CENTER FOR DRUG EVALUATION AND RESEARCH (CDER)

Food and Drug Administration

Briefing Materials Supporting CDER's Proposal to Withdraw Approval of Makena

Docket No. FDA-2020-N-2029

September 16, 2022

TABLE OF CONTENTS

TABLES

FIGURES

LIST OF ABBREVIATIONS

ACOG	American College of Obstetricians and Gynecologists
ACRHD	Advisory Committee for Reproductive Health Drugs
ANDA	Abbreviated New Drug Application
BPD	Bronchopulmonary Dysplasia
BRUDAC	Bone, Reproductive and Urologic Drugs Advisory Committee
CDC	Centers for Disease Control and Prevention
CDER	Center for Drug Evaluation and Research
CFR	Code of Federal Regulations
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel method
DUOG	Division of Urology, Obstetrics, and Gynecology
EPPPIC	Evaluating Progestogens for Preventing Preterm birth International Collaborative
FAERS	FDA's Adverse Event Reporting System
FDA	Food and Drug Administration
FDAMA	Food and Drug Administration Modernization Act of 1997
FDASIA	Food and Drug Administration Safety Innovations Act of 2012
GA	Gestational Age
FD&C Act	Federal Food, Drug, and Cosmetic Act
HPC	Hydroxyprogesterone Caproate
IVH	Intraventricular Hemorrhage
MFMU	Maternal Fetal Medicine Unit
NDA	New Drug Application
NEC	Necrotizing Enterocolitis
NEJM	New England Journal of Medicine
NISS	Newly Identified Safety Signal
NOOH	Notice of Opportunity for a Hearing
PTB	Preterm Birth
RCT	Randomized Controlled Trial
RDS	Respiratory Distress Syndrome
SEE	Substantial Evidence of Effectiveness
SHR	Bayesian Shrinkage Estimation
SMFM	Society of Maternal and Fetal Medicine
sPTB	Spontaneous Preterm Birth
VTE	Venous Thromboembolism

The Center for Drug Evaluation and Research (CDER or the Center) of the Food and Drug Administration (FDA or the Agency) hereby submits its briefing materials in support of the proposed withdrawal of approval for Makena, and approved generics referencing Makena.

I. EXECUTIVE SUMMARY

Makena has not been shown to improve neonatal outcomes from premature birth, is no longer shown to be effective for its approved use, and has known risks. The 1,708-person confirmatory trial designed to verify Makena's clinical benefit instead failed to show that Makena has any benefit to newborns. Data from this trial, taken together with other evidence, also fail to show that Makena reduces the risk of recurrent preterm birth. For these and other reasons detailed herein, Makena should be withdrawn from the market.

Preterm birth (PTB), defined as birth prior to 37 weeks of gestation, currently occurs in approximately 10% of all births and 8% of singleton pregnancies.[1] Premature birth is a significant public health problem because infants born prematurely 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 PTB. Although neonatal outcomes are the most relevant measurement of benefit in the treatment of PTB, gestational age (GA) of delivery may be considered as a proxy measure for neonatal health because it is related to the development of the fetus.

In February 2011, CDER approved Makena (hydroxyprogesterone caproate injection) under the accelerated approval pathway to reduce the risk of recurrent PTB in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth (sPTB). The approval was based on the results of the Maternal Fetal Medicine Unit (MFMU) Network Trial 17P-CT-002 (hereafter, Trial 002) conducted between 1999 and 2002, in which the treatment arm receiving hydroxyprogesterone caproate (HPC) 250 milligram (mg) injection had a statistically significantly lower rate of delivering prior to 37 weeks' gestation than the placebo arm. CDER approved Makena (HPC 250 mg/milliliter (ml) injection)[2] under the accelerated approval pathway on the basis that Makena's effect on this intermediate clinical endpoint,[3] while not itself

---

[1] Centers for Disease Control and Prevention, *Preterm Birth*, CDC.GOV, https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm (last visited Sept. 13, 2022). Martin JA, Osterman MJK. Exploring the decline in the singleton preterm birth rate in the United States, 2019–2020. NCHS Data Brief, no 430. Hyattsville, MD: National Center for Health Statistics. 2022, available at https://www.cdc.gov/nchs/data/databriefs/db430.pdf (last visited Sept. 14, 2022).

[2] Studies may use the terms 17P, HPC, 17-OHPC, or Makena. When this document refers to a study of Makena, we mean one that studied either Makena itself or a product that contained the same active substance, HPC, as Makena, at the same dose. Unless specifically noted, all publications referenced in this document evaluated HPC 250 mg/ml injection, the same dose as in Makena. Some of the publications describe population-based real-world evidence studies of HPC. The dosing for these studies was based on clinical guidelines that relied on Trial 002.

[3] Gestational age of delivery is an intermediate clinical endpoint. An intermediate clinical endpoint is a measurement of a therapeutic effect that can be measured earlier than an effect on irreversible morbidity or mortality that, in the context of accelerated approval, is considered reasonably likely to predict the drug's effect on irreversible morbidity or mortality or other clinical benefit. FDA Guidance for Industry, Expedited Programs for Serious

a direct measure of clinical benefit to neonates, was deemed reasonably likely to predict reduction of neonatal morbidity and mortality associated with PTB.[4] Trial 002 did not show that women at higher risk of PTB benefited more from Makena than women at lower risk for PTB. It also did not show that any appropriately identified subpopulation benefited more from Makena than any other subpopulation.[5]

As a condition of its approval under the accelerated approval pathway, the sponsor[6] was required to complete a postmarketing trial to verify and describe the clinical benefit of Makena on reducing the risk of neonatal morbidity or mortality from PTB among babies born to women[7] with a singleton pregnancy who had a previous singleton sPTB. This postmarketing trial, Trial 003, failed to show that Makena reduced the risk of neonatal morbidity and mortality from complications of sPTB, and failed to show that Makena had any effect on the endpoint that was the basis of Makena's approval—reduction in the proportion of women delivering prior to 37 weeks' gestation. Trial 003 also did not show any drug effect on any identified subpopulation, including those at higher risk for PTB. Similar to Trial 002, Trial 003 did not show differential drug benefit based on the presence of more or fewer risk factors associated with PTB.

Given the results of Trial 003, Makena should be withdrawn from the market. Approval of a drug that was approved under the accelerated approval pathway may be withdrawn if, among other reasons, the confirmatory trial fails to verify and describe the drug's expected clinical benefit, or the drug is not shown to be safe or effective under its conditions of use.[8] Both of these independent grounds for withdrawal are present here. Trial 003 failed to verify Makena's expected clinical benefit, and Makena is no longer shown to be effective for Makena's indicated

---

Conditions—Drugs and Biologics (May 2014) (Expedited Programs Guidance), available at https://www.fda.gov/media/86377/download, at 18. In the context of accelerated approval, both intermediate clinical endpoints and surrogate endpoints function as predictors of clinical benefit.

[4] Reducing the risk of delivery < 37 weeks gestation in and of itself is not clinically relevant without a reduction in adverse neonatal outcomes associated with PTB. References elsewhere in this document to the effect on this intermediate clinical endpoint as "the basis" for approval are intended to encompass both the effect on this endpoint and the evidence showing that this endpoint was reasonably likely to predict the clinical benefit of reducing neonatal morbidity and mortality associated with PTB.

[5] When CDER refers to an identified subpopulation, subgroup, or subset in the remainder of these briefing materials, it means one that can be identified using defined pre-randomization variables (e.g., Black compared to non-Black). Covis identified what they referred to as a "higher-risk population subgroup" (Sponsor's Response to NOOH at 26, available at https://www.regulations.gov/comment/FDA-2020-N-2029-0051). CDER does not consider this subgroup to be appropriately identified since the steps taken to subset the Trial 003 subjects for analysis included using site-level post-randomization variables including using the clinical trial sites' placebo-arm recurrent PTB rate and the number of subjects enrolled at the site. These are not identifiable characteristics that can be used to distinguish patients who may benefit more from Makena.

[6] At the time of approval, the sponsor of new drug application (NDA) 021945 was Hologic, Inc. Subsequent to the approval, the ownership of this application was transferred to KV Pharmaceuticals Company, and then it was transferred to Lumara Health, Inc., which was later purchased by AMAG Pharma. Covis Pharma GmbH, the current sponsor of NDA 021945, acquired the application from AMAG Pharma in March 2021.

[7] CDER uses the term "women" to refer to pregnant people in these briefing materials for consistency with the language used in the labeled indication for Makena and the questions presented.

[8] *See* section 506(c)(3) of the Federal Food, Drug, and Cosmetic Act (FD&C Act) and 21 CFR 314.530(a).

population (i.e., a woman with a singleton pregnancy and a prior singleton sPTB). If CDER had access to the data from both Trials 002 and 003 in 2006–2011, at the time the application was being considered, together with all the other evidence available today, it would have concluded that efficacy had not been shown, and it would not have approved the drug under either the accelerated (based on the intermediate clinical endpoint of gestational age of delivery) or traditional (based on neonatal morbidity/mortality outcomes) approval pathways.

Furthermore, there is no basis to retain Makena's approval but narrow its indicated population to a subgroup of "high-risk" patients, as Covis suggests, because the evidence does not show that this—or any—subset of Makena's indicated population responds more favorably to Makena than any other subset.

We recognize that premature birth is a significant public health concern with unmet medical need. We also recognize that Makena is the only currently approved treatment for a serious condition that disproportionately affects some of our nation's most at-risk women, children, and families. While there is a significant need for treatment for this condition, FDA has a responsibility to the American public, including to the people that Makena is intended to benefit, to ensure that the drugs that the Agency approves are both safe and effective. We are extremely disappointed that the available evidence does not demonstrate Makena's effectiveness. We are mindful of the significance of proposing to withdrawal Makena from the market. We made the determination that this was necessary only after careful consideration of the available scientific evidence. The benefit-risk calculus of a drug that lacks efficacy is necessarily unfavorable, as all drugs, including Makena, carry risks. Put simply, Makena's risks, which include, among other things, thromboembolic events (i.e., blood clots), are unacceptable in light of the lack of evidence of clinical benefit. Failing to withdraw Makena would mean maintaining FDA approval of a drug that, based on all available evidence, has not been shown to be more effective than, but is riskier than, no treatment. This would be a disservice to patients at risk for recurrent PTB and would undermine the accelerated approval pathway.

Furthermore, leaving Makena on the market would inhibit the gathering of high-quality evidence that would be needed to show its effectiveness. sPTB is a poorly understood syndrome that is ill-suited to non-randomized observational studies and unblinded randomized trials for generation of evidence of effectiveness. Therefore, only a prospective, randomized, double-blind, placebo-controlled trial would be adequate to overcome the negative trial result of Trial 003. The sponsor proposes to conduct such a trial enrolling approximately one to three thousand women with a prior sPTB < 34 weeks entirely or mostly from the U.S. But while Makena remains on the market, conducting another randomized, double-blind, placebo-controlled trial is infeasible in the U.S. As reflected by the difficulty enrolling U.S. women in Trial 003 after Makena was approved, patients are less likely to enroll in trial for a drug approved to treat a serious condition and risk receiving a placebo when they can simply receive the drug by *not* enrolling in such a trial.

Even if such a trial could be conducted, it could not be completed in a timely manner. Trial 003 took nearly a decade to complete. A new trial—which, as a practical matter, could occur in the

U.S. only if Makena is withdrawn from the market—would likely take at least as long. Retaining Makena's approval in the interim would mean that a drug indicated to treat a serious condition would likely remain FDA-approved for at least another ten years, even though the drug has not been shown to be effective. Additionally, Covis proposes to further study high-risk women, defined as having a prior sPTB < 34 weeks, and to 'enrich' the study for Black women. But while women with these characteristics are at higher risk for PTB, the evidence does not show that Makena works better (or that it works at all) for women with these characteristics.

Finally, contrary to Covis' position that retaining approval of Makena would be in the best interests of patient care given that Makena is the only FDA-approved drug for prevention of recurrent PTB, failing to withdraw Makena in the face of all of the available evidence that Makena is no longer shown to be effective would be a disservice to women at risk for recurrent sPTB and the children born to these patients. The continued marketing of Makena in the absence of demonstration of benefit incurs false hopes, the risks associated with treatment, and other burdens including unnecessary procedures and overutilization of healthcare resources. As one example, women receiving treatment with Makena receive up to 20 injections during pregnancy, each of which may require either a prenatal clinic visit, home health nursing care, or resources to train a lay person, in addition to the discomforts and risks (such as bleeding or infection) of the injection itself. Further, retaining approval of Makena would make a study in U.S. women infeasible to conduct, and would likely hinder research and development of other, perhaps more promising, treatments for reducing the risk of PTB.

Accordingly, approval of Makena should be withdrawn.

II. QUESTIONS PRESENTED

The hearing will cover four main questions. Those questions, and CDER's proposed answers, are as follows:

Question No. 1: Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal morbidity and mortality from complications of preterm birth?

Response: No. Trial 003, the required postmarketing confirmatory study, failed to verify Makena's purported clinical benefit: reducing neonatal morbidity and mortality from complications of PTB in women with a singleton pregnancy who have a history of singleton sPTB. The 2019 Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC)[9] voted unanimously that Trial 003 failed to verify the clinical benefit that CDER thought, at the time of accelerated approval in 2011, had been reasonably likely based on Makena's effect on

[9] In March 2022, the Agency renewed the charter for the Obstetrics, Reproductive and Urologic Drugs Advisory Committee (ORUDAC) (formerly known as the BRUDAC). The function of the Committee no longer includes osteoporosis and metabolic bone disease. *See* https://www.federalregister.gov/documents/2022/03/23/2022-05973/advisory-committee-obstetrics-reproductive-and-urologic-drugs-advisory-committee-renewal.

the intermediate clinical endpoint shown in Trial 002.[10] In fact, AMAG Pharma, Covis' predecessor, conceded that the confirmatory study failed to verify Makena's clinical benefit.[11] AMAG Pharma stated that "[t]here are no controlled trials demonstrating a direct clinical benefit, such as improvement in neonatal mortality and morbidity" and that Trial 003 "did not demonstrate a statistically significant difference between the treatment and placebo arms for the co-primary endpoints."[12] Covis has also acknowledged this failure to show clinical benefit by stating that "the findings from Trial 003 (PROLONG)[13] do not verify the clinical benefit of Makena on neonatal morbidity and mortality from complications of preterm birth."[14] Accordingly, the grounds for expedited withdrawal of approval under section 506(c)(3)(B) of the FD&C Act and 21 CFR 314.530(a)(1) for failure to verify clinical benefit have been met.

Question No. 2: Does the available evidence demonstrate that Makena is effective for its approved indication of reducing the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth?

Response: No. Considering all the available evidence, Makena is not shown to be effective at reducing the risk of PTB in women with a singleton pregnancy who have a history of singleton sPTB. Trial 003, a trial almost four times larger than Trial 002 and the only randomized, double-blind, placebo-controlled trial to directly measure Makena's purported clinical benefit to neonates, not only failed to demonstrate any such benefit,[15] but also failed to show any drug effect on gestational age at delivery, which was the basis of Makena's accelerated approval in 2011. This led 13 of the 16 2019 BRUDAC voting members to determine that, based on these two trials, there is not substantial evidence that Makena is effective for its approved indication.

Trial 002 had limitations, including that it did not plan to evaluate neonatal measurements as an efficacy outcome[16] and that there were questions about the generalizability of its study population to the proposed indicated population. Trial 002 was also the only adequate and well-

[10] The 2019 BRUDAC voted 16 (No) to 0 (Yes) on the question "Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal outcomes?" Summary Minutes, BRUDAC Meeting 6, 9 (Oct. 29, 2019), available at https://www.regulations.gov/document/FDA-2020-N-2029-0118 and attached as Appendix 2.

[11] AMAG Backgrounder, BRUDAC Meeting (Oct. 29, 2019) (page 18 of 91, Neonatal Composite Index) states: "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)." Available at https://www.regulations.gov/document/FDA-2020-N-2029-0114 and attached as Appendix 2.

[12] AMAG Pharmaceuticals Announces Topline Results from the PROLONG Trial Evaluating Makena, March 8, 2019, available at https://www.amagpharma.com/news/amag-pharmaceuticals-announces-topline-results-from-the-prolong-trial-evaluating-makena-hydroxyprogesterone-caproate-injection/.

[13] The sponsor refers to Trial 003 as the "PROLONG" study (Progestin's Role in Optimizing Neonatal Gestation).

[14] Letter from Rebecca Wood, Esq. dated July 1, 2022, at 1 (submitted to Makena docket, Docket No. FDA-2020-N-2029).

[15] A neonate refers to an infant in its first 28 days of life.

[16] Trial 002 collected individual neonatal morbidity and mortality outcomes, such as respiratory distress syndrome and necrotizing colitis, however, they were not prespecified for efficacy assessment.

controlled trial supporting Makena's approval. As a condition of Makena's approval under the accelerated approval pathway, the sponsor was required to conduct Trial 003, which was designed with extensive FDA input to achieve its objective of verifying (or not) Makena's clinical benefit and substantiating (or not) its treatment effect on gestational age of delivery. Trial 003 was a rigorous, well-designed, well-conducted trial that was adequately powered to detect efficacy – but despite this, it did not do so. In fact, if the findings of Trials 002 and 003 were submitted at the same time in an NDA seeking approval for Makena, CDER would conclude that there is not substantial evidence of Makena's effectiveness and would not approve the drug under any approval pathway—traditional or accelerated.

CDER has also reviewed results from other studies of the effect of HPC on sPTB. First, CDER considered the EPPPIC study (Evaluating Progestogens for Preventing Preterm birth International Collaborative),[17] a meta-analysis study of 31 randomized controlled trials that evaluated various progestogens, including progesterone and HPC, for different uses related to PTB. Among EPPPIC's 31 clinical trials, CDER focused on the five placebo-controlled clinical trials, including Trials 002 and 003, that were relevant to Makena in that they evaluated the effect of HPC on PTB in singleton pregnancy. These data did not show a beneficial treatment effect on PTB < 28 weeks, < 34 weeks, or < 37 weeks gestation in women with a prior sPTB[18] or a short cervix in the current pregnancy.[19]

[17] The EPPPIC Group. Evaluating Progestogens for Preventing Preterm birth International Collaborative (EPPPIC): meta-analysis of individual participant data from randomised controlled trials. The Lancet 2021;397 (10280):1183-1194.

[18] The gestational age cutoffs of < 28 weeks, < 34 weeks, and < 37 weeks were selected by the EPPPIC investigators in their meta-analysis. No specific gestational age cutoff has been determined to be optimal in assessing neonatal outcomes. As an example, the gestational age cutoffs used in Trial 002 were < 37 weeks, < 35 weeks, and < 32 weeks. Since 2014, the National Center for Health Statistics has reported birth data using the following definitions: "late preterm infants" are those born at a gestational age between 34 weeks 0 days and 36 weeks 6 days of gestation; "early preterm infants" are those born at less than 34 completed weeks gestation.

[19] A short cervix in mid-trimester of pregnancy is a risk factor for PTB, regardless of the patient's obstetrical history (i.e., whether she had a prior PTB). The diagnosis of a short cervix is made by vaginal ultrasound prior to 24 weeks gestation, but clinicians have used various thresholds of cervical length (< 1.5 cm, < 2.0 cm, or < 2.5 cm, etc.) to determine whether intervention is needed. The American College of Obstetricians and Gynecologists uses < 2.5 cm for patients with singleton pregnancies, regardless of past obstetric history. *See* Prediction and Prevention of Spontaneous Preterm Birth: ACOG Practice Bulletin, Number 234. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Obstetrics; Obstet Gynecol. 2021;138(2):e65.

Second, CDER also conducted keyword searches of the PubMed database and identified five published observational studies that evaluated the efficacy of HPC for PTB. [20, 21, 22, 23, 24] In addition to these observational studies, CDER reviewed three other double-blind, randomized, controlled trials (RCTs)[25,26,27] referenced by Covis, that evaluated the effectiveness of Makena in women at high risk for sPTB. These published studies varied in study designs and in study populations. Taken together, these studies do not suggest a beneficial effect of HPC in reducing the risk of PTB in the populations studied and bolster the determination that Makena is not shown to be effective.

Figure 1 shows the relative risk reduction with Makena compared to placebo/no treatment for preterm delivery at several gestational age cut-offs for the available randomized, placebo-controlled trials in Makena's target population (Trial 002 and Trial 003) and the published results from trials in women at high risk for PTB from HIV (Price) or multiple gestations (Caritis, Rouse),[28] and observational studies in women eligible to receive Makena (Hakim, Wang, Massa).[29] Trial 002 is clearly the outlier in showing a favorable effect of Makena. The confidence intervals in all other studies overlap 1.0, showing no statistically significant effect.

---

[20] Bastek JA, Adamczak JE, Hoffman S, Elovitz MA, Srinivas SK. Trends in prematurity: What do changes at an urban institution suggest about the public health impact of 17-alpha hydroxyprogesterone caproate? Matern. Child Health J. 2012;16:564–568.

[21] Nelson DB, McIntire DD, McDonald J, Gard J, Turrichi P, Leveno KJ. 17-alpha hydroxyprogesterone caproate did not reduce the rate of recurrent preterm birth in a prospective cohort study. Am. J. Obstet. Gynecol. 2017;216:600.e1–9.

[22] Massa K, Childress K, Vricella LK, et al. Pregnancy duration with use of 17-a-hydroxyprogesterone caproate in a retrospective cohort at high risk of recurrent preterm birth. Am. J. Obstet. Gynecol. MFM 2020;2:100219.

[23] Wang X, Garcia S, Kellom K, Boelig R, Matone M, Eligibility, utilization, and effectiveness of 17-alpha hydroxyprogesterone caproate (17OHPC) in a statewide population-based cohort of medicaid enrollees. Am. J. Perinatol. 2021.

[24] Hakim J, Zhou A, Hernandez-Diaz S, Hart J, Blair J. Wylie B, Beam A, Effectiveness of 17-OHP for prevention of recurrent preterm birth: a retrospective cohort study, Am. J. Perinatol. 2021.

[25] Price JT, et al. Weekly 17 alpha-hydroxyprogesterone caproate to prevent preterm birth among women living with HIV: a randomised, double-blind, placebo-controlled trial, The Lancet HIV. 2021; 8(10): e605–e613.

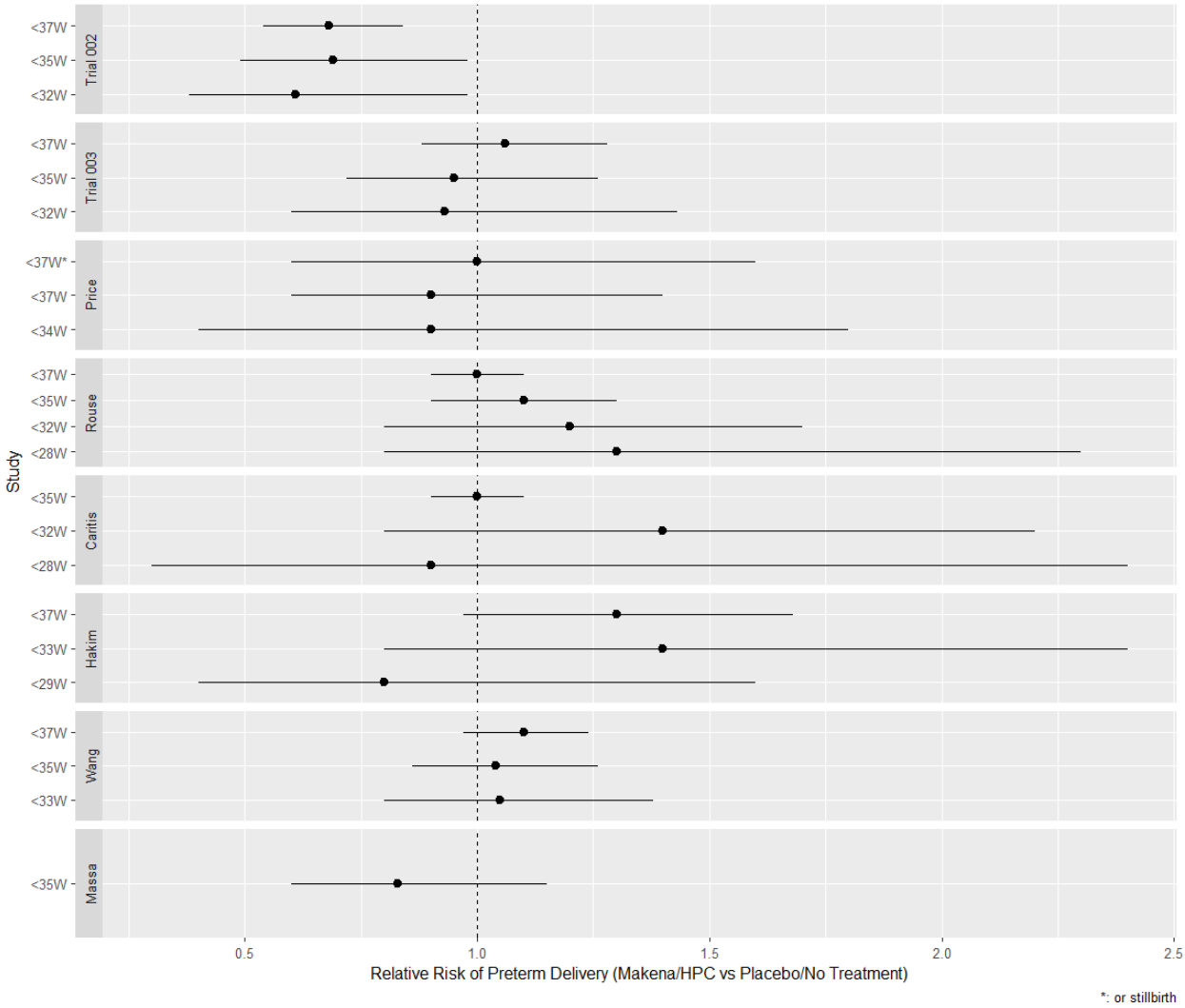
[26] Rouse DJ, et al. A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. N. Engl. J. Med. 2007;357:454–461.

