's important to remember the bias element that was in play in the U.S. Trying to do a clinical trial in the presence of an existing standard of care does bias your population that you put in, so I don't think we're seeing a generalizable population.

Dr. Sibai, would you like to comment on the patients that would be the most appropriate?

DR. SIBAI: Baha Sibai, UT Houston. There is really no doubt you have got degrees of risk and degrees of benefit, based on using this medication. Unfortunately, I as an obstetrician have to use a group of women who have a risk called prior preterm birth,

and I am using a prophylactic medication.

The number needed to treat in populations similar to what we see in Meis is about 5 to 6 in other women with prior spontaneous preterm birth. They might still have the benefit, however, the number needed to treat could be 25 or could be 50. However, considering the safety of the medication, as well as how bad it takes to have a baby born and go into a neonatal intensive care unit, it becomes extremely important for me to use all women with prior spontaneous preterm birth because at the present time, I do not have any person who responds.

To give you an example, we currently screen every woman for group B strep. At least 1 million women screened positive. We give all of these women antibiotics during labor, and only probably 100 or 200 will have group B strep. However, we don't know who is this person, so we give -- I think of this as 17P, having a baby with group B strep is catastrophic, but having a premature baby at 1 to 6 weeks is also catastrophic.

So really, we're talking about prophylaxis.

At the present time, I cannot tell you who will benefit or not. All I can tell you is there are women who will have a huge benefit, but at the end of the day, our risk factor has to be a prior spontaneous preterm birth.

DR. KROP: Dr. Miller, would you comment to -- Dr. Miller was an investigator actually in the PROLONG study.

DR. MILLER: Hugh Miller from Tucson, Arizona, maternal fetal medicine specialist who actually did participate in the PROLONG study. I accept your question. In my study site, we enrolled 22 patients; 15 of them got 17P, 7 got vehicle, and we had a 20 percent reduction.

So I think there were segments of the PROLONG population that did substantially benefit. We saw an over 20 percent reduction in preterm birth. But you do have to remember that the paradigm of treatment at the time that the PROLONG trial was being conducted was that this was the standard of care. There was no question about that among obstetricians, among maternal fetal medicine experts.

Our problem was that we didn't have an FDA-approved drug. as time advanced and with the accelerated approval in 2011, it became increasingly difficult to ask any patient to participate, both ethically for us, as Dr. Blackwell said. It became kind of unconscionable to subject patients to a 33 percent chance of not getting a drug that we all believed in. And as access improved, Medicaid patients -- again, my population represents 55 percent Medicaid. Once Medicaid had an FDA-approved drug to approve, all of my patients no longer would participate in this trial.

So I think the premise that this was a very skewed population has to be accepted, and it's why the study, in large part, was driven to another part of the world where the background risk of preterm birth is just completely different.

DR. LEWIS: Thank you. Dr. Orza?

DR. ORZA: I have two questions that go to the possibility, the feasibility of conducting an additional trial, and the first one is for Dr. Blackwell about slide CO-85 and CO-86, where you

encapsulate the statements from the SMFM and the ACOG.

Generally, the recommendations that come from clinical societies are accompanied by some indication of the strength of the recommendation and also the level of the evidence. Do you have that for either of these or whether there was any opinion in these guidelines as to what it would take for either of these societies to be in a position of equipoise and to require additional evidence?

DR. KROP: Dr. Blackwell?

DR. ORZA: First question.

DR. BLACKWELL: Hi. Sean Blackwell from UT Houston. I read the statements when they came out to the press just like everyone else. The statements, it's my impression that they are meant for interim guidance while experts and the society gain additional information. There is no strength related to the level of recommendation. There was no grade that we often use in our SMFM guidelines.

My interpretation and my understanding is that there's still a lot of work to be done to take the PROLONG results, and then combine them with other

trials, formally and statistically. and to potentially be able to take a deeper dive into looking at subgroups or other aspects.

With the PROLONG study just coming out within a week of this meeting, I think it probably takes our society some time to mull over the data, to have some vigorous debates, and to argue through it before I think our society could come up with a practice recommendation, in order to make sure we get it right and not have to go back after something is so essential that was in routine clinical practice.

DR. ORZA: My second question goes to the additional evidence and analysis that you referenced. The organization that I work for, PCORI, has funded an individual participant level data meta-analysis, which the protocol for it is published, but the results are currently undergoing peer review, and I'm not privy to those. But my question for your company is, have you contributed your data to that IPD meta-analysis?

DR. KROP: I can take that as the sponsor. We have not participated, and the reason being is that the study you're referring to was already completed by the

time we got the PROLONG data, so it was already almost under publication or in review. So we didn't; we weren't able to get that data in then.

DR. LEWIS: Thank you. Dr. Reddy?

DR. REDDY: Thank you for the clear presentations; a couple of clarifying questions. In comparing the Meis trial and the U.S. PROLONG population, it looks like the gestational age of the qualifying delivery, there's a 1 and a half week difference. Is that correct? For the U.S. PROLONG qualifying delivery, it's 32.5 it looks like, and for Meis, it's 30.6.

DR. KROP: Yes.

DR. REDDY: Okay. I just want to make sure.

DR. KROP: Yes.

DR. REDDY: There were differences. One and a half weeks at that gestational age and the risk of recurrence, that's a big difference to point out.

Then, I just wanted to ask about the trial and the sites again. There was a DSMB for the study for PROLONG?

DR. KROP: Yes, there was a DSMB. The DSMB

was charged with safety only, and they were looking at unblinded safety data, but they were not reviewing efficacy data.

DR. REDDY: So they didn't look at the rate of outcomes?

DR. KROP: No, they didn't. They add only the overall event rate in front of them. It was not unblinded. That was not the charge of the DSMB.

DR. REDDY: Okay. So until the end of the trial, there was no idea about the outcome rate.

DR. KROP: No, there was not.

DR. REDDY: Okay. And this is very basic. The vehicle was the same for both trials, right?

DR. KROP: The vehicle was exactly the same for both trials, and, yes, it was reviewed. When the approval originally of Makena was under review, there were comparability studies requested by FDA to assure that the product used in the Meis trial is similar to what we use now in the commercial product, which was used in PROLONG.

DR. REDDY: Thank you.

DR. LEWIS: Thank you. Dr. Jarugula?

DR. JARUGULA: Very nice and clear presentations from the sponsor. I just have a quick question, actually, to Dr. Sibai. I found the meta-analysis of 17P very interesting. It demonstrated 42n percent reduction with I think the analysis of five studies. I'm a clinical pharmacologist, so naturally inclined to know what is the dose used in these studies. I was wondering if you can share the doses used in these studies so we can reflect on the current dose being proposed or proposed for this 17P.

DR. KROP: I can have Dr. Sibai come up, but I would say that dose we used to select, I should say, for the PROLONG study was based on these studies, based on the LeVine, Johnson, and the Yemini study, as well as the Meis trial, all showing efficacy at the 250-milligram dose.

Dr. Sibai, do you have any additional --

DR. SIBAI: When we were designing the study, we had to rely on what's available. The 250-milligram dose was really used by several of these, and we relied on the study done by Johnson that was published in the New England Journal, which used the 250-milligram every

week.

DR. REDDY: Thank you.

DR. LEWIS: Thank you. Dr. Wade will have the last question.

DR. WADE: Thank you --

DR. WING: Thank you. In follow-up -- I'm sorry.

DR. LEWIS: I said Wade.

DR. WADE: Thank you. As a neonatologist on the committee, I'm interested in how you chose the neonatal morbidity composite index. That seems to be an unusual neonatal outcome to use. I'm just wondering about its validity and how you chose it.

DR. KROP: This was really chosen based on discussions with FDA at the time and in concert with some of the maternal fetal medicine experts as to what would be the most relevant outcomes to include. We obviously looked at a whole host of other I should say complications, as well as secondary endpoints, but those were the ones that were chosen for the composite. There's nothing validated, if that's what you're asking.

DR. LEWIS: Thank you. Dr. Wing, and then break.

DR. WING: Thank you, Dr. Lewis. This is actually a follow-up to your question. Do we know -- and I think the answer's probably no, but since the widespread use of 17P, have we actually seen a drop in the frequency of recurrent spontaneous preterm births, or are the numbers just too small to be able to track?

DR. KROP: Yes. It's too small to be able to track based on the CDC -- the statistics they put out every year on preterm birth, it wouldn't be detected. It's a too small subset.

DR. WING: And then, perhaps, does Dr. Owens know? As somebody who monitors these morbidities in her state, do you have data from Mississippi that might help us understand whether or not there's been good clinical impact?

DR. KROP: Dr. Owens?

DR. OWENS: Michelle Owens from Jackson, Mississippi. So the information or the data that I do have is, unfortunately, not available. I can see if we

might be able to get ahold of some of that data, but I can tell you that we have seen, with a concerted effort to expand within our 65 percent Medicaid-covered patient population -- to create, or eliminate, rather, all barriers to 17P. Subsequent to that initiative, we noticed an 18 percent decrease in overall preterm births within our state, and subsequent to that, received the Virginia Apgar Award from the March of Dimes as a result.

While there are clearly other things that we had also, other initiatives that were also underway during that time, it seemed very serendipitous that subsequent to increasing access for this large population of women who had historically had multiple barriers to receiving 17P, that once we were able to take that away, we saw this significant decrease that has been substantiated by our managed Medicaid plans, and that information has been made -- I know it's available publicly because it's been presented in public forums in the past. But I just don't know. We might be able to try to see if we can get ahold of that for you after the break, but I'm not sure that we'll be

able to get ahold of that information.

DR. LEWIS: Thank you. We'll now take an approximately 10-minute break. Panel members, please remember no discussion of the meeting topic during the break, amongst yourselves or with any member of the audience. We will resume at 10:40.

(Whereupon, at 10:29 a.m., a recess was taken.)

DR. LEWIS: Thank you, everyone. Let's now proceed with the FDA presentations.

FDA Presentation - Barbara Wesley

DR. WESLEY: Advisory committee members, representatives from AMAG, representatives from the FDA, and guests, I am Barbara Wesley, the primary medical reviewer for this new drug application or NDA. I am also a maternal fetal medicine health specialist, and before coming to the FDA, I had 23 years of clinical practice at urban academic medical centers and also had a little over two years as director of maternal child health in the city of Philadelphia.

This presentation will review the FDA considerations and analysis of pivotal studies 002

regarding accelerated approval, Makena, FDA actions, and postmarketing requirements. More specifically, my presentation will focus on pivotal Trial 002 supporting approval, including the findings in areas of controversy; the 2006 advisory committee meeting; the three actions taken by the FDA; and the postmarketing requirement for the confirmatory trial.

Trial 002 was funded by the National Institute of Child Health and Development and conducted by the Maternal Fetal Medicine Units Network from 1999 to 2002. The positive findings of hydroxyprogesterone caproate, or HPC, to reduce the risk of preterm birth was published in the New England Journal of Medicine in 2003. This trial is also known as the Meis trial. Then in 2006, a new drug application was submitted to the FDA for HPC 250 milligrams weekly.

The indication for HPC or Makena is to reduce the risk of preterm birth in pregnant women with a history of at least one spontaneous preterm birth. Makena is administered at a dose of 250 milligrams once a week, beginning between 16 week 0 days and 20 weeks 6 days gestation until week 37 or birth,

whichever occurs first. I would like to mention that this dose is the same dose that delalutin was approved for in 1956 for gynecologic indications.

The pivotal Trial 002 was a double-blind, placebo-controlled trial. They randomized subjects 2 to 1 to HPC or placebo. The primary efficacy endpoint was percent birth less than 37 weeks gestation. Additional endpoints requested by the FDA, after the trial's completion, and submission of the NDA, included percent birth less than 35 weeks and less than 32 weeks gestation, and a composite index of neonatal morbidities that was developed by the applicant.

The composite was based on the number of births of infants who experienced any one of the following: death, respiratory distress syndrome, bronchopulmonary dysplasia, grade 3 or 4 intraventricular hemorrhage, proven sepsis, or necrotizing enterocolitis.

As stated previously, the primary efficacy endpoint was the percent of preterm births less than 37 weeks. Of the 310 subjects treated with HPC, 37 percent delivered prematurely and 55 percent in the

placebo arm delivered prematurely. There was an 18 percent reduction in preterm births below 37 weeks. However, it is noteworthy that preterm birth rate of 55 percent in the placebo arm was considerably greater than the expected background rate of 36 percent in another Maternal Fetal Medicine Units Network study, the Home Activity Uterine Monitoring study, which was used to power this study.

Finally, I bring to your attention that the preterm birth rate of 37 percent in the HPC treatment arm was similar to the preterm birth rate of 36 percent in the placebo arm of that study. Sixty percent of the subjects in this study were black or African American. Therefore, data were broken down to black versus non-black. Although black Americans generally have a higher rate of preterm birth compared to other racial ethnic groups in the United States, there was no significant difference in the preterm birth rate by race in this trial.

In blacks, the placebo rate 52 percent. In non-blacks, the placebo rate was 59 percent. Therefore, this population with an overall placebo

preterm birth rate of 55 percent was high risk regardless of race. However, despite the high placebo rate of preterm birth, the median gestational age in the HPC arm was 37.5 weeks and 36.5 weeks in the placebo arm. Also, in both arms -- and this is not on the slide; I have other slides that we'll show this in more detail -- in both arms, the median birth weight was 2500 grams or more, so the median was not low birth weight. Therefore, most of the preterm births were late preterm births.

We were particularly interested in the preterm birth rate at gestational ages less than 35 weeks since birth at these lower gestational ages at that time were thought to be a more robust predictor of infant mortality or morbidity.

This slide lists the percentages of preterm births at selected gestational ages. Based on the adjusted 95 percent confidence interval, the upper limits of the confidence intervals with delivery at less than 32 and less than 35 weeks were close to zero, indicating the treatment effect of Makena was not much different than placebo at these gestational ages.

Also, I want to note the adjustments that were made because of interim analysis.

The ultimate goal of reducing the rate of preterm birth is to prevent neonatal and long-term morbidity and mortality associated with prematurity. The individual morbidities listed in this slide were grouped to form a composite index of morbidity. All infants with one or more of the listed morbidities were counted in the index. We have not provided p-values because these comparisons were post hoc analyses, event rates were low, and no adjustments were made for the multiple endpoints.

It should be noted that HPC did not consistently decrease the incidence of individual components of the index. Also, the most common outcome respiratory distress syndrome, which appeared to drive the difference between Makena and placebo for the composite index, is highly correlated with gestational age of delivery, and is therefore not independent of the primary outcome.

Overall, the lower percentage of infants in the HPC arm, 12 percent, compared to 17 percent in the

placebo arm, had one or more of the morbidities that comprise the composite index. However, the difference between the treatment arms was not statistically significant.

To summarize, the applicant sought approval for HPC based on findings from a single clinical trial and a surrogate endpoint less than 37 weeks gestation for infant mortality and morbidity. We were concerned that these findings may not be applicable to the general United States population. The recurrent preterm birth rate in the placebo arm was notably high, a majority of the subjects were black, and enrollment occurred from academic centers only, with one center recruiting 27 percent of the subjects, and that was the University of Alabama.

The main reason the FDA convened an advisory committee in 2006 for this application was to get their input on which gestational age at birth serves as a surrogate likely to reasonably predict infant mortality and morbidity from prematurity. Twenty-one members were present to vote, and the outcome of the vote was as follows: for preterm birth less than 37 weeks, 5

voted yes; for preterm birth less than 35 weeks, 13 voted yes; and for preterm birth less than 32 weeks, 20 voted yes.

In October 2006, the FDA determined that the NDA could not be approved. The primary deficiency was that evidence of efficacy based on a single trial that relied on a surrogate endpoint, deemed by most advisory committee members to be an inadequate surrogate, was not sufficiently robust evidence to support approval. The FDA determined that further evidence of efficacy in terms of direct benefit to the neonate or a surrogate, such as a preterm birth less than 35 weeks or less than 32 weeks, was needed.

The FDA also withheld approval in 2009 so the applicant could demonstrate they could conduct Trial 003. At this time, resulting from a publication in the Journal of Pediatrics, along with other publications, the American College of Obstetrics and Gynecology published committee opinion 404, which stated the following.

"Late preterm infants defined as infants born between 34 and 0-7ths and 36 and 6-7ths weeks are often

mistakenly believed to be as physiologically and metabolically as mature as term infants. They have higher rates of infant mortality and morbidity than term infants, and this is the largest population of preterm births."

In 2011, the applicant resubmitted the application, which upon review FDA determined that they resolved previous deficiencies. The application was approved under the accelerated approval regulations to reduce the risk of preterm birth and women with a singleton pregnancy who have a history of singleton spontaneous preterm birth.

The effectiveness of Makena was based on a persuasive improvement on the proportion of women who delivered less than 37 weeks gestation, a surrogate endpoint that FDA now deemed acceptable in light of the new data indicating higher rates of neonatal mortality and morbidity in late preterm births.

Trial 003 three was ongoing, and the applicant demonstrated that it could successfully be completed. As a condition of accelerated approval, the applicant was required to complete the confirmatory clinical

trial of HPC Trial 003 to verify the clinical benefit to neonates from the reduction in the risks of preterm birth.

I have now presented the complicated regulatory history of FDA's review, which culminated in 2011 in accelerated approval of Makena based on Trial 002. I will now turn our presentation over to my statistical colleague, Dr. Jia Guo, to discuss results from the confirmatory trial.

FDA Presentation - Jia Guo

DR. GUO: Good morning everyone. My name is Jia Guo. I'm the statistical reviewer from the Office of Biostatistics at CDER FDA. I'm going to present the efficacy results for Makena in confirmatory Trial 003. In my presentation, first I will provide an overview of Trial 003, including trial design, subject disposition, demographics, baseline characteristics, and efficacy results, followed by FDA's exploratory analysis and concluding remarks.

As you already heard from the applicant's presentation, Trial 003 was a multicenter, randomized, double-blind, placebo-controlled trial. Subjects were

randomized to Makena or placebo with a 2 to 1 ratio. The randomization was stratified by study site and gestational age. The trial design and eligibility criteria were very similar to Trial 002.

Trial 003 enrolled women who are at least 18 years old with a singleton pregnancy, and the gestational age was between 16 to 20 weeks with a history of singleton spontaneous preterm birth. Subjects who had a significant medical disorder, or had multifetal gestation, or with no major fetal anomaly or fetal demise were excluded.

Based on Trial 002 efficacy results, Trial 003 was adequately powered to detect a 35 percent reduction, from 17 percent to 11 percent, in the percentage of neonates with at least one neonatal composite index event and a 30 percent reduction, from 30 percent to 21 percent in the percentage of preterm births prior to 35 weeks.

Approximately 1700 subjects were randomized to receive either Makena or placebo. Almost all subjects completed the study, and 93 percent of subjects completed treatment. The intent-to-treat population

included all randomized subjects, and it was used for evaluation of preterm birth endpoints.

The liveborn neonatal population included all neonates of subjects in ITT population who were liveborn and have available morbidity and mortality data. There was a minor discrepancy on the sample size of liveborn population between the applicant's and FDA's analysis due to the mortality and the morbidity data change on 3 neonates. This discrepancy does not impact any conclusions in my presentation.

The Makena and the placebo groups were comparable across demographics and baseline characteristics. Overall, 88 percent of randomized subjects were white, 7 percent were self-identified black, and 5 percent of other races. Approximately 10 percent of randomized patients were single or without a partner.

Nine percent of subjects used substances, including alcohol, tobacco, and illicit drugs during pregnancy, and 15 percent of subjects had more than one previous spontaneous preterm birth; 391 subjects were enrolled from the U.S., which were about 23 percent of

the overall study population. Please note the size of the U.S. subpopulation in Trial 003 was not substantially less than the size of Trial 002, which had 463 subjects.

Trial 003 was designed to demonstrate efficacy on co-primary endpoints, the surrogate endpoint preterm birth prior to 35 weeks and the clinical endpoint neonatal composite morbidity and mortality index, which is a yes/no variable defined as yes if the liveborn neonate had any of the events listed on the slide.

There are two secondary efficacy endpoints. Preterm births prior to 32 weeks and prior to 37 weeks were of clinical interest. This table summarizes the analysis results for the co-primary and the secondary efficacy endpoints. The percentage of neonates who had at least one neonatal composite index event and the percentage of preterm births prior to 35 weeks were much lower than expected. The neonatal composite index was scored as yes in 5.4 percent and the 5.2 percent in liveborn neonates in Makena and the placebo groups, respectively, with a difference of 0.2 percent.

The percent of preterm births prior to 35

weeks was 11 percent and 11.5 percent in Makena and placebo groups, with an estimated treatment difference of minus 0.6 percent. The p-values for testing the difference between Makena and placebo were much greater than 0.05, meaning treatment differences were not statistically significant, and the estimated differences between treatment groups were close to zero for both co-primary endpoints. With respect to the two secondary endpoints of preterm births prior to 32 weeks and prior to 37 weeks, no Makena benefit was noted either.

The applicant conducted post hoc analysis to understand the lack of correlation between efficacy results observed in Trial 002 and Trial 003. Generally, FDA does not support subgroup analysis for inference of efficacy when the primary analysis result does not demonstrate efficacy. There are multiple reasons to not consider subgroup analysis to support establishing efficacy when treatment benefit in the overall population is not significant.

The major statistical reason is the inflation of type 1 error probability. That is the heightened

probability of incorrectly concluding treatment benefit. When such subgroup analyses are used to search for evidence of a benefit, there is the high probability that any observed favorable subgroup results are due to chance alone. Therefore, FDA considers such analysis for hypothesis-generating purpose only, generally.

Nevertheless, FDA reviewed the applicant's post hoc analysis results to explore whether differences in key design aspects of Trial 002 and Trial 003 might clarify the divergent efficacy results. FDA compared the two trials with respect to demographics, baseline characteristics, and the responses in the placebo groups, then conducted subgroup analysis.

Trial 002 and 003 were nearly identical in design. However, when comparing the demographics and the baseline characteristics, notable differences exist between the two trials with respect to five factors, including black race; history of more than one previous spontaneous preterm birth; single or without a partner; substance use during pregnancy; and less or equal

12 years of formal education.

This bar graph shows the percentage of each factor in Trial 002, Trial 003, and the U.S. subgroup in Trial 003, which are presented by the gray, blue, and orange bars. Compared to Trial 002, Trial 003 had a lower percentage of black subjects, as well as subjects who had more than one previous spontaneous preterm birth, who are single or without a partner, or who used substances during pregnancy, and also had a lower percentage of subjects who had lower education levels. The U.S. subgroup of Trial 003 falls in between Trial 002 and Trial 003.

Comparing the placebo group in the two trials, the percentage of neonates who had at least one neonatal composite index event and the percentage of preterm birth prior to 35 weeks were higher in 002 and lower in 003, with the percentage in U.S. subgroup of Trial 003 falling in between.

In the applicant's briefing document, the overall baseline risk of preterm birth was assessed across the two trials using a post hoc composite risk profile constructed by the applicant. The components

of this composite risk of five selected baseline factors was presented on an earlier slide, and show, again, here. Please note, black race and a number of previous preterm births are associated with higher rates of preterm births, but the other factors have not been consistently associated with an elevated risk of preterm births.

This bar graph demonstrates the percentage of subjects who had at least one of these factors. Trial 002 had the highest percentage, Trial 003 had the lowest percentage, and the U.S. subgroup of Trial 003 was in between. Based on all the comparisons between Trial 002 and Trial 003, the overall study population of Trial 003 appeared to be at a lower risk of preterm birth and neonatal events compared to Trial 002, and the risk of U.S. subgroup of Trial 003 falls in between.

FDA conducted subgroup analysis by region, race, and history of spontaneous preterm birth. For each of this subgroup analysis, the difference between Makena and the placebo groups was computed using two methodologies, a stratified Cochrane-Mantel-Haenszel

method and shrinkage estimation through Bayesian modeling.

The subgroup analysis using CMH method evaluates a particular subgroup category independently from other subgroup categories, and it relies only on the data from that category. The Bayesian shrinkage estimation analysis evaluates all subgroup categories jointly and borrows information across subgroup categories to reduce the variability of the estimates and to prevent random highs and random lows. Conclusions from these two subgroup analyses was similar, but we present results from both methods for completeness on the following slides.

Another analysis was conducted by the composite risk profile at baseline. This slide shows the subgroup analysis results by region for co-primary endpoints. The region was defined as U.S. and non-U.S. The upper part of the display is for the neonatal endpoint. The lower part is for the preterm birth prior to 35 weeks. The numbers in the parentheses after each region are the sample size of Makena and placebo groups in that region.

The second and third columns are for the percentage of subjects who had an event of each co-primary endpoint by treatment group, followed by the estimated percentage difference between Makena and the placebo using stratified CMH method and shrinkage estimation in the fourth and the fifth columns, respectively.

On the right is the plot of the point estimates with corresponding 95 percent confidence intervals. The X-axis is for the difference between Makena versus placebo. The middle vertical line is the reference line indicating no difference between treatment groups. The left side of the vertical line is favoring the Makena group and the right side is favoring placebo. The blue lines are for the overall population results. The green lines are for the subgroup results estimated using stratified CMH method, and the red lines are for the subgroup analysis results using shrinkage estimation.

As you can see, the confidence intervals for the treatment difference for both co-primary endpoints, in both the overall population and in the regional

subgroups, include zero, indicating no evidence of Makena benefit versus placebo, based on both analysis methods. Furthermore, all estimated differences between treatment groups are small and close to zero, with some estimates favoring Makena and others favoring placebo, and with the magnitude of the differences slightly smaller based on the shrinkage estimation method. In addition, there was no treatment by region interaction for each co-primary endpoint.

In summary, the Trial 003 subgroup analysis did not show Makena had a favorable treatment effect compared to placebo for either co-primary endpoint in either the U.S. or non-U.S. region, and the results do not provide support for regional differences, explaining the differences in results between Trial 002 and 003.

This slide shows the subgroup analysis results by region for the two secondary endpoints. Similarly, no evidence of a treatment effect was seen for the endpoints of delivery prior to 32 weeks or prior to 37 weeks in either the U.S. or non-U.S. region.

This slide shows the results by race, black

versus non-black. The estimates of the difference are close to zero with all confidence intervals including zero. This race subgroup analysis did not provide evidence that Makena had a treatment effect on either co-primary efficacy endpoint in the black or non=-black subgroups. Similarly, no evidence of treatment effect was seen for preterm birth prior to 32 weeks and prior to 37 weeks within race subgroups.

This slide presents the subgroup analysis results by the history of spontaneous preterm birth, which was categorized as had one or had more than one previous preterm births. This subgroup analysis did not provide evidence that Makena had a treatment effect on either co-primary efficacy endpoint in either subgroups.

This subgroup analysis did not provide evidence that Makena had a treatment effect on either of the secondary efficacy endpoints in either subgroups, defined based on history of spontaneous preterm births. We also conducted additional subgroup analysis by substance use during pregnancy, marital status, and education level.

The results show no evidence of a treatment effect for Makena versus placebo on all the four efficacy endpoints in this subgroup as well. In summary, Trial 003 does not provide any evidence that Makena had treatment benefit in a particular subgroup, based on the five factors that differentiate the study populations in the two trials.

We performed another analysis based on the applicant's post hoc composite risk profile as mentioned in a prior slide. Three groups were defined. The first group includes subjects who did not have any of the factors included in the composite; the second group includes the subjects who had at least one factor; and the third group includes subjects who had add these two factors.

The bar graph on the left is for the neonatal composite endpoint. The height of the bar represents the percentage of neonates in each treatment group for that race group. The difference between the blue bar and orange bar represents the treatment effect of Makena versus placebo for the neonatal composite endpoint in that risk group.

As we see from the bar graph, when the overall risk increases on the X-axis, the percentage of the neonates who had at least one neonatal composite index event in that treatment group, increases as well. However, the treatment effect of Makena versus placebo on this endpoint did not improve. In the group of subjects who had at least two factors, placebo was favored instead.

Similar results were seen for the preterm birth prior to 35 weeks, shown in a bar graph on the right. This analysis does not support the applicant's point that, overall, the lower risk of preterm birth or neonatal events in Trial 003 explains the lack of efficacy in Trial 003, given that no suggestion of efficacy was seen even in the groups with higher risk levels.

In summary, Makena did not demonstrate a statistically significant treatment effect versus placebo on the co-primary efficacy endpoints of gestational age at delivery and the neonatal composite index in Trial 003, and estimated differences versus placebo were close to zero. Furthermore, exploratory

analysis did not show evidence that Makena has treatment benefit within any specific subgroup in Trial 002.

Although the selected risk factors may have an impact on the overall percentage of subjects who had preterm births or neonatal composite events, there's no evidence in Trial 003 that these factors may impact the treatment effect.

This concludes my presentation. Next, my colleague Dr. Huei-Ting Tsai, will present drug utilization in the U.S..

FDA Presentation - Huei-Ting Tsai

DR. TSAI: Good morning. I'm Huei-Ting Tsai. I'm an epidemiologist at the Office of Surveillance and Epidemiology. The objective of my presentation is to provide an overview of hydroxyprogesterone caproate use in the U.S. to evaluate its public health impact. I will refer to hydroxyprogesterone caproate as HPC throughout my talk.

My presentation includes the result from two separate analyses. In each analysis, we estimated a number of patients with injectable HPC use and the

possible reason for the use. The first analysis estimated utilization of injectable HPC in U.S. outpatient setting. This analysis provides national estimates of HPC use among pregnant and non-pregnant patients using proprietary database available to FDA.

The second analysis evaluated injectable HPC use during the second or third trimester in pregnancies with live births, using a distributed Sentinel database. We conducted this analysis in Sentinel distributed database because it gives us information specific to these two trimesters of pregnancy, whereas the result of the first analysis does not.

I will first present the results of our analysis, the estimated injectable HPC use in U.S. the outpatient setting. This figure shows the estimated number of 15- to 44-year-old patients, regardless of pregnancy status, with a dispensed prescription of injectable HPC from U.S. outpatient pharmacies.

Our results show an estimated 8,000 patients received a dispensed prescription for injectable HPC in 2014, and then increasing to 42,000 in 2018. Of note, these results do not include bulk powder forms of HPC

typically used for compounding in pharmacy or clinics.

We also obtained diagnosis associated with injectable HPC use in 15- to 44-year-old women, using a database that captured monthly surveys from a sample of 3200 office-based physicians reporting on patient activity during one day a month. This dataset provides prescriber intended reason for drug use and our national estimates.

For HPC, an estimated of 50 percent of the reported diagnosis was for supervision for high risk of pregnancy of which 78 percent was specifically for supervision of pregnancy with a history of preterm labor. Of note, this diagnosis data do not provide information about history of preterm delivery, specifically; only a history of preterm labor.

Because progesterone has also been used for preventing preterm births, we also look at the possible reason for progesterone use. The data has showed that 14 percent of the reported diagnosis call for supervision of high risk of pregnancy, while female infertility was the most common diagnosis related to progesterone use.

The analyses have some limitations, but the estimated number of patients using injectable HPC came from retail and mail-order pharmacy setting and did not include estimates from hospital or clinical settings where this product may also have been used. We obtained diagnosis related to HPC from an office-based physicians survey. The survey data do not necessarily result in dispensed prescriptions.

In summary, while outpatient injectable HPC use increased over the extended time frame of 2014 to 2018, utilization of HPC was low. Further, the use of injectable HPC was largely associated with a diagnosis or history of preterm labor.

For the next action, I will present the results of our analysis, focusing on utilization of HPC during the second or third trimester of pregnancy only. We conducted this analysis using the FDA Sentinel distributed database. The Sentinel distributed database contains administrative claim data for most of the commercially insured patients. We included pregnancy with live births delivered during January 2008 through April 2019. We evaluated all product

forms of HPC and progesterone.

To understand possible reasons for injectable HPC use, we searched for the presence of three related obstetrical conditions to HPC use. The narrow definition includes any of the three conditions here: a preterm delivery but only in a prior pregnancy; a preterm labor but only in a current pregnancy; or cervical shortening only in a current pregnancy. In contrast, the broad definition includes the same three conditions as a narrow definition, but each condition was not restricted to either prior or current pregnancy.

We identify a total of 3.4 million live birth pregnancies in the Sentinel distributed database. This figures shows the number of pregnancies using HPC or any progesterone during the second or third trimester per thousand pregnancies over the time frame of 2008 to 2018.

The red line demonstrate that in 2018, injectable HPC was used in about 13 per 1,000 pregnancies. The number of pregnancies using injectable HPC increased over the study time frame,

although the use was low compared to the total number of pregnancies. The blue line represents the use of either HPC or progesterone during their second or third trimester, approximately 25 per 1,000 pregnancies, or less than 3 percent of live birth pregnancies in the Sentinel database.

This table shows the majority of pregnancies using injectable HPC had a related obstetrical condition. This data on the left column are our narrow and broad definition of a related or obstetrical condition. The next column over shows of pregnancies using injectable HPC, 73 percent and 98 percent had at least one related obstetrical condition by narrow and broad definitions, respectively.

This analysis has the following limitations. First, it's conducted among live birth pregnancies in the Sentinel distributed database, so it does not project nationwide use and may not be generalizable to women without a commercial insurance plan. Second, we did not examine the timing of a related obstetrical condition relative to injectable HPC use, so the presence of a related obstetrical condition may not

necessarily be the indication for injectable HPC use. Lastly, our data did not capture medications that are out of pocket, which may underestimate the use of injectable HPC.

In summary, we found modest use of injectable HPC during the second or third trimester of live birth pregnancies and a high percentage of pregnancies using injectable HPC during their second or third trimester, and at least one related obstetrical diagnosis recorded before or during the pregnancy.

Now, I would like to turn my presentation to my colleague, Dr. Christina Chang, to give a summary presentation from FDA's perspective. Thank you.

FDA Presentation - Christina Chang

DR. CHANG: Good morning. My name is Christina Chang, and, again, I am a clinical team leader in the Division of Bone, Reproductive, and Neurologic Products, and I will be giving the summary remarks on behalf of the FDA review team. Because both the applicant and my FDA colleagues have already presented quite a bit of information, I will stay with the key concepts that we think will be the most germane

to the panel's deliberation.

As a reminder of why the topic of today's meeting is of tremendous importance, we know that neonatal mortality and morbidity from preterm birth remains a significant public health concern. Preterm birth, defined as the delivery prior to 37 weeks of gestation, currently affects approximately 10 percent of all births in the United States.

To date, we do not have any drug products specifically approved by the FDA to reduce neonatal mortality and morbidity due to prematurity, and in clinical practice, progestogen, whether in synthetic forms or natural progesterone, have been used to reduce the risk of preterm birth. For women with a singleton pregnancy and who already have a prior spontaneous preterm delivery, current professional practice guidelines recommend starting progesterone treatment in the second trimester of pregnancy to reduce the risk of return preterm birth.

At this time, Makena is the only pharmacotherapy approved to reduce the risk of recurrent preterm birth. Based on its accelerated

approval, Makena's indication states that it is approved to reduce the risk of preterm birth in women with a singleton pregnancy who have a history of singleton, spontaneous preterm birth.

The data that supported the accelerated approval for Makena came primarily from a single clinical trial sponsored by the NICHD, Trial 002, which the applicant and FDA already reviewed in depth. As you recall, delivery at less than 37 weeks gestation was evaluated as the primary efficacy endpoint in Trial 002.

Now, moving on to Trial 003, I'll point out that in this confirmatory trial, two efficacy measures were assessed. One was the clinical endpoint, namely the neonatal outcomes, and the other a surrogate endpoint, which is delivery at less than 35 weeks gestation. Delivery at 35 weeks gestation was chosen as a co-primary efficacy measure because this trial was initiated in 2009, two years before the agency came to the conclusion that late preterm birth was also consequential in terms of neonatal outcome.

The second point I want to call your attention

to is the temporal distance between Trial 002 and Trial 003, with Trial 003 finishing 16 years after Trial 002 had been completed, and this illustrates the challenges in conducting large clinical trials in obstetrics, possibly because obstetrical practitioners tend not to deviate from existing clinical guidelines.

As you have already seen, Trial 003 was more than three times larger in size than Trial 002, with a U.S. subset in 003 almost approaching the entire 002 sample size. Makena did not differ from placebo for either the clinical endpoint of neonatal outcome or the surrogate endpoint by gestational age at delivery at 35 weeks. No difference between Makena and placebo was discernible for delivery at 32 weeks or 37 weeks gestational age.

In addition to the trial failing to meet its primary objectives, in no subgroup analyses that we conducted did we observe any difference between Makena and placebo, and those subgroups included race, previous number of spontaneous preterm births, and region U.S. versus non-U.S., as already discussed.

These findings bring us to the concept of what

constitutes a standard for regulatory approval. According to the regulations, all drugs, including those approved under the accelerated approval pathway, must demonstrate substantial evidence of effectiveness, and the regulations refer to evidence consisting of adequate and well-controlled investigations, including clinical investigations.

You'll notice that I highlighted here in red the phrase, "adequate and well-controlled investigations" with the word "investigations" in plural, because the agency has generally interpreted the regulation as referring to more than one clinical study being used to support approval, and here in the case of Makena, we now have two adequate and well-controlled clinical investigations.

There is Trial 002, showing convincingly, based on a surrogate endpoint, that Makena reduced the proportion of preterm birth before 37 weeks. But now we also have a much larger trial, 003, that evaluated not only a surrogate endpoint but a clinical outcome as well.

In Trial 003, the size of the U.S. subgroup,

which was 391, is almost as large as the entire cohort of Trial 002, which was 460. This larger trial, 003, also convincing, showed that Makena conferred no treatment benefit whatsoever. Importantly in Trial 003, Makena had no treatment effect based on the surrogate endpoint of delivery in less than 37 weeks gestation, the same endpoint that was positive in Trial 002.

Here's a schematic of the two regulatory pathways to obtain FDA's approval for a drug. On the left is the accelerated approval pathway, where the agency grants accelerated approval based on a surrogate endpoint that we believe reasonably likely to predict a clinical benefit.

The advantage of the accelerated approval pathway lies in providing patients earlier access to promising therapy without waiting for a large preapproval confirmatory trial. However, at the time of the accelerated approval, when the decision is granted, there's less certainty in being able to translate the observed treatment effect into clinical benefit. And because of the uncertainty, a

post-approval, confirmatory trial is required to verify the clinical benefit.

Contrast that to the traditional approval pathway on the right. Typically, we rely on a clinical endpoint that directly measures how a patient in question, in our case, the neonate, feels, functions, or survives. Alternatively, if the surrogate endpoint has been validated to actually predict clinical benefit, the surrogate endpoint can be used to support the traditional approval.

What could explain the conflicting results from these two adequate and well-controlled trials? At the minimum, we envision these three scenarios. In the first scenario, Trial 002 was falsely positive, and in the second scenario, Trial 003 was falsely negative. In the third scenario, the discrepancy is attributable to differences that we haven't explained; and if the panel has other hypotheses, we would be interested to hear them as well.

So having discussed the results from both trials and the possible reasons for conflicting findings, we're asking the panel to weigh in on the

questions of the day. With Makena, has substantial evidence of effectiveness been established?

As Dr. Nguyen showed this morning, we would like to hear the panel opine on two issues of concern. The first issue relates to the conflicting results, based on the surrogate endpoint, the gestational age at delivery. In Trial 002, less than 37 weeks gestation at delivery produced a positive result, but in Trial 003, the same surrogate endpoint produced a negative result, as did the less than 35 weeks delivery surrogate endpoint.

If the treatment effect, based on the surrogate endpoint of gestational age of delivery, is not substantiated, do we have substantial evidence of effectiveness to support approval? Furthermore, there is issue of concern number two; namely, the clinical benefit has not been verified. Here we have Trial 003 that did not show any improvement in neonatal outcome. Again, given this concern, can we conclude that there is substantial evidence of effectiveness to support approval?

With that, I'll conclude my presentation and

bring the FDA's overall presentations to a close. The FDA team stands ready to respond to any questions the panel might have, and we look forward to a productive discussion.

Clarifying Questions to FDA

DR. LEWIS: Thank you. We'll now take clarifying questions for the FDA. If possible, please indicate the person to whom your question is directed, and if possible, the slide number from the FDA. Please remember to state your name for the record before you speak. I'm going to start actually with Dr. Gillen.

DR. GILLEN: Thank you. This is a question pointed at Dr. Guo, and thank you for presenting the subgroup analyses. That would have saved me the long, labored question that I asked previously of the sponsor, which I think should have been presented there.

Just in completeness, I guess, I agree completely and wholeheartedly with the FDA's position on subgroup analyses, but I think what we're looking for here is the elimination of some of these pathways. I agree with you it's either a false positive, a false

negative, or it's some change in the distribution between the two subpopulations where we have effect modification.

So I guess in completeness of that, I know that you looked at the baseline risk factor sub analyses, but another way, possibly a more sophisticated and maybe slightly more efficient way to do that, is to, for lack of a better term, develop a propensity score for being in one study or the other, and then match or adjust on that propensity score.

Was that done? And if that was done, did it produce any similarities between the first trial and the PROLONG study?

DR. GUO: This is Jia Guo, statistician from FDA. We didn't do that propensity score analysis. We came up with this analysis using the composite risk profile, which was constructed by the applicant. So basically, we look at how many risk factors they have, kind of like generally define the risk groups, like no risk, and at least have one factor or two factors. I also look three factors, at least three factors. But of the subgroups, the size is too small, but the trend

is still the same. You don't see the benefit even with the risk increases.

DR. GILLEN: I understand that the subgroups become small as you do that. That's exactly why I'm asking about, somewhat, the weighted average, if you will, of all the composites as you go through for the propensity.

So the answer is we haven't looked at that, but as we've broken down the baseline risk factors, we don't see anything that would bring the two studies closer together in terms of the effect that was observed.

DR. GILLEN: Right, yes.

DR. LEWIS: Thank you. Dr. Orza?

DR. ORZA: My question is for the FDA clinical reviewers about study 003, in terms of study 003 was 10 to 20 years later than 002. And what we wind up with is lower than expected rates of premature birth in both groups.

Could that be due to the fact that these women were being seen every week, of which seems, even in a high-risk pregnancy, is unusual. So there were all

kinds of other aspects to their care. Could that be a factor for driving down both the premature birth and the negative outcomes in the babies?

DR. NGUYEN: Hi. Christine Nguyen, FDA. That's an excellent question. I would point out that the more intensive care usually occurs in all clinical trials, including 002 and 003. So I don't believe that there was, perhaps, a differential in the attention to the subject trials in 003 compared to 002.

DR. ORZA: There wouldn't be in terms of the attention paid, but 10 and 20 years later, do we know more or do we do different things in those encounters that could explain part of the difference between 002 and 003?

DR. NGUYEN: Christine Nguyen again. Again, this is why we have a prespecified protocol, and we did our best to keep the design and hopefully the conduct of those trials very similar, so that we can really try to isolate the effect of the drug itself and neutralize other factors, so to speak.

DR. WESLEY: This is Dr. Wesley. I'd like to just add that whatever changes occurred over time would

be equally distributed between the control group and the intervention group, so that would not be any different between those two arms.

DR. ORZA: Is there any way to test for that?

DR. WESLEY: Well, the purpose of a randomized-controlled trial is to eliminate those factors.

DR. ORZA: Right. I understand that, but if something in the randomization failed or the misclassification across groups was differential, that would affect it even if there was randomization.

DR. CHANG: Christy Chang, FDA. Could I also add that when 002 was being conducted, the participating centers were from the MFMU Network, and these are tertiary academic centers. So patients were receiving the highest level of intense monitoring they possibly could have.

DR. LEWIS: Thank you --

DR. NGUYEN: To answer -- I'm sorry. I don't think we answered your question. Christine Nguyen again. So that's why we look at the demographics and baseline factors between the two treatment arms, and

they were balanced, in actually both 002 and 003.

DR. ORZA: But not the factors of the clinicians or the centers, just of the patients. Is that correct?

DR. NGUYEN: Well, the centers that are invited and accepted to participate in the trial have to pass certain criteria, and they do have to follow the same protocol.

DR. GUO: This is Jia Guo, statistician. I just want to add one point, that in Trial 003, the randomization was stratified by site. I think any influence from the site could be evened out.

DR. LEWIS: Thank you. Dr. Bauer, and then Dr. Davis.