[27] Caritis SN, et al. Prevention of preterm birth in triplets using 17 alpha hydroxyprogesterone caproate a randomized controlled trial. Am. J. Obstet. Gynecol. 2009; 113(2):285-292.

[28] These three clinical trials were identified through a list of trials the sponsor provided as supportive of Makena's safety and through a search of the ClinicalTrials.gov data bank. CDER conducted an advanced search of ClinicalTrials.gov to identify double-blind, randomized, placebo-controlled trials of Makena or hydroxyprogesterone and their published studies. CDER limited its search to trials identified as Phase 3, completed, interventional studies, and then manually identified any double-blind, randomized, placebo-controlled trials with outcomes related to GA at delivery and/or neonatal outcomes in which intervention started at GA 15w-21w. CDER last confirmed these search results on September 15, 2022.

[29] These observational studies were identified through a PubMed keyword searches using the keywords "hydroxyprogesterone caproate," "effectiveness," "effect," "preterm birth." No other limitations were applied. Studies were selected if they were cohort or case-control observational studies (non-randomized) that evaluated the

Figure 1: Outlier is Trial 002: Forest Plot of Relative Risk of Preterm Delivery in RCTs and Observational Studies in the Indicated Population and RCTs in Non-Indicated High Risk Populations



- Trials 002 (Meis) and 003 (PROLONG): Randomized Controlled Trials (RCTs) for Makena's intended population
- Price, Rouse, Caritis: RCTs for women at high risk for PTB (HIV – Price; multiple gestations – Caritis, Rouse)
- Massa, Hakim, Wang: Observational studies with untreated concurrent comparator
- Studies that do not report relative risk (Bastek, Nelson) are excluded from this figure. Both studies found no difference in overall PTB rates comparing study periods before and after Makena's approval among Makena's indicated population.

Examining the studies together, the available evidence does not show that Makena reduces the risk of recurrent PTB. Accordingly, Makena is no longer shown to be effective, and the grounds

effectiveness of HPC treatment compared to non-use. After screening at the titles and abstract level, five studies met these criteria. CDER last confirmed these search results on September 13, 2022.

for expedited withdrawal of approval under section 506(c)(3)(C) of the FD&C Act and 21 CFR 314.530(a)(6) are met.

Question No. 3: Should FDA allow Makena to remain on the market? As part of that discussion, you may discuss:

- whether the benefit-risk profile supports retaining the product on the market;
- what types of studies could provide confirmatory evidence to verify the clinical benefit of Makena on neonatal morbidity and mortality from complications of preterm birth?

Response: No. Makena's benefit-risk profile does not support the product remaining on the market, given Makena's known risks and lack of demonstrated benefit.

Not only did the postmarketing confirmatory trial fail to verify clinical benefit to neonates, but it also showed no effect on the gestational-age endpoint that was the basis of the initial approval. Considering all evidence available today, the benefit-risk profile of Makena is unfavorable, for its indicated population and any identified subset of that population, due to its lack of demonstrated efficacy, as well as the risks associated with the drug. Adverse events associated with Makena can be severe or serious, such as thromboembolic events (i.e., blood clots) and depression. Any amount of risk is unacceptable without countervailing benefit. A recent study reporting increased cancer risk in the offspring of women treated with HPC, the active ingredient in Makena, highlights the uncertainty regarding the intergenerational safety of Makena for the children of women who took it during pregnancy.[30] Failing to withdraw Makena would maintain FDA approval of a drug that, based on all available evidence, has not been shown to be more effective than, but is riskier than, no treatment.

Nothing other than a prospective, randomized, double-blind, placebo-controlled trial could adequately verify clinical benefit, especially in light of the negative results of Trial 003. Inherent limitations to observational studies or externally controlled trials, whether retrospective or prospective, and limitations to randomized unblinded or non-placebo-controlled trials, in the context of Makena being the only approved pharmacotherapy for recurrent sPTB, preclude the use of these study designs to obtain reliable evidence of Makena's efficacy. Because sPTB and recurrent sPTB are unpredictable and poorly understood, it is impossible to identify ahead of time, and control for, all potential confounding patient characteristics in the absence of a randomized, placebo-controlled trial. Further, because professional societies continue to recommend Makena's use, or at least recommend that providers consider using Makena, patients who are not prescribed Makena in clinical practice (i.e., those who might serve as the control group in an observational study or externally controlled trial) will be substantially different in baseline characteristics than patients for whom Makena would be recommended per treatment

[30] Murphy CC, Cirillo PM, Krigbaum NY, Cohn BA. In utero exposure to 17α-hydroxyprogesterone caproate and risk of cancer in offspring. Am. J. Obstet. Gynecol. 2022 Jan;226(1):132.e1–132.e14.

guidelines. Thus, without randomization, a placebo control, and blinding, it would not be possible to conclude whether any treatment effect, if one is seen, could be attributed to Makena.

Question No. 4 (For vote): Considering your responses to the previous questions both in the discussions and votes, should FDA allow Makena to remain on the market while an appropriate confirmatory study is designed and conducted?

Response: No. For the reasons discussed in greater detail above, Makena should not remain on the market while a confirmatory study is designed and conducted. The only randomized, double-blind, placebo-controlled study that measured Makena's purported clinical benefits not only failed to show any such benefits, but also failed to show drug effect on the gestational age of delivery endpoint that was the basis of Makena's approval. Trial 003 was well-designed, well-conducted, and appropriately powered to assess Makena's drug effect on these outcomes. Furthermore, multiple subgroup analyses of Trials 002 and 003 did not identify any subgroup of subjects, such as those with certain risk factors or demographic characteristics, for which efficacy was consistently shown in both Trials 002 and 003. In the absence of demonstrated benefit, Makena's benefit-risk profile is plainly unfavorable – for both Makena's overall indicated population, and for any identified subset of that population.

If the sponsor wishes to study Makena in another RCT in the United States, the only potentially feasible way to do so is to first withdraw the drug from the market. While Makena remains on the market, conducting another randomized, double-blind, placebo-controlled trial is infeasible in the U.S., because most patients are extremely unlikely to risk being randomized into a placebo arm when they can ensure they receive Makena by *not* enrolling in such a trial. Even if such a trial could be conducted (presumably by enrolling largely or entirely women from outside the U.S.), it would likely take at least another decade, as seen for Trial 003, before results that could potentially alter the current negative benefit-risk calculus could be obtained, submitted, and analyzed. Given that Makena is no longer shown to be effective, but does carry risks, there is no public health justification for allowing Makena to remain on the market in the interim, and doing so would undermine the integrity of the accelerated approval program. We fully acknowledge the gravity of removing the only therapy currently approved to reduce the risk of recurrent sPTB. However, we have concluded that it is important to proceed to protect patients from the unnecessary risks of a drug that has not been shown to be effective.

We have long recognized the need for safe and effective treatments for reducing neonatal morbidity or mortality from PTB. The serious public health consequences of PTB and lack of other approved treatments for this condition factored into CDER's decision to grant accelerated approval to Makena in 2011. The accelerated approval pathway allows for earlier approval of drugs that could provide a meaningful therapeutic advantage over available treatments, including no treatment, for a serious or life-threatening disease or condition. The availability of expedited withdrawal for these drugs is an important tool to counterbalance the earlier approval provided by this pathway, given the degree of uncertainty of the surrogate or intermediate clinical endpoint's ability to predict clinical benefit. Here, the confirmatory trial not only failed to verify clinical benefit—it also failed to corroborate the drug effect on the gestational age endpoint that

was the basis of the accelerated approval. Accordingly, based on the available evidence, the drug is no longer shown to provide an advantage over no treatment, and should not remain on the market.

III. LEGAL AND REGULATORY FRAMEWORK

A. Overview of the Accelerated Approval Program

The accelerated approval pathway aims to expedite the approval, and therefore availability, of "new drug products that have been studied for their safety and effectiveness in treating serious or life-threatening illnesses and that provide meaningful therapeutic benefit to patients over existing treatments."[31]

The accelerated approval pathway has evolved since the Agency established it thirty years ago. In 1997, Congress codified the Agency's authority to act under the accelerated approval program in the Food and Drug Administration Modernization Act of 1997 (FDAMA),[32] adding section 506 to the FD&C Act. In 2012, Congress amended section 506 of the FD&C Act via the Food and Drug Administration Safety and Innovation Act of 2012 (FDASIA) to facilitate somewhat broader use of the accelerated approval pathway. Section 901 of FDASIA amended the FD&C Act to provide that FDA should take into account the "severity, rarity, or prevalence of the condition and the availability or lack of alternative treatments."[33]

For a drug granted accelerated approval, the sponsor is generally required to conduct postmarketing confirmatory trials to verify and describe the drug's predicted effect on clinical benefit.[34] FDA has the legal authority to use expedited procedures to withdraw approval of a product that has received an accelerated approval if, among other reasons, "a study required to verify and describe the predicted effect on irreversible morbidity or mortality or other clinical benefit of the product fails to verify and describe such effect or benefit," or "other evidence demonstrates that the product is not safe or effective under the conditions of use."[35]

The accelerated approval pathway, like the traditional approval pathway, is rooted in the fundamental requirement that a drug product must be both safe and effective as a condition of marketing approval in the United States. Under both pathways, there must be substantial evidence at the time of approval that the drug is effective for its proposed conditions of use.[36]

---

[31] 21 CFR 314.500. In 1992, FDA issued regulations to create the accelerated approval pathway to expedite the approval of promising new therapies for treatment of serious and life-threatening illnesses. FDA proposed the accelerated approval regulations in April 1992 and adopted them in December 1992. 57 Fed. Reg. 13234 (Apr. 15, 1992); 57 Fed. Reg. 53942 (Dec. 11, 1992).

[32] Public Law 105-115.

[33] Public Law 112-144; section 506(c)(1)(A) of the FD&C Act.

[34] Section 506(c)(2)(A) of the FD&C Act.

[35] Section 506(c)(3) of the FD&C Act.

[36] Section 505(d) of the FD&C Act requires "substantial evidence" of effectiveness, which means "evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by

For traditional approval, effectiveness is based on a measurement of clinical benefit or on a validated surrogate endpoint,[37] whereas accelerated approval is based on a drug's effect on a surrogate or intermediate clinical endpoint that is "reasonably likely . . . to predict [a drug's] clinical benefit."[38] Compared to traditional approval, the uncertainty under accelerated approval lies in the nature of the efficacy endpoint: the surrogate or intermediate clinical endpoint is *reasonably likely to predict* clinical benefit. Because of this uncertainty, the accelerated approval pathway is reserved only for products that treat a serious or life-threatening disease or condition and generally only where the new product provides a meaningful therapeutic benefit over existing treatments, including no treatment.

Under the accelerated approval pathway, products are subject to the expedited withdrawal provisions of 21 CFR 314.530(a)(1) through (6). If a postmarketing confirmatory study fails to verify clinical benefit,[39] "other evidence demonstrates that the drug product is not shown to be safe or effective under its conditions of use,"[40] or if any other grounds for withdrawal set forth in the statue and regulations are met, FDA may withdraw approval of the drug product. As explained in its initial rulemaking establishing the accelerated approval pathway, the Agency must have the authority to expedite the withdrawal process when a clinical benefit is not confirmed because "[o]therwise, the risk of continued exposure of patients with serious or life-threatening diseases to ineffective or unsafe drugs outweighs the potential benefits."[41]

In the thirty years since its creation, the accelerated approval pathway has been used primarily in settings in which the disease course is long and an extended period of time would be required to measure the intended clinical benefit of a drug.[42] The program has allowed for earlier approval of many new treatment options for patients with serious or life-threatening illnesses and conditions, in some cases where there are no available alternatives.

---

scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof"; *see also* Section 506(e)(2), which states in part, "Nothing in this section shall be construed to alter the standards of evidence under subsection (c) or (d) of section 505 (including the substantial evidence standard in section 505(d) of this Act)…[.] Such sections and standards of evidence apply to the review and approval of products under this section, including whether a product is safe and effective."

[37] Section 505 of the FD&C Act; Expedited Programs Guidance, at 17–19.

[38] Section 506(c)(1)(A) of the FD&C Act.

[39] 21 CFR 314.530(a)(1).

[40] 21 CFR 314.530(a)(6).

[41] New Drug, Antibiotic, and Biological Drug Product Regulations; Accelerated Approval, Proposed Rule, 57 Fed. Reg. 13234, 13239 (Apr. 15, 1992), available at https://www.govinfo.gov/app/details/FR-1992-04-15.

[42] *See* Expedited Programs Guidance, at 15.

B. Verification of Clinical Benefit

For drugs granted accelerated approval, FDA generally requires postmarketing confirmatory trials to verify and describe the clinical benefit of the product.[43] Generally, the confirmatory trial evaluates a clinical endpoint that directly measures clinical benefit. A clinical endpoint is a variable that directly measures a therapeutic effect of a drug—an effect on how a patient feels (e.g., symptom relief), functions (e.g., improved mobility, no need for supplemental oxygen, free from complications of intracranial bleeding), or survives.[44] A clinical benefit is a positive therapeutic effect that is clinically meaningful in the context of a given disease.[45] These postmarketing confirmatory trials must be completed with due diligence, to determine whether the predicted clinical benefit has been verified.[46]

C. Withdrawal Standard

Expedited withdrawal is an integral part of the accelerated approval framework. Section 506 of the FD&C Act provides FDA the legal authority to both accelerate the approval of NDAs and to expedite the withdrawal of approval of a drug product that was approved under the accelerated approval framework.[47] The pathway enables earlier access to a promising drug on the condition that sponsors conduct one or more studies to confirm whether the drug actually delivers its intended clinical benefit. If such studies do not verify benefit, the well-known consequence is that withdrawal is generally expected to follow.

FDA has the statutory authority to use expedited procedures to withdraw approval of a product that has received an accelerated approval if, among other reasons, "a study required to verify and describe the predicted effect on irreversible morbidity or mortality or other clinical benefit of the product fails to verify and describe such effect or benefit,"[48] or "other evidence demonstrates that the product is not safe or effective under the conditions of use."[49] Conditions of use include use for the indication described in the labeling, which includes the patient population.[50] FDA regulations provide that FDA may withdraw an accelerated approval when "[a] postmarketing

[43] Section 506(c)(2)(A) of the FD&C Act; *see also* 21 CFR 314.510.

[44] Expedited Programs Guidance, at 17.

[45] *Id.*

[46] Section 506(c)(3)(A) of the FD&C Act.

[47] Section 506(c)(3) of the FD&C Act.

[48] Section 506(c)(3)(B) of the FD&C Act.

[49] Section 506(c)(3)(C) of the FD&C Act.

[50] *See, e.g.*, FDA Guidance for Industry, Medical Product Communications That Are Consistent With the FDA-Required Labeling Questions and Answers (June 2018), available at https://www.fda.gov/media/133619/download, at 4–5.

clinical study fails to verify clinical benefit[,]"[51] or "[o]ther evidence demonstrates that the drug product is not shown to be safe or effective under its conditions of use."[52]

Regarding generic versions of drugs, FDA's legal framework establishes, with certain exceptions, that abbreviated new drug applications (ANDAs or generic drug applications) relying on a reference listed drug will be withdrawn when the agency has completed the withdrawal of approval of the listed drug, as further described below. Neither the withdrawal provisions of section 506(c) of the FD&C Act nor 21 CFR part 314 subpart H explicitly addresses the status of ANDAs that rely on a reference listed drug approved under section 505(c) pursuant to section 506(c) when approval of that reference listed drug is withdrawn. Section 505(e) of the FD&C Act includes certain grounds for withdrawal that could apply to an ANDA for reasons specific to the ANDA.[53] In addition, under section 505(j)(6) of the FD&C Act, FDA has the authority to withdraw approval of an ANDA when the listed drug it references was withdrawn for grounds described in the first sentence of section 505(e) or was withdrawn under section 505(j)(6) or which, as determined by the Secretary, has been withdrawn from sale for safety or effectiveness reasons. Thus, under section 505(j)(6), if an ANDA refers to a listed drug that has been withdrawn as described in the previous sentence, withdrawal of the ANDA under section 505(j)(6) will follow. FDA regulations (21 CFR 314.150 and 314.151) address withdrawal of ANDAs. Specifically, 21 CFR 314.151 addresses withdrawal of ANDAs when approval of the NDA for the reference listed drug is withdrawn. The regulations provide, in part, that the approval of an ANDA "identified in the notice of opportunity for hearing on the withdrawal of a listed drug will be withdrawn when the agency has completed the withdrawal of approval of the listed drug."[54]

IV. ACCELERATED APPROVAL OF MAKENA FOR REDUCING THE RISK OF RECURRENT PRETERM BIRTH IN SINGLETON PREGNANCY

On February 3, 2011, FDA approved NDA 021945 for Makena (hydroxyprogesterone caproate [or HPC] injection) under the accelerated approval pathway[55] to reduce the risk of PTB in women with a singleton pregnancy who have a history of singleton sPTB.[56] In support of efficacy, the NDA relied on data from a single trial, the MFMU Network Trial 17P-CT-002 (Trial 002),[57] in which, compared to placebo, Makena reduced the proportion of women

[51] 21 CFR 314.530(a)(1).

[52] 21 CFR 314.530(a)(6).

[53] Section 505(e) of the FD&C Act.

[54] 21 CFR 314.151(b)(3).

[55] Section 506(c) of the FD&C Act; 21 CFR part 314, subpart H.

[56] The February 2011 original approval letter for Makena is included in the action package for NDA 021945, which is available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945_makena_toc.cfm.

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

delivering prior to 37 weeks gestation, an intermediate clinical endpoint reasonably likely to predict clinical benefit to neonates.

Makena's indication, usage, dosage, and administration are described in the product's approved labeling:

> Makena is a progestin indicated to reduce the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth.
>
> …
>
> Administer intramuscularly at a dose of 250 mg (1 mL) once weekly
>
> Begin treatment between 16 weeks, 0 days and 20 weeks, 6 days of gestation
>
> Continue administration once weekly until week 37 (through 36 weeks, 6 days) of gestation or delivery, whichever occurs first (2.1)[58]

Since Makena's approval, FDA approved eight ANDAs for drug products referencing Makena as their referenced listed drug: (1) ANDA 211070, held by Eugia Pharma Specialties Limited; (2) ANDA 211071, held by Eugia Pharma Specialties Limited; (3) ANDA 210618, held by Slayback Pharma LLC; (4) ANDA 210877, held by Slayback Pharma LLC; (5) ANDA 208381, held by Sun Pharmaceuticals Industries Ltd; (6) ANDA 210723, held by American Regent, Inc.; (7) ANDA 210724, held by American Regent, Inc.; and (8) ANDA 211777, held by Aspen Pharma USA Inc.

A. The MFMU Network Trial ("Trial 002")

The Makena NDA relied on evidence from Trial 002 as primary support of efficacy and safety. Trial 002 was a randomized, double-blind, placebo-controlled trial designed with a planned sample size of 500 women to detect a 33% relative reduction (from 37% to 25%) in the rate of PTB < 37 weeks with 80% power. Initiated in 1999 and completed in 2002, Trial 002 ultimately enrolled 463 women with a singleton pregnancy and at least one prior singleton sPTB from 19 university-based clinical centers in the United States in the MFMU Network.[59] The primary efficacy endpoint was the proportion of pregnant women delivering < 37 weeks gestation, and secondary endpoints included the proportion of those delivering < 35 or < 32 weeks gestation. The data and safety monitoring board recommended stopping the trial at an interim analysis based on outcome data from 351 randomized participants. Women randomized up to that time remained in the trial, resulting in 463 women with outcome data. Trial 002 results showed that 37% of women given HPC delivered prior to 37 weeks gestation while 55% of women given

[58] *See* FDA-approved prescribing information for Makena (INDICATIONS AND USAGE, DOSAGE AND ADMINISTRATION), available at https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s013lbl.pdf and attached as Appendix 10.

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

placebo did so. The treatment effect was most evident for later PTB (delivery ≥ 35 weeks to < 37 weeks gestation). The proportions of women delivering at < 35 and < 32 weeks gestation were also lower among women randomized to Makena compared to placebo; however, the absolute treatment differences were smaller relative to the treatment difference for < 37 weeks. Efficacy results are summarized below:

Table 1: Efficacy – Proportion of Trial 002 Subjects Delivering at < 37, < 35, and < 32 Weeks Gestational Age (ITT Population)

Efficacy Outcome	HPC (Makena) (N=310)[1]	Placebo (N=153)	Absolute % Treatment Difference (95% CI)[2]	Relative Risk (95% CI)[2]
Birth < 37 weeks	37%	55%	-18% (-28, -7)	0.68 (0.54, 0.84)
Birth < 35 weeks	21%	31%	-9% (-19, -0.4)	0.69 (0.49, 0.98)
Birth < 32 weeks	12%	20%	-8% (-16, -0.3)	0.61 (0.38, 0.98)

Source: Adapted from Table 5 in Makena's prescribing information[60]

[1] Four Makena-treated subjects were lost to follow-up. They were counted as deliveries at their gestational ages at time of last contact (18^4, 22^0, 34^3, and 36^4 weeks).

[2] Adjusted for interim analyses; the final analyses use a nominal p-value of 0.0345 (Z-score = 2.1232) and the adjusted confidence intervals (equivalent to a 96.6% confidence interval) to preserve the overall Type I error of 0.05.[61]

Trial 002 was published in the New England Journal of Medicine (NEJM) in 2003, before the original sponsor first sought marketing approval in 2006,[62] and reported that HPC 250 mg/mL administered intramuscularly once weekly reduced the risk of preterm delivery in women with a prior sPTB.[63]

In Trial 002, the absence of a prespecified efficacy endpoint on neonatal outcomes, the ultimate clinical outcome of interest for treatment related to sPTB, and selection of gestational age of delivery < 37 weeks as the primary efficacy endpoint, posed challenges in our review of the Makena application (see section IV.B below).[64] In general, neonatal outcomes are optimal when

[60] *See* FDA-approved prescribing information for Makena, available at https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s013lbl.pdf and attached as Appendix 10.

[61] Statistical Review in the February 2011 action package for NDA 021945, available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945_makena_toc.cfm.

[62] *See* Advisory Committee on Reproductive Health Drugs (ACRHD) Meeting on Gestiva (August 29, 2016) (2006 ACRHD Meeting), Appendix 1. The original sponsor submitted the NDA for Makena in 2006. The 2006 Advisory Committee considered whether the primary endpoint of Trial 002, prevention of PTB prior to 37 weeks gestation, was an adequate surrogate for a reduction in fetal and neonatal mortality or morbidity. By a vote of 16 to 5, the 2006 Advisory Committee voted that this endpoint was not an adequate surrogate as a measure for the clinical benefit for Makena. The 2006 Advisory Committee also voted that prevention of PTB prior to 35 weeks gestation (13 Yes to 8 No) and prior to 32 weeks gestation (20 Yes to 1 No) would both be adequate surrogates for a reduction in fetal and neonatal mortality or morbidity."

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

[64] Trial 002 was designed, conducted, completed, and published in 2003 without any CDER involvement. According to the Cross Discipline Team Leader's review memo (February 3, 2011), a notable issue of concern was the primary endpoint of reduction in PTB < 37 weeks. As noted in this review, "…the Division believed that demonstration of treatment benefit should focus on reduction of [neonatal] morbidity and mortality, rather than on

delivery occurs at term, defined as delivering between 37 to 40 weeks gestation. PTBs, defined as delivery between 20 and prior to 37 weeks gestation, can be associated with significant neonatal morbidity and mortality. Because it is related to the degree of fetal development, gestational age at delivery could be considered a proxy for neonatal developmental health, to varying degrees. In general, the likelihood of adverse outcomes in the neonate born spontaneously prematurely increases with decreasing gestational age at delivery. However, the likelihood and severity of adverse neonatal outcomes do not correlate linearly with gestational age at delivery. For example, it is expected that prolonging pregnancy could result in improved neonatal morbidity/mortality in extremely premature infants (e.g., those born at < 28 weeks gestation). It is less clear whether prolonging pregnancy through medical interventions could result in improved neonatal morbidity/mortality in neonates born late preterm. Another factor adding to the uncertainty of the relationship between gestational age at delivery and neonatal outcomes is that the mechanisms and underlying causes of preterm labor and PTB are poorly understood. It could be that preterm labor leading to PTB may be triggered by an unrecognized toxic uterine environment, and medical interventions to prolong the pregnancy with pharmacotherapy may render a worse outcome to the neonate than if sPTBs were allowed to occur. Thus, there is uncertainty regarding the impact that prolonging pregnancy through pharmacotherapy, especially in later gestation, has on improving neonatal outcomes.