DR. BAUER: I have two quick questions, and I think the first one goes to Dr. Guo as well. That is that your analyses all used absolute risk, which is a perfectly valid measure of association, but it does make it a little bit difficult to compare that with what the investigators thought that they were going to get before the study, and that is their power calculation.

I'm just wondering if you verified the relative risk estimates that they have presented to us today, specifically the hazard ratio of 0.95 for the PTB less than 35 risk with a confidence interval of 0.71 to 1.26. The reason that I point that out is that the sponsor plans to exclude at least a 30 percent reduction in that outcome; therefore, the number of events really can't be used as an explanation for the fact that they didn't get positive results. In fact, they got the results that they estimated they would get based on their power sample.

So did you actually confirm those relative risk reductions?

DR. GUO: I didn't do the analysis, but we confirmed the data. The dataset we used is the same.

DR. BAUER: Okay.

DR. GUO: So the reason why --

DR. BAUER: There's no reason to think it would be wrong.

DR. GUO: -- yes.

DR. BAUER: Okay.

DR. GUO: The reason why we use absolute risk

reduction is because when you talk about relative risk reduction, it is relative to the placebo background rate. But the two trials have very different background rates. So when you do the comparison across the two trials using relative risk reduction, even though they may have the same relative risk reduction -- just assume -- it means very different for the absolute risk reduction, which tells you the percentage of patients that actually can benefit.

DR. BAUER: I understand. That definitely impacts the public health. And I'm just wondering if someone at FDA could actually comment on the meta-analysis that was discussed in the sponsor's slide CO-27, with a point estimate of 0.58 and confidence intervals that went from 0.38 to 0.9.

Did FDA look at that meta-analysis, and was that part of the data that was reviewed in terms of what's the prior probability of one of the trials being wrong, either 002 or 003?

DR. NGUYEN: Hi. Christine Nguyen again. We did not formally analyze this meta-analysis, and it was used as a concept for Trial 002. Given that we have

two adequately designed and powered studies, we wouldn't typically rely on something of lesser evidence, or let's say lesser strength of evidence such as a meta-analysis, particularly when you're looking at studies that were done in the '60s and '70s with very small sample sizes.

So I do not think that this meta-analysis would influence the way we interpret the evidence that we have today.

DR. WESLEY: One other comment. Dr. Wesley. Some of the indications for treating were very different in those studies. Some of them had cerclage and some of them had ruptured membranes. There were different scenarios and clinical scenarios, whereas these two trials were pretty much exactly alike.

DR. CHANG: Christy Chang from FDA. If I could also add to that, the CO-27, some of the studies were done evaluating preterm labor, not necessarily preterm birth, reduction risk.

DR. LEWIS: Dr. Davis, and then Dr. Reddy?

DR. DAVIS: Jon Davis from Tufts. Thank you for your presentations. I guess my question is, does

it really have to be that one is a false negative and one is a false positive? I think you have two well-designed, well-controlled, well-conducted clinical trials done 15 to 20 years apart, in different populations, in different countries, with different outcomes, and the data are what the data are.

Preterm birth has clearly been a holy grail that we've all worked for most of our careers to try to see if we can figure out. And maybe we don't understand exactly why the trials are different, and we can't demonstrate it statistically, but I suggest that they are.

You're probably aware there was a large, randomized, multinational trial of antenatal steroids done recently, and underdeveloped countries finding that the steroids not only didn't help neonatal morbidity and mortality, but made it worse. So we're not going to stop using antenatal steroids because it was a different trial and doesn't necessarily pertain to this.

I'm just curious how you're looking at that. In other words, since the second trial, 003, is more

recent, does that mean that it's more impactful? Should we be weighting these two trials differently? What are some of your thoughts about that?

DR. CHANG: Christy Chang, FDA. I'll turn the table back to you. That's what we want to hear from the panel.

DR. LEWIS: Thank you. Dr. Reddy, and then Dr. Smith.

DR. REDDY: I am trying to grapple with this data, having just delivered a 25-weeker on labor and delivery when I came on. This is really difficult, I agree. Both trials were well done, so what do we do with this data?

I wanted to go back to the gestational age of the qualifying pregnancy. I'd be very interested in understanding, between the Makena and the placebo group, the difference in additional days and weeks gained in pregnancy, because the MFMU did do a study of the Meis trial, and they showed 34 weeks and beyond, that those women who had an index pregnancy or qualifying pregnancy 34 weeks and beyond gained less time and the benefits were for women who are earlier

than 34 weeks.

So I'd like to see this data focusing on the PROLONG U.S. population, not the non-U.S. population, because as you showed, it's closer to the Meis trial population, the PROLONG U.S. population, except, like I mentioned before, there's a 1 and a half week difference in the qualifying pregnancy, and it's like around 32 weeks. For the Meis trial, it was 30.6, and the PROLONG U.S. trial was 32.5. That difference in morbidity at that gestational age, what we can hear from our neonatal colleagues is huge.

So I'd like to understand the days gained. I'm not a biostatistician, but how could we understand that between Makena and placebo in the PROLONG U.S. population, specifically?

Then another question I guess I have to ask is the primary outcome, preterm birth less than 35 weeks, in the PROLONG U.S. population, it looks like there is 11 percent difference. It's 15.6 versus 17.6 in the placebo group, so that's a 2 percent difference favoring Makena. So that's about an 11 percent difference. What would the sample size have to be to

demonstrate that difference? It's massive, but I'm just curious.

Then the last question is, did anyone ever talk about the UK and progesterone use? My impression is they don't use 17-OHPC; they use vaginal progesterone if they use anything.

Sorry, I kind of --

DR. NGUYEN: That's okay. Christine Nguyen again. Well, I can answer the UK question. We have not looked into the practice guidelines that the UK, number one, but there were not that many subjects enrolled from the UK, or if any, I'm not sure. As far as Trial 003, that certainly wouldn't affect the findings that we saw.

As far as looking at days prolongation in the U.S. subgroup, I have to ask my stats colleagues to see if we had done an analysis on that particular question.

DR. GUO: In addition to the five factors, the subgroups we presented here, I think also the applicant part, and we both looked at numerous other factors, including the gestational age at the qualifying delivery, and we couldn't find anything really

convincing that Makena showed efficacy results in that specific subgroup related with the gestational age at the qualifying delivery.

Back to the U.S. versus the non-U.S. question, you see that 2 percent difference, but the thing is that is a point estimate. You cannot rule out that is different from zero, so that's the problem.

DR. REDDY: No, I was asking what would the sample size be needed to do that?

DR. GUO: Another question is, to other experts here, if you plan another study, that 2 percent is what you want to expect to see in that trial. So that's back to the power issue. When people are saying the study is underpowered, you need to know is underpowered for what; what's the hypothesis?

Trial 003 is preplanned to see that 30 percent reduction, the relative risk, translate to 6 percent absolute difference on neonatal, but the study is not underpowered to detect that difference, but you are not really powering your study to detect your observed results.

DR. REDDY: Yes. I was focused just on the

U.S. PROLONG patients and their outcome of 35 weeks.

DR. NGUYEN: Right. This is Christine. I think it's fair to say that to adequately power a study, to look at a 2 percent difference, we would need to know a few factors, what's the baseline preterm rate, and that would drive some of it. But certainly, assuming everything being equal and based on the findings we saw from 003, it would require a very large trial. And I won't put a number on it, but I can tell you it's going to be huge.

DR. REDDY: Right. So then, back to the other question, you said you looked at the age of the qualifying delivery. You said there was no significant difference, depending upon the gestational age of the qualifying delivery. So did you just look at the cutoffs, 35, 32, 37, or did you do it looking at time of prolongation?

DR. GUO: Jia Guo from FDA again. You can refer to the two tables in the FDA briefing document, in the appendix. We presented all the subgroup analysis results that we have looked at. From there, we look at the gestational age of qualifying delivery

with 20 to 28 weeks, 28 to 32, 32 to 37, and 35 to 37. We couldn't find any convincing evidence.

Also, it's hard because we did a lot of post hoc subgroup analysis here, so it's really hard to -- sometimes you see -- just like I present on the slide, some evidence you see may be due to chance only because we have a really high probability of the type 1 error because there's no multiplicity control here. So even if you see some difference, that may be because it's just randomly -- it's just due to chance.

We are kind of looking for convincing, consistent evidence across the two trials and also across the two efficacy endpoints, together. We don't find any convincing evidence for the subgroup defined, based on the gestational age of qualifying delivery.

DR. LEWIS: Okay. One other person from the FDA; please state your name.

DR. BAER: This is Gerri Baer. I'm a neonatologist at the FDA, and I appreciate your question, and my mic just got cut. I'll address the endpoint question that you had about the date and the potential benefit in prolonged pregnancy by days, or

even a week.

One of the biggest challenges that we have struggled with internally is how to best measure this. If you prolong a pregnancy, as you know, at 24 weeks by a number of days, that might be a clinical benefit, but if you prolong that pregnancy at 34 weeks by a number of days, there might be a benefit, but it's a much smaller benefit.

So if we could look and say that prolonging pregnancy by 5 days, it was effective and that was a true effect, that would be fantastic, but it's not a straight forward endpoint, and we continue to deliberate on how to look at gestational age because of that.

DR. LEWIS: Thank you. Dr. Smith?

DR. SMITH: Brian Smith. My question is for Dr. Chang. I think just to clarify your last couple of slides, after accelerated approval of a molecule, is the ultimate goal of the confirmatory trial, where you say verification of clinical benefit, to show benefit for the surrogate endpoint, preterm birth, for which the molecule has the indication, or the clinical

endpoint neonatal morbidity?

DR. CHANG: I'm sorry. Could we pull up the last couple of slides from my presentation? I think it would be 12 and 13. Would it help if I go over the processes again?

Here again, I think Dr. Nguyen also mentioned this morning that we're grappling with two issues of concern here. The first issue is that from 002 and 003, we have different results based on gestational age at delivery, based on the surrogate endpoint alone. So now having reviewed these two clinical investigations, do we have enough to support substantial evidence for effectiveness, given the conflicting endpoint findings?

Next slide, slide 13. Now, with issue number two, clinical benefit was only measured in 003 and not in 002. So our question to you is, has the clinical benefit been verified as required by law?

DR. LEWIS: Dr. Shaw, final question.

DR. SHAW: This will be a verification question, and this will be for Dr. Chang. This was your slide 4, where I'm trying to understand your definition of substantial evidence of effectiveness.

And it seemed that you equated it with evidence that has to come from multiple clinical investigations. Is that the definition of substantial evidence? And if not, maybe you can clarify.

DR. NGUYEN: Hi. Christine Nguyen, FDA, and, actually, I'll take this question. That's another really good question. As written by law, when the Amendments Act of 1962 went through, that established the requirement to establish efficacy before approval because before 1962, all you needed was to show that your drug is safe enough.

The way that the law is written, we at FDA traditionally interpret that as requiring two adequate, well-controlled trials; so it's both the quantity and the quality of the trials. Now, the scientific principle behind the two trials is that they allow for independent substantiation of the drug's benefits, so substantial evidence.

That said, over the years, we have accepted -- or rather, we've considered trials from adequate and controlled single trials with persuasive findings -- and there are other criteria with that, but

I won't belabor that -- as substantial evidence. So the question is, we must require that you have two adequate and well-controlled trials, but when we do, we do need to take into account the data from both trials.

Does that answer your question?

(Dr. Shaw gestures yes.)

DR. LEWIS: Dr. Eke, last question.

DR. EKE: Thank you. So my concern was -- actually, I have a couple of them, but the one that concerned me the most was enrollment into Trial 003. After the advisory committee talked about this in 2006 and the FDA considered it and agreed to enroll patients into Trial 003, was there any kind of foresight that there were going to be problems with enrollment, given that when the drug gets approval, patient enrollment gets low, especially when societies endorse the medication?

Have there been other conditions in medicine, other trials, where subsequent trials did not enroll as much because of this situation? Because I feel it kind of played some role into why Trial 003 rolled out low in the U.S..

DR. CHANG: Christy Chang from FDA. I could try to answer some of that question from Dr. Eke. The second review cycle for Makena resulted in a not approval action, precisely because FDA had concerns about whether this trial could be feasible and could be completed successfully. So at the time of the 2009 action to not approve the application, we asked for the applicant to agree to enroll at least 10 percent of the total subjects from the U.S. and Canada, and also we needed them to show that the IRB approval could be obtained from at least 15 investigation sites.

Also, enrollment had to be greater than 15 subjects at any U.S. clinical sites. That was all built in, in a very thoughtful discussion at the time of the second review cycle, something that we did consider.

DR. LEWIS: Thank you. I know that some people have follow-up questions. There will be a little time after lunch to address those, as well as certainly some questions that begin to touch on things that are really discussion points, and we'll certainly build in lots of time for that.

We're going to now break for lunch. We will convene in this room in one hour, at 1:05, at which time we'll begin the open public hearing session. Please take your personal belongings with you at this time. Panel members, please remember no discussion of the meeting contents during lunch amongst yourselves, with the press, or any members of the audience. Thank you, and, panel members, there is a small conference room for us to have lunch.

(Whereupon, at 12:04 p.m., a lunch recess was taken.)

A F T E R N O O N S E S S I O N

(1:05 p.m.)

Open Public Hearing

DR. LEWIS: If people could take their seats, I'd like to begin the program again.

Both the Food and Drug Administration and the public believe in a transparent process for information gathering and decision making. To ensure transparency at the open public hearing, the FDA believes it is important to understand the context of an individual's presentation. For this reason, FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement to advise the committee of any financial relationship that you may have with the sponsor, its product, and, if known, its direct competitors.

For example, this information may include sponsor's payment of travel, lodging, or other expenses in connection with your attendance at this meeting. Likewise, FDA encourages you at the beginning of your statement to advise the committee if you do not have any such financial relationships. If you choose not to

address this issue of financial relationships, it will not preclude you from speaking.

The FDA and this committee place great importance in the open public hearing process. The insights and comments provided can help the agency and this committee in their consideration of the issues before them. That said, in many instances and for many topics, there will be a variety of opinions. One of our goals today is for this open public hearing to be conducted in a fair and open way, where every participant is listened to carefully and treated with dignity, courtesy, and respect. Therefore, please speak only when recognized by the chairperson. Thank you for your cooperation.

Would speaker 1 please step up to the podium and introduce yourself? State your name and any organization you are representing for the record. Welcome.

DR. ALADDIN: I'm Meena Aladdin, a health researcher at Public Citizen's health research group, and I have no financial conflicts of interest. Public Citizen strongly urges the committee to recommend that

the FDA withdraw approval of Makena from the market, as there is a lack of substantial evidence that the drug is effective. Public Citizen has petitioned the agency to take such action.

During the initial review of the NDA for Makena, the lead FDA statisticians strongly recommended against the drug approval, noting the following regarding the single, seriously flawed, premarket, phase 3 clinical trial. From a statistical perspective, the level of evidence from study 17P CT002 is not sufficient to support the effectiveness of 17P. The primary reason is the absence of a second confirmatory study. Study 17P CT002 was not designed for drug approval. The statistician further says the results of the analyses of the 32- and 35-week endpoints suggests that false positive rates could be as great as 1 out of 40.

The PROLONG trial was a well designed, appropriately powered clinical trial, the design of which was mutually agreed upon by both the sponsor and FDA. It did not suffer from the multiple flaws seen in the premarket trial. Most importantly, the PROLONG

trial failed to show a statistically significant treatment effect for Makena on any primary or secondary endpoint.

The FDA concluded, in summary, Trial 003 did not demonstrate a treatment benefit of Makena on reducing the neonatal composite index or the rate of spontaneous preterm birth prior to 35 weeks gestation, and nowhere is there evidence of a treatment benefit on the rate of spontaneous preterm birth prior to 37 weeks or 32 weeks gestation.

Furthermore, the FDA concluded that the unplanned exploratory subgroup analyses conducted by the sponsor do not provide convincing evidence of efficacy over placebo with any subpopulation, and there is no statistically significant interaction between Makena and any of these risk factors.

Maintaining approval of Makena in the absence of any demonstrated clinical benefits would make a mockery of more than a 50-year FDA legal standard, requiring substantial evidence of a drug's effectiveness. Therefore, Public Citizen strongly urges the committee to recommend that the FDA withdraw

approval of Makena from the market, as it fails to provide any clinical benefit. Thank you.

DR. LEWIS: Thank you. Speaker number 2, please.

DR. URATO: Hello. I'm Dr. Adam Urato. I'm an obstetrician/gynecologist and the chief of maternal fetal medicine at Metro West Medical Center in Framingham, Massachusetts, and a co-petitioner with Public Citizen. I have no financial conflicts of interest.

I'm here today to strongly urge the FDA to withdraw approval of Makena, based on the recent definitive findings that it is ineffective for preventing preterm birth. As a clinician, I counsel patients with prior preterm birth regularly. I have delivered lots and lots of babies in my career, many of whom were premature.

Preterm birth is a major problem caused by many different factors, but this drug is not the solution. Approval of this drug was based on a single study that had many significant flaws, relied on a surrogate efficacy marker, and did not show meaningful

clinical benefit. Furthermore, the FDA mandated postmarket study, the PROLONG trial, showed Makena to be ineffective in preventing preterm birth. This makes continued use of this drug indefensible.

I must add here that it was noted today that the American College of OB/GYN and Society of Maternal Fetal Medicine have recently made statements supporting Makena. It should be noted that these groups are funded by AMAG Pharmaceuticals.

Proper counseling of patients involved reviewing risks and benefits of Makena. The risks are injection site reactions, possible increased risk in pregnancy complications, including stillbirth, and unknown long-term adverse effects from in utero exposure. And benefits, the drug has no proven benefits. I'm certain that when patients are properly counseled, they would never agree to be injected with it.

I would also like to highlight that the drug is a synthetic hormone that crosses the placenta and enters into the fetus during development. It enters cells in the fetal brain, the reproductive organs, and

throughout the body. The long-term effects of a fetal exposure to synthetic hormones are not known, but we have been down this road before.

Diethylstilbestrol, DES, was used by millions of women across three decades. Fetal exposure to this synthetic hormone resulted in severe and terrible long-term health effects for many who were exposed. Part of the tragedy of DES is that despite how it was promoted to the public, the drug was not effective in preventing abortion, miscarriage, and preterm birth.

The lesson we learned from DES was clear. We would never again expose pregnant women and their developing babies to a synthetic hormone that did not have good evidence of proven effectiveness, and yet, 50 years, we're making that same mistake. History will judge us poorly if we do not pull this drug from the market and if we continue injecting this synthetic hormone into pregnant women. Thank you for allowing me to speak to you today.

DR. LEWIS: Thank you. Speaker number 3, please.

DR. FOX-RAWLINGS: Thank you for the

opportunity to speak today on behalf of the National Center for Health Research. I am Dr. Stephanie Fox-Rawlings, the center's research manager. Our center analyzes scientific and medical data to provide objective health information to patients, health professionals, and policy makers. We do not accept funding from drug or medical device companies, so I have no conflicts of interests.

The mortality and morbidity associated with preterm birth is a serious issue, which puts children at risk for long-term developmental problem. Treatments that decrease risk for preterm birth and improves neonatal outcomes are needed, but any drug given for this purpose must accomplish this purpose without undue risk.

Based on the evidence being discussed today, there is not consistent evidence that Makena actually does this. When the FDA approves a drug, even if it's based on accelerated approval, there's a lot of pressure to keep it on the market regardless of postmarket data, but in this case, there's no evidence that this drug decreased neonatal death or morbidity,

which are the most important outcomes and the outcomes required for full approval.

Although the first study showed a statistically lower rate at birth before 37 weeks, from 55 percent 37 percent, that could still have occurred by chance. In the confirmatory study, the rate of births before 35 weeks was 11 percent instead of 11.5 percent, and a similarly small difference for births before 37 weeks, both of which were not statistically significant and would not have been sufficient merit approval. At the same time, there were almost twice as many stillbirths for babies whose mothers took Makena, 2 percent versus 1 percent in the first trial and 1 percent versus half a percent in the confirmatory trial.

FDA's reputation depends on admitting when a promising new treatment is later found to be not so promising. The purpose of an advisory committee meeting is to provide objective advice to encourage FDA to stick to the science and admit when there is not evidence that the benefits outweigh the risks for a product, such as the case with Makena.

At most advisory committee meetings, the sponsors recruited clinicians and/or patients to speak on behalf of their product. As scientists, physicians, and patient and consumer representatives, please keep in mind that just because a patient has a good outcome after using a medical product, it does not mean that the medical product caused that good outcome.

As you already know, randomized, double-blind, controlled clinical trials give us a much more accurate assessment of whether a product works than just antidotal information, however heartbreaking or compelling. Makena may possibly reduce preterm births for some pregnant women who have previously had a spontaneous preterm birth, however, with the conflicting results in the two studies, the sponsor needs to determine if there is a subgroup of pregnant women who are likely to have benefits that outweigh the risks, and if so, to be able to define that group for an indication.

But the benefit also has to be clinically meaningful. The sponsor needs to demonstrate a clinically meaningful impact for neonates, such as

improved survival or health outcome. Unless the sponsor can do these two things, approval for this product should be rescinded. Thank you.

DR. LEWIS: Thank you. Speaker number 4, please.

DR. HILL: Good afternoon. I'm Dr. Washington Hill from Sarasota, Florida, and I've practiced OB/GYN or MFM 55 years. AMAG supported my travel and hotel, but not my time or my opinion. Preterm birth is a significant problem in the U.S., especially in African Americans.

In 2003, Meis reported it could be reduced through weekly injections of 17P. Subsequently approved and marketed as Makena for patients with prior spontaneous preterm birth. Last year, ACOG reaffirmed patients with this indication should be offered 17P, now a current clinical guideline. Last Friday, ACOG reaffirmed again it is not changing these recommendations.

17P should not go away because of PROLONG, as it has been a part of the OB/GYN's care prevention of preterm birth for years, resulting in less preterm

birth, especially in African Americans disproportionally affected and at significant risk, as Dr. Owens pointed out this morning.

The populations of these studies were markedly different. Putting a finer point on it, demographics matter, as pointed out in the Meis study conclusion. Her study included the highest of the high risk for preterm birth: black, under stress, or unmarried, smokers, underweight, history of previous preterm birth, and no prenatal care; far different than PROLONG patients, who were predominantly neither American, or African American, but European and without social determinants of health, so important in causing preterm birth.

Let's not eliminate this effective intervention from our preterm birth prevention toolbox because of PROLONG, a non-comparable, negative trial. If we do that, we would be ignoring results of the landmark positive Meis study, the 2019 positive meta-analysis, and over 15 years of positive clinical use showing safety and efficacy in reducing preterm birth. We would also be doing less than we could for

our patients with prior spontaneous preterm birth.

Makena is the only FDA-approved treatment for patients with prior spontaneous preterm birth and needs to be available for us doing all we can to prevent preterm labor and preterm birth. There is insufficient evidence and data today for its removal. We need 17P, as pointed out Friday and today by SMFM, so we can make the best decision with our patients and choose what is in their best interest. Thank you for your time.

DR. LEWIS: Thank you. Could we hear from speaker 5, please?

DR. BARTON: Good afternoon. I'm John Barton, a maternal fetal medicine specialist in private practice in Lexington, Kentucky. For disclosure, AMAG Pharmaceuticals has agreed to pay for my travel expenses to this meeting. I did not, however, have a financial arrangement concerning my presentation, nor do I have a financial interest in the outcome of this presentation.

I've been in practice in our community hospital for 27 years. Three of the greatest problems in current obstetrical care are hypertension,

hemorrhage, and prematurity. Over the past five years, obstetrical societies have made great end roads in reducing complications from hypertension and hemorrhage. Prematurity, however, remains a significant clinical problem.

Several of our previous treatments for prematurity prevention have been withdrawn from use, including ritodrine, terbutaline, and prolonged IV magnesium sulfate therapy. Intramuscular 17-alpha hydroxyprogesterone has been shown to be beneficial in reducing the recurrent risk of spontaneous preterm delivery as one of the few approved interventions to reduce the incidence and burden of spontaneous preterm delivery in our patients and on our healthcare system.

In my office electronic medical record, I have a standard counseling note for patients with a history of a previous spontaneous preterm delivery. I state that a spontaneous preterm delivery in a previous pregnancy is well documented as placing the current pregnancy at risk for prematurity. I then discuss some of the specific theories as to why 17P may result in reduced rate in preterm delivery.

Finally, based on the literature and some of my own previous publications concerning 17P therapy, I affirmed that women who are candidates for this therapy should have progesterone supplementation initiated between 16 and 24 weeks gestation and continued through 36 weeks gestation.

Finally, in providing an analogy, in protocols to reduce infection in hospitals, patients transferred with an IV or to have their IV removed and replaced once are performed under known sterile conditions.

From a clinical standpoint, it's important, however, not to remove a good IV until you've replaced it with one of equal or better quality. Similarly, as a practicing physician at a community hospital, I believe we should be reluctant to remove FDA-approved 17P therapy unless we have another therapy of equal or greater ability to reduce the recurrence, risk, and burden of spontaneous preterm delivery. Thank you.

DR. LEWIS: Thank you. Speaker 6, please.

MS. OSMAN: Good afternoon. My name is Robin Osman. Danielle Boyce asked me to read her testimony on her behalf. She planned to be here today, but

unfortunately had a last-minute issue arise, and had to stay home to care for her premie today. This is her testimony.

"Good afternoon. My name is Danielle Boyce. I'm here to share my personal perspective. I have been on an FDA advisory committee and have served as an FDA patient representative. I have been in your shoes and appreciate the weight of the decision that you need to make. I consider it my civic duty to participate because I have a premie.

"I want to share with you my belief that pregnant women should have access to Makena if they are at risk for having another preterm birth. My son Charlie was born in 2010 at 34 weeks after a significant struggle with preterm labor.

"When Charlie was born, I was under the impression that 34 weeks was no big deal. That is the public perception, but that is not the case. Despite his decent birth weight, 5 pounds 8 ounces, Charlie had many of the conditions of prematurity, including respiratory distress syndrome, jaundice, breastfeeding challenges, and temperature regulation problems. We

faced a 10-day NICU stay.

"The long-term consequences of Charlie's premature birth continue to this day. He developed infantile spasms, a catastrophic form of epilepsy, has had two brain surgeries, autism, and has profound cognitive impairment. He was born at 34 weeks, but I will take care of him for the rest of his life.

"I did not take the decision to have another child lightly. I reviewed the safety and efficacy evidence on my own. I have a master's in public health with a concentration in epidemiology and spoke to top maternal and fetal medicine doctors. I asked for their clinical experience. All agreed that I should take Makena.

"I took their advice, and to my amazement, 34 weeks came and went, and I was still pregnant; then 35, 36, and 37 weeks. With each day that went by, all I could think of was the organ development, weight gain, and all the other benefits of keeping him cooking one day at a time. In May 2017, I had a full-term, 7-pound baby boy named Nash. I remember looking down at his perfect little face in the delivery room and saying,

'Thank God I took those shots.'

"I don't know for sure that it was Makena that gave me a full-term baby, but given the lack of side effects, I would never forgive myself if I hadn't done everything that I could possibly do to prevent preterm birth. If I ever have another child, I will be devastated if I do not have the means of potentially preventing another premature birth. Thank you very much for your time. I wish you the best in your deliberations."

DR. LEWIS: Thank you. Speaker 7, please.

DR. NORTON: Thank you. Good afternoon. My name is Dr. Mary Norton, and I'm a practicing perinatologist and director of maternal fetal medicine at UCSF. I'm here representing the society for maternal fetal medicine as past president and current chair of the publications committee. I have no conflicts of interest to disclose.

We all know that preterm birth is a major public health problem, that prior preterm birth is a significant risk factor, and 17P has been used in an attempt to decrease the risk of recurrence. In 2003,

Meis, et al. reported a 34 percent reduction in recurrent preterm birth in women given 17P and also demonstrated reductions in some neonatal complications.

After the Meis publication, ACOG and SMFM have recommended progestogens for women with a prior spontaneous preterm birth. In 2017 SMFM reaffirmed a recommendation that pregnant women with prior spontaneous preterm birth receive weekly 17P. However, as we've heard today, the PROLONG study found no benefit of 17P compared with placebo in reaching either their primary outcomes.

An important difference between PROLONG and Meis involve the study populations. As we have heard over the course of the day, PROLONG patients had a much lower baseline risk, and this complicates interpretation of the results. Both Meis and PROLONG found no increase in congenital anomalies or evidence of teratogenic effects. Long-term outcomes are unknown, although long-term adverse effects have not been reported.

Preterm birth is clearly a complex disorder. While factors such as race and the number and

gestational age of prior preterm births are associated with recurrence, specific criteria to quantify risk, the interaction between risk factors, and optical management of at-risk women are not well understood. Patient level criteria to determine potential response to 17P have not been confirmed.

Based on the evidence of effectiveness of 17P demonstrated in the Meis study, which is the trial with the largest number of U.S. patients, SMFM believes that providers should continue to have access to 17P for women at high risk of recurrent spontaneous preterm birth. The risk-benefit discussion with such women should incorporate shared decision making, taking into account the lack of short-term safety concerns, but uncertainty regarding benefit.

We recognize that 17P is associated with significant healthcare costs, discomfort from the injection, and extra patient visits, and that long-term potential maternal and neonatal effects are unknown. The lack of benefits seen in PROLONG raises questions regarding the efficacy of 17P, and SMFM recommends that additional studies are needed to determine if there are

populations or subgroups in which 17P may provide a benefit. We are aware of ongoing studies, including the large IPD meta-analysis discussed today, and will continue to closely follow advances in this area to assure optimal care for women and provide guidance for maternal fetal medicine subspecialists. Thank you.

DR. LEWIS: Thank you. Speaker 8, please.

MS. CHIAVERINI: Hello. My name is Amelia Chiaverini. I will be reading the testimony of Anabel Jimenez-Gomez, as she couldn't be here today.

"I support Makena for families that are considering using it. I really wanted to be here in person because Makena helped me bring home the baby that my husband and I so wanted and prepared for. After losing my first baby at 20 weeks to preterm birth, it was critically important to me to do everything I could to make it to full term.

"My first pregnancy was a rough one. When I was 20 weeks along, I was feeling lower back pain and was really uncomfortable. After an ER visit, the doctor said a UTI was the cause of my discomfort. I was prescribed antibiotics and muscle relaxers. Within

24 hours, I got a lot worse and ended up back in the hospital. I went into preterm labor.

"Our baby girl was stillborn. The whole birth was a very traumatic experience, which I still have nightmares about. The doctors ran tests but couldn't find an exact cause for my preterm birth. They asked, 'Did you hurt yourself? Did you fall, lift something heavy?' They couldn't pinpoint exactly what caused it. It was really stressful to both my husband and I.

"About five months later, I found out I was pregnant again. We were scared and wished we had waited a little longer. My doctor told me we would take different precautions because my pregnancy was considered high risk. I had biweekly doctor visits with a different goal for each appointment. The main goal was to make it to 20 weeks, so my doctor suggested Makena.

"At first, I was terrified to try something new. She gave us statistics and also let us know that other women had gone through similar experiences. This gave us hope, so we decided to try it out. The medical team was really good at teaching my husband to

administer the shots. He administered them for me at home once a week for 16 weeks. They were painful, but looking back, I realized it was all worth it.

"I delivered my baby boy, Mateo, at 39 weeks and 5 days, which was just 2 days before his due date. The delivery was a little less stressful, but I had an amazing team that could take care of me and calm my nerves the entire time. It took 2 days of labor, but Mateo finally came out in a smooth delivery. He was 8 pounds even, 20 and a half inches long.

"Even though it was scary to lose my first baby and then go through my second pregnancy, I'm really glad that we did, and have Mateo today with the help of Makena. I didn't know if it would work or not, but I was willing to try anything that could help me carry a pregnancy to full term. Makena had a significant impact on us.

"I believe Makena can help a lot of women carry their rainbow babies to full term safely. I recommend it to women who have gone through a similar experience as mine. Thank you for listening to my story. Anabel Jimenez-Gomez."

DR. LEWIS: Thank you. Speaker 9, please.

DR. MOLEY: Hi. I'm Dr. Kelle Moley. I'm the chief scientific officer and senior vice president of the March of Dimes. Before this, I was at Washington University in St. Louis as a practicing OB/GYN for 30 years.

On behalf of the March of Dimes, I'm pleased to provide comment on the state of maternal and child health in the U.S.. March of Dimes, a nonprofit, nonpartisan organization fights for the health of all moms and babies. We advocate for policies to protect them. We work to radically improve the health care they receive. We pioneer research to find solutions, and we empower families with programs, knowledge, and tools to have healthier pregnancies.

March of Dimes does not offer recommendations on medical treatments, however, we do rely upon the leading medical societies and organizations, such as ACOG and SMFM to make such recommendations. March of Dimes then supports and communicates these to all stakeholders.

We do this all because today in America, we

face an urgent maternal and infant health crisis. Approximately every 12 hours, a woman dies due to complications resulting from pregnancy, and more than 50,000 others experience dangerous complications that could have killed them, making our country among the most dangerous places in the developed world to give birth.

For women of color, the dangers of giving birth or even more acute. Black mothers are more than three times as likely to die from pregnancy related to complications as white peers. But this crisis isn't only about moms; it's also about their babies. It's about the continuum of care for all moms and babies as their health is intertwined. In fact, the U.S. prematurity rate may have increased for the fourth consecutive year. Each year in the U.S., 22,000 babies die; that's 2 babies every hour, and approximately 1 in 10 babies are born preterm.

Preterm birth increases from 9.63 percent in 2015 to more than 10 percent in 2018. In a few days, on November 1st, we will mark the start of Prematurity Awareness Month, and November 4th will be the

nationwide release of the March of Dimes report card, which highlights the collective factors that contribute to maternal and infant mortality and morbidity. The report card grades the nations, all states, and the District of Columbia and Puerto Rico, based on the latest data on preterm birth rates, and spotlights the issues contributing to poor health.

March of Dimes' mission is to fight for the health of all moms and babies. Consistent with our mission, when an evidence-based intervention like 17P becomes available, our overwhelming interest is to increase access so that all eligible women receive it no matter what their income or insurance status. For many years, we've advocated for access to 17P for all eligible women due to the evidence about its effectiveness in reducing preterm birth. We've educated women and providers about the importance of 17P.

In conclusion, the U.S. needs to be aggressively paying attention and looking for ways to solve the national maternal and infant health crisis of increasing preterm birth rates. We stress the need for

more therapies, more solutions, more devices, and everything possible to address the birth crisis we're experiencing.

Therapeutics for preterm births such as 17P and all future therapies should be available so that physicians can use their discretion to prescribe them to the correct subset of patients with these complex and multifactorial conditions.

The accelerated approval pathway is critical to achieving this goal, as preterm birth disproportionately affects underserved populations in the U.S. We applaud the FDA's history of continuing effectiveness therapies of preterm birth as worthy accelerated drug approval, and trust this will continue to be its practice.

It's essential that the U.S. do everything possible to ensure that moms and babies are healthy. We thank you for the opportunity to comment during today's meeting. March of Dimes stands at the ready to serve as a resource to this committee.

DR. LEWIS: Thank you. Speaker 10, please.

MS. JOHNSON: My name is Allison Johnson. My

travel is being reimbursed by AMAG Pharmaceuticals, however, I'm not being compensated for my time, and this testimony is my own.

I'm a mom to three beautiful little boys. In July of 2018, my third son Andrew joined our family, and I credit Makena with helping to bring him into our lives. But in order to tell my story around Makena, I need to take you back to the birth of our second son Teddy.

My water broke at 34 weeks 6 days with Teddy. It was a very complicated delivery. The doctors tried for nearly 40 minutes to first get a spinal, then epidural in place for my repeat C-section. Both were unsuccessful, which eventually led to me being put under general anesthesia. His birth was traumatic, and this is a story that I wait to tell my pregnant friends until after they've given birth. But I know we were lucky. Teddy was born at 5 pounds, 12 ounces, and he thankfully had no complications. He required some early intervention services up until the age of 2, but now he's a healthy, thriving, and rambunctious 4 year old.

Following Teddy's birth, if you had asked my husband and I whether we were done having kids, I almost always said yes. I'd been told almost right away that once you have a spontaneous preterm birth, your chances of having another are much higher. However, my husband and I knew in our hearts that our family wasn't complete. There was still a missing piece, but I was nervous about another pregnancy.

So my husband and I decided to meet with my doctor, who was confident that I could have a successful pregnancy if we chose to have another child. She explained to us that in order to help with preterm birth, there was an injection, Makena, that she would recommend. My husband and I talked through our options following that appointment, and we decided to try to expand our family once more.

A few months later, I was pregnant with Andrew, and I began the Makena injections as prescribed. My husband learned from the nurse how to administer them at our home, and each week, from 16 weeks to about 35 weeks, he helped give me those shots in our upstairs bathroom, and it actually became

a family affair. Sometimes our two other boys wanted to help, too, and they were in charge of the band-aids.

I was fully prepared for Andrew to arrive before my scheduled C-section date. I had my bags packed and ready to go by 32 weeks, but it never happened, and he was born at a healthy 8 pounds, 1 ounce. He had made it to full term, and I thank Makena for helping us to get there.

I'd like to ask that the FDA take my experience into consideration when you evaluate Makena and its effectiveness. While I wasn't in either of the clinical trials discussed earlier today, Makena helped me and my baby, and I hope that you will give that hope and chance to other anxious and excited families as well. Thank you.

DR. LEWIS: Thank you. Speaker 11, please.

MS. JOHNSON: So again, my name is Allison Johnson, and I will be reading the testimony of Glory Joseph.

"This is my story and my most recent encounter with Makena. Through the use of Makena injections, I was able to deliver a healthy baby girl. Because of

the success I had my husband and I have decided that we will be using Makena again once we decide to become pregnant. Because I was unable to present today, I have attached some photos of my beautiful family, including Grace Marie Joseph, whom we often refer to as our Makena baby, which I will be sharing with you today.

"With my first ever pregnancy, everything seemed to be going well, but too soon into my pregnancy, I started experiencing painful contractions. I went to the ER. All tests were normal. Ultrasound had shown a viable fetus. I was discharged home with undiagnosed, unknown cause for my symptoms to experience premature rupture of membranes shortly, 4 days later, without any known cause.

"The loss came just a week after we had announced the pregnancy and made it public. It was almost shameful to have to go and tell people we weren't pregnant anymore. I'm fortunate to have a very supportive family and friends who helped me get through it, but it was definitely a tough time. I'd get emotional seeing other pregnant women or other babies

around the time we had delivered.

"My husband and I both really wanted to build a family, so we decided to try again. In the back of my mind, I was scared I couldn't carry a full-term pregnancy. We knew we wanted another child, but it was scary. When I became pregnant again, I asked my general OB to refer me to a high-risk specialist because of my history. She agreed, and I saw the specialist at 12 weeks.

"She told me that there was a medication we could try once I reached 15 weeks, Makena. I discussed it with my husband and family and did my own research. There didn't seem to be many side effects, so I decided I may as well try it and see if it worked. Once I got to 16 weeks, it was both scary and exciting. I knew there was hope once I started taking Makena, but I wondered if the shop would even work for me.

"The major side effect that I experienced was pain at the site of the injection. With the combined continuous prenatal care, plus weekly Makena up to 36 weeks, I was able to deliver a healthy, beautiful, baby girl, Grace Marie, at 37.4 weeks. She weighed

7 pounds 10 ounces.

"I would highly recommend Makena to any other mothers like me who had preterm births. Thank you for this opportunity to share my story. I truly support Makena. Glory Joseph."

DR. LEWIS: Thank you. Speaker 12, please.

DR. JACKSON: Hi. I'm Marc Jackson. I'm an MFM and the vice president for education at the American College of Obstetricians and Gynecologists. We represent more than 58,000 physicians and other partners dedicated to advancing women's health. I have no personal financial relationships to report, but in 2019, AMAG provided a grant to ACOG to support medical student projects, but not our practice activities or our clinical guidance.

In the time since we submitted our written comments to the committee, the PROLONG trial, Trial 003, has been published. This multinational RCT of patients with a prior preterm birth found no difference in recurrent preterm birth prior to 35 weeks or the neonatal composite outcome between women treated with 17 hydroxyprogesterone caproate or placebo.

Several comments about the study need to be made. Although the study design was similar, the PROLONG study 003, as executed, was fundamentally different from the MFMU trial, 002, that was published back in 2003. This is evidenced by the large difference in the baseline preterm birth rates less than 37 weeks, 23 percent versus 55 percent.

Thus, the study population in Trial 003 was a lower risk population than in 002, and substantially so. Differences in the 002 and the 003 populations, with respect to the number of prior preterm births, smoking rates, social, ethnic, and racial differences, and national differences in healthcare delivery, makes plain at least some of the discrepancy. Because of these differences, a head-to-head comparison of the two trials is inappropriate.

Despite the PROLONG study's findings, the results do not indicate that the initial U.S. based Trial 002, the MFMU trial -- they do not indicate that it was wrong or that its conclusions are misleading in some way. Rather, the data from Trial 003 should be examined as part of the body of literature on

placebo-controlled trials using 17-OHP in preventing preterm birth.

It is that broader examination of the literature that should be used to determine whether there is substantial evidence of effectiveness, not the recent Trial 003 alone. Until a comprehensive analysis can be done, ACOG will continue to recommend that physicians offer 17-OHP to pregnant women with a prior preterm birth.

We will continue to monitor this topic and to evaluate additional data and analyses when they're published, and we'll address new findings in the review process for our clinical guidance as needed. Continued access to 17-OHP is important for our patients, and ACOG respectfully encourages this committee to table any decision on whether to withdraw drug approval until a complete meta-analysis using patient-level data from all the available studies can be done. Thanks for the opportunity to speak.

DR. LEWIS: Thank you. Speaker 13, please.

MS. CHIAVERINI: Thank you for giving me time to speak today. Again, my name is Amelia Chiaverini.

I am being reimbursed by for my travel expenses by AMAG because I wanted to personally tell you about my experience with Makena. I believe this product must be available to women that face similar situations to prevent further emotional and financial stress. I am taking time away from my responsibilities as a mother and wife to be here today. It is that important to me.

In January 2011, I went into preterm labor. I was given several medications to help me and my baby. Unfortunately, after 5 days, I was in labor again and was rushed to the operating room for an emergency C-section. On February 2nd, my first son was born at 27 weeks, 1 day, weighing only 1 pound 14 ounces. It was a terrifying experience.

I briefly saw Duncan before he was transported to a children's hospital. He was so tiny, and the tubes seem to engulf him. My room was near the waiting area to reduce the constant reminder of his absence from the maternity ward. Duncan spent 3 and a half months in the NICU. He received many medical interventions, including oxygen, phototherapy, feeding tubes, PICC line, blood transfusions, and a surgery.

I had to get past all these issues to focus on giving Duncan care and breast milk. The emotional toll was much more difficult to overcome. Here are some memories that stick with me: finding out that a young mother I was talking with had experienced the NICU two times previously; hearing the anguished cries of grief from a mother because her child had died while I quietly held my tiny boy and cried for her and for me; and the worst day, March 21st, when the staff had to manually resuscitate Duncan. Though it was stressful for me and my family, we made it through. Duncan came home on May 19th weighing 8 pounds 1 ounce.

Before my next pregnancy, my husband and I talked with my obstetrician about preventing preterm birth. He told us about Makena. Together, we decided it was a great option for us because it did not come from a compound facility. By receiving the shots, I felt empowered. I was doing all I could to help my baby, and it also eased my stress. On December 12, 2013, Donovan was born at 38 weeks 6 days, weighing 6 pounds 7 ounces. I believe Makena made his full-term birth possible.

There are many women with similar stories that need Makena to help prevent preterm birth, which could also reduce their emotional and financial stress that preterm birth creates. Makena should be available to these women as it was for me. Thank you again for letting me tell my story with Makena.

DR. LEWIS: Thank you. Speaker 14, please.

DR. RANDELL: Good afternoon. My name is Dr. Michael Randell. Thank you for allowing me to speak to you today during the public hearing on Makena and 17P. In my brief comments, I will focus on my concerns if the FDA decides to withdraw Makena from the market. I do not have any conflicts. AMAG Pharmaceuticals has paid my travel to be here, but I have not been compensated for my time.

I am an OB/GYN in Atlanta, Georgia. I'm a fellow of the American College of Obstetricians and Gynecologists and a diplomat of the American Board of Obstetrics and Gynecology. I've been in private practice for more than 24 years following my training. I've delivered thousands of babies and have managed preterm labor, including using progesterone for

pregnancy prolongation in my patients with a documented history of a previous spontaneous birth at less than 37 weeks of gestation.