B. Considerations During Initial Approval of Makena Based on Trial 002

The approval of Makena required three review submissions and complicated considerations. In the first review cycle in 2006, CDER assessed that Trial 002 could not support traditional approval because it was not designed to and did not evaluate the drug's efficacy on the clinical benefit of interest—neonatal morbidity/mortality from PTB—and that its primary efficacy endpoint of gestational age of delivery < 37 weeks was not an adequate surrogate. At CDER's request during the review cycle in 2006, the sponsor conducted a post-hoc analysis of neonatal outcomes (neonatal composite index). The composite index was based on the number of infants experiencing one of the following: death, respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), grade 3 or 4 intraventricular hemorrhage (IVH), proven sepsis, or necrotizing enterocolitis (NEC).[65] Although there was a lower proportion of subjects in the 17-HPC group (11.9% vs. 17.2% in the placebo group) who experienced at least one event of the composite endpoint, this difference was not statistically significant. Members of the 2006 Advisory Committee for Reproductive Health Drugs (ACRHD) 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 result of the post-hoc neonatal morbidity/mortality composite index.[66]

---

increasing the gestational age at delivery without any associated clinical benefit." *See* February 2011 action package for NDA 021945 available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945_makena_toc.cfm.

[65] Although Trial 002 collected data on these individual neonatal outcomes, they were not prespecified for efficacy assessment.

[66] 2006 ACRHD Meeting on Gestiva (August 29, 2016), Appendix 1.

In the first and second review cycles, there was uncertainty about whether Trial 002's primary endpoint (PTB < 37 weeks) was clinically relevant for accelerated approval. Members of the 2006 ACRHD opined that gestational age of delivery < 37 weeks was not reasonably likely to predict clinical benefit to the neonate, but gestational age of < 35 weeks could be adequate.[67] CDER determined the findings for the gestational age < 35 weeks in Trial 002 were not robust enough to approve under accelerated approval based on only a single adequate and well-controlled study. When the NDA was resubmitted in the third review cycle in 2010, new data had emerged that infants born "late preterm" (between 34 and < 37 weeks gestation) are at higher risk of adverse neonatal outcomes than term infants.[68] This new evidence led CDER to determine that gestational age < 37 weeks was an acceptable endpoint reasonably likely to predict clinical benefit to the neonate from complications of PTB. Accordingly, CDER reconsidered the data from Trial 002 for this endpoint as the basis of accelerated approval.

Covis notes that CDER has previously described reduction in PTB < 32 or < 35 weeks as a "well-established surrogate" or "established surrogate" and as such, Trial 002 should have supported traditional approval based on the drug effect on these two gestational age thresholds. We disagree. According to the 2011 division director's summary memo,[69] "[a]t the end of both the first and second review cycles, DRUP [Division of Reproductive and Urological Products] reviewers were focusing on the clinical finding of a reduction in preterm births at < 35^0 weeks as a possible basis for approval under [accelerated approval]."[70] In fact, CDER's advice that the clinical endpoint of neonatal composite index be a co-primary endpoint along with gestational age of delivery < 35 weeks in Trial 003 reflects its assessment that the latter endpoint was not validated for traditional approval. Further, Makena's effect on gestational age at delivery either < 35 weeks or < 32 weeks in Trial 002 was not statistically persuasive enough to support substantial evidence of effectiveness based on a single clinical trial.

Although Makena was ultimately approved, there were several review issues that are relevant to this hearing. First, there was only one adequate and well-controlled trial (Trial 002) for Makena's effect on gestational age of delivery < 37 weeks. No other robust clinical evidence existed to substantiate or corroborate the drug's efficacy at the time of the 2011 approval. As explained in section III., there must be substantial evidence at the time of approval that the drug is effective for its proposed conditions of use. Traditionally, FDA has interpreted substantial

---

[67] The 2006 ACRHD considered whether the primary endpoint of Trial 002, prevention of PTB prior to 37 weeks gestation, was an adequate surrogate for a reduction in fetal and neonatal mortality or morbidity. By a vote of 16 to 5, the 2006 ACRHD voted that this endpoint was not an adequate surrogate as a measure for the clinical benefit for Makena. The 2006 ACRHD also voted that prevention of PTB prior to 35 weeks gestation (13 Yes to 8 No) and prior to 32 weeks gestation (20 Yes to 1 No) would both be adequate surrogates for a reduction in fetal and neonatal mortality or morbidity. 2006 ACRHD Meeting on Gestiva (August 29, 2016), Appendix 1.

[68] Engle WA, Tomashek KM, Wallman C, Committee on Fetus and Newborn. "Late-preterm" infants: a population at risk. Pediatrics. 2007 Dec;120(6):1390-401.

[69] Summary Review in the February 2011 action package for NDA 021945, available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945_makena_toc.cfm.

[70] The phrase "35^0 weeks" refers to 35 weeks and zero days.

evidence of effectiveness (SEE) as clinically and statistically significant findings from at least two adequate and well-controlled trials.[71] Having at least two adequate and well-controlled trials ensures independent substantiation of experimental findings and strengthens a conclusion of effectiveness. Conclusions based on two high-quality trials will generally be more reliable than those based on a single comparably persuasive study. In certain circumstances, however, it may be appropriate to rely on one adequate and well-controlled trial to provide substantial evidence of effectiveness. Based on Trial 002, CDER determined there was substantial evidence of Makena's effect on gestational age of delivery < 37 weeks, and this effect was reasonably likely to predict clinical benefit. Nonetheless, a positive finding from a single adequate and well-controlled trial, even if well-conducted and considered sufficient for providing substantial evidence, may have unknown biases or may reflect a chance finding that would not be present in a second adequate and well-controlled trial.[72] Indeed, a condition of FDA's approval of Makena under the accelerated approval pathway was that the sponsor conduct a second trial (Trial 003).

Second, certain aspects of Trial 002 posed concerns about generalizability of the trial's findings. These concerns included that subject recruitment was exclusively from the 19 MFMU academic centers (with 27% of subjects from a single academic center, University of Alabama), and that 43% of Black subjects came from a single center, University of Alabama. Also, the rate of recurrent PTB in the placebo arm (55%) in Trial 002 was exceptionally high, given the expected rate from a MFMU trial in high-risk women used to power Trial 003 was 37%.[73,74,75] Both CDER and members of the 2006 advisory committee questioned whether the high placebo rate in Trial 002 might have been a factor in the efficacy demonstrated for Makena. Although these issues were adequately addressed for purposes of Makena's accelerated approval, the study population of Trial 002 was nonetheless not representative of Makena's indicated population—an issue that the confirmatory trial could address. Trial 003 was therefore designed to provide data from a broader population by using approaches such as ensuring that no site enrolled more than 15% of the total number of subjects and that at least 10% of the total sample size came from the U.S. and Canada. An international trial for Trial 003 was planned because it was expected that after Makena was approved and commercially available in the U.S., enrolling such a large, placebo-controlled trial only in the U.S. was not possible. CDER was reassured, at the time of

[71] FDA Draft Guidance for Industry: "Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products" (December 2019), available at https://www.fda.gov/media/133660/download. The SEE standard applies to both the traditional and accelerated approval pathways. Compared to traditional approval, the greater uncertainty under accelerated approval lies in the nature of the efficacy endpoint – but in both cases, SEE must be shown.

[72] *Id.*

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

[74] February 2011 action package for NDA 021945, available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945_makena_toc.cfm.

[75] *See* 2006 ACRHD Meeting on Gestiva (August 29, 2006), Appendix 1.

Makena's approval, that the sponsor had recruited from a diverse group of U.S. sites, including academic centers, military medical centers, and private practices, for Trial 003.

Makena was eligible for accelerated approval in part because PTB is a serious condition for which no other FDA-approved treatment is available. Through the accelerated approval pathway, Makena was allowed to enter the market despite remaining uncertainty regarding clinical benefit because Trial 002 demonstrated that it had an effect on rates of PTB < 37 weeks, an endpoint reasonably likely to predict neonatal outcomes. Therefore, Makena's approval was conditioned on the sponsor's commitment to conduct a second adequate and well-controlled trial to verify Makena's predicted clinical benefit.

V. BASIS FOR PROPOSAL TO WITHDRAW APPROVAL OF MAKENA AND ITS GENERICS

As explained in section III.C, FDA may withdraw approval of a drug product approved through the accelerated approval pathway pursuant to section 506(c)(3)(B) of the FD&C Act and 21 CFR 314.530(a)(1) if a postmarketing confirmatory study fails to verify clinical benefit. In addition, FDA may withdraw approval of the drug product pursuant to section 506(c)(3)(C) of the FD&C Act and 21 CFR 314.530(a)(6) if other evidence demonstrates that the drug product is not shown to be safe or effective under its conditions of use. As described in section III.A, a drug's effectiveness must be supported by substantial evidence. If there is no longer substantial evidence that a drug product approved through the accelerated approval pathway is effective, then FDA may withdraw the drug product's approval pursuant to this second ground of not being shown to be effective for its approved use.

Both of these independent grounds for expedited withdrawal of approval are met here. First, the required postmarketing clinical trial, Trial 003, failed to verify the predicted clinical benefit of reducing neonatal morbidity and mortality from complications of PTB. Second, given the results of Trial 003, Makena is no longer shown to be effective at reducing the risk of recurrent sPTB. That is, the available evidence—including Trials 002, 003, and additional evidence discussed below—does not provide substantial evidence of Makena's effectiveness. Either of these bases on its own is sufficient for withdrawal of approval under the statute and regulations. Taken together, they make the case for withdrawal especially compelling.

Other factors also weigh in favor of withdrawal of approval. First, notwithstanding Covis' arguments to the contrary, there is not substantial evidence that Makena is effective for its approved use in its indicated population, and there is also not substantial evidence that Makena is effective for any identified subset of women. Accordingly, there is no basis to retain Makena on the market and narrow its labeling to "high-risk" women, as Covis suggests.

Next, particularly in light of the negative result of Trial 003, only a prospective, randomized, double-blind, placebo-controlled trial could potentially provide sufficient evidence of Makena's effectiveness, but such a trial is infeasible in the U.S. so long as Makena remains approved. That is, leaving Makena on the market is a critical impediment to the sponsor's proposal to conduct

another placebo-controlled trial of the drug in a study population comprised predominantly of women in the U.S. Although a trial predominantly or entirely outside the U.S. may be feasible, it would take many years to complete, and there is no public health justification for leaving Makena on the market in the meantime.

Furthermore, withdrawal is needed to maintain the integrity of the accelerated approval pathway. If Makena remains on the market while the sponsor conducts another randomized, double-blind, placebo-controlled trial of the size and rigor that could potentially be sufficient to provide evidence of Makena's effectiveness in the face of the negative results from Trial 003, it would likely take at least another 10 years to complete the trial. This would mean that Makena would remain approved for at least 20 years without confirmation of clinical benefit and now (post-Trial 003) without substantial evidence of effectiveness for its approved use. If CDER cannot withdraw drugs approved under the accelerated approval pathway when multiple grounds for withdrawal exist and the prospect for demonstrating effectiveness is at best remote, this would frustrate Congress' purpose in providing for expedited withdrawal of drugs approved under this pathway.

Lastly, we acknowledge the significant need for treatment for PTB and the public health urgency to address PTB and the devastating effects it has on children and their families. CDER would welcome an effective treatment for PTB. However, Makena is not shown to be effective, and absent such evidence, it is important that the Agency withdraw Makena from the market. At-risk women and their children are entitled to the same assurance as everyone else that the FDA-approved drugs they take are safe and effective. Not withdrawing Makena after it is no longer shown to be effective would amount to a failure to provide this assurance.

A. Trial 003 Failed to Verify the Expected Clinical Benefit of Makena on Neonatal Morbidity and Mortality from Complications of Preterm Birth

Trial 003 failed to verify the expected clinical benefit of Makena on neonatal morbidity and mortality from complications of PTB.

Initiated in 2009 and completed in 2018, Trial 003 was an international, randomized, double-blind, placebo-controlled trial specifically designed to verify Makena's clinical benefit in women with essentially the same eligibility criteria as those of Trial 002. That is, the inclusion and exclusion criteria had very minor differences (e.g., slight change in allowable gestational age at randomization). The study populations in both trials met the criteria to receive Makena as currently labeled for its approved use. Trial 003 was intended both to provide evidence of Makena's clinical benefit for neonates resulting from complications of PTB and to substantiate Makena's effect on the intermediate clinical endpoint of gestational age of delivery (as this effect had only been shown in one clinical trial, Trial 002, at the time of Makena's approval in 2011).

Therefore, Trial 003's co-primary efficacy endpoints[76] were (a) delivery < 35 weeks gestation and (b) a neonatal morbidity/mortality composite index.

Trial 003 was initiated in the U.S. and Canada in 2009 to ensure at least 10% of the 1,700 planned subjects (at least 170 subjects) would be from the U.S. and Canada before expanding to Europe. After Makena's approval in 2011, enrolling U.S. subjects became increasingly difficult, and recruitment relied on sites outside the U.S., including in Russia and Ukraine. Despite these challenges, Trial 003 randomized a total of 1,708 women from nine countries, compared to 463 women in Trial 002. Russia, Ukraine, and the U.S. were the three highest enrolling countries, randomizing 621 (36%), 420 (25%) and 391 (23%) subjects, respectively.[77,78] While Trial 002 enrolled exclusively from the 19 university-based clinical centers in the United States in the MFMU Network, with the University of Alabama recruiting 27% of the study population and 43% of the Black women, the U.S. portion of Trial 003 enrolled from 41 centers from diverse sites, including academic centers, military medical centers, and private practices, and consisted of 29% Black women (113 of 391 U.S. subjects). Trial 003 provided data for 1,651 liveborn neonates. The Makena and placebo groups were balanced across all reported demographics, baseline characteristics, and obstetrical characteristics in the current and previous pregnancies.

Trial 003 failed to demonstrate a statistically significant treatment effect on the co-primary endpoints (proportion of women delivering prior to 35 weeks and neonatal composite index). Also, no differences between Makena and placebo were seen in the secondary outcomes of delivery < 32 or < 37 weeks (< 37 weeks was the primary efficacy endpoint in Trial 002 that formed the basis for accelerated approval), as shown in Table 2 below. There was no evident numerical separation between Makena and placebo looking at the point estimate of the treatment difference for these efficacy endpoints.

[76] Multiple primary endpoints become co-primary endpoints when it is necessary to demonstrate an effect on each of the endpoints to conclude a drug is effective.

[77] The number of U.S. women enrolled in Trial 003 (N=391, Trial 003 U.S. subgroup) more than doubled CDER's recommendation that at least 10% of subjects (170 subjects) be from the U.S. and Canada and was close to the number of U.S. women enrolled in Trial 002 (N=463).

[78] The majority of U.S. participants in Trial 003 were enrolled before Makena's approval.

Table 2: Efficacy – Confirmatory Trial 003 Efficacy Results[79]

Efficacy Outcome*	Makena (N=1130)	Placebo (N=578)	Treatment Difference[§] (95% CI)	Relative Risk[§] (95% CI) ****	Statistically significant?
Neonatal composite index**	5.4%	5.2%	0.2% (-2.0, 2.5)	1.05 (0.68, 1.61)	No
Birth < 35 weeks	11%	12%	-0.6% (-3.8, 2.6)	0.95 (0.71, 1.26)	No
Birth < 32 weeks	5%	5%	-0.4% (-2.8, 1.7)	0.92 (0.60, 1.42)	No
Birth < 37 weeks***	23%	22%	1.3% (-3.0, 5.4)	1.06 (0.88, 1.28)	No

*Co-Primary endpoints: Neonatal composite index, Birth < 35 weeks. Secondary endpoints: Birth < 32 weeks, Birth < 37 weeks. Missing data were not imputed.
**Proportion of neonates experiencing at least one event of the composite index; N=1651 (Makena missing data on 39 neonates, Placebo missing data on 18 neonates); defined as Yes if the liveborn neonate had any of RDS, BPD, Grade 3 or 4 IVH, NEC, proven sepsis, death.
***Primary efficacy endpoint of Trial 002.
**** AMAG's slide CO-53 and Blackwell et al. list a RR for the neonatal composite index of 1.12 (0.72, 1.72), different from CDER's calculations. CDER calculated its results using the livebirth flag and the neonatal composite index variables. The Makena arm had 59 of 1091 liveborn neonates with an index event, and the Placebo arm 29 of 560 liveborn neonates had an index event.
[§]Cochran-Mantel-Haenszel (CMH) method stratified by gestational age at randomization; For treatment difference: p-value = 0.84 (neonatal composite index), p-value=0.72 (birth < 35 weeks).

Accordingly, the required postmarketing confirmatory study failed to verify clinical benefit. Pursuant to section 506(c)(3)(B) of the FD&C Act and 21 CFR 314.530(a)(1), FDA may withdraw approval of Makena on this basis alone. More simply put, the answer to the question of "*Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal morbidity and mortality from complications of preterm birth?*" is "*No.*"

B. Makena is Not Shown to Be Effective for Its Approved Indication

In addition, the available evidence demonstrates that Makena is not shown to be effective at reducing the risk of recurrent PTB in women with a singleton pregnancy who have a history of singleton spontaneous PTB, or improving neonatal outcomes.

Trial 003 not only failed to demonstrate Makena's benefit to the neonate, but it also failed to show any effect on the intermediate clinical endpoint of gestational age at delivery that was the basis of Makena's accelerated approval (see Table 2). The sponsor agreed, stating that Trial 003 "did not demonstrate a statistically significant difference between the treatment and placebo arms for the co-primary endpoints."[80] The BRUDAC met on October 29, 2019, to consider the findings of Trial 003 and discuss the evidence regarding the effectiveness of Makena in reducing

[79] CDER's Proposal to Withdraw Marketing Approval of MAKENA® (hydroxyprogesterone caproate injection); Notice of Opportunity for a Hearing (NOOH) (October 5, 2020) at 6, available at https://www.regulations.gov/document/FDA-2020-N-2029-0001.

[80] AMAG Pharmaceuticals Announces Topline Results from the PROLONG Trial Evaluating Makena, available at https://www.amagpharma.com/news/amag-pharmaceuticals-announces-topline-results-from-the-prolong-trial-evaluating-makena-hydroxyprogesterone-caproate-injection/.

the risk of recurrent PTB and improving neonatal outcomes. All 16 voting members of the BRUDAC concluded that the findings from Trial 003 failed to verify the clinical benefit of Makena on neonatal outcomes from complications of PTB.[81] The BRUDAC further concluded, by a vote of 13 (no) to 3 (yes), 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 PTB.[82]

Covis asserts that Trial 002 provides substantial evidence for Makena's effectiveness, and that Trial 003's failure to show Makena's effectiveness does not "invalidate" the Trial 002 results because Trial 003 was a flawed study.[83]

CDER agrees that, at the time of approval, there was substantial evidence of Makena's effectiveness based upon a single adequate and well-controlled trial. Makena was approved for use in all women with a singleton pregnancy who had a prior singleton sPTB as the identifying risk factor for recurrent PTB. However, despite essentially the same enrollment criteria defining the same patient population, Trial 003 did not substantiate Makena's treatment effect on gestational age seen in Trial 002. To reach the conclusion that there is substantial evidence of effectiveness, where we have multiple trials, we seek substantiation of benefits across those trials to have confidence that a drug effect seen in one trial is indeed a true effect. Plainly this is not the case here. Accordingly, we agree with Covis that Trial 003 does not "invalidate" Trial 002 – but it does conclusively fail to substantiate the drug effect seen in Trial 002. Two adequate and well-controlled trials assessing the same drug for the same condition, where the first has positive results and the second has negative results, generally constitutes sufficient information to conclude that there is a lack substantial evidence of effectiveness. It is not necessary to conclude that the negative study "invalidates" the positive study to determine lack of substantial evidence of effectiveness. Therefore, considering the results of Trials 002 and 003 together, there is not substantial evidence of effectiveness for Makena's indicated use. Where, as here, the available evidence does not meet the "substantial evidence" threshold, the drug is not shown to be effective.[84]

Accordingly, Makena is not shown to be effective at reducing the risk of PTB in women with a singleton pregnancy who had a prior singleton sPTB. Pursuant to section 506(c)(3)(C) of the FD&C Act and 21 CFR 314.530(a)(6), FDA may withdraw approval of Makena on this basis. More simply put, and as described further in sections D. and E., the answer to the question of "*Does the available evidence demonstrate that Makena is effective for its approved indication of*

[81] The 2019 BRUDAC voted 16 (No) to 0 (Yes) on the question "Do the findings from Trial 003 verify the clinical benefit of Makena on neonatal outcomes?" *See* Summary Minutes, BRUDAC Meeting 6, 9 (Oct. 29, 2019), available at https://www.regulations.gov/document/FDA-2020-N-2029-0118 and attached as Appendix 2.

[82] Summary Minutes, BRUDAC Meeting 6, 9 (Oct. 29, 2019), available at https://www.regulations.gov/document/FDA-2020-N-2029-0118 and attached as Appendix 2.

[83] CDER addresses Covis' arguments concerning the supposed "flaws" of Trial 003 in section V.D.

[84] Under section 505(d) of the FD&C Act, FDA may not approve a drug if "there is a lack of substantial evidence that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof."

reducing the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth?" is "*No.*"

C. Makena's Efficacy Has Not Been Demonstrated In Any Subgroup

Makena is not shown to be effective for its approved indication of reducing the risk of PTB in women with a singleton pregnancy who have a history of singleton spontaneous PTB. There is also not substantial evidence that Makena is effective for any identified subset of the women for whom it is approved.

1. Overview

CDER agrees with Covis that the populations of Trials 002 and 003 differed in certain prognostic factors (e.g., demographics and socioeconomic factors) for PTB. The relevant question regarding Makena's efficacy, however, is whether these prognostic factors are also effect modifiers—that is, whether pregnant women with one or more of these factors (e.g., being Black, having more than one prior sPTB) respond better, or worse, to Makena than those without these factors (e.g., not being Black, only one prior sPTB) – and therefore, whether the differences in these prognostic factors, and not Makena, could be responsible for the conflicting results between Trials 002 and 003. Accordingly, the sponsor and CDER conducted additional analyses of Trial 003 to assess: (1) whether there was a subpopulation (within Makena's approved population) in whom Makena showed evidence of therapeutic effect, and (2) whether the factors implicated in modulating the risk of recurrent PTB and that differed between Trial 002 and Trial 003, such as Black race, modify the response to Makena (i.e., are effect modifiers) that could explain the discrepant efficacy findings between Trials 002 and 003. Whereas Trial 002 showed a treatment effect for most if not all subgroups evaluated, Trial 003 did not show a treatment effect for any subgroup analyzed. Therefore, CDER did not identify a subpopulation within Makena's approved population for whom there was a consistent treatment effect of Makena on the endpoint of gestational age of delivery in both Trial 002 and Trial 003. This lack of consistency of treatment effect suggests that the treatment effect seen in Trial 002 could be a false positive. Regarding the risk factors that differed between the two trials, CDER determined that there was no evidence that these risk factors were effect modifiers, and they did not explain the differences in the efficacy findings between the two trials.[85] Put differently, among pregnant women with a history of sPTB, there was no evidence that those with one or more factors respond to Makena differently than those who did not have these factors. Furthermore, in Trial 003, subjects did not respond to Makena compared to placebo irrespective of the number of risk factors. CDER's interaction[86] and subgroup analyses by individual factor and as a composite variable did not demonstrate that these factors were effect modifiers; these factors did not lead to

[85] CDER Decisional Memorandum, NDA 021945 Makena (hydroxyprogesterone caproate) (Oct. 5, 2020), available at https://www.regulations.gov/document/FDA-2020-N-2029-0089 and attached as Appendix 4.

[86] An example of a statistical interaction of one of these factors would be if a drug's treatment effect was different by race.

differential treatment response to Makena.[87] That is, having multiple prognostic factors may put a woman at higher risk of PTB, but there is no evidence these risk factors make her respond better to Makena than someone at lower risk.

Regarding post-hoc subgroup analyses, to the extent Covis relies on post-hoc subgroup analyses to support its assertion that there is substantial evidence of Makena's effect, either on gestational age at delivery or clinical benefit in the neonate, such reliance is significantly weakened scientifically by the analyses' post-hoc nature. Generally, it is not scientifically appropriate to rely on subgroup analyses to support an inference of a drug's efficacy when the primary analysis result does not demonstrate such efficacy.[88,89] A critical statistical concern with reliance on post-hoc subgroup analyses, among other reasons, is the increased likelihood for Type 1 error (i.e., a false positive, based on finding a difference between drug and placebo when there really is none) that results in the heightened probability of incorrectly concluding treatment benefit. When such subgroup analyses are used to search for evidence of a drug's benefit, there is a higher probability that any observed favorable subgroup results are due to chance alone. Therefore, CDER generally considers such an analysis only for the purpose of generating a hypothesis that can be tested further in a subsequent trial. CDER used post-hoc analysis results solely to explore the potential explanations behind Trial 002's and Trial 003's divergent efficacy results and to generate hypotheses for future trials.

To show Makena's lack of demonstrated efficacy across subgroups, we first present subgroup analyses of Trial 002 data to address Covis' arguments that Black women and women whose qualifying sPTB occurred earlier than 34 weeks respond more favorably to Makena. Then, we present subgroup analyses of Trial 003 data,[90] and we then discuss the lack of trends across Trials 002 and 003.

2. Trial 002 Time-to-Event Analyses Do Not Demonstrate That Pregnancies in Black Women and Women who had a Prior sPTB < 34 Weeks Responded Better to Makena

Time-to-event analyses conducted for Trial 002 do not demonstrate that Makena had a greater effect in Black women (compared to non-Black women) or women who had a prior sPTB < 34 weeks (compared to women who did not have a prior sPTB < 34 weeks).

Covis has asserted that Black women in Trial 002 experienced a benefit from Makena in the reduction of pre-term birth events in an earlier gestational timeframe than non-Black women. Covis has also stated that women who had a prior sPTB < 34 weeks are more likely to

---

[87] *See* Section V.3.

[88] FDA 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.

[89] FDA Draft Guidance for Industry, Multiple Endpoints in Clinical Trials (January 2017) https://www.fda.gov/media/102657/download.