While preterm birth affects about 10 percent of births in the United States, Georgia's preterm birth rate is higher than the national average. Therefore, preventing preterm birth in my patients has been a major focus of my Atlanta practice. I began using 17P in 2008 following the recommendation of ACOG and the Society for Maternal Fetal Medicine that stated, "Progesterone supplementation for the prevention of recurrent preterm birth should be offered to women with a singleton pregnancy and a prior spontaneous preterm birth due to spontaneous preterm labor or premature rupture of membranes."

Last Friday, ACOG announced it is not changing its clinical recommendations at this time, and it continues to recommend offering 17P.

In each pregnancy, there are two patients, the mom and the baby. This precious package requires OB/GYN to provide their patients with the safest and highest quality of care. I was always concerned with

having to obtain compounded 17P that is not made under FDA-approved conditions, so when Makena was approved, I immediately began prescribing Makena instead of compounded 17P. I've observed several of my patients not have another preterm delivery when using Makena, and I saw it improve neonatal outcome. In my experience, Makena is effective. I've seen the benefits.

Few physicians understand the difference between compounded and FDA-approved medications. In 2014, I wrote an article, Risks and Liabilities of Prescribing Compounded Medications. In this article, I stated, "The potential for patients to suffer serious harm from substandard medications prepared by compounding pharmacies is very real."

Healthcare professionals should be aware of the potential liability to which they expose themselves whenever they prescribe or administer compounded products. Patients injured through the use of compounded medications that do not meet FDA requirements for safety, efficacy, or quality may file lawsuits against the pharmacy, alleging product

defects, as well as against the prescribing physician and medical facility, alleging professional negligence. That is breach of the applicable standard of care.

While understanding the PROLONG study showed that Makena is no better than placebo in preventing preterm birth, I don't believe that this study will change the current standard of care to prescribe 17P to pregnant women at risk. If the FDA decides to withdraw Makena, which I strongly urge the FDA not to do, OB/GYNs will return to using compounded 17P, potentially placing their patients and themselves at significant risk.

Few physicians have the training or experience to suitably evaluate a compounding pharmacy's ability to maintain an accepted technique and consistency of drug concentrations, or to investigate how the pharmacy ensures the potency and purity of their active pharmaceutical ingredients and finished products.

FDA regulation serves an extremely important role in keeping America's drug supply safe. Therefore, I believe that for now, it is in the best interest of patients and my profession that the FDA does not

withdraw Makena. Thank you very much.

DR. LEWIS: Thank you. Speaker 15, please.

DR. CARITIS: Hello. My name is Steve Caritis. I am a professor of obstetrics and gynecology in reproductive sciences at the University of Pittsburgh, and a specialist in maternal fetal medicine. I have a few comments that I hope the committee will find useful in their deliberations.

First, I'd like to establish my credentials. My colleague, Dr. Venkataramanan, who you see up there, and I have published 27 research papers on 17-hydroxyprogesterone caproate, which I will refer to as 17-OHPC, including the first paper on the assay of 17-OHPC and the first pharmacokinetic and pharmacodynamic studies of 17-OHPC in both Singleton and twin gestations. These studies were supported by the Maternal Fetal Medicine's Units Network and the Obstetrical Fetal Pharmacology Research Centers. None of these studies were supported by industry.

Our research that is most relevant to your deliberations is our pharmacodynamic study of 17-OHPC in women with singleton gestation. In that secondary

analysis of data from the MFMU Omega 3 study, we reported concentrations ranging from 4 to 56 nanograms per mL; that's on the left there. That is despite the subjects all receiving an identical dose of 250 milligrams weekly.

The figure on the right indicates a linear relationship from these same data between log transform 17-OHPC plasma concentrations and the rate of preterm birth. Clearly, those women with higher concentrations had lower rates of preterm birth. These data suggest 17-OHPC efficacy for preterm birth reduction.

The possibility that a higher concentration of 17-OHPC might be associated with lower rates of preterm birth led us to initiate a prospective study within the Obstetrical Fetal Pharmacology Research Centers. We will randomize 300 women with a prior preterm birth across 5 university centers to either 250- or 500-milligram weekly doses of 17-OHPC. This will provide a pharmacodynamic analysis of 17-OHPC that may assist in establishing a pharmacologically based dosing regimen.

Despite FDA approval of 17-OHPC in 1956 and the recent approval of Makena, a dose-ranging study had

not been reported; neither had a dose or concentration response study been reported for 17-OHPC and the rate of preterm birth. The weekly dose of 250 milligrams for preterm birth prevention is not based on any pharmacologic data or principle, confounding any meaningful assessment of drug's efficacy.

In the way of disclosure for myself and Dr. Venkat [ph], the 17-OHPC for this study that I referred to earlier is being provided by AMAG Pharmaceuticals without charge to the OPRC. The data obtained and publication rights are retained by the investigators. In addition, we are also negotiating to perform a study for AMAG, comparing intramuscular and subcutaneously administered 17-OHPC. Thank you.

DR. LEWIS: Thank you. Speaker 16.

DR. THOM: Good afternoon. My name is Elizabeth Thom, and I do not have any financial relationships with the sponsor. I'm a research professor of biostatistics statistics and bioinformatics from George Washington University biostatistics center, and the center has been the data coordinating center for the NICHD MFMU networks since

the beginning of the network, and as such, I was involved in the Meis study, and I was the principal investigator of the coordinating center and oversaw the conduct of the trial.

The data coordinating center was responsible for assisting with the development of the protocol, creating the data, the case report forms, providing the data management system, monitoring protocol adherence, and doing weekly editing and auditing. I believe that we did a good job because we were very familiar with obstetrics and obstetrical trials. So overall, I think the data were very good quality and the protocol adherence was good.

I was actually present at the interim monitoring meeting when the Data and Safety Monitoring Committee recommended early termination of the study, and I have no doubts that the trial was truly positive. The data had been consistent at the previous interim look, and I'm pleased of that, and although the outcome rate was higher than expected, the women who agreed to the trial were at very high risk.

To change subjects, in the last few years, I

have also been a member of the Secretariat for individual participant data meta-analysis funded by the PatientCenter.com Research Institute, which was referred to earlier today, and that is comparing vaginal progesterone, oral progesterone, and 17-OHPC with control or with each other. It is known as EPPPIC.

As a member of the Secretariat, I helped design the overall study, but I have had no involvement in the actual analysis. The meta-analysis itself was conducted by an independent but very well respected group in the UK. None of the members of that team have been a part of a previous progesterone trial or progesterone meta-analysis and were considered to be unbiased.

This is the largest and most comprehensive individual participant data meta-analysis to date. They looked at 30 trials in about 10,000 women, and about half of them were trials of 17-OHPC. They included 84 percent of the data of randomized trials in 17-OHPC. Those that weren't included are mainly small, unregistered, or single center. The results have not

been published, so I can't talk about that, but I believe that these data are important and should be taken into consideration.

Finally, on a personal note, I was the mother of a preterm baby of 32 weeks gestation, and although it was 5 years ago, I can tell you the experience never goes away. After my son was born, we had several difficult years; and although it was not nearly what some families go through, it certainly factored into my decision not to have another child, as 17-OHPC was not available then, and if it had been, things might have been different.

So on both a scientific and personal level, I ask that the FDA panel and the FDA do not negate the results of the Meis trial by the results of the PROLONG study, but consider the fact that the original trial is more relevant to the U.S. population, that high-risk women might very well benefit from 17-OHPC, and to take into account the results of the EPPPIC meta-analysis when it becomes available. I believe that 17-OHPC should be an option for high-risk women with a prior preterm birth and shared decision making between the

doctors and women who could potentially benefit from it. Thank you.

DR. LEWIS: Thank you. Would the final speaker please approach the podium?

(No response.)

Clarifying Questions to Applicant or FDA

DR. LEWIS: Okay.

We have time for some clarifying questions for the FDA and the sponsor by the committee members.

Dr. Gillen, I think you're up first. You had a question left over from this morning.

DR. GILLEN: Yes, thank you. My question is primarily to Dr. Wesley, and it's really around clarification of the 37-week endpoint that was used in the first study. As you'll recall and was stated earlier, in that 2006 advisory committee meeting, there was pretty strong consensus that the 37-week was not a quote/unquote, "adequate surrogate," adequate surrogate I presume meaning satisfying the Prentice criteria.

So what was stated about that -- and this is really a follow-up, to some degree, to Dr. Shaw's question about substantial evidence for efficacy. Part

of that is the quality of the endpoint and the clinical relevance of the endpoint, I would argue.

The question is, when you described the timeline about new information coming out on the 37-week endpoint as, quote/unquote, "becoming an adequate surrogate," how does that impact our view of what is substantial evidence for efficacy, as described by the sponsor, to be honest, in their presentation?

What's the FDA's point of view?

I'm trying to get a feel for where you are on the 37-week endpoint and what the timeline was, because it seems like the PROLONG study was already underway at the time that you had made that decision that the 37-week now is, quote/unquote, "adequate."

Can you fill me in on this?

DR. WESLEY: Well, it's somewhat difficult because nobody knows exactly the best surrogate to use for this. At the time when the data came out -- and it wasn't just a publication; it was also states made a law that you couldn't induce somebody before 39 weeks, if you recall. You're not a clinician, but 39 weeks, you had to wait to induce somebody because of the

morbidity occurring in the late preterm birth.

So because the results were so persuasive at 37 weeks, even though they weren't at 32 and 35, we decided to give it a chance and go ahead and do the provisional approval. It's not clear exactly, but I wanted to show a slide to show you the population in 002.

Can you pull up slide 20? It is an older population of preterm births, and that might be why, because you had so many more of them in that population, you see the median -- I don't look at means, but the median preterm birth rate in the treatment arm was 37 and a half weeks, and in the placebo arm, it was 36 and a half weeks; only one week difference.

It seems as though because the population was older in that thing, it might have been affected. I don't know. This is not written in stone with us. We keep looking. We keep looking at the literature, we keep up with changes, and we make decisions based on that. That's the best I can say.

DR. GILLEN: My question is somewhat pointed

to your slide 14, which says, "FDA concluded that delivering at less than 37 weeks of gestation was an adequate surrogate endpoint." Is that still the position of the FDA? I'm just trying to get -- if we're asked to come back and judge the first study based upon its merits, which we already did once in 2006 -- I happened to be there. So now if we're asked to judge it again, I want to know where the FDA stands on this as an endpoint.

Given what I'm reading here, is that the official stance of the FDA?

DR. WESLEY: There is no official stance. We decided at that time, with the people there, to do that -- to use that gestational age. But I can't say there's an official stance. I mean, it's something that we keep evaluating all the time.

DR. NGUYEN: Hi. Christine Nguyen, FDA. Let me try to address your question. You're asking whether, in 2019, we would consider the gestational age of delivery less than 37 weeks an adequate surrogate endpoint for accelerated approval, and the answer would be yes.

DR. LEWIS: Thank you. Dr. Orza?

DR. ORZA: I have some questions about the safety side. In their comments and also in their petition, Public Citizen commented on and did some analysis of the rate of stillbirths, which was higher in both studies in the treatment group. I was wondering what FDA's analysis of that had shown.

Also, the sponsor recommended to describe data that they had on the long-term effects, out to an average of, I think they said 4 years. And I was wondering if the FDA had analyzed those data and what your conclusions were.

DR. CHANG: Hi. Christy Chang from FDA. Your first question was about the safety findings from both 002 and 003. You're correct that from the 002 study, there appears to be a signal in increasing early fetal loss and early infant deaths from study 002. But in study 003, based on our review, it appears that the incidences for these findings were similar in both treatment groups. Furthermore, the 003 study was designed to rule out a twofold increase in adverse neonatal outcome, and was shown in 003.

DR. ORZA: They were similar overall, but specifically for stillbirths, they were higher in the treatment group in both studies, and that was what the Public Citizen analysis referred to. There was also a concern about where in the 16- to 20-week window the treatments were started, and they seemed to suggest that there was a difference between early in that window and late in that window, potentially, on the rate of stillbirth.

Did you do similar analyses?

DR. WESLEY: Can you pull up slide 24? This shows the two studies, and if you look at stillbirths, you have a 2 percent rate in the treatment arm of 002 and zero percent of the placebo arm. Then in 003, you have a 1 percent stillbirth rate and a 0.5 percent.

So these are very small numbers. The percentages are not that dramatically different. No, we didn't really look at the time of starting of the drug and the relationship of stillbirth because the numbers are so small, it would be hard to really do that analysis, but that is something that's worth considering in the future.

DR. LEWIS: Thank you. I think sponsor wanted to say something to that point.

DR. ORZA: And also the long-term data, the long-term safety data.

DR. KROP: We evaluated the stillbirth rate very carefully and had an independent maternal fetal medicine physician, who was blinded, to review the details. I'd like to call up Dr. Sibai who reviewed these himself.

DR. SIBAI: Baha Sibai, UT Houston. I reviewed the data for both the Meis trial as well as the PROLONG. For the PROLONG, this was blinded. For the Meis study, I had the data because it's already published and available. I looked through every one of these, and as you see from here, from the PROLONG study, there was only one unexplained. For the others, I identified 11 factors.

The way I did it, I used the publication from the stillbirths, which is the NICHD network, where they had several factors there. I evaluated maternal, fetal, placental, cord abnormalities in making my decision. And it is reassuring to see that, really, in

either one of these studies, there was no signal that 17P increases stillbirth.

DR. LEWIS: Thank you. Dr. Davis?

DR. WESLEY: Was there a question on long-term follow-up?

DR. LEWIS: I'm sorry. That's right. I apologize.

DR. WESLEY: Can you pull up slide 30 and 31? The follow-up of children on 003 is not complete, so I'll just show you the results of 002. This is a screening. The ASQ scores are screening for developmental problems. If you look at the treatment arm and the placebo arm -- and remember, this is a 2 to 1 ratio, so they had to look at percent -- you see that the treatment arm had 27 and a half percent positive screens; the placebo arm 28 percent positive screens.

Can you bring up slide 31? These are the people with a positive screen who also had a diagnosis of developmental delay. Those in the treatment arm had 2.6 percent developmental delay -- no, I'm sorry -- 6.7 percent developmental delay. Those in the placebo arm, 9.8 percent.

So there really isn't much difference -- this is a safety study only, between the treatment and the placebo arm -- when it came to screening and developmental delay. If you look at the percentages now, there are some differences, but they're not that significant.

DR. DAVIS: How old were these children?

DR. WESLEY: They're about 18 months old.

DR. DAVIS: And do you know why they used this test versus a Bayley, which is more --

DR. WESLEY: That was used in terms of the diagnosis, yes. The Bayley is more diagnostic and not a screen, so it was used for the diagnosis.

DR. LEWIS: Before you get to your question, Dr. Davis, is this the entire population of 003, or --

DR. WESLEY: No. This is only 002. Because it was not set up beforehand, if you look at slide number 28, it tells you how many. Fourteen of the original 19 study sites in 002 were able to participate. This was post hoc set up and done, so you didn't get everybody, but it had a good percent. Eighty percent of the mothers who participated in the

study had this screen and diagnostic testing.

DR. NGUYEN: Hi. Christine Nguyen. Let me just clarify, the infant follow-up for 003 is ongoing, and the results are blinded. So we're not able to show you those results, and I believe there are data on about 200 children.

DR. LEWIS: Just one more. I will get to your next.

So this is 14 of the original study sites children were eligible to participate. Was there a good distribution of sites throughout the country or were they skewed in terms of a preponderance of one study site?

DR. WESLEY: From my recollection, it was fairly widely distributed. These are 14 sites that were able -- but they were in different parts of the country. There was no particular segregated group of them, no.

DR. LEWIS: Dr. Davis?

DR. DAVIS: Thank you. Jon Davis from Tufts. The definitions of your neonatal morbidities were a little perplexing, so in other words -- and it may be a

moot point because the rates were so low and the average delivery time was 37 weeks, so that's why you may not have had very many. But certainly some of the definitions were bronchopulmonary dysplasia, which was defined as oxygen use for 28 days, which I think I stopped using about 20 years ago.

So I didn't know how those were drafted and whether those are viable, and whether we should be relooking at the definitions and potentially reanalyzing the data with more updated definitions.

I had one more question.

DR. CHANG: Christy Chang from FDA. Some of these may be better addressed by the company. If we could pull up Dr. Sibai's slides from CO-38.

DR. NGUYEN: I'd like to remind the committee that this neonatal index was based on data of when 002 was conducted, so this is 1999. It is about 20 years old. When we proceed with a confirmatory trial, we like to be as consistent as possible with the trial that gained initial approval. So I think that's one explanation.

DR. WESLEY: These definitions were developed

by the Maternal Fetal Medicine Network Units, not by us.

DR. CHANG: I'm wondering if Dr. Sibai has any more comments about this slide, which shows the long-term neonatal follow-up on the babies, whose mothers participated in 002.

DR. KROP: Dr. Sibai, do you want to go up and comment?

DR. SIBAI: Do you want me to comment on this or there's a question? Sorry.

DR. CHANG: I'm just wondering if you had any comments, any additional comments, besides what you already talked about this morning. Based on what the slide has shown, of all the infants that were enrolled in the follow-up study, there didn't appear to be any differences in motor development.

DR. SIBAI: Correct. I would like to point out that, really, the median age at follow-up was 48 months, and you can see the 75th percentile. The other thing I want to emphasize, really, there was no gender differences, which was one of the endpoints. We looked at 12 points for masculinity and 12 points for

femininity in this evaluation, and there was no significant difference.

In regard to the question about BPD, this is really the definition that was used in the neonatal research network among the various studies.

DR. DAVIS: My final question to FDA is, in your market scan data, we've been told you can't do another trial because everyone's using this already, and it's an established treatment. I was curious if we actually know -- most neonatal trials, we can see that 85 percent, 90 percent of our mothers have gotten antenatal steroids before the babies deliver.

Do we have any idea what the market use is? I'm not sure if you would know or maybe the sponsor. How many of these mothers who actually have had a previous preterm birth are receiving the medication? Because it was my sense that it was still relatively low throughout the United States. So whether that really does preclude doing another study, I wasn't sure.

DR. TSAI: This is Huei-Ting Tsai from FDA. Can you clarify? Are you asking the utilization among

the people using the injectable HPC, how many have the preterm delivery?

DR. DAVIS: Yes. So in other words, if we're being told that this is now standard of care being used widely throughout the United States and would preclude doing another study, is that true? I mean, are 80 or 90 percent of all the mothers who are now pregnant, who have had a previous preterm delivery, are they receiving 17P?

DR. TSAI: If we look at slide 10 I think for the Sentinel -- for the drug use slide, slide 10 in drug use slide, FDA drug use slide, but you probably have the information, basically in the Sentinel analysis, it does include the Market Scan data, and that's a major data planner. You can refer the data we got from the Sentinel analysis to see how the use might be in Market Scan.

DR. NGUYEN: Can you pull up drug utilization slide 10, please?

DR. TSAI: Slide 10 in drug use presentation.

DR. NGUYEN: The next FDA slide.

Christine Nguyen. To answer your question, we

have to know the universe of all eligible women in the U.S., and then figure out how many of those receive Makena. So I'm not sure -- well, Market Scan, we will not be able to get the information on that denominator.

DR. KROP: We do have some data on utilization that was from a chart review. I don't know that that would be helpful in your question. It was a thousand patients that we went back and tried to get the denominator that you're referring to. And what we found was, based on that, those were all indicated patients, that about 75 percent of them were taking 17P. This was in 2017.

I'm sorry. I don't know why it's not coming up. But it included both 17P compounded, as well as 17P Makena. The combination was 75 percent, the vast majority of that being Makena, and then there was some off-label use of vaginal progesterone in about 10 percent of patients, and about 15 percent of patients were not being treated.

DR. LEWIS: Okay. Dr. Hunsberger, go for it.

DR. HUNSBERGER: I just had a question for the applicant. They were discussing why, potentially,

another study couldn't be done maybe as a randomized study between another treatment. On slide 83, you put up different treatments and said, well, none of these are beneficial, but if you look at the odds ratio, that's pretty much the odds ratio or the relative risk you saw in your study.

So it's not quite consistent to say the PROLONG study or we should approve this, when these are given as evidence of not being beneficial, and maybe also a discussion of why you couldn't do a randomized study between one of these treatments.

DR. KROP: I'd like to call up Dr. Blackwell to address that question.

DR. BLACKWELL: Thank you. Sean Blackwell from UT Houston, Houston, Texas. I think, certainly, any group of trialists can do a trial. The question is on whether or not it would be informative for this particular question. Certainly, we could do a comparative trial, a randomized-controlled trial of 17P to any therapy. The question is, would it be informative based on the information that we have already?

This is three large placebo-controlled trials, adequately powered with a very high-risk patient population similar to the Meis study, again, different than what I would describe in a PROLONG population, that showed no difference related to treatment effect. Certainly, it's possible to do a trial. The question is whether or not it would be informative and confirmatory. That was the point that I was making in my presentation.

DR. LEWIS: Thank you. I think at this point, we do have a lot of material to get through this afternoon in terms of discussion, and some of the points that are bothering people perhaps you'll have an opportunity to air those concerns. At this point, let's take a 5-minute break, 5 minutes. We'll reconvene at 2:30.

(Whereupon, at 2:25 p.m., a recess was taken.)

Questions to the Committee, Discussion, and Voting

DR. LEWIS: We will now proceed with the questions to the committee and panel discussion. I'd like to remind the public observers that while this meeting is open for public observations, public

attendees may not participate, except at the specific request of the panel.

We will have three discussion questions and three voting questions. Some of them have subparts. We'll start with the first discussion question. If you have a comment to offer, please raise your hand to be recognized.

Discussion question 1, discuss the effectiveness of Makena on recurrent preterm birth and neonatal morbidity and mortality. Dr. Shaw?

DR. SHAW: Hi. Thank you. I guess this is a comment and potentially discussion, that the sponsor might like to respond to this comment. I can refer, actually, to Jia Guo's slide number 3, which has the Trial 003 study design. When I think of the effectiveness of Makena, we have these two trials. I've heard a couple people talk about Trial 003 as a well-powered, well-designed trial. But when I look at the trial design that's on Guo's slides, number 3, that was powered based on a baseline rate that did not apply.

I understood earlier that the DSMB did look at

overall event rates, lumped, and they would have known early on that the baseline rate was off; that instead of the expected 17 percent for the neonatal composite index, they were seeing a background rate of about 5 percent, so a third. And the same thing for the reduction of the preterm birth; instead of the background rate of 30 percent, they were seeing something maybe lumped at around 11.

Over the 9 years that enrollment took place, I'm sort of confused as to why that might not have been -- it must have been evident that it was no longer set up to be a confirmatory trial. It was underpowered. It was terribly underpowered.

So I feel like I can only consider the evidence of the first trial in terms of a trial that was adequately powered to detect efficacy. So we're sort of sitting in a very similar place in the sense of one adequately powered trial. That's basically just a comment.

DR. LEWIS: Others, discussion?

DR. NGUYEN: May I respond to that comment? Christine Nguyen.

DR. LEWIS: Yes.

DR. NGUYEN: When we power a confirmatory trial, the best evidence we go on is the treatment effect that we see in the approval trial. We can't predict in advance what the results of the confirmatory trial would be. I mean, you can't look into the future. I can't answer why the data were not reviewed formally and assessing about event rates and what have you.

But it doesn't make 003 not an adequate and well-controlled trial. It was powered based on the best available evidence. So again, when we're looking at 003, we're trying to find a drug effect, so I think it's important to look at all the data in front of us.

DR. SHAW: Absolutely. I think speaking from what I -- and I might have misunderstood, but a lot of times DSMBs, we have to monitor event rates because we all do the best we can. And frequently, especially when we go into a new population, we need to realize we may have powered on the wrong thing, and generally background event rates would be considered, and maybe it wasn't. But that's still a piece of the trial, and

its hindsight could be 20/20, but it's just something to be aware of.

We can't refer to that -- you did the best you could, and that's not in question, but this was a trial powered for a different population than the one it was inevitably --

DR. NGUYEN: So I would comment that the eligibility criteria was the same as 002. So the intention there is that you enroll the same population. And again, we can't predict in advance what the results will look like for 003.

Another thing I would also clarify is we approved Makena based on the findings of 002, so we expect the treatment effect to be similar. So we're not looking at a totally different population or somehow looking for different outcomes. We're looking for a verification of the drug's effect.

DR. LEWIS: Okay.

DR. GUO: Jia Guo from FDA. I have a comment on that.

Could you please get my slide 27? Go back one to 26. When we talk about a power of the study, that's

a very important concept at a design stage. We know the power is the conditional probability, but at that time we have an expectation of the treatment effect we will observe in this trial.

We're not talking about the retrospect -- when people say the study and the power, we commonly think about the retrospective calculated power based on the study results.

DR. SHAW: I'm sorry. I just want to be clear that that was not my question about retrospective power. It's just understanding a baseline rate used for the power.

DR. GUO: Yes. And if you look at Trial 003 results and look at a confidence interval based on applicant's relative risk reduction, you see for the neonatal composite index, the relative risk reduction, actually, for the neonatal is positive 12 percent, and the confidence interval, the lower bound, is minus 28 percent, which actually does not cover that 35 percent, what they expect to observe in the study. So in that way, this study is not underpowered to detect their original plan for the relative risk reduction.

DR. LEWIS: Okay. If we could show the discussion point again, and I think Dr. Reddy was next, the first discussion question for the committee.

DR. REDDY: Just to build on what Dr. Shaw said, they did not look at the event rate. I just wanted to make sure -- the DSMB for 003, because I asked that question.

DR. SHAW: There were two different answers, actually. It was confusing.

DR. REDDY: When I asked, one of my first questions was, for 003, did they at any point go to the DSMB about the event rate or to the FDA because the event rate was lower than expected, and the answer was no.

DR. KROP: [Inaudible - off mic] -- charged to look at efficacy and did not comment to us about event rates. That was not their charge for the committee.

DR. SHAW: But I was confused because at one point, I thought I heard you say the overall rate was looked at, not the efficacy, which would be by arm.

DR. KROP: I think they knew the overall rate,

but that was not -- I mean, they weren't telling the sponsor you're underpowered; you need to go do something. I think at this point, this is a rare disease, and the idea that even if we were powered to go do 3500 patients, it wouldn't have even been possible. It would be another 10-year study. So I'm not sure whether that would help the situation.

DR. REDDY: I wanted to clarify that. But in terms of question 1, to me, the focus is preterm birth. I think it's an important outcome because we know preterm birth gestational age is directly related to neonatal morbidity/mortality. So I think, to me, I'm focusing on preterm birth and gestational age at delivery because we know that is directly related to morbidity and mortality.

Then for me, I'm interested only in the 003, the U.S. portion. I feel the other portion is not applicable to us here in the U.S. So given being focused on 002, which was a well-done RCT of American population and U.S. PROLONG, which more reflects the U.S. population, I think there is evidence that Makena is effective.

DR. LEWIS: Dr. Bauer?

DR. BAUER: I'm going to be the devil's advocate here because I'm going to take just the opposite. I'm going to suggest that actually 003 was actually the more properly done trial, and that you can't just ignore the fact that the trial enrolled people at a lower risk. In fact, the right question is, was there any evidence that the drug had differential effect in the lower risk people as opposed to the higher risk?

Both in 003 and in 002, there was no evidence that the drug had any better or any worse effect, depending on what the baseline risk was. It's a very important issue that Dr. Shaw brought up about the event rate because if you're studying a lower risk population, you have less of a likelihood to show a meaningful difference. But remember that the power calculation for 003 said that they wanted to find a 30 percent or greater reduction in the risk of their primary endpoint. In fact, their confidence intervals excluded that interval.

So I would not argue that that was an

underpowered trial. In fact, I'm going to take just the opposite. I think that there are questions about the much older trial. Really, an event rate that's almost twice in the placebo group of what you would expect, based on other populations, to me is not yet explained, and there are also differences in randomization that we can't account for, particularly that purports to women that had more than one preterm labor. So I think we could call into question the validity of actually 002 as much, or in my opinion more than 003.

DR. REDDY: I understand your concerns. I'm worried about 003 in terms of the neonatal morbidity and mortality was so low. We can't poo-poo we do not know the underpinnings of preterm birth in this country. We heard about all these risk factors, but even if you count for all these risk factors, there's still an elevated rate controlling for all these things.

Really, Ukraine and Russia to base majority of patients in 003, it makes me feel very uneasy because they had a very low rate. I want my neonatology

colleagues to comment on the extremely low rate from very preterm births in this study.

DR. LEWIS: I know Dr. Davis is up next, but if somebody wants to quickly comment on Dr. Reddy's observation? Is there a neonatologist in the house?

DR. DAVIS: I think we agree that the primary reason to use this drug is to prolong pregnancy and minimize neonatal morbidity and mortality. None of that was shown in either trial because the rates overall were quite low.

We as neonatologists see the bulk of our morbidity and mortality in babies delivered less than 30 weeks gestation. I think most NICUs in the United States have survival rates well over 90 to 95 percent in babies over 30 weeks gestation, and we have the most concerns and see the most severe illness in preterm infants who are delivered less than 28 to 30-weeks gestation.

Most of our neonatal trials studying major morbidity and mortality are limited. Usually we go from 23 to 29 weeks gestation, and we don't enroll anyone over that because the rates of complications get

much lower, and then you can't get enough patients and power your trials properly.

So I would suggest that even if you were to do another study, the rates here are so low that you could never power a study to find a significant difference, at least in my mind from looking at these data. If you look at the deliveries at less than 28 weeks gestation, which is what we really worry about the most, if anything, it was slightly higher in both 002 and 003 in the Makena group. It doesn't look like it was statistically significant, but there was certainly no benefit.

What it suggests, we've talked about the multifactorial nature of preterm delivery, and it may be that more mothers at less than 28 or 30 weeks have inflammation, infection, et cetera, Which we tend to see after delivery, and maybe the pathogenesis is somewhat different at older gestational ages. But I think from this standpoint, the rates are incredibly low, and if you're using the drug in order to improve neonatal outcome, you can't demonstrate that.

I do agree that late preterm infants do have

higher rates of long-term morbidity and mortality, but the question then, which we talked about earlier, if you're getting us from 36 weeks to 36 and five-sevenths, is that a meaningful clinical outcome that you're going to be able to demonstrate a significant difference in that 6-day period, and is the risk of injecting this medication -- and I feel better about seeing the 4-year follow-up that there is no obvious signal of any differences, but does the risk potentially outweigh the benefits of that extra 5 or 6 days when you're talking at somewhere around 36 to 37 weeks?

I would have a really, really difficult time either designing that trial or figuring out how to interpret those data.

DR. LEWIS: Thank you. Dr. Gillen?

DR. GILLEN: Thank you. I'll take what I would consider to be the easier one first on this, and that, no, I don't believe that effectiveness for neonatal morbidity and mortality has been established. I think gestational age has been and is a surrogate here for neonatal morbidity and mortality.

There have been changes in evolutions in what we would define as an adequate surrogate, depending upon the time frame for the gestational age at the time of birth, but neither study has demonstrated, in my mind, anywhere close to efficacy on neonatal morbidity and mortality.

Now, with respect to preterm birth, I agree wholeheartedly with Dr. Bauer in that there are still questions remaining about the placebo control rate in the first study. It's an anomaly that has yet to be explained as to why it was so high, and the observed rate at less than 37 weeks was effectively around where previous studies, placebo arms, were sitting, and that has not been explained.

If one is going to say that the reason that there's a lack of replication, which this is the underlying argument here, and this is where I began my very first question of the day, is because there's a difference in the patient populations, I have yet to see one subgroup where the two started to be compatible with one another.

Even in a data-driven world, we can't find one

subgroup where there's effect modification or evidence of that effect modification that's sitting here. Cutting it by U.S. population, black versus non-black population, that is yet to be demonstrated to me. So I believe that even with respect to preterm birth at this point, that there is fairly weak evidence, I would argue, in terms of effectiveness.

DR. LEWIS: Anyone else? Question 1?

(No response.)

DR. LEWIS: So on the question of effectiveness of Makena on neonatal morbidity, there seems to be no one commenting that Makena does affect neonatal morbidity and mortality on recurrent preterm birth. There's some range of opinion in terms of whether you should value 002 or 003 more so; or whether either of them show effectiveness.

Dr. Lindsay?

DR. LINDSAY: I just wanted to weigh in on the issue of the efficacy of Makena recurrent preterm birth, and I really wanted to ask a question based on a couple of things I've heard about the independent patient meta-analysis data that's going on.

My question is -- and this is just a general comment -- when we get the results from independent patient meta-analysis, will that trump the results of what we get from the randomized clinical trials?

One speaker made the comment that maybe we should wait for our deliberations until we have those results, and I would agree. I have to be candid. I've been prescribing the medication for a number of years, but in terms of looking at the evidence and looking at the data, it's really kind of hard to say that it's been very effective if you look at the data very critically.

I'm just asking is that meta-analysis going to be a tiebreaker, or I wanted someone to kind of make a comment about whether the independent data meta-analysis will trump the results of these two well-conducted, randomized-controlled trials, because that would help me in my deliberations.

DR. LEWIS: Well, that's a good question, and it kind of does feed into our discussion question 2 about a confirmatory trial, if that's to be designed. So I think, if you don't mind, we'll kind of fold that

in.

Oh, I'm sorry. Go ahead, FDA.

DR. JUNG: Hi. My name is Dr. Taehyun Jung from FDA, Office of Biostatistics. I authored the meta-analysis of the two published studies in the briefing document. The FDA reviewed two published studies. One is a published in the American Journal of OB/GYN in 2018, authored by Romero, et al. This study used vaginal progesterone, and the dose was ranging between 90 to 200 milligrams daily. There were 5 studies that was used for meta-analysis, and that was administered by intravaginal.

This study was limited because the study population was different from study 003. The Romero study had spontaneous preterm birth, but it was only 30 percent. All of the subjects had 100 percent short cervix that was defined as cervical length less than 25 millimeters. And the Romero study didn't use the approved dose, that is 250 milligrams weekly.

Also, the authors conducted a post hoc analysis on U.S. and non-U.S. white population and black population. The white population showed a higher

risk reduction compared to the black population. The black population showed a relative risk of 0.86, but it crossed the reference line, so there was no difference. the U.S. population and both non-U.S. showed significant risk reductions, but the U.S. population had a higher risk of preterm birth compared to the non-U.S.

DR. LEWIS: I'M sorry. Could you just clarify that again? So you're talking about vaginal progesterone in a meta-analysis? Was Makena in this?

DR. JUNG: The study published in 2008 was using vaginal progesterone only.

DR. LEWIS: Vaginal only. Okay. Thank you.

DR. KIM: I'm Clara Kim from Office of Biostatistics. I just wanted to clarify that the meta-analysis that Dr. Jung is talking about is the one that's included in the backgrounder. I think the patient-level meta-analysis that you're referring to, we haven't gotten a chance to review it. So how much we rely on that, I think that would be a review issue.

DR. NGUYEN: So if I may provide some guidance, we rely on the most robust strength of

evidence when making our decision. So unless we think that the individual patient data meta-analysis, which I suspect is going to be a little more heterogeneous than the two adequate and well-controlled prospectively designed trials, it will be hard for us to think that would trump the very robust evidence from the two trials we have in front of us.

So I can't answer it for sure, but you just kind of eyeball the robustness of the evidence that are generated from the two different analyses, that that would sort of guide how we handle those data.

DR. LEWIS: Dr. Orza?

DR. ORZA: One possibility I think that could come out of the IPD meta-analysis -- and again, I haven't seen the results either; I'm not privy to those -- is that it might not contribute to these questions specifically, but it might identify, for example, a legitimate comparator to get us out of the jam of having to use a placebo.

DR. LEWIS: Dr. Eke, did you have a comment as well on this question? No?

Okay. Are we ready for question 2? Question

2, if a knew confirmatory trial were to be conducted, discuss the study design, including control, doses of the study medication, efficacy endpoints and feasibility of completing such a trial.

Don't all speak at once. Yes?

DR. JARUGULA: As the industry representative here, I'd just like to comment. Having seen the evolution of this development, the study 003, how long it took to complete the study, given the recommendations of the societies and also about the ethics of using placebo in this, I think it would be extremely hard for any company to conduct such a study. You've seen that study 003 background rates were much, much lower than anticipated, and yet we tend to use that study as a basis to utilize the findings of the other study.

So I don't know. I'm still conflicted on that. But leaving that aside, I think conducting another's study, a well-controlled, double-blind study would be extremely difficult. I would venture to ask the committee and others to discuss other possibilities here, either finding a subpopulation or any other

possibilities.

DR. LEWIS: Dr. Gillen?

DR. GILLEN: Possibly controversial thinking out loud here, but the sponsor has very clearly articulated that they don't believe that another study would be feasible given the fact that accelerated approval was already granted, and it is very hard to recruit from the same patient population. I would conjecture maybe that accelerated approval was potentially given too quickly in this case and has convoluted this problem.

I guess a question for some of my clinical colleagues around the table is, if approval was withdrawn, could this study be done, and done appropriately, with a representative patient population to attempt to confirm, if you will, Trial 002, which is what the purpose of 003 was, and what I've been told is that could not be done because of the changing patient population and the difficulty of recruiting.

I'm not really giving an answer here on the feasibility, but I understand the logistical difficulties, and I think we've been conditioning upon

the fact that the accelerated approval is granted and will stay granted. And I think we need to think about the two hypotheticals to say, what if it wasn't there, could we do an adequately controlled trial and actually get to an answer?

DR. LEWIS: That's kind of what we're asked to talk about in question 3. What are the potential consequences?

Dr. Orza, and then Dr. Wing.

DR. ORZA: I'm having trouble articulating this idea, so bear with me. But in study 003, I'd like to see data about a control group, what was going on out there with women at high risk for premature birth outside of the study to understand what the baseline might have been because the women in this study weren't just getting an injection of placebo. They were getting weekly attention and care. And it could be that because both of them got that, regardless of whether or not they got the drug, that that actually is the answer to why the rates were so low, both in the placebo group and in the control group.

So we might have in fact discovered the way to

make this better, completely independent of the drug. So I would like more information about what was going on outside of the trial to try to understand better what was going on inside of the trial, and to help us think about what the next study should look like.

DR. LEWIS: Thank you. As I understand it, in 002, though, the same thing, their placebo group also got weekly attention. No? Yes, they did.

DR. ORZA: Right, kind of setting that aside because I don't know what happened there.

DR. LEWIS: Oh, okay. Dr. Wing?

DR. WING: So my thoughts are all over the map, so please bear with me. I'm going to talk to issues related to both questions 2 and 3. I'm going to leave an open-ended question, first, for people who are more informed than myself, which is one of the elements of question 2, which, is 250 milligrams of this drug the right dose? And it's perhaps what we're seeing in the differences of these trials related to the dosing.

I'm going to throw another variable in here, in the discussion, because I really am going to stir it all up, is whether or not the timing of administration

of these drugs also affected the results and can account for the discrepancies in the two trials. So that's me as a clinical trialist talking about design.

I think feasibility, we're going to bash it around quite a bit. I think the ethics of doing a placebo-controlled trial when this drug has had FDA approval is a non-starter, at least in my opinion. It's just not going to happen.

So then we have to go to the alternative, then, which is if you pull the approval of the drug and say we're going to conduct the trial, then you've got to consider the legal implications, which the FDA I think has argued, at least in my mind, appropriately that that would be an okay thing to do. But there will be clinical and political consequences of that because, clearly, the clinical consequences, as a clinician, we're desperate as MFMs. Perhaps, I'm less desperate now because I've walked away from the bedside, but we don't have anything that's really good; just stop this problem that causes insufferable pain. So we succumb to emotion as a result of that.

I think Sean said it best, that the clinical

response out there in the field is going to be that our brethren will start prescribing other versions of progesterone, whether it's vaginal, or oral. or some other compounded injectable, and they may all at once; that that could happen or they could put in more cerclages that were unnecessary. So in that regard, I think we're also looking at other ethical implications here, where we're doing harm where we shouldn't be.

As physicians, we take these oaths to do good and also do no harm, so I think we have to ask ourselves what good are we really doing here? Then I think the political implications are clearly, we know that there are disadvantaged populations in this country, and we have data. The black and white says that the 17P somehow prevented some recurrent preterm birth in a disadvantaged patient population. That to me stands above all else in considerations of these trials.

DR. LEWIS: Dr. Hickey, a new confirmatory trial?

DR. HICKEY: Well, I'm going to say Dr. Wing stole much of my thunder --

(Laughter.)

DR. WING: I didn't mean to.

DR. HICKEY: -- pretty much all of it. I would agree we are fairly desperate in terms of finding solutions for people, and that was, I think, our difficulty in the PROLONG trial when you try to enroll a patient and say we have a potential preventative agent for you or you can roll the dice and do placebo. So I think feasibility of a placebo arm is almost nonexistent.

I do like Dr. Caritis' idea of looking at different dosing agents, and that would probably be my goal, would be to do dosing, but also to really follow the PK/PD and see if we see is there a threshold level that we need to reach in women; because I can tell you, looking at our practices versus other practices, that people really ramp up that use of progesterone when it's not working beyond that recommended dose, and they do see benefits, so they keep doing it.

So clearly, I think there's some anecdotal evidence that perhaps looking at dosing may be part of our issue, and I'm really hoping that some of the

individualized data helps us pull out that subgroup that really is going to be the beneficiaries of this work.

DR. LEWIS: Thank you. Dr. Reddy?

DR. REDDY: I agree, A placebo-controlled trial cannot be done in this country given everything that's been said. Patients, they'll go to compounding. They'll use other means to try to decrease their risk of preterm birth. But we definitely need more evidence. So even if we can't do an RCT, I agree with PK/PD studies, dosing studies. There have been studies where they use 500 bid in France and found, in fact, it did not work; it did not decrease. So there is some literature out there.

I think the EPPPIC meta-analysis that was mentioned, we need a well done IPD of Makena, not vaginal progesterone. If a trial is desired, there are some options. You could have a control group using vaginal progesterone; it's not great. Also the UK, like I mentioned, I don't think they're using Makena, so that's another population.

If there's some way to gather more

information, so a registry of patients who've had previous spontaneous preterm birth, the data that was presented, it was previous preterm birth. So the question was how come only 39 percent of women are getting Makena if they've had a previous preterm birth? So 30 to 40 percent of preterm births are iatrogenic; they're not spontaneous. So we need high quality data, which we're lacking, so the eligible women, an and observational study.

As physicians, as a clinician, we have to counsel patients. We have to incorporate this PROLONG information. And it is going to change counseling because there is evidence. We have to incorporate that level of uncertainty. We can't be this clearly decreases the rate of preterm birth by a third; now, it has to be nuanced based on other factors.

DR. LEWIS: Thank you. Dr. Drake?

DR. DRAKE: Matthew Drake for the Mayo Clinic. Unfortunately, I also think this is an unfeasible trial unless you can, a priori, identify a group that is going to have a 55 percent risk of preterm birth. If you can't, a priori, identify that group, which it

sounds like it's probably going to be hard to do, then I think it's going to be essentially impossible to do this.

One thing we haven't really heard about is whether this -- maybe we did, but I don't recall hearing it, whether 17P undergoes any metabolism and whether that's different between any patient populations; whether it is or isn't metabolized faster in an African American population, versus a Caucasian population, versus an Italian population, versus anything like that.

Some presented from the audience, looking at pharmacodynamic/pharmacokinetic data, but whether that metabolism is important and leads to differences in the level of 5 up to 56 that they measured is, I think, perhaps very important and may underlie some of these findings. So if there was a way of identifying and addressing some of those issues, it could be important.

DR. LEWIS: Thank you. Ms. Ellis?

MS. ELLIS: Hi. Thank you. I came to this meeting. I'm the patient representative. I'm the only one at this table without an advance degree or any

degree at that moment, but what I do have is a personal history of preterm labor, and I was able to, with things that are not approved anymore and bed rest, bring my second daughter to deliver at 38 weeks. Then she herself has had a preterm labor. So my grandson, we've had some early intervention and difficulty.

So this is a topic very near and dear to my heart, so I'm trying to bring in the personal, human element as we talk about this. Reading through the briefing materials, the statistical considerations were just really over and above what I could comprehend, and I came here seeking clarity and more confused than I was when I showed up, as I'm sure many people here are.