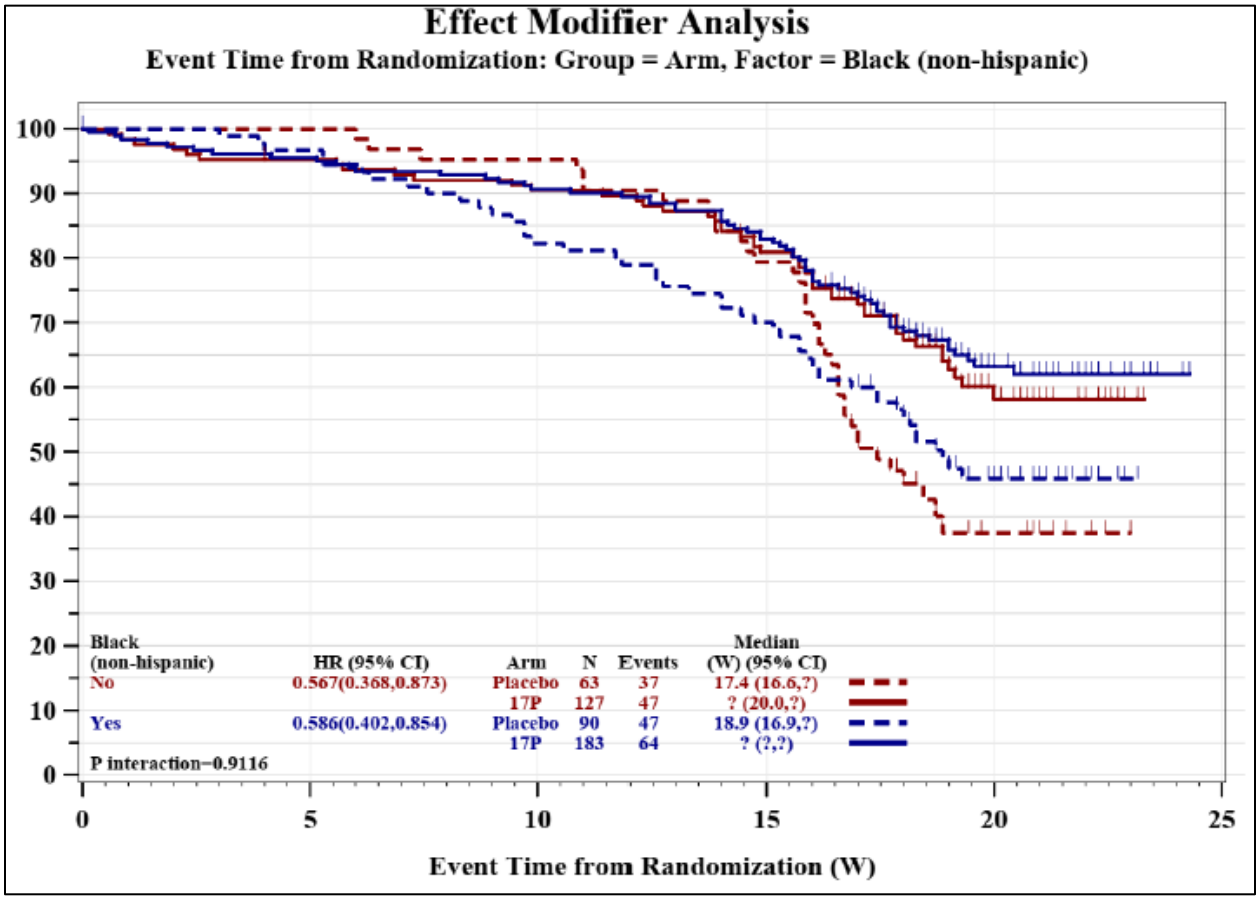
[90] CDER Statistical Review NDA 021945/S-023 available at https://www.regulations.gov/document/FDA-2020-N-2029-0189 and attached as Appendix 3 contains additional Trial 003 subgroup analyses.

experience a future PTB compared to those women who did not have a prior sPTB < 34 weeks. Based on this, Covis has suggested that Makena may have a more beneficial effect in women who had a prior sPTB < 34 weeks. These assertions about Makena's efficacy, though, are not supported by Covis's or CDER's analysis.

First, although pregnant Black women are more likely to have a PTB, the treatment effect of Makena is not shown to be different for Black and non-Black women, as demonstrated by the models that test the interaction of race and treatment effect and other statistical assessments. Careful analysis of the figures created by Covis and submitted to CDER (in July and August 2022), as well as those generated by CDER in prior statistical reviews and below, reveal critical flaws in Covis' assertions based on its time-to-event analysis.

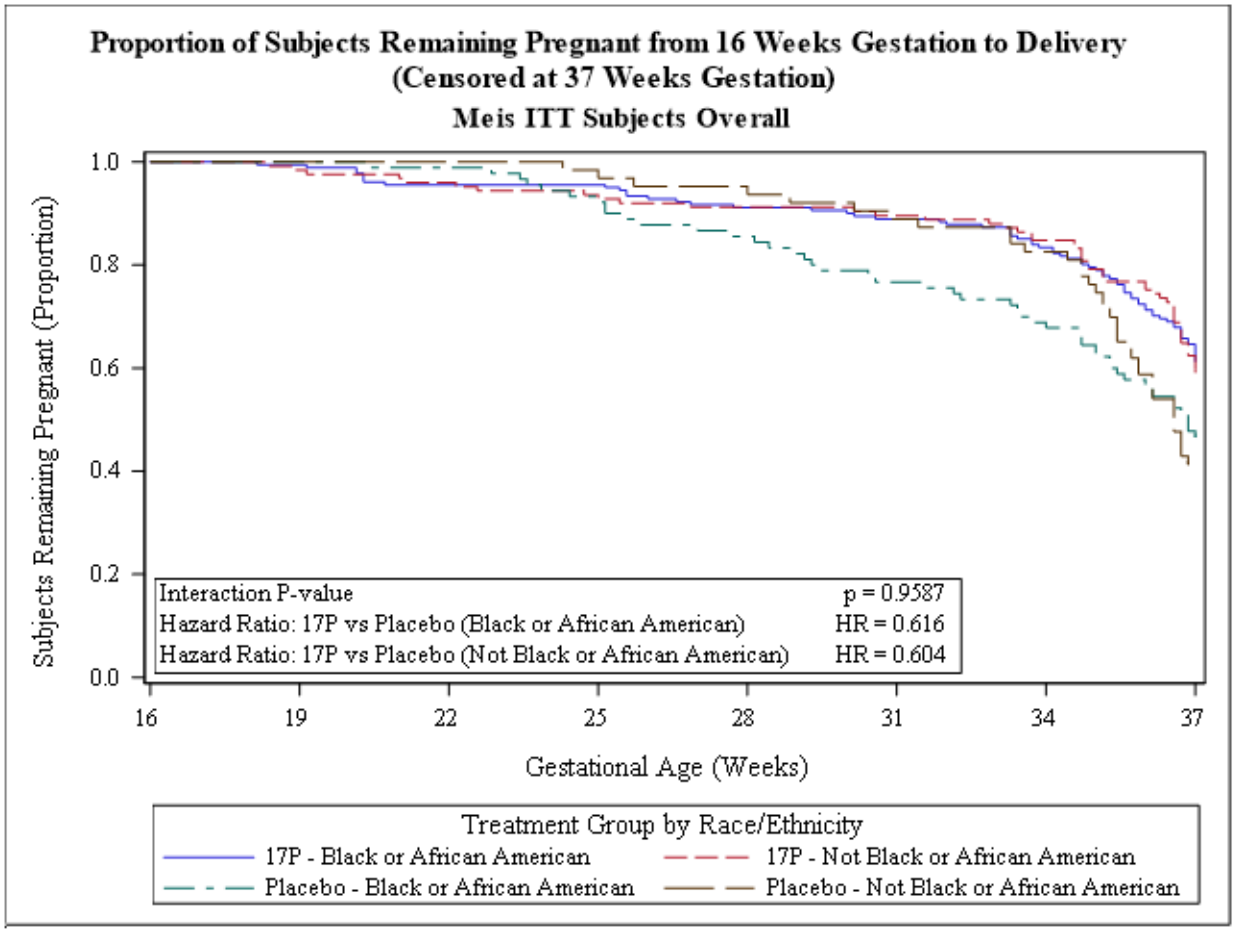
In Trial 002, when time-to-delivery is examined by racial subgroups, the shapes of the curves may visually appear to show differences between the subgroups. The apparent difference is seen when looking at curves without confidence bands, using time from the date of randomization (Figure 2). Figure 3 shows a similar analysis from Covis where time is based on gestational age. Drawing inferences from these visual differences is misleading. Overall, relying only on a visual inspection of the shape of these simple curves, Black subjects appeared to receive a greater benefit from Makena than non-Black subjects in certain time ranges on the curves. However, this relationship is inconsistent, because all women randomized to Makena appear to have a *higher* rate of early losses than did women randomized to placebo; this is true overall and in subgroups defined by race. More fundamentally, visual inspection alone is not a scientifically appropriate method to reach conclusions about whether the simple curves represent true differences between Black and non-Black women in Makena's effect on PTB. When considering the large uncertainty in these findings, as demonstrated by the wide confidence intervals and the notably insignificant p-value of the interaction of the effect with race, there is no evidence of a difference in Makena's effect for Black women from their time-to-event analyses. This large uncertainty is present when assessing time to delivery using gestational age (Figure 4) or time since randomization (Figure 5).

Figure 2: No Interaction: Covis Effect Modifier Analysis of Time since Randomization to Delivery Comparing Black and Non-Black Subjects by Treatment Arm (Trial 002)



Source: Covis (submitted to CDER 17 July 2022)

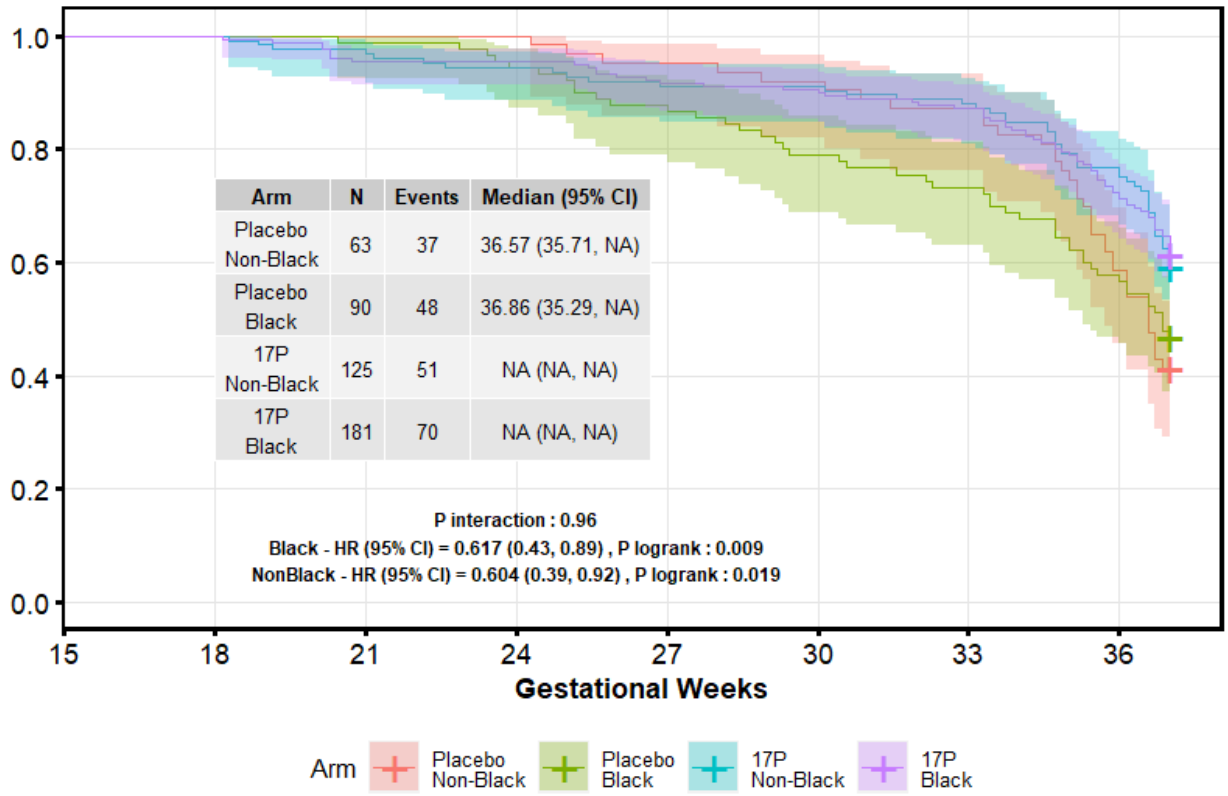
Figure 3: No Interaction: Covis Effect Modifier Analysis of Event Time based on Gestational Age Comparing Black and Non-Black Subjects by Treatment Arm (Trial 002)



Source: Covis (submitted to CDER 12 August 2022)

Figure 4: No Interaction: Covis Effect Modifier Analysis of Event Time based on Gestational Age Comparing Black and Non-Black Subjects by Treatment Arm with CDER Confidence Bands (Trial 002)

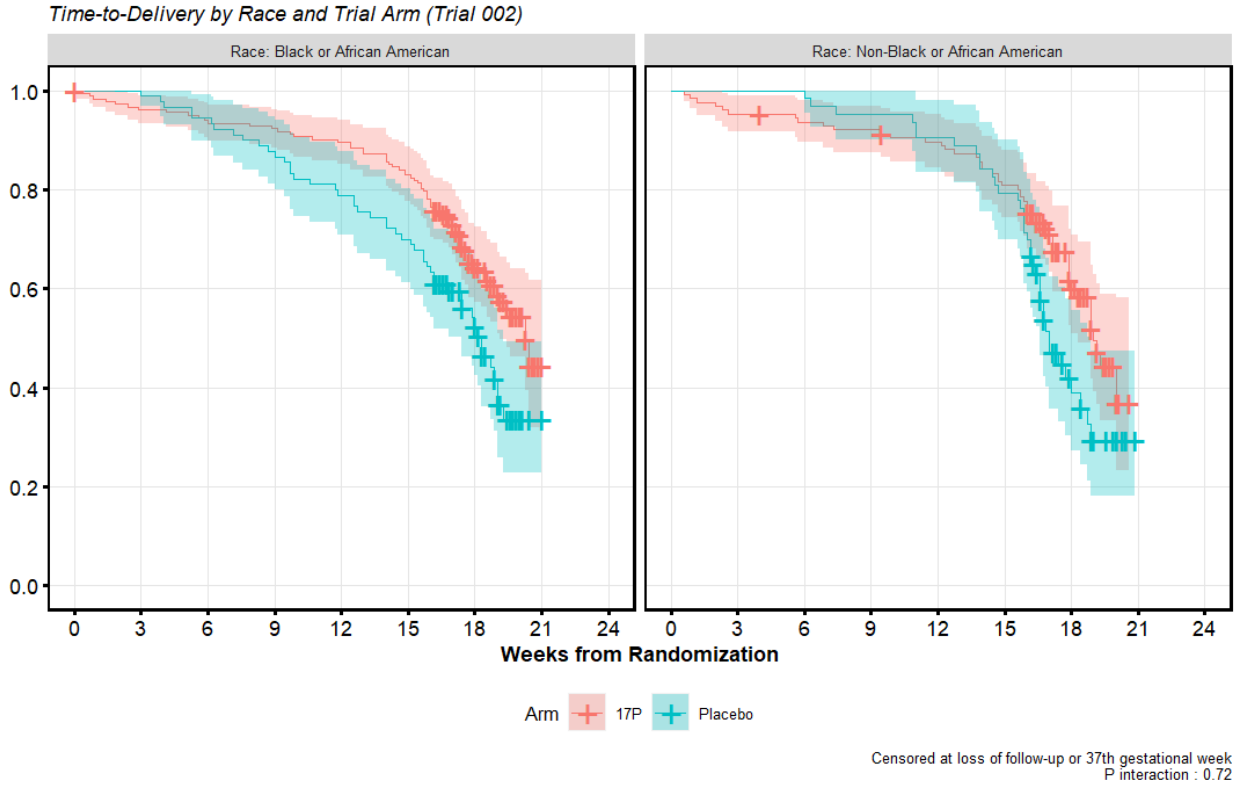
Figure 6



Arm	N	Events	Median (95% CI)
Placebo Non-Black	63	37	36.57 (35.71, NA)
Placebo Black	90	48	36.86 (35.29, NA)
17P Non-Black	125	51	NA (NA, NA)
17P Black	181	70	NA (NA, NA)

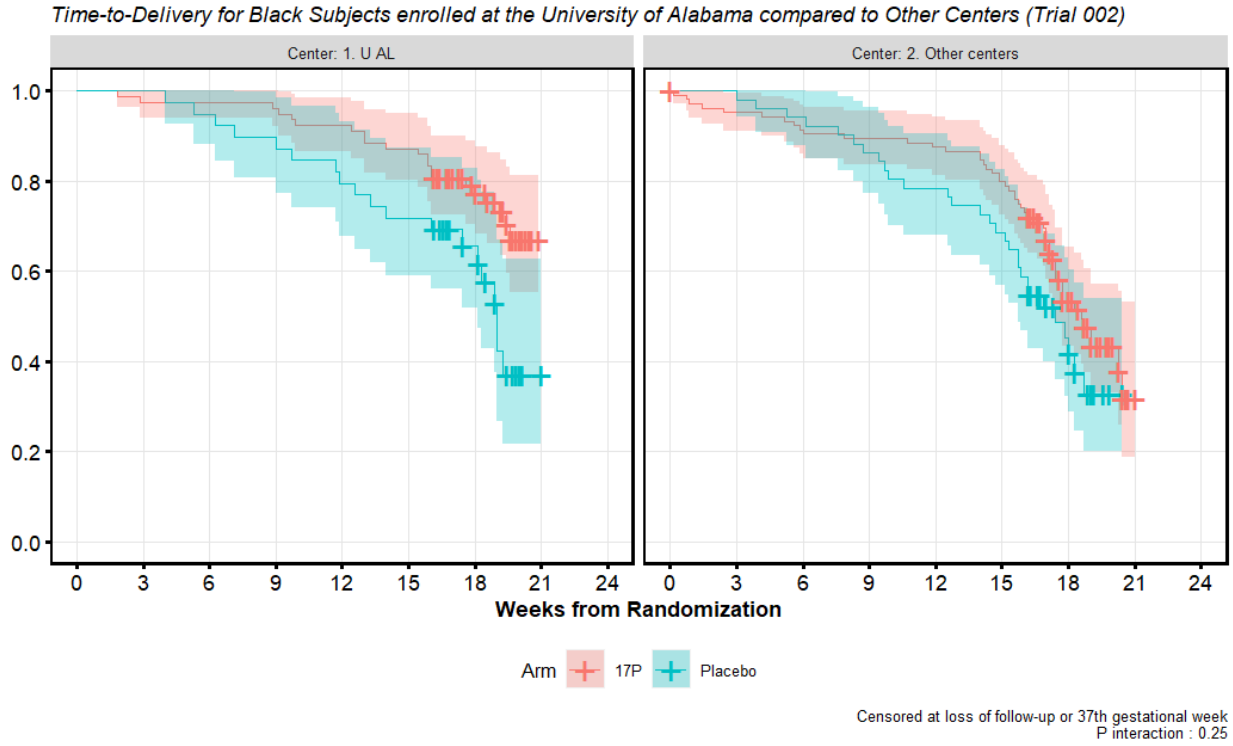
Source: CDER (based on Figure 6 submitted by Covis to CDER on 12 August 2022)

Figure 5: No Interaction: Time-to-Delivery since Randomization by Race and Trial Arm (Trial 002)



In addition, some of the visual differences in the time-to-event curves cannot be reliably attributed to race as they may be influenced by study center, particularly how this factor overlaps with a subject's race and the gestational age at time of enrollment at the study center. First, certain study sites in Trial 002 contributed a greater proportion of subjects by race. While almost 60% of subjects enrolled in Trial 002 were Black, the enrollment of Black subjects was not uniform across study sites. The highest enrolling center, University of Alabama (Center 8), accounted for 27% of all subjects enrolled in the study and 43% of all Black subjects. Almost all subjects (93%) enrolled at the University of Alabama center were Black. Second, in addition to enrolling 43% of all study subjects who were Black, the University of Alabama accounted for 44% (72/164) of all study subjects who enrolled prior to 18 weeks gestational age compared with 18% (54/199) of all study subjects who enrolled after 18 weeks of gestational age. As shown in Figure 6, the shape of the time-to-event curve for Black women at the University of Alabama study site is noticeably different than the curve for Black women at all other sites, although the confidence intervals are wide. Based on these study enrollment patterns, the time-to-event curves observed for Black and non-Black women might be influenced by study site (the University of Alabama), gestational age at enrollment (e.g., enrollment prior to 18 weeks gestational age), or both, as well as other unknown factors.

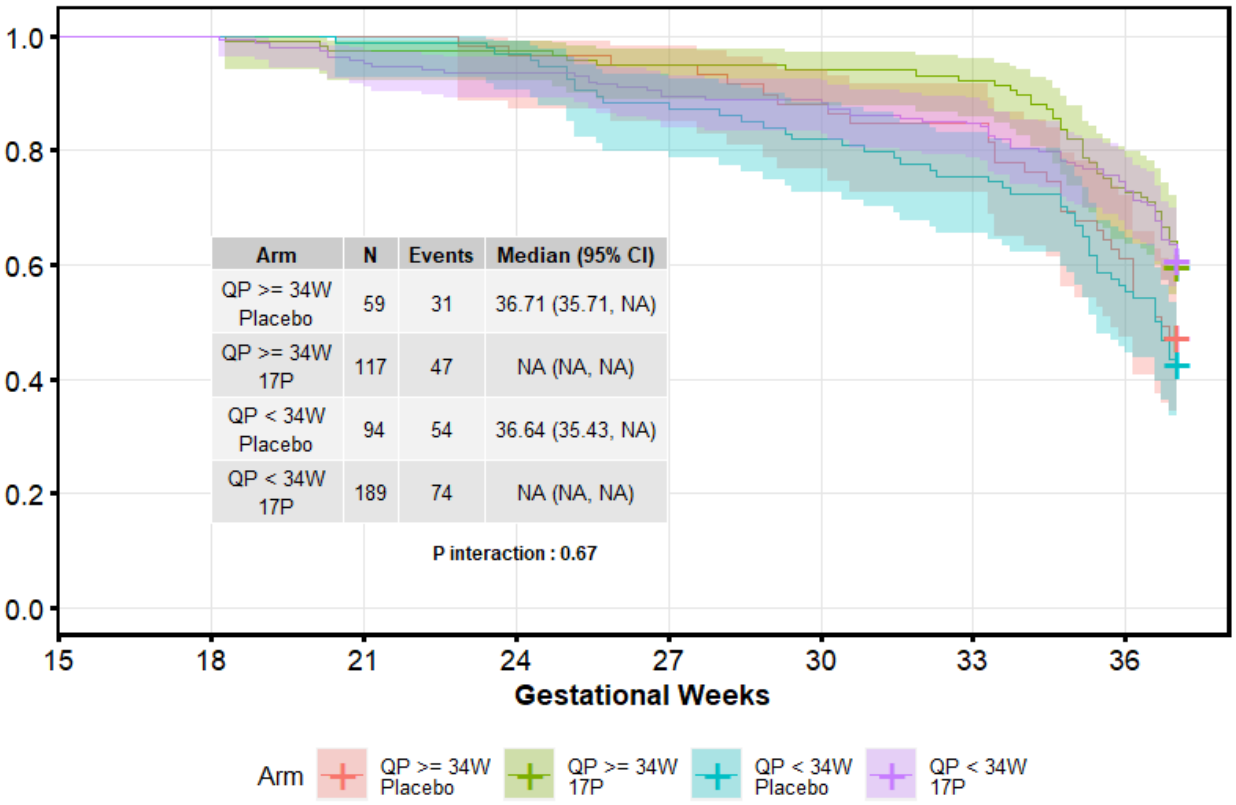
Figure 6: Time-to-Delivery Since Randomization for Black Subjects Enrolled at the University of Alabama Compared to Other Centers (Trial 002)



Second, Covis has asserted that women who experienced a prior sPTB < 34 weeks are at increased risk for a future PTB and suggested that Makena may have a more beneficial effect in these women. Covis has further asserted that women who experience a full-term birth (FTB) after a sPTB are less likely to experience a future PTB. To further understand these assertions in the Trial 002 population, CDER also assessed Makena's effect in subsets of subjects with qualifying sPTBs < 34 weeks and compared them to subjects with qualifying sPTB 34 to < 37 weeks, overall and in the subgroup of subjects who did not have a full-term birth between the qualifying sPTB and trial randomization.

Figure 7: No Interaction: Covis Effect Modifier Analysis by Gestational Age (GA) of Qualifying sPTB. Event Time Based on Gestational Age with CDER Confidence Bands (Trial 002)

Figure 3a with Treatment Arm



Arm	N	Events	Median (95% CI)
QP >= 34W Placebo	59	31	36.71 (35.71, NA)
QP >= 34W 17P	117	47	NA (NA, NA)
QP < 34W Placebo	94	54	36.64 (35.43, NA)
QP < 34W 17P	189	74	NA (NA, NA)

Source: CDER based on Figure 3a submitted by Covis to CDER on 12 August 2022

Based on the results for the interaction term (p=0.67) in the time-to-event model (Figure 7), it does not appear that a qualifying pregnancy delivered prior to 34 weeks gestational age leads to a more favorable drug effect. In addition, for women in this subgroup who did not experience a FTB after a sPTB, CDER's analysis did not find that they received greater benefit from Makena (e.g., interaction term p-values were not statistically significant).

In addition, CDER has not identified *any* subgroup in Trial 002 that can be consistently defined and that has consistently positive results across these analyses, i.e., for statistically significant effect modification, including when compared to results in the same subgroup in Trial 003 (discussed further in Section V.C.4). Further, it is incorrect to say that a more favorable drug effect has been shown for any identified subgroup when the interactions the sponsor tested were not found to be statistically significant.

3. Trial 003 Subgroup Analyses Do Not Support Efficacy of Makena in Any Identified Population

The sponsor and CDER conducted a series of exploratory subgroup analyses to understand the potential reasons for the negative findings in Trial 003. Table 3 shows the subgroup analyses pre-specified in Trial 003's protocol. CDER also conducted post-hoc additional analyses.

Table 3: Trial 003 Pre-Specified 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. Six of the 1708 subjects did not have complete information provided in the Pregnancy Analysis dataset.

These analyses did not identify a subgroup in which Trial 003 provided evidence of efficacy for Makena. Trial 003 did not demonstrate effectiveness of Makena for women in the U.S. as compared to women outside the U.S., women with cervical length less than 25 mm as compared to longer, non-Hispanic women of Black race as compared to non-Black race, or for any of the other subgroups pre-specified for Trial 003. Of note, Covis has proposed several additional potential definitions for subgroup populations and has used many more general terms such as "high-risk." General phrases are not useful because they cannot be implemented as part of inclusion and exclusion criteria for a clinical trial or used at the point of care to make a treatment decision with someone who is currently pregnant. For CDER to consider a potential subpopulation, it needs to be defined by clear variables available at baseline in a trial and variables that prescribers can evaluate prior to advising the use of Makena to an individual patient, and CDER would need to see efficacy in the subpopulation in prospectively planned analyses across multiple independent data sets. Post-hoc exploratory subgroup analyses may be biased and are therefore, for hypothesis-generating purposes.[91]

[91] FDA Guidance for Industry, E9 Statistical Principles for Clinical Trials (September 1998), available at https://www.fda.gov/media/71336/download.

In addition to the subgroup analyses proposed by the sponsor in the protocol and in 2019, CDER conducted additional exploratory analyses of Trial 003 using logistic regression models for each co-primary efficacy endpoint with treatment, region, each of the factors in Table 3, and their interaction with treatment. CDER also created post-hoc subgroups that were composed of multiple (not just single) known risk factors, especially the factors that differed between Trials 002 and 003, to evaluate whether a "higher" risk subgroup might benefit from Makena. These analyses do not provide evidence of efficacy of Makena over placebo in any subpopulation in Trial 003, and there was no statistically significant interaction between Makena and any of these factors. In other words, these analyses did not demonstrate any difference in Makena's effect between subgroups with these factors. Analogous analyses in the Trial 003-U.S. subgroup produced results similar to those for the overall Trial 003 population.

Subgroup analysis using the Cochran-Mantel-Haenszel (CMH) method evaluates a particular subgroup category independently from other subgroup categories and relies only on the data from the subjects in that particular category. Bayesian shrinkage estimation (SHR) analysis evaluates all subgroup categories jointly and borrows information across subgroups to reduce the variability of the estimates and prevent random highs and random lows. CDER conducted subgroup analyses using both the CMH and SHR methods. Neither method revealed evidence of an effect within a specific subgroup in Trial 003, nor was there evidence of differences in effects across subgroups. Below, we provide details of subgroup analyses based on potential risk factors. Additional subgroup analyses can be found in CDER's 2019 Advisory Committee Briefing Document[92] and the Statistical Review.[93]

i. Trial 003 Analyses Do Not Show a Differential Treatment Response Among Subgroups

1. REGION (U.S. compared to non-U.S.) Subgroup: The region subgroup analyses for Trial 003 were important because the sponsor suggested the differences in efficacy findings between Trial 002 and Trial 003 may be attributable to the study population being U.S.-women-only in Trial 002, compared to a population in Trial 003 that included non-U.S. women. CDER therefore focused by region (U.S., non-U.S. subgroups) its analyses of Trial 003's co-primary efficacy endpoints and two secondary endpoints of interest, one of which was the surrogate endpoint for Trial 002 (gestational age of delivery < 37 weeks), using stratified CMH and SHR analyses.[94,95]

---

[92] FDA Backgrounder, BRUDAC Meeting (Oct. 29, 2019) available at https://www.regulations.gov/document/FDA-2020-N-2029-0108 and attached as Appendix 2.

[93] *See* CDER Statistical Review NDA 021945/S-023 available at https://www.regulations.gov/document/FDA-2020-N-2029-0189 and attached as Appendix 3; *see also* CDER Clinical Review, NDA 021945/S-023, available at https://www.regulations.gov/document/FDA-2020-N-2029-0185 and attached as Appendix 9.

[94] FDA Slides, 14–20, Figures 1–7, BRUDAC Meeting (Oct. 29, 2019) available at https://www.regulations.gov/document/FDA-2020-N-2029-0101 and attached as Appendix 2.

[95] Shrinkage analyses were re-run increasing the number of iterations from 60,000 to 6,000,000. See Appendix 6 for Figures 1–6 that include these results. Results were similar.

Figure 8: Region Subgroup: No Evidence of Treatment Effect on the Neonatal Composite Index or the Proportion of Trial 003 Subjects Delivering < 35 Weeks Gestational Age

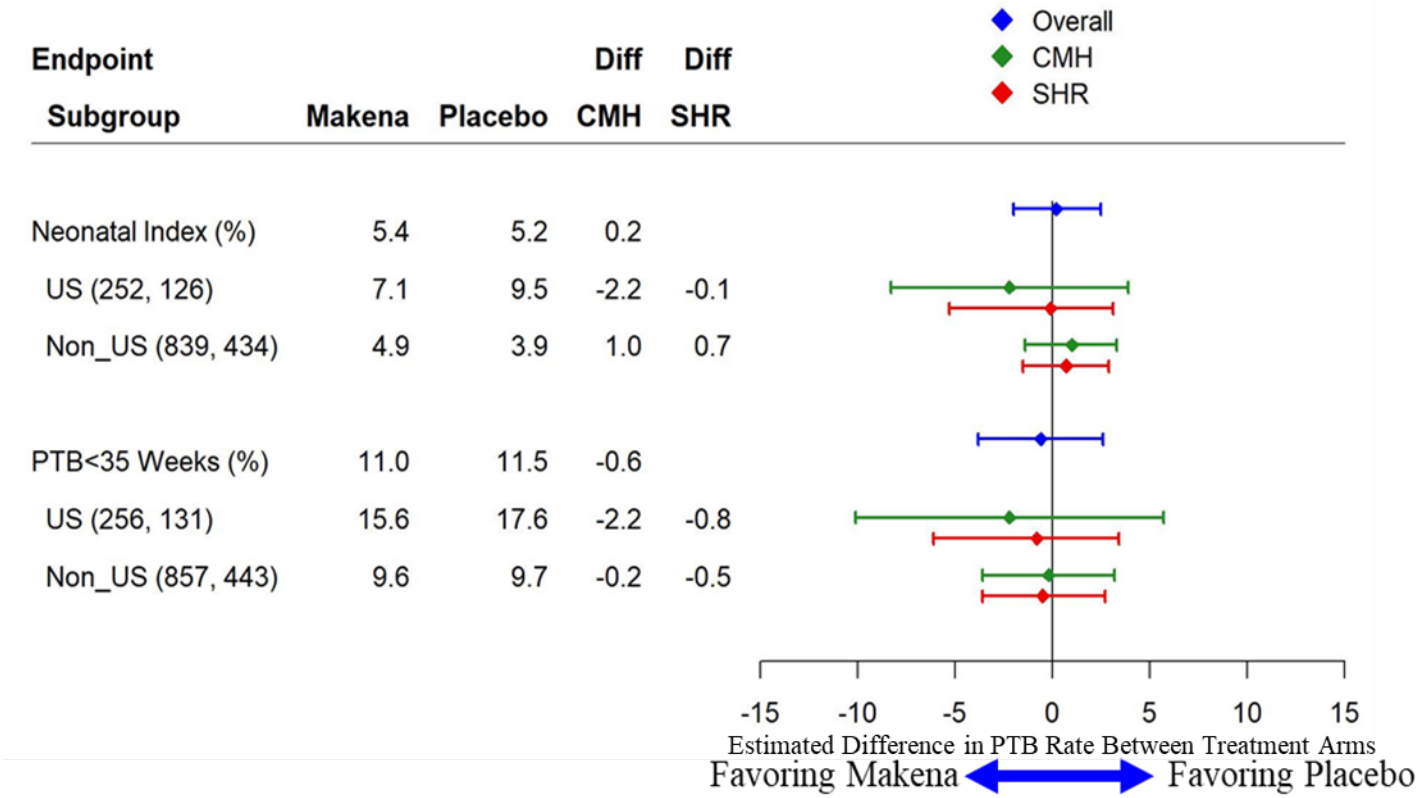
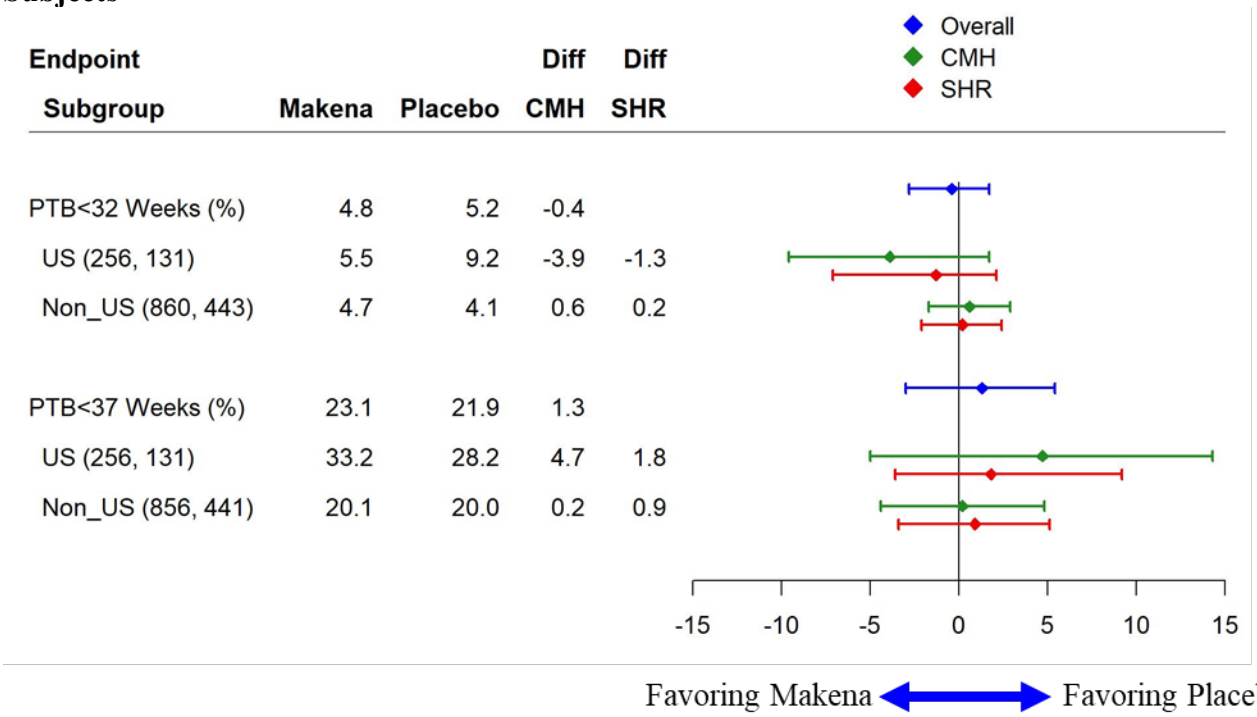


Figure 9: Region Subgroup: No Evidence of Treatment Effect on the Proportion of Trial 003 Subjects Delivering at < 37 and < 32 Weeks Gestational Age in Either US or non-US Subjects



The confidence intervals for treatment difference for these efficacy endpoints in both the overall Trial 003 population and in the regional subgroups of U.S. and non-U.S. include zero, indicating no evidence of Makena's superiority over placebo based on either analysis method. CDER's analysis found no positive trend favorable to Makena in the Trial 003 U.S. subgroup. The point estimates across the four endpoints (neonatal composite index, PTB < 32, < 35, < 37 weeks) are not consistently favorable for Makena. PTB < 37 weeks is unfavorable for Makena, which is notable since PTB < 37 weeks was the endpoint that supported the 2011 approval and the endpoint upon which Makena had the largest treatment effect in Trial 002. In addition, there was no evidence of a treatment effect by region interaction for each co-primary endpoint in Trial 003; that is, there was no evidence that being from the U.S. or outside the U.S. made a difference to a woman's response to Makena.

2. RACE (Black compared to non-Black) Subgroup: The sponsor's assertion that race played a role in the differences in the efficacy outcomes between Trial 002 (59% (n=273) Black subjects) and Trial 003 (7% (n=114) Black subjects)[96] is not supported by the subgroup analyses for either trial, as discussed below. The treatment differences for the co-primary endpoints and secondary endpoints of interest for Trial 003 overall and within subgroups by race (Black, non-Black) are close to zero, with all confidence intervals including zero.[97] This race subgroup analysis did not

[96] Of note, in the U.S. subgroup of Trial 003, 29% of the subjects were Black.

[97] *See* Figures 3 and 4.

provide evidence that Makena had a treatment effect in Black or in non-Black women. And in Trial 002, Makena's treatment effect was similar between Black and non-Black subjects. Therefore, race subgroup analyses in Trial 002 and Trial 003 indicate there is no evidence of a differential treatment effect of Makena based on race. Put differently, there is no evidence Black race is associated with a better treatment response to Makena.

Figure 10: Race Subgroup: No Evidence of Treatment Effect on the Neonatal Composite Index or the Proportion of Trial 003 Subjects Delivering < 35 Weeks Gestational Age in Either Black or non-Black Subjects

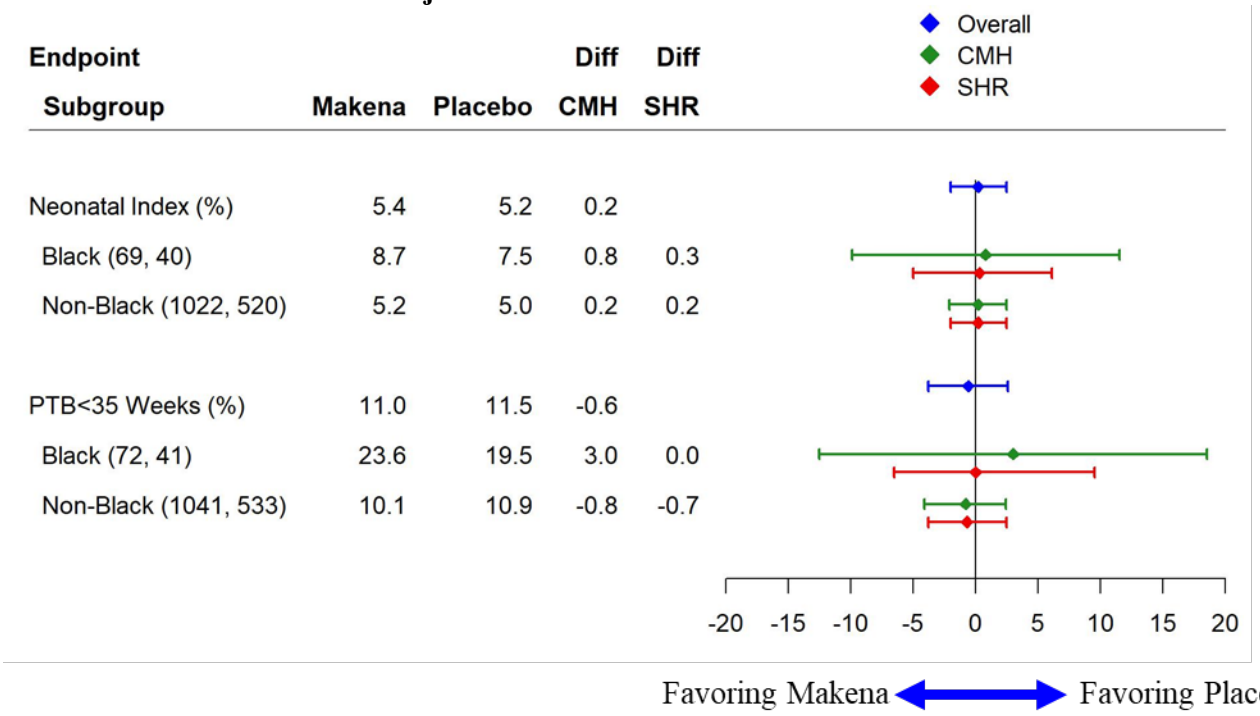
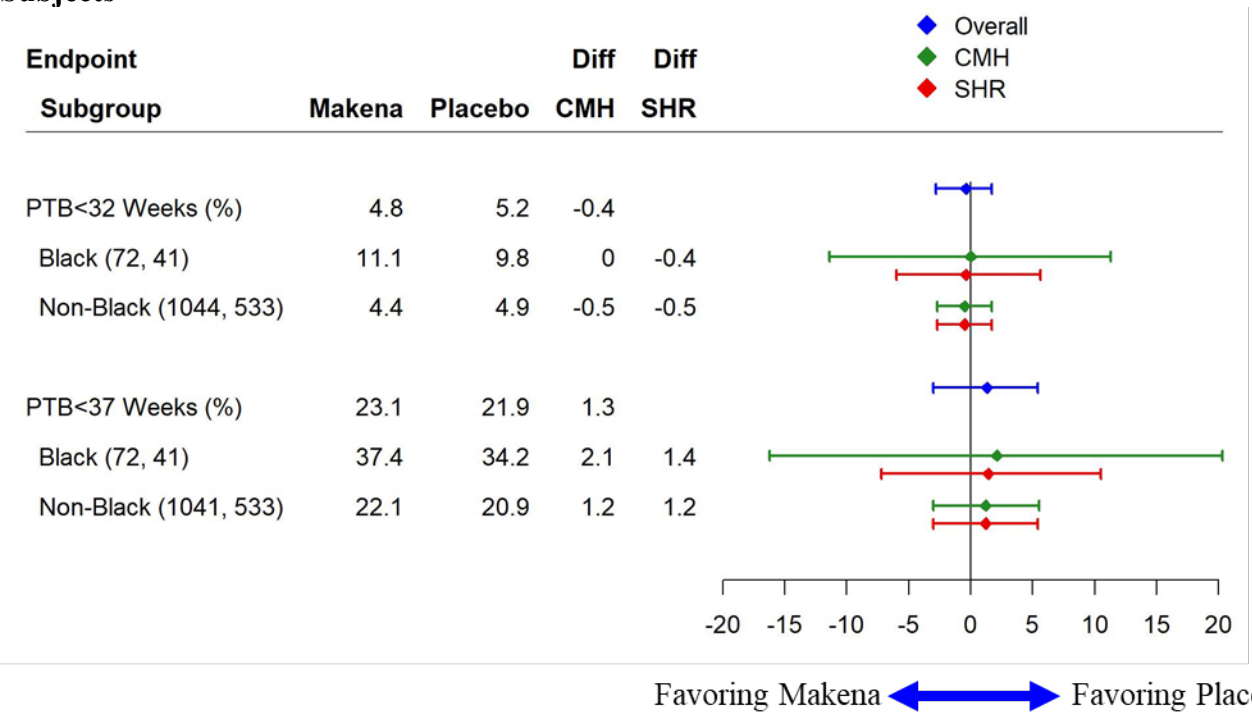


Figure 11: Race Subgroup: No Evidence of Treatment Effect on the Proportion of Trial 003 Subjects Delivering at < 37 and < 32 Weeks Gestational Age in Either Black or non-Black Subjects



3. NUMBER OF PRIOR sPTB (1 compared to >1) Subgroup: CDER also explored whether there was evidence of a differential treatment effect among women with different numbers of prior spontaneous PTBs. The risk of recurrent sPTB increases with increasing number of prior sPTB. This subgroup analysis in Trial 003 did not provide evidence of a treatment effect within either of the subgroups, nor did it provide evidence that the number of prior sPTBs modifies the treatment effect, for either the co-primary endpoints (Figure 12) or the secondary endpoints (Figure 13). In Trial 002, the treatment effect on gestational age of delivery < 37 weeks was independent of the number of prior sPTB. Therefore, the subgroup analyses in Trial 002 and Trial 003 indicate there is no evidence of a differential treatment effect of Makena based on having one or more than one prior sPTB.

Figure 12: Number of Prior Singleton sPTBs Subgroup: No Evidence of Treatment Effect on the Neonatal Composite Index or the Proportion of Trial 003 Subjects Delivering < 35 Weeks Gestational in Subjects With 1 or >1 Prior sPTBs

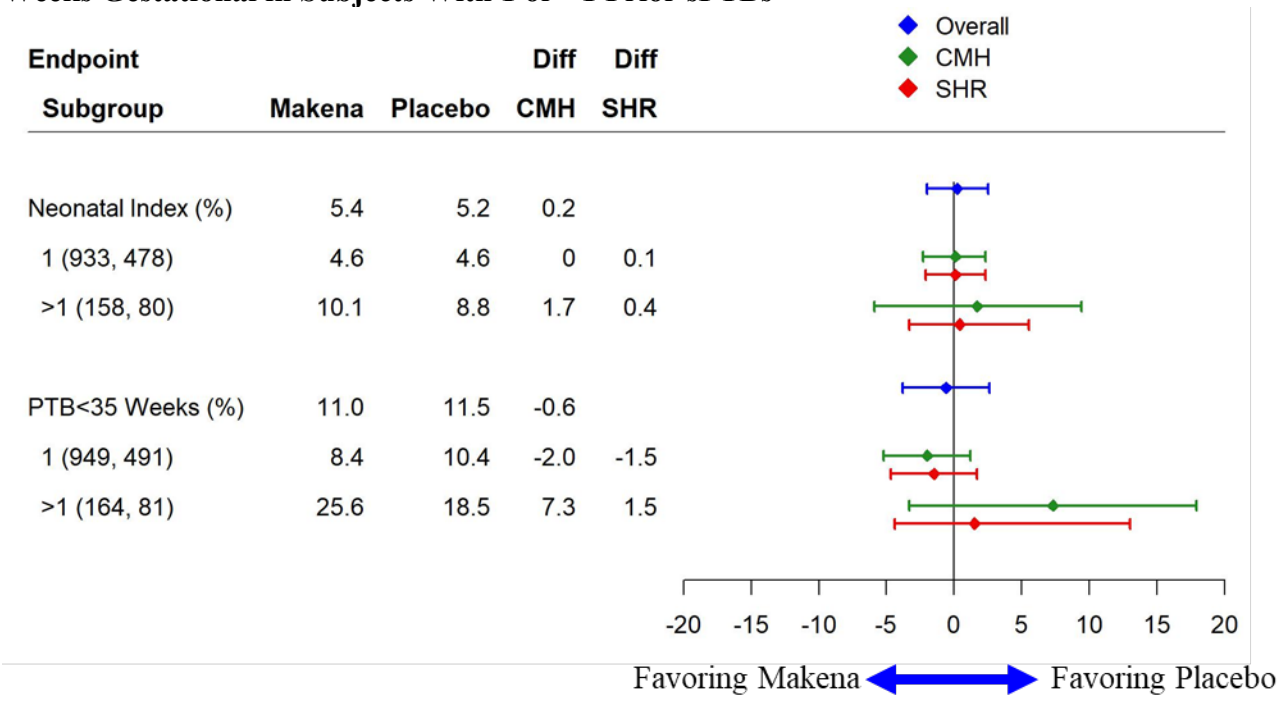
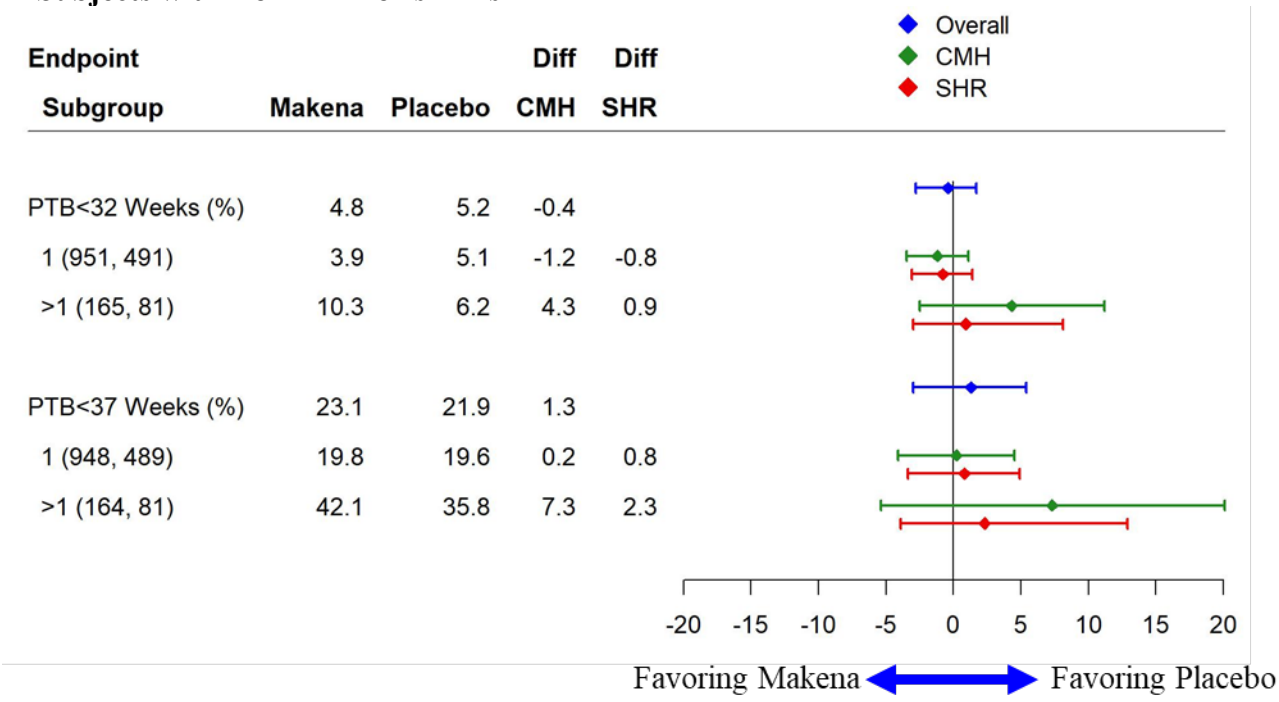


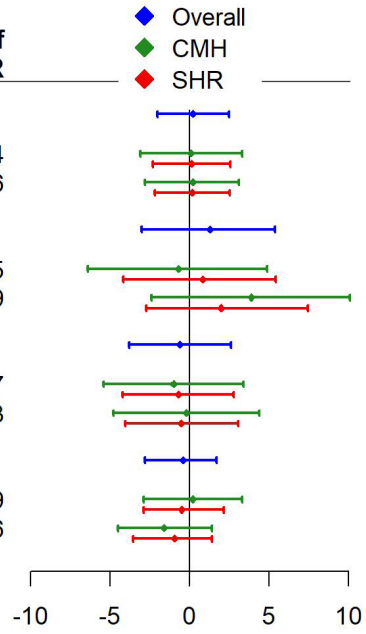
Figure 13: Number of Prior Singleton sPTBs Subgroup: No Evidence of Treatment Effect on the Proportion of Trial 003 Subjects Delivering at < 37 and < 32 Weeks Gestational Age in Subjects with 1 or >1 Prior sPTBs



4. GESTATIONAL AGE of Prior sPTB Subgroup: The sponsor hypothesized that women who had an earlier sPTB, in particular < 34 weeks, fell under the classification of "high risk." This subgroup analysis in Trial 003 did not provide evidence of a treatment effect within either of the subgroups, nor did it provide evidence that one or more prior sPTBs earlier than 34 weeks modifies the treatment effect, for either the co-primary endpoints or the secondary endpoints (Figure 14). There are several ways to consider defining this population. For example, a subject in Trial 003 may have had the qualifying sPTB for Trial 003 at 35 weeks, and she also may have had an earlier sPTB < 34 weeks. The results of CDER's analyses classifying subjects by only their trial qualifying sPTB and also by any prior sPTB < 34 weeks were similar. The subgroup analyses in Trial 002 and Trial 003 indicate there is no evidence of a differential treatment effect of Makena based on having a prior sPTB < 34 weeks.