This trial seems to me to be about time. Whether or not that time actually is clinically meaningful is something that's kind of debatable here as well. And something that Dr. Reddy said earlier today was about what's missing for me is for the people who have had a previous preterm labor, how did this drug help them

get more time?

I mean, as a whole group, we've got those

results, but what are the results if people are starting this at different times? So we don't know -- it's hard to tie everything together. So if there were some kind of registry or something, that you brought up, having this information might be useful going forward. Thank you.

DR. LEWIS: Thank you. Dr. Davis?

DR. DAVIS: I would agree that it's going to be impossible to do the same trial for a third time, nor since the first two trials didn't have dramatic impact on neonatal outcome, I don't know that I would want to do that. But if there are opportunities to enrich the population that you're studying -- and I think Mat mentioned before was appropriate -- maybe one previous preterm delivery alone is not adequate to predict, in a meaningful way, the impact of preterm delivery.

We now have an obesity epidemic that's different between the two studies. We have a more substance use problem than we had before. And maybe you're identifying high-risk populations and doing it in a way that, okay, you had a previous preterm

delivery at less than 35 weeks, that's one point; less than 28 weeks, that's two points; you're African American, and that's a point; you're obese, that's a point; your smoking history, that's a point.

Maybe there's a way of enriching that population so you can get to a much higher risk group because maybe that will have an impact at that stage. And I do like the idea of either a dose escalation trial, which then might preclude use of a placebo, or potentially a placebo trial with a different population and a different trial, but I definitely would not necessarily do the same trial over again.

DR. LEWIS: Thank you. Dr. Eke?

DR. EKE: Thank you. I kind of wear three hats, being an MFM, a clinical pharmacologist, as well as a clinical trialist. I keep scratching my head because looking at what we have facing us right now, I could not agree more with my colleagues, it's going to be very difficult another trial, basically looking at the logistics, and the ethical as well as the legal aspects to this.

What we have left would be to see how to get

that subset of patients who benefit from this drug. I believe that there are some people who benefit; not everyone, some who do benefit from the drug, and our job should be to look for those patients to give this drug to.

Dr. Caritis talked about the dose response, which I totally agree with. When he discussed that idea a couple of years ago, I was on board with it as well. I was surprised that there was no PD aspect done for this drug, so that is one aspect.

An aspect, which no one has talked about, which Dr. Drake kind of mentioned briefly, is the pharmacogenetics of this drug. Tracy Manuck, who is at UNC, there are two landmark papers that she's published. One of them, she actually used samples from patients from the Meis trial.

She went back, collected samples from these patients and looked at their genetics. Is there something within these patients that actually make them respond more, which she called responders versus non-responders. That study showed that some people that actually responded more, they had some genes that

were over-represented versus those that were not.

So that is something as well we could look at, and see patients who really need this drug, and whether we can say a patient who gets this drug will be African American, has these kind of genes, blah, blah, blah, and that will kind of help us streamline whichever kind of study we need to do in the future.

DR. LEWIS: Thank you. Dr. Smith?

DR. SMITH: Sure, just a comment. Neonatologists are guilty of this, but it seems a little bit late in the drug development pathway to be talking about trying to find the right dose of the medicine after two huge randomized-controlled trials. I also worry about the feasibility, especially if you start looking at randomizing against a non-FDA approved therapeutic approach. If anything, that group is going do a little bit better than maybe placebo, and your sample size is just going to have to be that much bigger.

DR. LEWIS: Dr Shaw?

DR. SHAW: Hi. Yes. I guess I just wanted to comment on the potential design if we could do a trial

for further study. I feel like I'm hearing discussion of what might be an observational study, some kind of pragmatic study of people or registry. But I would say that a study in which we want to gain information can't be observational. I think these two well-controlled trials showed us when we equated the care on the two arms, we couldn't see a difference between black and white or education, high or low

So if we can't see any large differences in these pretty big groups of well-studied people, I'm not sure how we could imagine using regression and adjust our way out of the obvious confounders if they're going to be in an observational study. So I don't have confidence that we'll get clarity from a study that's not a controlled study or some kind of observational registry.

DR. LEWIS: Anyone else? Yes? Dr. Wade?

DR. WADE: Before we move on to question 3, I would just second what others have said, but I do believe there is lots of exposure out there. We saw that in the Sentinel review, so it would at least steer us to how much we're going to work towards a

randomized-controlled trial if we looked at the observational data. We haven't heard anything specifically about all this. exposure leading to any reductions in preterm birth, so it seems like that exposure data is out there, whether or not we've looked at it on a state-by-state basis, or not.

Then I agree with everyone that we are trying to figure out who this highest risk population is, and in reviewing about the progesterone levels and how there is this broad variation of progesterone levels, almost 10-fold across women that were receiving 17-OHPC, it feels like there may be some more information there about what's driving the variation. Is that something inherent to the patient or is it something inherent to the dose of the drug? So there may be more information there that we could tease out.

Lastly, I looked at table 22 in the appendix, which looked at the U.S. subset of Trial 003, comparing Makena to placebo in all these different high-risk stratification groups. Although, I'm sure these differences are not necessarily statistically significant, the earliest gestational age of the prior

preterm birth being in the 0 to 20 weeks or 20 to 28 weeks, that seems like a huge risk factor. The Makena group actually had more.

So there isn't even a balance of -- when my eyes go to what are the highest risk women in these groups using Trial 003 U.S. subset, the Makena is not performing well in what I'm drawn to as my highest risk groups. So I think there still is really a lot more work to be done to even figure out how to design what the next step would be.

DR. LEWIS: Thank you. Dr. Hunsberger?

DR. HUNSBERGER: I just have to say I agree with Dr. Shaw. I don't know how we'd figure anything out without a randomized study. And especially after listening to this whole discussion, I'm in equipoise, and I guess I wonder how the clinicians are kind of not in equipoise given we have these two randomized studies where they give very different results. How do counsel a patient given this data and not be in equipoise?

So to me, it seems like you have to have a randomized study to figure this out. I just think the data doesn't help us right now.

DR. LEWIS: Thank you. Dr. Reddy?

DR. REDDY: Well, to answer the point about being a clinician, unfortunately, in OB, that's a lot of what we have to do. A lot of the medications we use have not been studied in pregnancy. Even something as basic as chronic hypertension in pregnancy, we're like, well, you could be on meds, but there is no evidence that that works. In fact, quality evidence, the American College of OB/GYN says you should be taken off your medicines.

So I think we've gotten used to that. I think the PROLONG data is important, and it will be incorporated, and it will be explained, there's this one trial that shows this, there's another trial that shows that, and what the level of certainty is.

But one thing Michele Orza said, that now it's been bothering me for the past few minutes, is you were talking about weekly visits, the Ukraine and Russia, what else do they do? Do they put in cerclages, monitor the cervix every week? I have no idea what else they're doing for these women, so it may not be a study of just that medication, of just Makena, because

the way they practice is completely different than here. Even in the neonatal outcomes, what we call NEC, at least in the Maternal Fetal Medicine Units Network, there are strict definitions. The data is rigorously collected, but I'm not sure what happens in those countries.

DR. LEWIS: Thank you. Anyone else?

(No response.)

DR. NGUYEN: Dr. Lewis -- I'm sorry; Christine Nguyen -- I just want to remind everybody the clinical practice can vary, especially when we have so many sites. Please remember that there is a protocol in place to standardize practices. For example -- and I don't have details for the protocol -- certainly, I can't imagine Russia putting a cerclage and not the U.S. So just to let you know, there's a protocol in place that's standardized the care as much as possible.

DR. REDDY: Well, I think that's really important to ask then, was their standardized management? Probably not. Can someone from PROLONG answer about the management?

DR. KROP: Yes. I'd like to call up

Dr. Blackwell.

DR. BLACKWELL: Hi. Sean Blackwell from Houston, Texas. The research protocol for PROLONG specified research procedures, but clinical care was at the discretion of the treating attending clinicians. So there was not a standardized protocol for things such as screening for transcervical length; the management if there was a short cervix, and the nature or degree of tocolysis, or other obstetrical management options. It would be the randomization process, they would account for that, but the research protocol -- much in the same as in the Meis study, we did not standardize clinical protocol related to these obstetrical interventions.

DR. KROP: I think it's important to remember -- you brought up the differences between Russia, Ukraine, and the United States -- there is a very different healthcare system. It's a universal healthcare system. There's a social safety net that exists in those countries that doesn't exist here, and there is also preventive measures that are put in place that are far more extreme than we have in the United

States. They have nurses go out to patients' houses. They have pre-pregnancy counseling and getting patients on vitamin early. In the U.S., we of course have a bias in the other direction of putting on these healthier patients into the study just because of the existing standard of care.

DR. LEWIS: Thank you. Well, maybe I'll just weigh in that it's not just what the doctors do, it's what the society is like. A single pregnant woman in the United States is not necessarily the same as a single pregnant woman in the Ukraine or Europe: what kind of family support they have, what kind of neighborhood support they have, how much they have to work to make a living, food security, and housing security. All of those things I think have bearing.

Anybody else on question 2?

(No response.)

DR. LEWIS: Okay. Question 2. I think that there is pretty much agreement about the feasibility of completing a randomized-controlled trial being extremely difficult, as some feel that that's the only valuable data, really, that we're going to get, that an

observational data kind of study is not going to be helpful; and several people weighing in on the importance of getting pharmacokinetic data, which we really don't have, and that perhaps some sort of comparative trial with other kinds of progesterone could be a type of study design that might be useful, being a feasible thing.

In terms of other kinds of ways to design the study, maybe looking at an enriched population of high-risk patients as they exist today. We have a much more obese patient population than we did before. Substance use rates are different. Other ways to identify a group that might be helpful or might benefit from the drug, pharmacogenetic studies, dose-response studies; that, really, we just don't have data at this point that might help us understand the differences between the outcomes in study 002 and 003.

DR. GILLEN: At least from my standpoint --

DR. LEWIS: Sorry.

DR. GILLEN: -- the infeasibility of a randomized-controlled trial, what I am seeing is that's conditional upon the current accelerated approval still

being in play. I think the dynamic changes dramatically if you pursue removal of that approval. So that's me personally; I'm seeing that.

DR. LEWIS: Sure. So that could be, in fact, one of the potential consequences of withdrawing Makena on patients, and a clinical practice, one could be it's feasible, then, to do a placebo-controlled trial.

Does that reflect your view?

(Dr. Gillen gestures yes.)

DR. LEWIS: Okay. So we'll move on to question 3, which I just sort of summarized some of what you said a couple of times, discuss the potential consequences -- a very important point -- of withdrawing Makena on patients and on clinical populations, clinical practice. Let's have more of a discussion there.

Dr. Orza?

DR. ORZA: Just a technical question. It was referenced that if this were taken off the market, that people would be compounding it anyway. How does that work?

DR. LEWIS: FDA?

DR. NGUYEN: Christine Nguyen. This is where we need your input, particularly patients who are caring for pregnant women and how they're counseling their patients, based on the data from the two trials.

DR. LEWIS: Ms. Ellis?

DR. ORZA: I didn't understand that. My question was if this is -- so it's the withdrawal of this specific drug, but legally people are still allowed to compound it? Is that how it works?

DR. NGUYEN: I'll give you a very brief answer. Under certain circumstances, hydroxyprogesterone caproate, so the active ingredients, may be compounded. But that's pretty much all the details that I can provide regarding compounding. I think it does answer your question.

MS. ELLIS: So my follow-up question to Dr. Orza's is, do we have any data or any idea of what was the compounding usage prior to the accelerated approval, from the 2006 meeting when people were discovering that this might be helpful to the approval in 2011?

DR. NGUYEN: Christine Nguyen again. If I may

just remind the audience, I understand the compounding issue is important, however, it is not before the committee today, so that is not something we could be prepared to discuss.

MS. ELLIS: I'm just curious because one of the questions is what happens if approval is withdrawn, and it just is something that makes sense that it might happen. So I was just curious about that time frame, if we anything, if anybody knows anything about what was happening.

DR. LEWIS: I'll give FDA a minute or I'll give sponsor a minute. Are you ready? Go ahead.

DR. TSAI: Huei-Ting Tsai, FDA. Can we put up slide 22 in drug use, slide 22? This slide, the brown color shows the form of HPC use. If we look at usage before 2008 through 2011, in our data, the Sentinel analysis showed around less than 5 pregnancies per thousand pregnancies used the compounded HPC during the second or third trimester.

DR. KROP: So in 2005, there was a survey done of 572 maternal fetal medicine practitioners, and 67 percent of the respondents use progesterone at that

time to prevent preterm birth. This is before Makena was on the market, so this is obviously all compounding. Then there was a 2007 survey done of 345 OBs that showed 74 percent recommended or offered progesterone, and 92 percent of users began recommending it within three years of the Meis trial. There were two publications. One was by Nest in AJOG, and one was by Henderson in AJP.

DR. LEWIS: And that was any progesterone or that was HP?

DR. KROP: It doesn't specify. I think it was 17-hydroxy.

Dr. Sibai, can you comment on that?

DR. SIBAI: In the study that I mentioned about 5,400 women, every single one of them received the compounded. Makena wasn't approved by that time. In addition, during this time, I received a grant from the CDC to study responders, and we used the compounded. So if Makena is not available, I assure you every physician in the United States will find every way possible to use the compounded, or much worse, they're going to see start offering cerclage to

these women, which in my opinion is going to be catastrophic.

DR. LEWIS: Thank you. Dr Hickey?

DR. HICKEY: I was just going to say, clinically, when Makena was first approved, the price point also wasn't at an appropriate level for some people if they were paying out of pocket, so people continued to use the compounding form. And that would be, my expectation, if this was taken off the market and is not approved, then people are going to look for that equivalent wherever they can find it. Based on what we know with safety and poor outcomes, compounding pharmacies are not regulated, and I think that poses a serious health risk. But people will look for progesterone wherever they can find it. They won't just say, I'm not going to treat you.

DR. LEWIS: Dr. Lindsay?

DR. LINDSAY: Yes, I would second that comment. For years in our state, Makena was not approved, and you're going to see patients who are going to present with a history of preterm labor were using the compound. I think if it disappears tomorrow,

that would be the same course that we would take. We would be giving patients compounded 17-OHP.

DR. LEWIS: Dr. Shaw?

DR. SHAW: I'm thinking about this question about the potential consequences of withdrawing, so I'm thinking of the population that bears the higher burden of preterm birth, mainly a disadvantaged population that tends to be lower education, lower economic status, perhaps self-pay insurance. This is a population that we're seeing -- we have two trials now for which we're debating the efficacy results in. We're concerned about 002. We can't explain the really high background rates from the placebo. We have 003. There's a lot we can't explain there.

We're going to tell this disadvantaged population that this evidence is good enough for you. In some ways, if we can turn this political piece around and argue that side of the story, how do we give this population the best chance at hard scientific evidence? Because I can tell you, people are terrible at judging risk. It's an emotional decision. You can have the conversation, but you're going to take that

population that's not used to doing math and you're just going to start throwing statistics at them, and they're just going to not hear most of that.

So one consequence of withdrawal is a huge signal for concern. We're not sure. A consequence of not withdrawing is keep doing what you're doing; everything's fine. So I think the consequence of withdrawing allows for a deeper dive into this question. It's just not going to be possible. There is at least one, I think, advantage for this population, the very vulnerable, premature babies who aren't going to be able to weigh their options independently. So I think it's really important to think about the vulnerability of this population.

DR. BAUER: I agree with that; excellent and well said. I would argue also that this is going to be an opportunity, if it is withdrawn, for the professional societies to really look at their responsibility, and ethical responsibility, not only to their patients but to their members to really say, in fact, at least according to the FDA, it was inadequate evidence to say that we're doing net benefit for this.

There is an ethical responsibility not to provide ineffective treatments to a large proportion of the population, and then feel good that we've done everything we could do. In fact, it sounds like to me -- and again this is not my field, but there must be lots and lots of things that we don't understand about this disease because the rates vary so much over the world.

So that just suggests some of them are probably endemic to our society, but maybe there are others that can't be. I think this is an opportunity for us to really point that out. Again, I would hope that the professional societies would lead the way as opposed to opposing it.

DR. LEWIS: Ms. Ellis, and then Dr. Orza?

MS. ELLIS: I think what's missing here for me is just solid information that would help me vote with confidence. I think the only way to get that information -- it's very uncomfortable to say this; I feel like it's the Kobayashi Maru -- is to do a trial that stratifies, that is taking a lot more into consideration. And the only way to get that trial is

for this drug to be withdrawn. However, it's a great deal of discomfort because of the women who have access and who will not have access for whatever time it takes to get that going.

So whatever the usage was in 2006 for people going off and getting it on their own, it's going to be more because of social media and mommy blogs. People are going to be talking about this. So whatever path is taken going forward, I hope that we consider the gap. And for people who are in need or at high risk for preterm labor while things are happening, that somehow something is put in place so that they don't fall through this gap.

DR. LEWIS: Dr. Reddy?

DR. REDDY: I'd argue against withdrawing it. There are subsets of this population, very high-risk patients who probably do benefit from it, women who had more than 2 preterm births; women who have delivered below 28 weeks. So I don't think withdrawing it just to do a trial makes any sense.

I think, though, it's clear -- I think everyone agrees we need to do more research and get

better information on which patients could it be a benefit for. I think we're going to just have to -- the professional organizations, the best thing they can do is help us in counseling patients properly and getting them the right information, which they can do a good job with. But I think withdrawing it would be a disaster because it would be unethical for the patient populations who could benefit the most from it.

DR. LEWIS: So we do have an opportunity to vote, so it's not that you have to weigh in yes or no, but we are thinking of potential consequences, trying to get the views out there before we actually make up our minds,

Did you have a comment, Dr. Gillen? No?

DR. GILLEN: I always have a comment.

(Laughter.)

DR. GILLEN: I do, actually.

(Laughter.)

DR. GILLEN: I think certainly the way I view my job, as a public health practitioner and a clinical trialist, is to increase the prevalence of truly beneficial drugs. I think our job is to not only give

patients choices, but to give them well-informed, empirically driven choices that we can stand behind. I think that the horse has been let out of the barn on this, and we need to pull it back in. And the only way that we can pull it back in and get to an answer on this is by having a randomized clinical trial. The only way I see that happening is to remove that approval.

There's no other way to build upon that, and we are at a place right now, you can see it on this committee, in my mind, that we don't have an answer. I mean, we hear words like "it probably works in a subset of a population" or "this works in a subset of a population." I have not seen that subset of a population yet. It has not been quantified.

DR. LEWIS: Thank you. Anybody else on consequences of withdrawing Makena for patients and clinical practice?

DR. SHAW; This is just a clarification. Dr. Reddy. I wasn't sure about if there was a study we were referring to in terms of women who have more than 2 preterm births. You said that those, we know that

works. Was that coming from a different study than we saw today or -- just to get clarity.

DR. REDDY: There's a paper about the index pregnancy, the qualifying pregnancy. So the earlier the qualifying pregnancy, the more beneficial the effect of Makena; so that's published. In terms of women with 2 preterm births, that needs to be analyzed. That, I don't know. Those women are very high risk. Those are women who, if you counsel them, having counseled women like that, you tell them the data. You can tell them about the PROLONG study. They will take it because of the fact that there's one study that shows that there could be a benefit to them.

But I feel like we do have a lack of information. I would like to see an IPD with Makena only, not vaginal progesterone, and then also prolongation and pregnancy in both groups, based on what their index pregnancy delivery was.

DR. HUNSBERGER: Just to clarify, on the paper that you were discussing, was that from the 002 study or was that from the 003 study?

DR. REDDY: No, 002.

DR. HUNSBERGER; Okay. Thanks.

DR. LEWIS: So in terms of potential consequences of withdrawing Makena on patients in clinical practice, I think Dr. Wing summarized some of that under the prior discussion, political consequences in terms of some of the high-risk pregnancies among groups of minority races, low socioeconomic status, and emotional consequences. Patients really are in a desperate situation in that setting. They may have had a friend who's used it or they just feel like they want to do everything for their pregnancy.

One other hard consequence, of course, other types of progesterone will certainly be used, and we had a lot of discussion around what those constitute, primarily compounded forms of the medication. We don't know what the price point of those is going to be, and, of course, the risk-benefit status in terms of lack of not necessarily common practices creating a quality product.

So on the positive side, consequences of withdrawing the drug could be the opportunity to get higher quality data, avoid unknown risks from Makena

use, which certainly long term, we don't have a lot of data on, and the opportunity for professional societies to take the lead in creating better quality evidence going forward.

We now have three voting questions to start to look at. If there's no further discussion on the question, we'll begin the voting process. We will be using an electronic voting system for this meeting. Please press the button on your microphone that corresponds to your vote. You'll have approximately 20 seconds to vote. Please press the button firmly. After you've made your selection, the light may continue to flash. If you're unsure of your vote or you wish to change your vote, please press the corresponding button again before the vote is closed.

We're going to go around the room for these voting questions and ask each person to weigh in. If you just are agreeing with the last person, you don't have to state everything the last person said. You can just say I agree with the last person, but I will ask for a rationale from each person.

The first voting question is question 4 from

your booklet, do the findings of Trial 003 verify the clinical benefit of Makena on neonatal outcomes? And provide a rationale for your vote. You have the option of yes, no, or abstention.

(Voting.)

MS. BHATT: The voting results, zero is yes; no, 16; abstain is zero.

DR. LEWIS: Thank you. I'm going to start on my left with Dr. Eke, and we'll go around the room.

DR. EKE: Thanks. We've seen the data presented over and over again, here today. Based on what we see on both the 17-OHPC group and the placebo group, there was no evidence that there was increased benefit for the unit.

DR. LEWIS: Thank you. Dr. Hickey?

DR. HICKEY: I concur.

DR. LEWIS: Dr. Lindsay?

DR. LINDSAY: I concur.

DR. REDDY: I concur.

DR. WING: I concur.

DR. DRAKE: Agree.

DR. LEWIS: This is easy.

DR. BAUER: Yes, I agree.

DR. SHAW: Agree.

MS. ELLIS: I concur.

DR. ORZA: I concur.

DR. GILLEN: Agree.

DR. HUNSBERGER: Agree.

DR. SMITH: Agree.

DR. WADE: Agree.

DR. DAVIS: Agree.