Figure 14: Any Prior sPTB < 34 Weeks Compared to No Prior sPTB < 34 Weeks: No Evidence of Treatment Effect on Neonatal Composite Index or the Proportion of Trial 003 Subjects Delivering at < 37, < 35, and < 32 Weeks Gestational Age

Endpoint Subgroup	Makena (%)	Placebo (%)	Diff CMH	Diff SHR
NCI[1] (interaction p-value : 0.86)	5.4	5.2	0.2	
Any sPTB <34w	6.5	6.4	0.1	0.14
No sPTB <34w	3.6	3.4	0.2	0.16
<37w (interaction p-value : 0.30)	23	22	1.3	
Any sPTB <34w	23.9	24.6	-0.7	0.85
No sPTB <34w	21.9	18	3.9	1.99
<35w (interaction p-value : 0.57)	11	12	-0.6	
Any sPTB <34w	12.1	13.1	-1	-0.7
No sPTB <34w	9	9.2	-0.2	-0.53
<32w (interaction p-value : 0.50)	5	5	-0.4	
Any sPTB <34w	6.2	6	0.2	-0.49
No sPTB <34w	2.6	4.2	-1.6	-0.96



[1] NCI: Neonatal Composite Index

The sponsor also hypothesized that intervening FTBs reduced the risk of PTB among women with a prior sPTB < 34 weeks in Trial 003, even though Trial 002 had a higher rate of FTBs between the qualifying sPTB and trial randomization than Trial 003. CDER found no trends of effectiveness in the population with the qualifying sPTB < 34 weeks and no FTB since the qualifying sPTB. Of note, in this set of analyses, CDER analyzed the subset of women whose qualifying sPTB was < 34 weeks, not the subset of women with any prior sPTB < 34 weeks.

Table 4: Qualifying sPTB < 34 Weeks With no FTB Since the Qualifying sPTB: No Evidence of Treatment Effect on Neonatal Composite Index, or the Proportion of Trial 003 Subjects Delivering at < 37, < 35, and < 32 Weeks Gestational Age

Efficacy Outcome	HPC (Makena) (N=597)	Placebo (N=300)	Treatment Difference[§] (95% CI)	Relative Risk[§] (95% CI)
Neonatal composite index*	6.4%	6.5%	-0.1% (-3.6, 3.3)	0.98 (0.57, 1.67)
Birth < 37 weeks	22.9%	24.0%	-1.0% (-6.9, 4.9)	0.96 (0.75, 1.23)
Birth < 35 weeks	11.6%	12.4%	-0.9% (-5.5, 3.6)	0.92 (0.64, 1.34)
Birth < 32 weeks	6.0%	6.4%	-0.4% (-3.8, 2.9)	0.93 (0.54, 1.60)

*Proportion of neonates experiencing at least one event of the composite index
Subset of the Trial 003 subjects who have had no FTB since the qualifying sPTB [regardless of time] and qualifying flag sPTB < 34.
[§]Cochran-Mantel-Haenszel (CMH) method stratified by gestational age at randomization

Table 5: Qualifying sPTB 34 to < 37 Weeks With no FTB Since the Qualifying sPTB: No Evidence of Treatment Effect on Neonatal Composite Index, or the Proportion of Trial 003 Subjects Delivering at < 37, < 35, and < 32 Weeks Gestational Age

Efficacy Outcome	HPC (Makena) (N=392)	Placebo (N=226)	Treatment Difference[§] (95% CI)	Relative Risk[§] (95% CI)
Neonatal composite index*	4.5%	3.2%	1.3% (-1.8, 4.5)	1.42 (0.60, 3.34)
Birth < 37 weeks	25.3%	21.0%	4.3% (-2.6, 11.2)	1.20 (0.89, 1.64)
Birth < 35 weeks	11.4%	11.2%	0.1% (-5.1, 5.3)	1.01 (0.64, 1.61)
Birth < 32 weeks	4.1%	4.5%	-0.4% (-3.7, 3.0)	0.92 (0.42, 1.99)

*Proportion of neonates experiencing at least one event of the composite index
Subset of the Trial 003 subjects who have had no FTB since the qualifying sPTB [regardless of time] and qualifying flag sPTB 34 W to < 37 W.
[§]Cochran-Mantel-Haenszel (CMH) method stratified by gestational age at randomization

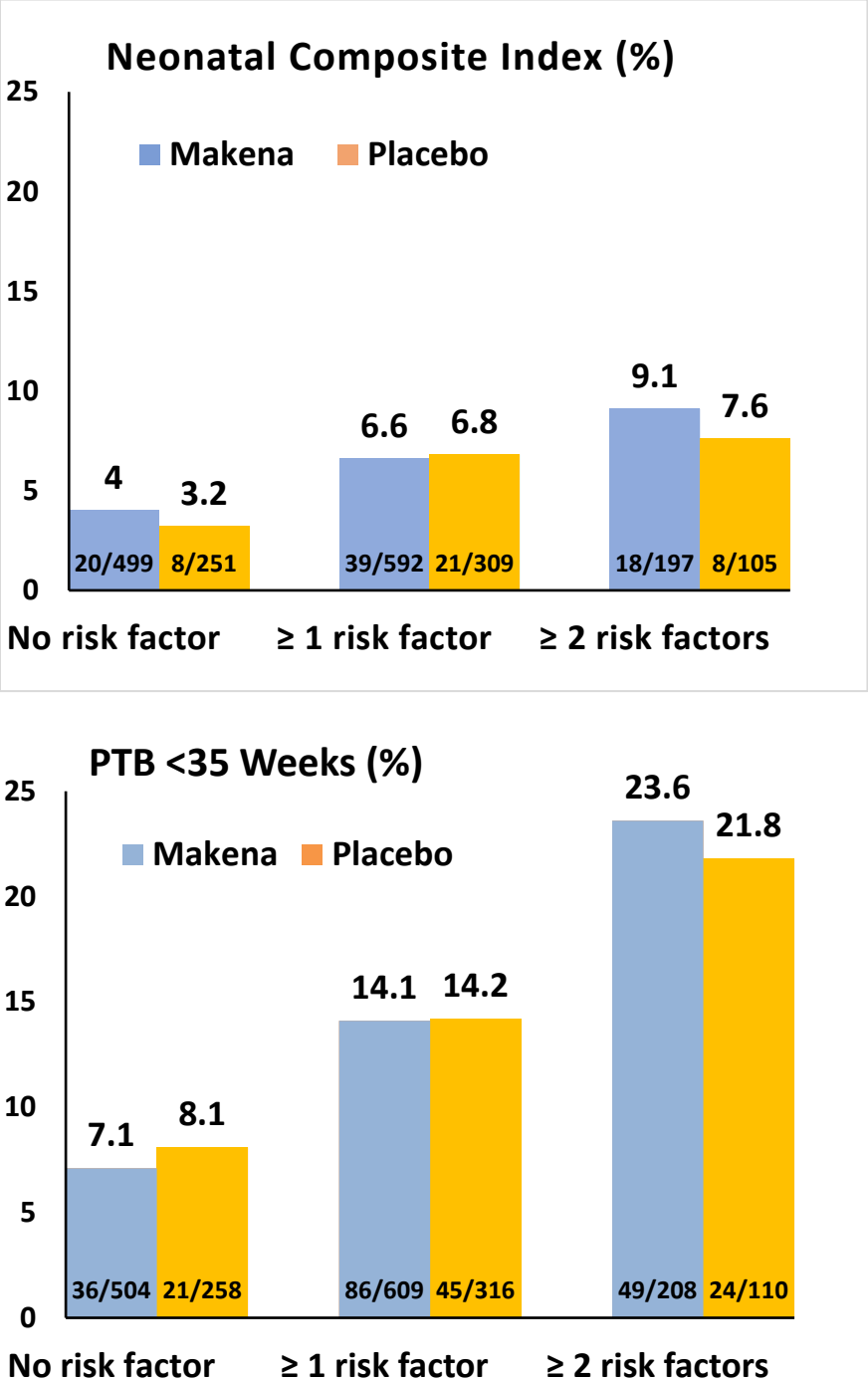
CDER also conducted subgroup analyses of subjects with a qualifying sPTB < 34 weeks within the last 5 years, subjects with any sPTB < 34 weeks in the last 5 years, and each of these subgroups excluding those subjects with any FTB at any time in their history. None of these analyses identified a subgroup for which Trial 003 demonstrated a trend of effectiveness for Makena.

5. RISK LEVEL of PTB Subgroup: The sponsor hypothesized that Makena may be effective in a "higher" risk patient population seen in Trial 002 and not in what the sponsor claims is the "lower" risk population seen in Trial 003. They identified five risk factors for PTB (history of > 1 prior sPTB, Black race, substance use in current pregnancy, ≤ 12 years of education, unmarried with no partner) that were more prevalent in the study population of Trial 002, which therefore may represent a higher risk population, compared to that of Trial 003. CDER undertook subgroup analyses for the efficacy endpoints, defining subjects as having none, one, or at least two of these five risk factors identified by the sponsor. There was no suggestion of efficacy even

in the groups with "higher" risk levels.[98] Although these five risk factors may impact the rate of PTB or neonatal composite index, there is no evidence based on subgroup analysis in Trial 003 that these factors have an impact on the treatment effect of Makena. Furthermore, compared to placebo, Makena did not have an effect for women at lower or higher risk of recurrent PTB.

[98] In fact, among participants with two or more risk factors, there was a small numerical imbalance not favoring Makena.

Figure 15: "Composite" Risk Level Subgroup: No Evidence of Treatment Effect in any Risk Group Defined Using Five Risk Factors Selected by the Sponsor (Trial 003)



Risk factors: Black race, history of more than one sPTB, single/without partner, substance use in pregnancy, ≤12 years education

6: FULL TERM BIRTHS Subgroup: The sponsor also claims that more subjects in Trial 003, compared to Trial 002, had full-term pregnancies between their qualifying sPTB and the trial,

and that could lessen the risk of future PTB in these trial subjects. In Trial 002, 101 (out of 463) or 21.8% of subjects had a qualifying sPTBs followed by full-term births. Although the number of subjects with intervening full-term pregnancies is higher in Trial 003 (N=198), the proportion of subjects (198 out of 1708 or 11.6%) who had an intervening FTB was lower in Trial 003 than Trial 002 (Table 6). Analyses do not demonstrate that intervening FTBs explain the discrepant efficacy findings between Trials 002 and 003.

Table 6: Trial 002 Higher Proportion of Subjects Compared to Trial 003 of Early Qualifying* sPTB Followed by Full-term Births for Both One Full-term Birth and Two or More Full-term Births

Qualifying preterm births followed by full-term births	Trial 002			Trial 003		
	n	n/101	n/463	n	n/198	n/1708
One full-term birth	75	74%	16.2%	155	78%	9.1%
Two or more full-term births	26	26%	5.6%	43	22%	2.5%
Total	101	100%	21.8%	198	100%	11.6%

*Qualifying sPTB was determined for the analysis in the following manner: for Trial 002 the latest sPTB was considered the qualifying sPTB (there was no qualifying flag in the dataset); for Trial 003 the qualifying flag in the dataset was used.

ii. There Are No Trends Pointing to Efficacy in Trial 003

In response to CDER's proposal to withdraw Makena, the sponsor asserted that "there are trends in the [Trial 003] data that point to efficacy in patient populations at higher risk of preterm birth."[99] To support this argument, the sponsor presents a post-hoc analysis it asserts is "to review efficacy in the higher-risk population subgroup." For this analysis, the sponsor selects U.S. clinical sites with 10 or more enrolled women and a placebo group PTB rate of ≥20%.[100] The sponsor claims that, among this "higher-risk" subgroup, the relative risk for PTB < 35 weeks was 0.52 (95% CI 0.29, 0.92).[101] It is not clear which gestational age of delivery this rate is defined at, how and why 20% was selected as "high risk" as opposed to other possible rates (e.g., 30%, 50%), and why a cut-off of 10 enrolled women was used. Such post-hoc exploratory subgroup analyses may be biased, in danger of being over-interpreted, and are difficult to interpret.[102]

Additionally, the sponsor claims that data from the U.S. subgroup of Trial 003 "show a trend toward reduction of preterm birth for U.S. [Trial 003] patients" based on differential rates of

[99] Submission Of AMAG Pharmaceuticals, Inc. in Response to the Food And Drug Administration's Notice of Opportunity for a Hearing and Proposal to Withdraw Approval Of MAKENA® (hydroxyprogesterone caproate injection) (Sponsor's Response to NOOH) at 17, available at https://www.regulations.gov/comment/FDA-2020-N-2029-0051.

[100] Sponsor's Response to NOOH at 26, available at https://www.regulations.gov/comment/FDA-2020-N-2029-0051.

[101] *Id.*

[102] FDA Guidance for Industry, E9 Statistical Principles for Clinical Trials (September 1998), available at https://www.fda.gov/media/71336/download.

PTB < 32 weeks between the control and treatment arms of U.S. women in Trial 003.[103] Although the CMH-derived point estimates of a treatment difference in the U.S. subgroup of Trial 003 slightly trends toward Makena based on the pre-term birth < 32 weeks endpoint, the confidence interval for the effect includes zero. Furthermore, when considering trends, it is important to consider the trend in all the endpoints in the U.S. subgroup, including that the trend for PTB < 37 weeks shows Makena *increasing* rates of PTB for U.S. subjects in Trial 003.[104] When looking at these highly related endpoints (PTB < 32 weeks and < 37 weeks), consistency of treatment effect across the endpoints is important. The data from Trial 003, even when subjected to multiple post-hoc analyses, do not produce such trends. As detailed above, across the many subgroup analyses of Trial 003 conducted by both CDER and the sponsor, there is no trend in favor of Makena.

4. There is No Demonstrated Efficacy in Any Identified Subgroup Across Trials 002 and 003

CDER's interaction and subgroup analyses of Trial 003 (i.e., analyses based on region, race, number of prior sPTB, gestational age of prior sPTB, composite risk based on the sponsor's identified five factors, and prior FTBs) did not show a differential response to Makena, compared to placebo, between groups with or without these factors or between groups with lower and higher baseline recurrent PTB risk. If Makena were truly effective in a "high-risk" population, especially at the magnitude and consistency seen in the overall Trial 002 results, one would expect to see a treatment effect in at least some subgroups of Trial 003. But, as shown above and in CDER's other documents, no such signal emerges from the data. When looking across the Trial 003 high-risk subgroups using multiple different analysis methods, CDER did not identify any statistically significant results, and the point estimates were consistently near zero or favoring placebo. Both in Trial 003 and in Trial 002, there was no evidence that the drug had any better or any worse effect dependent on the subject's baseline risk. Further, neither CDER nor the sponsor has found a well-defined consistent subgroup across either study where there was a beneficial effect.

D. There is No Basis to Discount Trial 003's Results or Elevate Trial 002's Results

Covis argues that Trial 003's purported "flaws" are such that its negative results should either be discounted or ignored entirely, and that, as a result, there is *still* substantial evidence of Makena's effectiveness based on Trial 002 notwithstanding the failure of the much larger Trial 003 to show any favorable drug effect, either in the overall indicated population or any identified subset of it.

The sponsor asserts that the Trial 003 population had different socioeconomic status factors and other risk factors linked to lower rates of PTB (e.g., substance use, education level, race) than the

[103] Sponsor's Response to NOOH at 26–27.

[104] FDA Backgrounder, Table 8 at 33, BRUDAC Meeting (Oct. 29, 2019), available at https://www.regulations.gov/document/FDA-2020-N-2029-0108 and attached as Appendix 2.

Trial 002 population, particularly among the subjects outside the U.S. enrolled in Trial 003.[105] Specifically, Covis argues that, in part due to unintentional selection bias, Trial 003 enrolled a low-risk population, resulting in a low rate of PTB in Trial 003's participants.[106] But Trial 003's population was not particularly "low risk," and in any event, differences in the characteristics of women in Trials 002 and 003 did not drive the different outcomes of those trials.

The sponsor also asserts that Trial 003's "observed power was inadequate to detect event rate differences in premature births and therefore also neonatal outcomes."[107] In fact, Trial 003 was adequately powered to detect the expected efficacy based on the results of Trial 002 and was reasonably powered to detect substantially lower-than-expected efficacy; it simply did not do so.

Next, Covis argues that the assessment of gestational age of prior qualifying pregnancies biased Trial 003 outcomes. Also, Covis has previously proposed combining the findings from Trial 002 and the U.S.-only subgroup of Trial 003 to demonstrate Makena's efficacy. Finally, Covis argues that Trial 002 could have supported traditional approval. These arguments are also unavailing, for the reasons described below.

1. Trial 002 Does Not Better Represent Makena's Indicated Population Than Trial 003

Covis asserts Trial 003 does not apply to Makena's approved population in the U.S. because Trial 003 enrolled a "low" risk population with a "lower" risk profile. Covis claims, therefore, that Trial 002 better represents Makena's target population. CDER disagrees on all points.

Regarding risk factors for PTB, CDER agrees with Covis that the study populations in Trial 002 and 003 differed. These differences, however, do not justify discounting the results of Trial 003. Rather, Trial 003 generally encompasses Makena's indicated population. We do not minimize the Trial 002 results; however, we note that concerns were raised by CDER and at the 2006 Advisory Committee Meeting about the generalizability of the study population to the U.S. population.[108,109]

The U.S. subgroup in Trial 003 better reflects the general target population of Makena than Trial 002 because its subjects were not highly concentrated at one investigational site. In Trial 002, an unusually large proportion of subjects (27%) were enrolled at a single clinical investigation site, the same site that enrolled 43% of the trial's Black subjects, which can create challenges when assessing generalizability of the study data.[110] Trial 003 was much larger than Trial 002, and the

[105] Sponsor's Response to NOOH at 17.

[106] *Id*. at 17–24, 65.

[107] *Id.* at 25.

[108] February 2011 action package for NDA 021945 available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945_makena_toc.cfm.

[109] *See* 2006 ACRHD Meeting on Gestiva (August 29, 2006), attached as Appendix 1.

[110] The largest center, the University of Alabama, enrolled 27% of all subjects in Trial 002 and 43% of all Black trial participants.

size of Trial 003's U.S-only subgroup was close to that of Trial 002. The U.S.-only subgroup in Trial 003 was enrolled at 41 centers, with the largest recruitment site enrolling only 12% of the U.S. subjects (46 of 391 U.S. subjects).

Broadly, CDER disagrees with Covis' assertion that Trial 003 enrolled a "low-risk" population, as reflected in the "low" recurrent PTB rate. The rate of recurrent PTB among subjects in the placebo arm in Trial 003 aligned more closely with the background risk of recurrent PTB in the U.S. population than the rate of recurrent PTB among subjects in Trial 002. In the U.S. subpopulation of Trial 003, the rate of PTB < 37 weeks in the placebo arm was 28%, consistent with a population at high risk for recurrent PTB (see below), and the trial's overall rates of PTB < 37 weeks in the placebo arm (22%) are not inconsistent with U.S. rates for recurrent PTB. While the risk of recurrent PTB in women with a prior sPTB can vary greatly, the notably high placebo recurrent PTB rate of 55% in Trial 002 is not representative of background risk of recurrent PTB in the U.S. population. A population-based cohort study evaluated the rate of recurrent PTB in the second pregnancy after having a PTB in the first pregnancy by racial distribution in U.S. women from Georgia. The overall recurrent rate for PTB (all births < 37 weeks) was 20% in White women and 26% in Black women. The study also evaluated the recurrence rate specifically in those with a prior sPTB < 32 weeks, an early gestational age considered to be major risk for having recurrent PTB. Among 1,023 White women, the recurrent rate for another PTB < 32 weeks was 8% and for PTB 32-36 weeks was 20%. Among comparable 1,084 Black women, the recurrent rates were 13% (another PTB < 32 weeks) and 23% (PTB 32-36 weeks).[111] A 1996 study by the MFMU Network reported, "[a]mong 378 patients with a prior spontaneous preterm birth or spontaneous abortion in the second trimester (gestational age range: 18–36 weeks), the rate of recurrent spontaneous preterm birth (< 35 weeks) varied between 14% and 15%, in contrast to the 3% rate of spontaneous preterm birth among 904 parous women with a previous uncomplicated term delivery."[112] Another study by the MFMU Network, published in 1999, reported women with any prior sPTB carried a 2.5-fold increase in risk of sPTB before 37 weeks gestation in the current gestation (22% vs. 9%).[113] A reputable reference source for high risk obstetrics indicates the recurrence rate increases in women of all races as the number of prior preterm births increases, with a nearly twofold rise for each prior preterm birth.[114] According to the Centers for Disease Control and Prevention (CDC),

---

[111] Adams MM, Elam-Evans LD, Wilson HG, Gilbertz DA. Rates of and factors associated with recurrence of preterm delivery. JAMA. 2000;283(12):1591–1596

[112] Goldenberg RL, Mercer BM, Meis PJ, Copper RL, Das A, McNellis D. The preterm prediction study: fetal fibronectin testing and spontaneous preterm birth. NICHD Maternal Fetal Medicine Units Network. Obstet Gynecol. 1996;87:643–48.

[113] Mercer BM 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. 1999;181(5 Pt 1):1216.

[114] Hyagriv N. Simhan; Vincenzo Berghella; Jay D. Iams. "Prevention and Management of Preterm Parturition." Creasy & Resnik's Maternal-Fetal Medicine, Principles and Practice 8th Edition, edited by Robert Resnik; Charles J Lockwood; Thomas R. Moore; Michael F. Greene; Joshua A. Copel; Robert M. Silver, Elsevier, 2018, 679–711.

the rate of PTB in singleton pregnancy in the U.S. is approximately 8.5%.[115] Assuming that having a prior sPTB increases the risk of having another PTB approximately two-fold,[111] the recurrence rate after one prior sPTB would be approximately 17%.[116] In summary, the recurrent PTB rates of < 37 weeks gestation for Trial 003 overall (22%) and for the U.S. subgroup (28%) are within the range of those reported for women in the U.S. with a prior singleton PTB.

Covis claims a major factor contributing to the lower rate of recurrent PTB in Trial 003 was that a smaller proportion of the trial's subjects had a prior sPTB < 34 weeks, an important risk factor for recurrent PTB, as compared to Trial 002. In Trial 002, 67% of women had one or more prior sPTBs < 34 weeks, and in Trial 003, that number is 61%. When looking at the sPTB that qualified them for the trial (the most recent sPTB), 61% of women randomized in Trial 002 and 57.5% of women randomized in Trial 003 had a qualifying PTB < 34 weeks. CDER's analysis showed that the distributions of the gestational ages of the qualifying sPTB and all prior sPTB are similar between Trials 002 and 003. In fact, the distributions of the gestational ages of all prior births shows a lower average gestational age of delivery for Trial 003 subjects (25.6 weeks GA) compared to Trial 002 subjects (29.1 weeks).

Table 7: Trial 002 and Trial 003 Subjects Have Similar Distributions of Gestational Age at Prior sPTB Deliveries

Trial 002						
Distribution of GA (weeks)	**Min**	**25% Percentile**	**Mean**	**Median**	**75% Percentile**	**Max**
Qualifying sPTB	20	28	30.8	32	35	36
All prior sPTB	20	27	30.5	32	35	36
All prior births	2	24	29.1	33	37	42
Trial 003						
Distribution of GA (weeks)	**Min**	**25% Percentile**	**Mean**	**Median**	**75% Percentile**	**Max**
Qualifying sPTB	20	28	31.4	33	35	36
All prior sPTB	20	28	31.3	33	35	36
All prior births	2	11	25.6	31	36	42

Importantly, differences in risk factors between subjects in Trial 002 and 003 do not account for the differences in the trials' efficacy findings. While Covis argues that the differences in risk factors between the patient populations in Trials 002 and 003 means that Trial 003 should not carry equal weight, there is no evidence that these risk factors drove the different efficacy results

[115] Martin JA, Osterman MJK. Exploring the decline in the singleton preterm birth rate in the United States, 2019–2020. NCHS Data Brief, no 430. Hyattsville, MD: National Center for Health Statistics. 2022, available at https://www.cdc.gov/nchs/data/databriefs/db430.pdf (last visited Sept. 14, 2022).

[116] *Id*.

between the two trials. Specifically, the analyses of Trial 002 indicated that the drug effect on delivery at < 37 weeks gestation was not differential based on varying maternal risks of experiencing recurrent PTB. Similarly, there was no differential drug response in Trial 003: no treatment effect was seen with Makena in any subgroup analyzed. In both trials, analyses did not find that the presence of risk factors made a difference in treatment response in either Trial 002, where a treatment effect was demonstrated, or Trial 003, where no treatment effect was seen.

As described in Section V.C., there is nothing to suggest a consistent, meaningful, favorable response to Makena compared to placebo that is durable across a subgroup in both Trial 002 and Trial 003. In addition, there is not a known biological or pharmacological reason to expect subgroup differences in treatment effect compared to placebo.

2. Trial 003 Was Not Underpowered

CDER disagrees that Trial 003 was inadequately powered. In fact, Trial 003 offered enough precision to rule out clinically relevant effects on sPTB < 37 and rule out the 35% reduction (or larger) in neonatal composite index that the trial was designed to demonstrate.

Contrary to Covis' arguments, Trial 003 enables analyses with appropriate power to examine key questions of efficacy including whether Makena conferred its expected clinical benefit. A claim that a study is "underpowered"[117] refers to the risk that the study may fail to show an effect of a drug where there actually is one. The study power is the probability that the study will detect a treatment effect of a specified magnitude (and not just any effect) when such effect exits; power is planned for and prespecified as part of designing a study.

The sponsor's power calculation for Trial 003[118] was designed to find a relative 30% or greater reduction in the risk of PTB < 35 weeks, the same reduction observed in Trial 002 for this endpoint. It also was designed, based on Trial 002, to find a relative 35% or greater reduction in the neonatal composite index which, according to the sponsor of Trial 003, "…was chosen to represent a clinically significant reduction."[119] Additionally, by powering the co-primary endpoints together at 90% if the outcome measures were highly correlated (and 88.2% if the outcome measures were independent), the neonatal composite index endpoint standing alone, and the PTB < 35 and < 37 weeks endpoints standing alone, had very high power for the hypothesized effect sizes and background (placebo) rates in the Trial 003 design.[120]

---

[117] Sponsor's Response to NOOH at 24–25.

[118] AMAG Briefing Document, BRUDAC Meeting (October 29, 2019) (page 55 of 91), available at https://www.regulations.gov/document/FDA-2020-N-2029-0114 and attached as Appendix 2.

[119] Blackwell SC, et al. 17-OHPC to prevent recurrent preterm birth in singleton gestations (PROLONG study): a multicenter, international, randomized double-blind trial. Am. J. Perinatol. 2020;37(02):127–136, available at https://www.thieme-connect.com/products/ejournals/html/10.1055/s-0039-3400227; *see also* Sponsor's NOOH Response at Attachment 31.

[120] For example, the design of Trial 003 was such that there was less than a 2% chance of failing to detect a 30% reduction in PTB < 35 weeks (i.e., 98% power).

In general, CDER does not support post-hoc power estimates (power calculated for trial analyses after they are performed, and where the results are known) because they can be misleading. However, in an attempt to understand the different findings between Trials 002 and 003, CDER looked at the power for the observed placebo event rates and the observed sample sizes in Trial 003. Because Trial 003 was powered to detect an expected change from 17% to 11% on the neonatal index (35% relative reduction) seen in Trial 002, Trial 003 had a high amount of power, around 90%, to detect changes for other endpoints, such as PTB < 37 weeks, with a placebo event rate of percentages as low as in the high teens to the low twenties. Therefore, Trial 003 had 90% power to detect a 30% relative reduction in an endpoint with a placebo rate of 21.9%, which was the observed placebo-arm event rate for PTB < 37 weeks in Trial 003. In other words, if Makena conferred a 30% relative reduction in the risk of delivering < 37 weeks (as reported in Trial 002), Trial 003 had 90% power to detect such a treatment effect even at the lower placebo rate of 21.9%. Despite being adequately powered to detect a 30% reduction in PTB < 37 weeks, Trial 003 did not succeed in doing so. Furthermore, Trial 003 still had reasonable power (> 70%) to detect a 25% relative reduction and the hypothesized 30% relative reduction with placebo PTB event rates as low as 15%.

It is important to look at the Trial 003 results and consider how much uncertainty there is, and is not, around the estimated effects (e.g., as reflected by the confidence intervals). The confidence intervals for the analysis of Trial 003 excluded the hypothesized relative reductions in PTB < 35 weeks (30%) and the neonatal composite index (35%) that the trial was designed to detect.[121] They also exclude many small clinically meaningful effects. For instance, the lower bounds of the 95% confidence interval for the relative risk (RR) and absolute treatment difference for birth < 37 weeks in Table 7 were 0.88 and -3.0, respectively. That is, Trial 003 excluded (based on the statistical interpretation of 95% confidence intervals) a relative risk reduction of > 12% and an absolute risk reduction of > 3% for this endpoint with Makena treatment. In other words, the precision of Trial 003 was good enough to rule out even small changes in the relative and absolute risk of PTB.[122]

[121] AMAG Slides, BRUDAC Meeting at CO-53, 54 (Oct. 29, 2019), available at https://www.regulations.gov/document/FDA-2020-N-2029-0117, and attached as Appendix 2.

[122] When a value is not in a 95% confidence interval, we often conclude that the value is not plausible or is unlikely to be the true value and can be "ruled out."

Table 8: Trial 003 Results Exclude Clinically Meaningful Effect Sizes

Efficacy Outcome	Trial 002 RR (95% CI)	Trial 003 RR (95% CI)	Trial 002 Treatment Difference (95% CI)[123]	Trial 003 Treatment Difference (95% CI)
Neonatal composite index*		1.05 (0.68, 1.61)[124]		0.2% (-2.0, 2.5)
Birth < 35 weeks	0.69 (0.49, 0.98)	0.95 (0.71, 1.26)	-9% (-19, -0.4)	-0.6% (-3.8, 2.6)
Birth < 32 weeks	0.61 (0.38, 0.98)	0.92 (0.60, 1.42)	-8% (-16, -0.3)	-0.4% (-2.8, 1.7)
Birth < 37 weeks**	0.68 (0.54, 0.84)	1.06 (0.88, 1.28)	-18% (-28, -7)	1.3% (-3.0, 5.4)

*In Trial 003, proportion of liveborn neonates experiencing at least one event of the composite index; N=1651 (Makena missing data on 39 neonates, Placebo missing data on 18 neonates); defined as Yes if the liveborn neonate had any of RDS, BPD, Grade 3 or 4 IVH, NEC, proven sepsis, death. A similar variable was created for Trial 002 and 35/295 (12%) of neonates born to Makena subjects and 26/151 (17%) of neonates born to Placebo subjects had one or more of the listed morbidities. The finding was not statistically significant.
**Primary efficacy endpoint of Trial 002

Looking at Figure 1 and Table 8, it is clear that:

- The entire confidence interval for the relative risk (0.54, 0.84) for the Trial 002 PTB < 37 weeks endpoint is excluded from the confidence interval (0.88, 1.28) for the same endpoint in Trial 003. This is also true for the treatment differences. If there were a true treatment effect of Makena for the indicated population, one would expect at least some overlapping of CIs.
- In Trial 003, the lower bound of the confidence intervals on the relative risks for PTB < 35 weeks and PTB < 37 weeks are above 0.70 and rule out the planned-for 30% relative reduction in both rates.
- The Trial 003 results excluded even a 4-percentage-point reduction in the rate of PTB < 32, < 35, and < 37 weeks. Based on the Trial 003 results, if there is a treatment effect on the rate of recurrent PTBs, it likely is modest and nowhere near the treatment effect size upon which Makena was approved in 2011.
- Although the neonatal composite index was not an endpoint in Trial 002, looking at Trial 003's results based on the sponsor's relative risk analysis of the neonatal composite index:
    - The sponsor's and investigator's point estimate for the relative risk is +12%[125] and CDER's estimate is +4%, both of which are unfavorable for Makena.
    - The lower bound of the sponsor's and CDER's confidence intervals are above 0.65, which excludes the planned-for relative 35% reduction.

[123] *See* FDA-approved prescribing information for Makena, available at https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s013lbl.pdf and attached as Appendix 10.

[124] AMAG Slides, BRUDAC Meeting at CO-53 (Oct. 29, 2019), available at https://www.regulations.gov/document/FDA-2020-N-2029-0117, and attached as Appendix 2; *see also* Blackwell, et al. list a RR of 1.12 (0.72, 1.72), different from CDER's calculations. CDER calculated its results using the livebirth flag and the neonatal composite index variables. The Makena arm had 59 of 1091 liveborn neonates with an index event, and in the Placebo arm 29 of 560 liveborn neonates had an index event.

[125] *Id.*

For these reasons, CDER concludes that Trial 003 was appropriately powered, even given post-hoc power estimates, to find a difference between the treatment arms and at least partly verify Trial 002.

3. The Determination of Gestational Age in the Qualifying Pregnancies in Trial 003 Was Reliable

Covis argues that the Trial 003 protocol did not require what Covis deems to be appropriate methodology to verify the prior qualifying gestational age. Covis asserts that, in the last decade, providers in the United States have consistently used fetal crown-rump and femur lengths as assessed by ultrasound to determine gestational age, whereas such practice is not common outside the United States, especially not in Russia and Ukraine, which enrolled 61% of Trial 003 subjects combined. Covis suggests that differences in obstetrical ultrasound practice would impact the assessment of gestational age for the subjects in Trial 003 who were enrolled at sites not using the up-to-date methodology even if the trial protocol was followed.

CDER disagrees with Covis that the assessment of gestational age of prior qualifying pregnancies would bias the outcomes of Trial 003. One inclusion criterion in the trial protocol called for "[d]ocumented history of a previous singleton spontaneous preterm delivery…defined as delivery from 20^0 to 36^6 of gestation following spontaneous preterm labor or pPROM [preterm premature rupture of membranes]."[126] The gestational age of the qualifying pregnancy was determined either by patient report or obtained directly from the medical record, and cross-checked by neonatal birth weight. If the neonate weighed more than 3300 grams at birth (90^{th} percentile for 36 weeks gestational age), that delivery would not qualify as preterm. Since the gestational age of the previous preterm delivery (the qualifying pregnancy) was a pre-randomization variable, upon randomization, the proportions of Russian, Ukraine, or U.S. subjects between the Makena group and the placebo group should be balanced. Accordingly, the gestational age of qualifying pregnancies should be comparable between the two treatment groups. The protocol-specified validation of the neonatal weight would be applicable to all subjects in the trial, irrespective of country of enrollment. Indeed, the Clinical Study Report notes that the two "[t]reatment groups were well balanced with respect to obstetrical characteristics and previous pregnancies."[127] Although Covis notes that the standard of care in Ukraine recommended ultrasound in the second trimester, not first trimester, Covis provided no data to demonstrate that, in Trial 003, a higher proportion of Russian and Ukrainian subjects had qualifying pregnancies in which the preterm neonates had significantly higher than expected birthweights compared to the proportion of U.S. subjects.

[126] Study Protocol, Protocol Number 17P-ES-003, Version 6.0, Version Date 06 April 2016, Section 5.1 Inclusion Criteria, available at https://clinicaltrials.gov/ProvidedDocs/29/NCT01004029/Prot_000.pdf.

[127] *Id.* at Section 2, Synopsis, Summary and Conclusions.

4. Selective Pooling of Data is Scientifically Inappropriate

Covis has previously proposed combining the findings from Trial 002 and the U.S.-only subgroup of Trial 003 to demonstrate Makena's efficacy.[128] But there is no scientific basis on which to select the U.S.-only subgroup from Trial 003 when subgroup analyses based on region (U.S. compared to non-U.S.) did not show differential treatment effect between women enrolled at sites in the U.S. and outside the U.S. in Trial 003. In the case of Makena, where there are two adequate and well-controlled trials that were appropriately conducted, such post-hoc analysis of selectively combining findings from these two trials is not scientifically valid. FDA's general position for the demonstration of efficacy is to evaluate the evidence from trials individually to provide independent substantiation of experimental results.[129]

5. Trial 002 Cannot Support Traditional Approval

Covis suggests that two of the gestational age endpoints in Trial 002 (PTB < 32 and < 35 weeks) could be viewed as validated surrogate endpoints and thus Trial 002 alone should have supported traditional approval of Makena, without a requirement that the sponsor conduct a confirmatory trial, rather than accelerated approval. CDER disagrees. The NDA for Makena relied only on one adequate and well-controlled trial (Trial 002) as substantial evidence of effectiveness. CDER determined that the reductions in PTB < 32 weeks or < 35 weeks, although statistically significant but with an upper bound of the 95% confidence interval for the relative risk near 1.0, were not statistically persuasive enough to support approval based on the findings from a single trial. Therefore, contrary to Covis' assertions, Trial 002 alone could not have supported traditional approval based on findings of gestational age of delivery < 32 weeks or < 35 weeks.

Finally, even if, hypothetically speaking, Trial 002 could have supported traditional approval—and CDER disagrees that it could have—that would not change CDER's conclusion that, in light of the data available today, there is not substantial evidence of Makena's effectiveness. Trial 002 was not, at the time of the accelerated approval, sufficient to support traditional approval. Furthermore, we have more information now than we did in 2011. With the entirely negative results from Trial 003, it is clear that the current evidence does not provide substantial evidence of effectiveness, and this conclusion is further supported by the negative results from three RCTs in related conditions of increased risk for sPTB and from five observational study analyses. As explained below, Trial 002 is the outlier among the studies relevant to Makena's effectiveness – it could well be a false positive. Considering the entire body of evidence regarding Makena's efficacy available today, CDER can only conclude that it falls short of the threshold for showing substantial evidence of effectiveness.

---

[128] Sponsor's Response to NOOH at 26–29.

[129] FDA Guidance for Industry, Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products (May 1998), available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/providing-clinical-evidence-effectiveness-human-drug-and-biological-products.

E. The Evidence Does Not Demonstrate Makena's Effectiveness for Its Indicated Population or Any Subgroup

Based on the available evidence, Makena is no longer shown to be effective for its approved use—either for its indicated population or any subset of that population. The randomized, placebo-controlled trials specifically designed to evaluate Makena's approved use in the intended population—Trials 002 and 003—provide the most rigorous and relevant evidence to evaluate Makena's efficacy for its approved use. Makena's accelerated approval was based on one adequate and well-controlled trial (Trial 002). The treatment effect shown in Trial 002 on the intermediate clinical endpoint of delivery < 37 weeks gestation appeared independent of race, number of prior preterm deliveries, and gestational age of the prior PTB. Trial 003, which was almost four times larger than Trial 002, failed to show that Makena reduced the proportion of women delivering prior to 37-, 35-, or 32-weeks' gestation. Furthermore, exploratory subgroup analyses of Trial 003 did not provide evidence of a treatment effect in any identified subgroup analyzed, including those with risk factors that differed between Trials 002 and 003. There was no consistent evidence of treatment effect within an identified subpopulation across Trials 002 and 003 (e.g., by race, number of prior sPTB). CDER has generally interpreted the substantial evidence of effectiveness standard as requiring 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. If the collective findings of Trials 002 and 003 were submitted at the same time in an NDA seeking approval, CDER would conclude that there is not substantial evidence of effectiveness of Makena for reducing the risk of recurrent PTB—for either the overall population of women with a singleton pregnancy and a history of singleton sPTB, or for any identified subset of that population.

Lastly, review of published evidence potentially relevant to the efficacy of Makena did not demonstrate Makena was effective in reducing the risk of PTB.

CDER thoroughly explored all available evidence of effectiveness. CDER's analysis included other available data, in addition to Trials 002 and 003, on the effect of HPC (the active ingredient in Makena) on singleton sPTB. These data included CDER's review of (a) the EPPPIC meta-analysis,[130] (b) five observational studies with HPC in Makena's indicated population, (c) three randomized, placebo-controlled trials described by Covis as supporting safety of HPC in other groups at high risk for PTB, and (d) three additional studies Covis points to as supporting the Trial 002 results. After review, CDER concluded these studies do not support a finding that Makena is effective in reducing PTB, as enumerated below.

a. EPPPIC (Evaluating Progestogens for Preventing Preterm birth International Collaborative) is an individual participant data meta-analysis of RCTs that evaluated the effect of

[130] *See* Addendum to Makena Decisional Memorandum, EPPPIC (Jan. 14, 2022), available at https://www.regulations.gov/document/FDA-2020-N-2029-0187 and attached as Appendix 5. The EPPPIC study is available at https://www.pcori.org/research-results/2017/evaluating-hormone-treatments-women-increased-risk-preterm-birth-%E2%80%93-epppic.

various progestogens (vaginal progesterone, oral progesterone, and injectable HPC) compared to control (placebo or no-intervention) or to each other administered during pregnancy in preventing PTB (first occurrence or recurrent PTB). EPPPIC included a total of 31 RCTs, consisting of 11,644 women with singleton or multifetal gestations, with or without a history of prior PTB, and with or without a short mid-trimester cervical length.[131] Among EPPPIC's 31 clinical trials, CDER focused on the five trials (two of which were Trials 002 and 003) in the EPPPIC meta-analysis that have some relevance to Makena, because these trials evaluated HPC in singleton pregnancies and were placebo-controlled. Other than Trials 002 and 003, two of the five trials included non-Makena indicated patients, such as pregnant women with a shortened cervix in the current pregnancy but no prior birth, or dosing different than that of Makena (HPC 500 mg rather than the HPC 250 mg in Makena). The remaining trial had known drug quality issues potentially impacting drug potency and efficacy. In addition, CDER's review identified important issues in data analysis and results interpretation of the EPPPIC study. Specifically, the investigator's conclusion of beneficial effect in reducing PTB< 34 weeks was based on an effect estimate that was not statistically significant (HR= 0.83, 95% CI= 0.68-1.01). Second, the claim of treatment effect among "high-risk" women (those with short cervix) is not evident because the analyses were based on a small subset (n=81, 2.7%) of total study populations in five trials, with 70% of them treated with an HPC dose regimen twice that of the approved Makena dose, and without suitable statistical adjustment for multiplicity after assessing multiple subgroup analyses. Even given these caveats, the meta-analysis of these five HPC trials did not show a statistically significant finding on the main outcomes of delivery prior to 37-, 34-, or 28-weeks' gestation, perinatal deaths, or serious neonatal complications.

b. CDER also identified five observational studies that provided exploratory evidence about HPC's effect for its approved use using keyword searches[132] of the PubMed database.[133,134] CDER concluded that none of these studies showed evidence that HPC reduced the rate of PTB.

---

[131] The EPPPIC Group. Evaluating Progestogens for Preventing Preterm birth International Collaborative (EPPPIC): meta-analysis of individual participant data from randomised controlled trials. The Lancet 2021;397 (10280):1183-1194.

[132] These observational studies were identified through PubMed keyword searches using the keywords "hydroxyprogesterone caproate," "effectiveness," "effect," "preterm birth." No other limitations were applied. Studies were selected if they were cohort or case-control observational studies (non-randomized) that evaluated the effectiveness of HPC treatment compared to non-use. After screening at the titles and abstract level, five studies met these criteria. CDER last confirmed these search results on September 13, 2022.

[133] As discussed above in CDER's proposed answer to Question 3, observational studies like these are of limited relevance to the overall body of evidence of efficacy for Makena. As explained in Section V.G., we believe only a randomized, double-blind, placebo-controlled study could potentially provide sufficient evidence of efficacy.

[134] The product and dose evaluated in these studies is as follows: Bastek (HPC, 250 mg), Hakim (Makena, 250 mg), Massa (HPC, dose and product unspecified), Nelson (HPC 250 mg compounded in sesame oil), Wang (HPC, dose and product unspecified).

Two of these studies (Bastek and Nelson) were conducted at large urban institutions that served obstetrical populations.[135, 136] These two studies compared the institutions' PTB rate when HPC was standard of care to the institutions' PTB rate prior to HPC becoming the standard of care (historical controls), and found no difference between overall PTB rates in the time periods compared among Makena's indicated population. Nelson et al. found that the overall rate of recurrent PTB for the entire cohort treated with HPC (25%) was comparable to the expected rate observed in historical obstetric population (16.8%). The Bastek data upon which Covis relies to support its claim of efficacy was only in a subset of women who went on to have PTB, not the overall population eligible for Makena. Covis' reliance on the Bastek study is particularly problematic, as this study used a pre-post cross-sectional design that compared the institutional rate of PTB before and after Makena became the standard of care at the institution. This study failed to provide the robust control of confounding required to allow for conclusions on treatment effect. Bastek et al. only included a limited number of variables to control for differences between populations in the two time periods, while failing to address the impacts that may have resulted from other changes over time, such as staffing and improvements for practitioner experience. The third study (Massa et al.) was a retrospective cohort study comparing HPC users to non-users in an academic tertiary care center with a patient population similar to Trial 002.[137] This study also failed to demonstrate that HPC was effective for use to prolong pregnancy up to 35 weeks. The fourth study (Hakim et al.) was a retrospective cohort study conducted in an insurance claims database and again failed to show benefit of HPC to prevent pre-term birth.[138] Finally, a retrospective cohort study (Wang et al.), undertaken among Medicaid enrollees in Pennsylvania, assessed effectiveness of HPC on prevention of recurrent PTB and admission to neonatal intensive care.[139] This study found no evidence of benefit in prevention of PTB and found no difference in the risk of admission of the neonate into intensive care between treated and untreated mothers. Although observational studies have limitations in drawing causal inferences, the lack of evidence of effectiveness across these 5 studies provides further support for the conclusion that Makena is not effective in the indicated population.

c. Covis referenced three randomized, placebo-controlled trials in patient populations at high risk for PTB for reasons other than having a prior sPTB to support Makena's safety (CDER

[135] Bastek JA, Adamczak JE, Hoffman S, Elovitz MA, Srinivas SK. Trends in prematurity: what do changes at an urban institution suggest about the public health impact of 17-alpha hydroxyprogesterone caproate? Matern. Child Health J. 2012;16:564–568.

[136] Nelson DB, McIntire DD, McDonald J, Gard J, Turrichi P, Leveno KJ. 17-alpha hydroxyprogesterone caproate did not reduce the rate of recurrent preterm birth in a prospective cohort study. Am. J. Obstet. Gynecol. 2017;216:600.e1–9.

[137] Massa K, Childress K, Vricella LK, et al. Pregnancy duration with use of 17-a-hydroxyprogesterone caproate in a retrospective cohort at high risk of recurrent preterm birth. Am. J. Obstet Gynecol. MFM 2020;2:100219.

[138] Hakim J, Zhou A, Hernandez-Diaz S, Hart J, Blair J. Wylie B, Beam A, Effectiveness of 17-OHP for prevention of recurrent preterm birth: a retrospective cohort study, Am. J. Perinatol. 2021.

[139] Wang X, Garcia S, Kellom K, Boelig R, Matone M, Eligibility, utilization, and effectiveness of 17-alpha hydroxyprogesterone caproate (17OHPC) in a statewide population-based cohort of medicaid enrollees. Am. J. Perinatol. 2021.

also evaluated them for efficacy).[140,141,142] Price et al. described a randomized, double-blind, placebo-controlled trial of 800 pregnant women with HIV at the University Teaching Hospital and Kamwala District Health Center in Lusaka, Zambia. Because women living with HIV face a substantially increased risk of PTB, this trial's objective was to evaluate whether HPC reduced the risk of PTB. Women with HIV who had a viable singleton pregnancy less than 24 weeks gestation could participate, but were excluded if they had or planned to have in situ cervical cerclage and history of sPTB, among other criteria. The primary outcome was a composite of PTB < 37 weeks or stillbirth (fetal death diagnosed before delivery or delivery of a neonate without signs of life). The results of this study showed no statistical or discernible numerical difference in PTB < 37 weeks or stillbirth, occurring in 9% in both groups. Rouse et al. and Caritis et al. reported findings from trials conducted in the U.S. by the MFMU Network, the group that conducted Trial 002. Both trials were double-blinded, placebo-controlled, randomized trials of HPC in women with multiple gestations, an established risk factor for PTB.[143] In the study by Rouse et al., 661 healthy women with twin gestations were randomized to weekly intramuscular injections of 250 mg or matching placebo, starting at 16 to 20 weeks gestation and ending at 35 weeks. This study found that HPC did not reduce the rate of PTB in women with twin gestations, with preterm delivery or fetal death before 35 weeks occurring in 41.3% of pregnancies in women receiving HPC and 37.3% of those in the placebo group. Caritis et al. randomized 134 healthy women with triplets to weekly intramuscular injections of 250 mg or matching placebo. Treatment was started at 16 to 20 weeks gestation and ended at 35 weeks. The results showed a relative risk of 1.0 (95% confidence interval 0.9-1.1). This study also concluded that HPC did not reduce the rate of PTB.

d. Covis points to three additional studies (two secondary analyses of RCT data, and one observational study) that it claims support the Trial 002 results. After review, CDER concluded these three studies do not support a finding that Makena is effective in reducing PTB.