DR. LEWIS: Thank you. So the committee's unanimous on that question, no evidence of neonatal benefit.

Question 5. Based on the findings from Trial 002 and 003, is there substantial evidence of effectiveness of Makena in reducing the risk of recurrent preterm births? And please provide a rationale for your vote; yes, no, or abstain.

(Voting.)

MS. BHATT: The results for question 5, yes is 3; no is 13; and abstain is zero.

DR. LEWIS: Okay. We'll do the same thing, but this time, each person please state your name into

the microphone for the record when you provide the rationale for your vote.

Dr Eke?

DR. EKE: Thanks again. So I voted based on what we have with us, which is the FDA definition of substantial benefit, which based on what we have defined, Trial 003 does not meet that standard.

DR. HICKEY: Kim Hickey. I voted no because I felt the data in the study populations were disparate, and you couldn't come to a conclusion that both had substantial supporting evidence.

DR. LINDSAY: Michael Lindsay. I voted no for the similar reason. If you combine the two trials, there is no substantial evidence there is effectiveness.

DR. REDDY: I guess I have a lot to talk about. I voted yes. Substantial I guess is subjective, though, I feel that there is evidence, based on 002 clearly, and then in 003, if you focus on the U.S. PROLONG trial and the primary outcome, although the difference of the benefit was small, that's why I voted yes, taking it all together.

DR. WING: I'm Deborah Wing. I voted no for reasons previously stated.

DR. DRAKE: Matthew Drake. I also vote no for reasons previously stated. Unfortunately, the 003 trials is just not confirmatory for what was nicely seen in 002.

DR. LEWIS: Thank you. I voted yes, basically, the same reasons as Dr. Reddy.

DR. BAUER: Doug Bauer. I voted no, much for reasons that have been already stated, but I was also impressed with the consistency of the subgroup analysis across both studies, which showed no consistent subgroup where there was an effect. I was also swayed by the fact that 002 is a 20-year old trial, and I didn't feel like we were able to really understand the dynamics of that trial as well as we were able to pick apart 003.

DR. SHAW: I think Dr. Bauer stated a lot of my reasons for voting no, and just really not being able to identify the patients reliably as to which ones you would counsel to take this versus not.

MS, ELLIS: Annie Ellis. I voted yes. I felt

that Trial 002 was still very compelling, although Trial 003 was not confirmatory.

DR. ORZA: Michele Orza. I voted no for similar reasons that have already been stated.

DR. GILLEN: Daniel Gillen. I voted no for reasons I've previously stated and those that have been also stated around the room.

DR. HUNSBERGER: Sally Hunsberger. I voted no, and I'd like to just affirm Dr. Bauer's comments in just that the consistency of the negative findings in the subgroups really swayed me.

DR. SMITH: Brian Smith. I voted no for the previously stated reasons.

DR. WADE: Kelly Wade. I voted no for the same reasons, and agree a lot with Dr. Bauer.

DR. DAVIS: Sean Davis. I voted no. While I, too, believe the results in 002 and do think this was a viable and quite important trial, it wasn't confirmed in 003. And in both trials, there was a lack of any detectable impact on the neonates, which is really what anyone really cares about.

DR. LEWIS: Thank you. Okay. Next question.

This is where it gets complicated.

(Laughter.)

DR. LEWIS: So FDA approval, including accelerated approval of a drug, requires substantial evidence of effectiveness, which is generally interpreted as clinically and statistically significant findings from two adequate and well-controlled trials, and sometimes from a single adequate and well-controlled trial.

For drugs approved under the accelerated approval pathway, based on a surrogate endpoint, the applicant is required to conduct adequate and well-controlled, post-approval trials to verify clinical benefit. If the applicant fails to conduct such a post-approval trial or if such trials do not verify clinical benefit, FDA may, following an opportunity for a hearing, withdraw approval.

Should the FDA, A) pursue withdrawal of approval for Makena; B) leave Makena on the market under accelerated approval and require a new confirmatory trial; C) leave Makena on the market without requiring a confirmatory trial? You're going

to provide rationale for your vote, including the following:

Vote A if you vote to withdraw approval. That may be appropriate if you believe the totality of the evidence does not support Makena's effectiveness for its intended use, and under those circumstances discuss the consequences of Makena's removal if not previously discussed in discussion point 3.

Vote B, require a new confirmatory trial. That may be an appropriate vote if you believe the totality of evidence supports Makena's effectiveness in reducing the risk of preterm birth, but there is no substantial evidence of effectiveness on neonatal outcomes, and you believe a new confirmatory trial is necessary and feasible.

Discuss how the existing data provides substantial evidence of effectiveness of Makena in reducing the risk of preterm birth, based on surrogate endpoint of gestational age at delivery, and also discuss key study elements, including study population, control, doses, and efficacy endpoints of the new confirmatory trial, if not previously discussed under

discussion point 2, and approaches to ensure successful completion of such a trial.

Vote C, leave Makena on the market without a new confirmatory trial. That may be appropriate if you believe Makena is effective for reducing the risk of preterm birth and that it is not necessary to verify Makena's clinical benefits in neonates. Discuss how the existing data provides substantial evidence of effectiveness of Makena in reducing the risk of preterm birth and why it is not necessary to verify Makena's clinical benefits in neonates.

Do people need a little extra time to digest this before they vote? Dr. Reddy?

DR. REDDY: So when it says trial, does it mean specifically RCT or does that mean research, further research?

DR. LEWIS: FDA, please, weigh in.

DR. NGUYEN: Hi. Christine Nguyen, FDA. So when we're talking a trial here, we are looking for a trial that will provide the robust evidence needed to verify the clinical benefits of Makena. That's the overall objective.

DR. LEWIS: Is that a randomized trial or not? Is it some other kind of study --

DR. NGUYEN: Sure.

DR. LEWIS: -- because we talked about other kinds of studies.

DR. NGUYEN: Yes. Certainly a randomized trial would be the design that we would think about, but, obviously, we are always open to other ideas that can achieve the same objective.

DR. LEWIS: When you say randomized trial, do you mean randomized placebo-controlled trial?

DR. NGUYEN: Same answer as previously. Here, we're trying to verify the benefit of the drug. So however that trial could be set up to help us identify the effect of the drug to the extent possible. So again, I think traditionally we think of a randomized-controlled trial, but is that the only trial? And if any of you have creative ideas of other trials that can give us the same information.

DR. REDDY: Sorry. I think this is an important point. Let's say you vote C, does that mean that the sponsor would not have to do any more

research?

DR. NGUYEN: Correct, as far as verifying the drug's benefit.

DR. REDDY: So if you want further research done, then that's B, but you're saying it has to be the trial. We talked about various research ideas.

DR. NGUYEN: Yes, so let me just clarify B. There are two things that need to be considered for B. So when we're talking about considering the new confirmatory trial is necessary and feasible, it's necessary if you believe that Trial 003 was significantly flawed in such a way that the results either should be discounted or the results are not usable, so that we actually need another trial. It's not because we can't figure out or we don't have all the explanations of the results.

So that's the first one. And B would also reflect the fact that you think a trial is feasible, and such a trial should provide robust evidence to verify the clinical benefit of Makena. So I will stick my neck out there and say probably a PK/PD won't verify the clinical benefit of Makena.

DR. CHANG: This is Christy Chang from FDA. Could I also add another point of clarification? If you're contemplating a confirmatory trial with an active comparator, because nothing is approved by the FDA for the same indication, how do we make that comparison?

DR. LEWIS: Dr. Orza?

DR. ORZA: I believe for comparative effectiveness studies, there is not a requirement that it be FDA approved, but only that it be in widespread use. So if it were possible to identify a comparator that wasn't widespread use, that would be, I think from a funder's point of view, acceptable. Whether it would be acceptable to the FDA is another question.

DR. NGUYEN: Christine Nguyen, FDA again. Our task is to ensure that the drugs we approve have substantial evidence of effectiveness and usually compare to a placebo. We do not usually accept as an active comparator, if I may use that term. That has not been demonstrated to be safe and effective for the intended use because we don't know how to interpret the results.

If Makena performs, say, the same as vaginal progesterone, is it because neither are working, or are they both working? We can't really interpret the results.

DR. ORZA: So it might not help the FDA, but it might help the clinical community.

(Pause.)

MS. ELLIS: There's no abstain button.

(Laughter.)

DR. LEWIS: There's no button, but you can abstain.

(Laughter.)

DR. LEWIS: Dr. Lindsay?

(No audible response.)

(Voting.)

MS. BHATT: For question 6A is 9; B is 6; and C is zero.

DR. LEWIS: Thank you. Let's go in the opposite direction just for variety's sake here. So we'll start with Dr. Davis.

DR. DAVIS: I was interested, as I mentioned previously, on a trial to try to better define a higher

risk population of mothers at risk of delivering preterm that potentially could have a more significant impact on neonatal outcome. I think those would be the ways that I would approach it with potentially dose escalation and other pharmacokinetics and pharmacometrics, and looking at dosing levels, and serum levels, and outcomes.

I recognize FDA's need to have a second confirmatory trial. I am concerned about putting the genie back in the bottle when it becomes standard practice and you have every major obstetrical organization supporting the continued use. I might suggest to FDA that they work with the sponsor to more narrowly limit the label and potentially indicate the non-confirmatory nature of the trial, though limited benefit to neonates, and the potential of limiting it to a higher risk population until another trial is done.

DR. LEWIS: Thank you. Dr. Wade?

DR. WADE: I voted no. I followed the outlined requirements of the accelerated approval process and what was outlined at the task at hand for

003, which I did not think verify -- unfortunately didn't verify the findings as 002. I am significantly worried about the consequences of that decision, though. and I think we could all spend a lot more time thinking about how to accelerate through another trial to get the data that we desperately need to safely treat women.

DR. LEWIS: Dr. Smith?

DR. SMITH: Brian Smith. I voted for option A. I would echo the comments made by Kelly Wade. I would also add that I heard one of the concerns with withdrawal of the molecule was that OBs would use unproven therapies like vaginal progesterone or cerclage, and to me I think the consideration there is that OBs have an obligation to their patients to do no harm.

DR. LEWIS: Thank you. Dr. Hunsberger?

DR. HUNSBERGER: Sally Hunsberger. I voted A. I just don't believe the totality of the evidence supports this, and I think this might be the only way to do a study where we will actually get the data that we need. And I think we really need data to understand

what's going on.

DR. LEWIS: Thank you. Dr. Gillen?

DR. GILLEN: Dan Gillen. I definitely think that there are many, many repercussions to the withdrawal, and I don't make that choice lightly, but for me it's a logical process of elimination. I do not believe that substantial evidence has been established, given the results of the two studies. And by the sponsor's own admission, they believe that we can't trust the second study because the first study was on the market and leads to a bias population, which means that if you're going to do an honest assessment of this drug, it would have to be removed.

DR. LEWIS: Dr. Orza?

DR. ORZA: Michele Orza. I voted B, although I felt that my votes on questions 4 and 5 inexorably led to a vote of A. So I am voting B with a couple of conditions. I'm assuming that the clinical societies will, as Dr. Bauer rightly suggested, lead the way. The new evidence is still under consideration by them. The IPD meta-analysis, which could be updated with the new data on Makena, has yet to be released, and they

will have to take that into consideration.

I think if they are moved to a position of equipoise so that a randomized, placebo-controlled, hopefully also with an active comparator -- if one is identified and can be done. then I think you can leave it on the market. But if that doesn't happen, then I think the FDA does need to withdraw it in order to make that study possible, because I do think that more compelling confirmatory evidence does need to be generated. I'm very compelled by Dr. Shaw's point about saying that this level of evidence is good enough for some people.

DR. LEWIS: Ms. Ellis?

MS. ELLIS: Yes. My heart wanted to vote C because mothers want nothing more than to have healthy babies, and the longer that we can keep them growing with our protection, the better. But I was prevented from doing so because choice B had the word "feasible," and if it's all false -- if one part's false, it's all false. So I could not vote that way.

I also had to consider the regulatory framework with which we are here and with which we

function, and accelerated approval requires confirmation. And this vote, depending on what the decisions are made later on, may prevent my own daughter from accessing this drug. However, I got lucky with my second pregnancy, using something we don't use anymore and bed rest. And I think that mothers and babies shouldn't have to rely on luck. We need data. Thank you.

DR. SHAW: Pamela Shaw, and I voted A, and I spent most of the day knowing I had to answer this question, thinking about this particular question. And if there's any way I could have chosen B -- but I can't think -- I'm thinking noninferiority, is there a active comparator? No. I just cannot think of a feasible trial, so picking B, to me, is just going to prolong this painful process even longer. So I'm thinking A was the best practical choice for finding something that will work in neonatal infants as fast as possible.

DR. BAUER: Doug Bauer. Unfortunately, I also voted for A with a lot of trepidation, probably from the patient standpoint, which I think Ms. Ellis just eloquently summarized for us. But also, I really feel

for the providers who are in the trenches, that are going to have to answer to their patients that are just demanding something for something. It's really an awful condition that we have no other choice for. But I really feel in the long run that removal of the drug is the right thing to do, and at least we'll have some possibility that then there'll be a properly done trial to finally answer the question.

DR. LEWIS: I voted B, reluctantly. I almost wanted to abstain because I think that the data are conflicting, and it's certainly not terribly persuasive one way or the other. I think that we would definitely benefit from additional data. I don't know that -- it's not going to be the quality of a randomized, placebo-controlled trial. I think it will shed some light, though, on perhaps understanding a population for whom this might be beneficial and ways that the drug's usefulness can be limited in some way, the labeling can be limited in some way that would help us find a better population who could use it.

DR. DRAKE: I'm Matthew Drake. I also voted for A. I think it's a very challenging situation we've

been tasked with. I feel for those patients. I feel for the practitioners who will have to deal with them. But ultimately, I tried to be objective and just look at the efficacy requirements as spelled out by the FDA, and I just, unfortunately, didn't think that those were met. So for that reason, I vote A.

DR. WING: I'm Deborah Wing, and I struggled with my vote, and I voted A. I put on my clinician scientist hat and looked only at the data, and I do not believe there is substantial evidence of effectiveness based on my read of both of the trials and listening to the deliberations today and through this afternoon. I fully appreciate and have experienced the agency's requirements to adequately powered, appropriately designed trials to move products out onto the market.

I agree with Dr. Gillen. I think this drug likely got to market a little bit early, so we are hamstrung because of lack of results in a validation trial that was spread across the world. Obviously, one of the things we try to do when we impart our clinical trials to the world is generalize them. We actually generalized Makena and got negative results, which is,

I think, not what we anticipated, but we do the science because we don't know. We asked a question and we didn't get an answer; we didn't get an answer we anticipated.

I'll come back to the ethical principles of doing good and doing no harm. I think the doing good here is continuing to ask the questions and asking are we doing good by the patients. And I think the only way by which to get the results of a confirmatory trial is to actually do another placebo-controlled trial.

As hard as that might sound, I know that the societies, the agency, and the sponsor will work together to try to figure out how to cover the gap we just created for the clinicians, and hopefully for the patients, because this is what we call in business, a big hairy audacious problem, and we have to put heads together and do something differently. But I'm not convinced that leaving Makena on the market as is, is the right thing to do.

DR. REDDY: I voted for B because I see A as untenable. I think withdrawing it from the market, you're not going to have a randomized-controlled trial.

It will be very difficult because, still, we are obligated to tell patients what the evidence is there.

002, the fact that it's 20 years old, I don't think that makes a difference because spontaneous, preterm delivery hasn't changed. It was a well done randomized-controlled trial. Why the rate was so high in the placebo group; who knows? But on the surface of it, it's a very supportive trial, and then you take 003, and, to me, it's apples and oranges.

The U.S. subgroup, there wasn't a significant difference. I get that. We can talk about power and the risk of it, but I do not think our RCT, a placebo randomized-controlled RCT will be done in the U.S. Patients are very smart. They have the information as physicians. I cannot say, oh, it's not FDA approved, so I'm not going to recommend it or I'm not going to discuss it, because all the medicines we use in pregnancy are not FDA approved. What we do is we counsel patients, and that's what we'll continue to do.

So I didn't vote for A because I think it's a big step backwards. I think by voting for B, we're getting additional information. I would only vote for

A if I thought the medicine was a danger, there was a safety issue, and I think 003 has resolved that. And at the least, I'm very happy about that, and I thought had no use whatsoever. So I think A is a vote for -- there's not going to be an RCT. Patients will not -- and physicians also. It's going to be very difficult to get patients into an RCT, placebo RCT.

DR. LINDSAY: Michael Lindsay. I voted for B. I agonized over this decision when I got the background information. I've been reading it over the last couple of weeks, and it was really clear that the evidence was conflicting, and I knew it was going to be conflicting today.

The reason why I chose B is I agree with Dr. Reddy. I didn't think A was really a valid choice. In terms of a clinician, I think one of the things that I struggle with is tomorrow I'm going to be seeing patients, and I have to give them some guidance of what they can do when they've had preterm delivery. I realize that this information is conflicting, and when you counsel people, you offer them the information, and then they make a choice.

I realize that doing another randomized-controlled trial may be the ideal way to kind of resolve the problem, but in the real world, as clinicians, we don't deal with idealism every day; we sort of deal with reality. I agree there probably are some subpopulations that are impacted in a positive way by this medication. We just haven't identified them, and I think that that would be one of the directions that I would encourage the FDA to pursue, encouraging investigators.

I think the reality, though, is as we let the genies out of the bottle and people know that there are medications that have been used for patients who had preterm deliveries, they're going to still want to get access to those medications. Clinicians like myself who've been out there for decades and have used compounding medications are going to give their patients compounding medications, and that's a reality.

So I think by following the rules -- and I say this to my trainees. I know the rules. I haven't followed them consistently, and I think this is an exercise that we really need to follow the rules, and

I'm not against that. But I think you also need to know the consequences, is that the problem is not going to go away, and people are going to seek other treatments and there'll be other methodologies of treatment.

DR. HICKEY: Hi. Kim Hickey. I also voted for B. I thought the idea of removal of the drug was, just like Dr. Reddy said, not feasible, and much like Dr. Lindsay said, our patients know it's there, and if I don't find them some sort of progesterone, they'll find someone who will. So I think doing the RCT placebo-controlled trial is not going to be feasible, and I feel there is a subset that have benefited from this.

I think it will be hard to look at someone who had a preterm delivery that had a term delivery on Makena, and then tell her, but it doesn't work, because we can all agree, and we all have, that the data's conflicting, and we don't like things about each trial. But to just toss it out and say we're going to go back to ground zero and put people at risk from potential compounded 17P, I don't think is worth it.

DR. EKE: I voted for B. Just like most of us said here, I struggled with this for days. Since I got the notification to go through this, I read through both trials. I struggled. The clinical trialist in me would say go for A, but when I look at the totality of the evidence, and especially what the consequences of this is going to be to all my patients and for people to take care of, if I look at what we have currently for treating -- this is not being sentimental, it's just looking back at why I voted for B. If we look at what we have, this is the only pharmacotherapy we have for preterm birth that has been shown to work in some populations.

The next thing, if we withdraw totally, people will be placed in cerclages, which studies have shown increases preterm birth in this population, and there are no other pharmacotherapies out there, so we'll see patients scrambling to get this. And I just worry about what that will be.

So why I looked at that, it was we keep this while we get -- I want to see a trial that will tell me which patients would benefit from this drug because I

know and I believe that there are populations or patients that will benefit from this drug. I want to see those populations. I want to see an increased -- or a better outcome in units. Those were the things that kind of drove me to vote for B. Thanks.

DR. LEWIS: Before we adjourn, are there any last comments from FDA?

DR. WESLEY: This is Barbara Wesley. I'd like to make one clarification about who makes what rules. The FDA doesn't make the rules. The Congress makes rules about the statutory requirements. We carry out the rules. I think Congress consults with the Institute of Medicine, if I'm not mistaken. But they make the rules and set the statutory requirements. We carry them out. I just want to clarify that because I think sometimes that gets confusing.

DR. LEWIS: Thank you all for your attention and your -- I'm sorry. Dr. Nguyen, yes?

DR. NGUYEN: Actually, Dr. Lewis, I have the last comments.

DR. LEWIS: Sorry.

DR. NGUYEN: I would like to add, on behalf of FDA, we really thank everyone here today. We thank the applicant for their excellent presentation and their professionalism. I'd like to thank, obviously, all the FDA review staff who have worked tirelessly and very quickly to bring this to a meeting, and certainly our presenters. I'd like to acknowledge team members who worked very hard behind the scene, Christina Chang, who is our team leader and our two project managers, and Kalesha Grayson and Jeannie Roule.

Certainly last but not least, I want to express our gratitude to all of our AC staff members and all of you sitting at the table today. We appreciate how difficult this was for you, and it was very difficult for us as well. We also appreciate our decisions will affect each individual patient and their families. We're not just looking at facts, but we do owe it to the public to do the right thing, which is to put out drugs that are safe and effective, and we need to consider both.

So thank you very much again. Thank you, Kalyani. Thank you, Dr. Lewis, and we'll see some of

you back tomorrow morning, so thanks.

Adjournment

DR. LEWIS: Yes. Thank you all for a productive day. Thanks to the FDA, sponsors, and of course the public for their contributions. We appreciate it. We are adjourned. Panel members, please take your personal belongings. The room will be cleaned at the end of today. Any material left on the table will be disposed of. Please leave your name badges, though, on the table; that I do want to remind you. So we're now adjourned. Thank you.

(Whereupon, at 4:26 p.m., the meeting was adjourned.)

October 29, 2019, Meeting of the Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) - Webcast Recording

The Center for Drug Evaluation and Research (CDER) provided a live webcast of the October 29, 2019, meeting of the Bone, Reproductive and Urologic Drugs Advisory Committee.

A recording of the webcast can be found at the following address:

- Start of Meeting to Morning Break: https://collaboration.fda.gov/pfdj6tbjng8i/
- Morning Break to Lunch Break: https://collaboration.fda.gov/pkmoqz9f1alj/
- Lunch Break to Afternoon Break: https://collaboration.fda.gov/pktdgjodgvx6/
- Afternoon Break to End of Meeting: https://collaboration.fda.gov/pel10yotagt7/

The webcast was broadcast using Adobe Connect. You can make sure your computer has the correct plug-ins to view the webcast at this web site:

https://collaboration.fda.gov/common/help/en/support/meeting_test.htm