The first study (Caritis et al.)[144] evaluated the relationship between HPC plasma concentrations and the rate of recurrent PTB in women with singleton gestation and receiving HPC 250 mg. The authors reported that women in the lowest quartile blood concentration of HPC had a higher risk of recurrent PTB than those in the second to fourth quartiles. However, we note there was no concentration-response among the three higher quartiles and without a placebo control, it was not possible to elucidate a drug effect, if one existed. The second study (Manuck et al.)[145]

---

[140] Price JT, et al. Weekly 17 alpha-hydroxyprogesterone caproate to prevent preterm birth among women living with HIV: a randomised, double-blind, placebo-controlled trial, The Lancet HIV. 2021; 8(10): e605-e613.

[141] Caritis SN, et al. Prevention of preterm birth in triplets using 17 alpha hydroxyprogesterone caproate a randomized controlled trial. Am. J. Obstet. Gynecol. 2009;113(2):285–292.

[142] Rouse DJ, et al. A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. N. Engl. J. Med. 2007;357:454–461.

[143] Sponsor's Response to NOOH at 38.

[144] Caritis SN, et al. Relationship between 17-alpha hydroxyprogesterone caproate concentration and spontaneous preterm birth, Am. J. Obstet. Gyencol. 2014; 210(2):128.e1–6.

[145] Manuck TA, Stoddard GJ, Fry RC, et al. Nonresponse to 17-OHPC for recurrent spontaneous preterm birth prevention: clinical prediction and generation of a risk scoring system. Am. J. Obstet. Gynecol. 2016;215:622.e1–8.

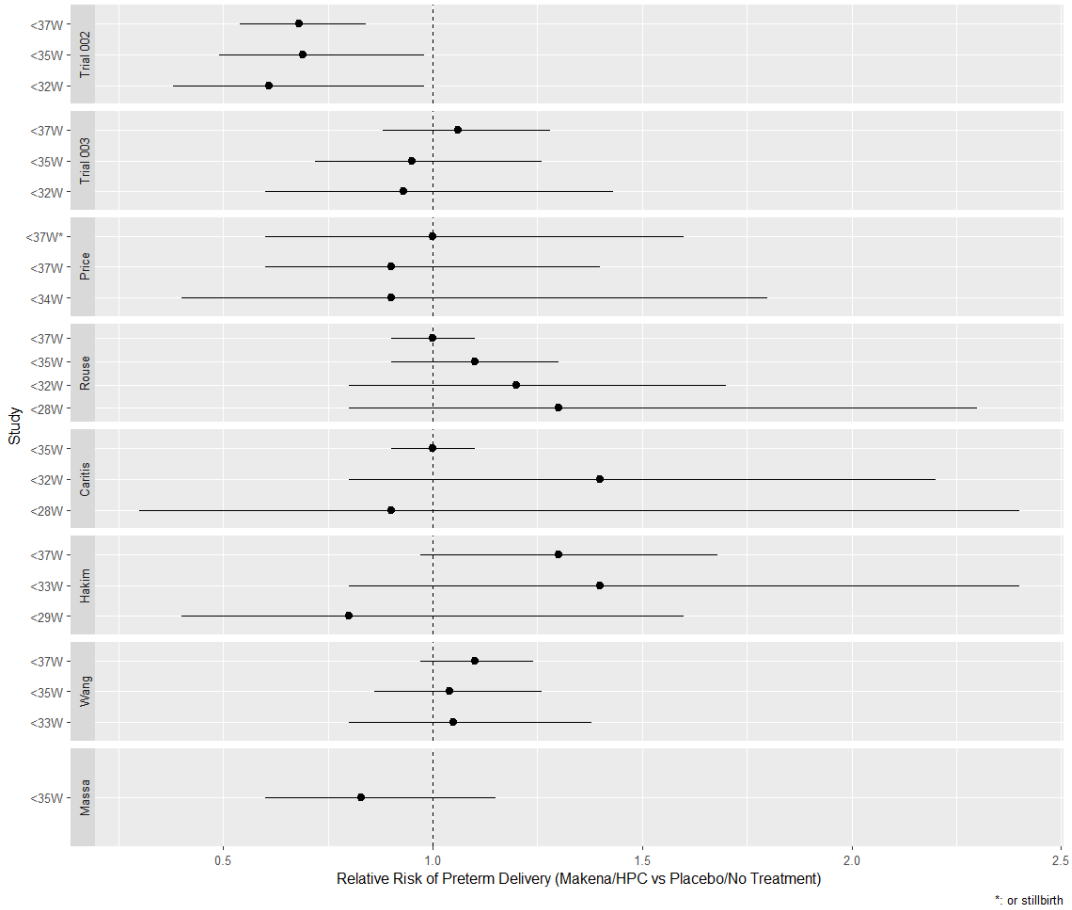
analyzed an RCT of omega-3 supplementation where all 754 women in the trial received background HPC 250 mg at the same dosing schedule as Makena. Secondary analyses classified women as HPC "responders" if they delivered at a gestational age at least three weeks later than the woman's prior PTB. Of the women in the trial, 21% were classified as "non-responders" and 79% were "responders."[146] These results do not provide support for Makena's efficacy, because the proportion of "responders" and "non-responders" in subjects not receiving HPC was not determined (i.e., there was no placebo-control for HPC), rendering these results uninterpretable. The majority of women with prior sPTB do not have recurrent PTB—so these results, in fact, could be compatible with the natural history of women with prior sPTB. Finally, the third study (Carter et al.)[147] was a retrospective cohort study using MarketScan claims data that found both early HPC initiation and compliance are associated with reduced rates of PTB, contrary to the similar subgroup analyses in Trials 002 and 003 which failed to find this effect. As with the other observational studies, important confounders were not controlled, such as health behaviors. Higher risk of PTB in less adherent subjects is likely to be due to many other reasons besides drug effect.

Figure 16 below shows a forest plot of the relative risk reduction with Makena (or HPC) compared to placebo (or no treatment) for preterm delivery at several gestational age cut-offs for the available randomized, placebo-controlled trials in Makena's target population (Trial 002 and Trial 003) and the published results from trials in women at high risk for PTB from HIV (Price) or multiple gestations (Caritis, Rouse), and observational studies in women eligible to receive Makena/HPC (Hakim, Wang, Massa). Trial 002 is clearly the outlier in showing a favorable effect of Makena. The confidence intervals in all other studies overlap 1.0, showing no statistically significant effect.

[146] The definitions of "responder" and "non-responder" in this study are misleading because the natural history of recurrent PTB in a woman with a prior sPTB is not well-understood.

[147] Carter EB, Cahill AG, Olsen MA, Macones GA, Tuuli MG, Stout MJ. Practical considerations with 17-Hydroxyprogesterone caproate for preterm birth prevention: does timing of initiation and compliance matter? J Perinatol. 2019;39(9):1182–1189.

Figure 16: Outlier is Trial 002: Forest Plot of Relative Risk of Preterm Delivery in RCTs and Observational Studies in the Indicated Population and RCTs in Non-Indicated High Risk Populations



- Trials 002 (Meis, N=463) and 003 (PROLONG, N=1,708): RCTs for Makena's intended population
- Price (N=800),[148] Rouse (N=661),[149] Caritis (N=134)[150]: RCTs for women at high risk for PTB (HIV – Price; multiple gestations – Caritis, Rouse)
- Massa (N=861),[151] Hakim (N=4,422),[152] Wang (N=4,781)[153]: Observational studies with untreated concurrent comparator
- Studies that do not report relative risk (Bastek,[154] Nelson[155]) are excluded from this figure. Both studies found no difference in overall PTB rates comparing study periods before and after Makena's approval among Makena's indicated population.

[148] Price JT, et al., Weekly 17 alpha-hydroxyprogesterone caproate to prevent preterm birth among women living with HIV: a randomised, double-blind, placebo-controlled trial, The Lancet HIV. 2021; 8(10): e605–e613.

[149] Rouse DJ, et al. A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. N. Engl. J. Med. 2007;357:454–461.

[150] Caritis SN, et al. Prevention of preterm birth in triplets using 17 alpha hydroxyprogesterone caproate a randomized controlled trial. Am. J. Obstet. Gynecol. 2009; 113(2):285–292.

[151] Massa K, Childress K, Vricella LK, et al. Pregnancy duration with use of 17-a-hydroxyprogesterone caproate in a retrospective cohort at high risk of recurrent preterm birth. Am. J. Obstet. Gynecol. MFM 2020;2:100219.

In sum, no RCTs (other than Trial 002), observational studies, or meta-analyses reviewed by CDER provide reliable evidence of an effect of HPC on reducing PTB. Data from these studies do not provide support for the efficacy of Makena in reducing the risk of recurrent PTB.

F. Makena's Benefit-Risk Profile is Unfavorable and Supports Removing the Product from the Market

Covis contends that the Agency should maintain approval of Makena while the company conducts further studies because the available data show that Makena has a favorable benefit-risk profile.[156] The opposite is true; the benefit-risk profile of Makena is unfavorable due to its lack of demonstrated efficacy and the risks attributable to the drug. As discussed in section V.E, the available evidence fails to demonstrate that Makena improves neonatal outcomes from complications of PTB or that Makena reduces the prevalence of PTB among Makena's indicated population or any subset of that population. In the absence of benefit, any amount of risk makes a drug's benefit-risk profile unfavorable, and all drugs carry risks.

Makena's medical risks include thromboembolic events (i.e., blood clots), allergic reactions, depression, decreased glucose tolerance, fluid retention that may worsen maternal conditions such as pre-eclampsia, and injection site adverse reactions.[157] Covis cited the Schuster et al.[158] study to support its argument that there is no increased risk of thromboembolic events for women receiving HPC. However, this reliance is problematic because Schuster et al. had no adjustment on confounders, which affects the validity of the study findings. Given the labeling for contraindication for venous thromboembolism (VTE) risk, women at risk for VTE were more likely to be non-users, biasing the results towards a finding of no increase in risk. Further, no baseline characteristics are presented to allow for FDA to evaluate comparability between the cohorts. In the absence of demonstrated effectiveness, treating pregnant women with Makena exposes them only to risk. Accordingly, Makena should be withdrawn from the market.

---

[152] Hakim J, Zhou A, Hernandez-Diaz S, Hart J, Blair J. Wylie B, Beam A, Effectiveness of 17-OHP for prevention of recurrent preterm birth: a retrospective cohort study, Am. J. Perinatol. 2021.

[153] Wang X, Garcia S, Kellom K, Boelig R, Matone M, Eligibility, utilization, and effectiveness of 17-alpha hydroxyprogesterone caproate (17OHPC) in a statewide population-based cohort of medicaid enrollees. Am. J. Perinatol. 2021.

[154] Bastek JA, Adamczak JE, Hoffman S, Elovitz MA, Srinivas SK. Trends in prematurity: what do changes at an urban institution suggest about the public health impact of 17-alpha hydroxyprogesterone caproate? Matern. Child Health J. 2012;16:564–568.

[155] Nelson DB, McIntire DD, McDonald J, Gard J, Turrichi P, Leveno KJ. 17-alpha hydroxyprogesterone caproate did not reduce the rate of recurrent preterm birth in a prospective cohort study. Am. J. Obstet. Gynecol. 2017;216:600.e1–9.

[156] Sponsor's Response to NOOH at 17, 52.

[157] *See* FDA-approved prescribing information for Makena, available at https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s013lbl.pdf and attached as Appendix 10.

[158] Schuster M, et. al. 17-alpha hydroxyprogesterone caproate and risk for venous thromboembolism during pregnancy, J. of Matern. Fetal Neonatal Med. 2021;1–2.

In addition to the risks associated with Makena's use, keeping Makena on the market incurs false hopes and unnecessary healthcare utilization. Even in women who do not experience adverse effects of Makena, its use requires weekly injections. The recommended dosing regimen calls for Makena injection to be initiated between 16- and 20-weeks' gestation, and its administration is continued once weekly until week 37 of gestation or delivery, whichever occurs first. Thus, a woman who starts treatment with Makena at 16 weeks of gestation could receive up to 20 injections in her pregnancy if she delivers at or after 37 weeks. For many women, Makena's treatment regimen requires prenatal clinic visits or home health nursing care for injections.

Further, a recent study by Murphy et al. reporting increased cancer risk in the children of women treated with HPC, the active ingredient in Makena, highlights the uncertainty regarding the intergenerational safety of Makena for the children of women who took it during pregnancy.[159] Contrary to Covis' arguments,[160] a potential intergenerational cancer risk associated with the active ingredient in Makena is relevant to the benefit-risk profile of Makena.[161]

As the Society of Maternal Fetal Medicine (SMFM) stated in its July 2020 statement regarding Makena, it is important to consider that the "long-term potential maternal and neonatal effects [of HPC] are unknown."[162] The SMFM's statement is consistent with CDER's assessment of the Murphy Article. In the absence of demonstrated benefit, any risk (even one as remote as an intergenerational effect) cannot be dismissed.

In recognition of the seriousness of this potential risk, CDER has evaluated a Newly Identified Safety Signal (NISS) for Makena regarding the risk of cancer in children who took HPC during pregnancy.[163] CDER closed the evaluation phase of the NISS with a classification of the safety signal as indeterminate, rather than refuted, and a recommendation for active monitoring. In other words, following CDER's comprehensive assessment, the findings are inconclusive with regard to the association of the risk with the medicinal product of interest.[164] This study highlights potential intergenerational safety concerns that merit further surveillance.

---

[159] Murphy CC, Cirillo PM, Krigbaum NY, Cohn BA. In utero exposure to 17α-hydroxyprogesterone caproate and risk of cancer in offspring. Am. J. Obstet. Gynecol. 2022 Jan;226(1):132.e1–132.e14.

[160] Letter from Rebecca Wood, Esq. dated June 9, 2022.

[161] Covis also states that there have been no reports of cancer or tumors from 17-OHPC in the FDA's Adverse Event Reporting System (FAERS) database. It is unclear how Covis arrived at this conclusion. CDER's Division of Pharmacovigilance conducted a search of the FAERS database on July 24, 2022 for reports involving HPC and cancer in the offspring of women who took HPC during pregnancy. A total of three potential cases of interest were identified.

[162] SMFM Statement: Use of 17-alpha hydroxyprogesterone caproate for prevention of recurrent preterm birth (July 2020), available at https://www.smfm.org/publications/280-smfm-statement-use-of-17-alpha-hydroxyprogesterone-caproate-for-prevention-of-recurrent-preterm-birth.

[163] Notification of Newly Identified Safety Signal to Covis Pharma Gmbh, June 9, 2022, Appendix 7.

[164] MAPP 4121.3, "Collaborative Identification, Evaluation, and Resolution of a Newly Identified Safety Signal (NISS)", available at https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/cder-manual-policies-procedures-mapp.

Failing to withdraw Makena would permit continued FDA approval of a drug that has, based on all available evidence, not been shown to be more effective than, but is riskier than, no treatment.

The sponsor argues that Trial 003 reaffirmed Makena's safety, and therefore the "the overall benefit-risk profile for Makena supports that the drug should remain on the market and available to patients, while further study is conducted on efficacy in higher-risk subpopulations."[165] Specifically, the sponsor claims that "the integrated dataset [of Trials 002 and 003] demonstrates *no difference* between Makena and placebo with respect to the specific risks identified by FDA,"[166] and asserts that "there should be no question that [Makena's] maternal and fetal safety has been confirmed in [Trial 003]."[167] CDER disagrees with the suggestion that there is no risk with Makena compared to placebo injection. The approved drug labeling for Makena lists warnings relating to, among other things, thromboembolic events, allergic reactions, decreased glucose tolerance, fluid retention, and depression,[168] and these adverse reactions are well-known with progestins.[169] After review of Trial 003 and considering the safety findings from Trial 002, CDER concluded, "Although the number of fetal and neonatal deaths are too low to draw definitive conclusions, … this safety outcome appear[s] to be similar between placebo and Makena. Otherwise, the safety profile of Makena remains unchanged." In other words, although Trials 002 and 003 did not reveal that Makena was associated with a specific safety risk with regard to fetal and neonatal deaths, the safety findings from Trial 003 also did not alter CDER's understanding of the other adverse reactions known to be associated with Makena, as described in its labeling. At a minimum, there are risks associated with Makena injection compared to no treatment (no injection) given injection-site adverse events such as swelling, pain, and infection. Further, although safety is carefully evaluated before a drug is approved, as the Murphy et al. study illustrates, sometimes safety issues emerge only after approval when the drug's use is expanded from the clinical trial setting to the general population, or after longer time horizons permit observation of longer-term or intergenerational effects. It is simply not possible to conclude that Makena, an intramuscular injection containing an active drug ingredient associated with known adverse reactions, administered repeatedly throughout pregnancy, would be as safe as no treatment at all.

---

[165] Sponsor's Response to NOOH at 5.

[166] *Id.* at 36.

[167] *Id.* at 40.

[168] *See* FDA-approved prescribing information for Makena, available at https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s013lbl.pdf and attached as Appendix 10.

[169] *See, e.g.*, FDA-approved labeling for Endometrium (NDA 022057) at https://www.accessdata.fda.gov/drugsatfda_docs/label/2008/022057s001lbl.pdf; *see also* FDA-approved labeling for Prometrium (NDA 020843) at https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/019781s017,020843s011lbl.pdf.

G. The Only Study That Could Provide Sufficient Evidence of Makena's Clinical Benefit is a Randomized, Double-Blind, Placebo-Controlled Trial

Randomized clinical trials are considered the gold standard for the evaluation of a drug effect, and the only randomized, double-blind, placebo-controlled trial to assess Makena's clinical benefit to neonates—Trial 003—failed to show that Makena conferred clinical benefit to neonates and failed to show that Makena reduced recurrent PTB. Particularly in light of Trial 003's failure to verify clinical benefit, only a prospective, randomized, double-blind, placebo-controlled clinical trial can provide sufficient evidence to establish substantial evidence of effectiveness of Makena.[170] At various timepoints following the 2019 BRUDAC, the sponsor proposed to conduct observational studies[171] because alternatives to an RCT have been discussed as a path to provide confirmatory evidence, including at the 2019 BRUDAC. CDER strongly disagrees that the non-RCT designs will be adequate to verify clinical benefit and establish the safety and effectiveness of Makena.[172,173]

Given that professional societies continue to recommend use of Makena, or at least recommend that providers consider the use of Makena, women in an untreated control group in a non-randomized trial or an observational study are likely to differ in characteristics that impact the risk of PTB. These characteristics could be known and measurable (such as mother's age, smoking history), known but unmeasurable (such as access to care, reasons for non-adherence), or completely unknown. The presence of these potential residual confounding factors can easily be an alternative explanation for any observed reduction in PTB. A non-randomized study might attribute an improvement in sPTB rate and neonatal outcome to Makena when the true reason for the difference was not due to medication, but was actually due to reasons unrelated to Makena,

---

[170] Despite Trial 003 already being underway with adequate enrollment from the U.S. and Canada, it took almost 10 years to complete Trial 003. Because enrollment in the U.S. became more challenging after Makena's approval, recruitment was expanded to Russia and Ukraine.

[171] *See, e.g.*, Letter from Rebecca Wood, Esq., dated March 30, 2022. In a meeting request dated February 19, 2020, the sponsor proposed to conduct two observational studies – a retrospective study (using "real world evidence from electronic health records") and a prospective study to "confirm clinical benefit and/or define the population in whom the drug may be useful." On March 11, 2020, CDER denied the meeting request as premature, because CDER was still considering the views expressed at the 2019 BRUDAC. On May 20, 2020, the sponsor submitted a meeting package containing the same proposal of two observational studies. CDER declined to meet with the sponsor given the inadequacies of the proposal to provide confirmatory evidence of Makena's effectiveness.

[172] Although FDA regulations (21 CFR 314.126(b)(2)(v)) contemplate the use of a historical control as one type of control in an adequate and well-controlled clinical trial that can provide substantial evidence of effectiveness, "historical control designs are usually reserved for special circumstances" such as for "diseases with high and predictable mortality" and "studies in which the effect of the drug is self-evident." These special circumstances in which we would consider historical control designs are absent with regard to Makena.

[173] The ICH E10 Choice of Control Group and Related Issues in Clinical Trials guidance Covis referenced states that "[a]n externally controlled trial should generally be considered only when prior belief in the superiority of the test therapy to all available alternatives is so strong that alternative designs appear unacceptable and the disease or condition to be treated has a well-documented, highly predictable course." Trial 003 failed to substantiate Makena's effect on reducing the risk of PTB seen in Trial 002 and failed to verify clinical benefit to the neonate. Therefore, there is not a prior belief in the superiority of Makena, let alone a prior belief so strong as to render alternative designs unacceptable.

including factors that we do not yet know contribute to PTB. Therefore, an observational study would not be adequate to reliably demonstrate benefit for Makena. Because sPTB and recurrent PTB are unpredictable and not well understood, it is impossible to identify ahead of time, and control for, all relevant confounding patient characteristics in the absence of a randomized, double-blind, placebo-controlled trial.

The sponsor appears to have changed its position from as recently as March 30, 2022,[174] and now agrees with CDER that another RCT is needed.[175] Although CDER disagrees with Covis that such a study could be conducted entirely or primarily in the U.S. while Makena remains approved, as explained in section V.H, and the design of such a study is beyond the scope of this hearing, we stand ready to engage with the sponsor on an additional RCT after the drug is withdrawn.

H. The Sponsor's Proposed Randomized, Placebo-Controlled Trial is Infeasible While Makena Remains Approved

The sponsor asserts FDA should maintain the approval of Makena while an additional randomized, placebo-controlled trial is undertaken to verify the benefit of Makena. The sponsor recently proposed a future randomized, placebo-controlled trial in "high risk" subjects, which it defines in part as patients who have had a prior sPTB at a gestational age < 34 weeks, to verify and describe the efficacy of Makena in reducing PTB and improving neonatal outcomes. Further, the sponsor plans for this trial to be conducted completely or largely in the U.S. because, although CDER does not agree, the sponsor contends that Trial 003's results are not generalizable to U.S. women because approximately 75% of the study population were non-U.S. and are "lower risk" than those in Trial 002. Despite acknowledging the significant recruitment challenges for Trial 003 in the U.S. after Makena's 2011 approval, Covis argues that it would be easier to enroll another randomized, placebo-controlled trial in the U.S. if Makena remains FDA-approved—for an indication that would be under study—than if approval were withdrawn. This argument is without merit.

First, from a practical perspective, CDER anticipates that the sponsor would likely face the same or even greater challenges in enrolling subjects, particularly those at "highest" risk for recurrent PTB, as the sponsor proposed, for another randomized, placebo-controlled trial than it did for Trial 003. The sponsor asserts that recent questions about Makena's efficacy may facilitate the recruitment and conduct of the new trial. CDER disagrees. If Makena remains approved, it would be considered safe and effective for its approved use and may continue to be the standard

[174] *See, e.g.,* Letter from Rebecca Wood, Esq., dated March 30, 2022, at 2 (Stating "…Covis has spent significant time and resources developing its proposal for additional studies that can be conducted to confirm Makena's clinical benefit and to further identify populations that would benefit most from its treatment. The proposal includes RWE-based studies (including a prospective external control study) and discusses under what circumstances these approaches may supplement or, if needed, replace randomized controlled trials (RCT)." (Emphasis Added).

[175] *See* Sponsor's July 17, 2022, Preliminary Briefing Materials, at 7, 87–95; Sponsor's Appendix to Preliminary Briefing Materials, at 1–19 (introducing an RCT protocol in these preliminary briefing materials).

of care. It is extraordinarily unlikely that any appreciable number of women at risk for recurrent PTB would choose to enroll in an RCT and risk receiving a placebo so long as they can simply receive Makena, a product approved by FDA for over a decade, by *not* enrolling in such a trial. Similarly, it seems unlikely that most providers would recommend that patients at risk for recurrent PTB, particularly patients at high risk, chance not receiving the only FDA-approved treatment option by enrolling in a placebo-controlled RCT over having a prescription for Makena in hand. Even with information regarding the negative results of Trial 003, providers regard FDA approval as indicating that the data (which is not limited to Trial 003 here) demonstrate substantial evidence of effectiveness. It is a tall order to ask them to recommend patients enroll in a placebo-controlled trial despite continued FDA approval of Makena.

Second, the survey results Covis uses to support its position are deeply flawed. The sponsor presented survey findings from healthcare providers and women in whom Makena may be indicated to support its position that there is greater willingness to participate in a randomized, placebo-controlled trial if Makena remains approved. But the survey findings do not support this position. For example, in Q19, providers were asked, "How likely are you to recommend a pregnant patient enroll in a placebo-controlled study comparing the efficacy of a product vs placebo when the product has been approved by FDA?" The results indicate that 78% of providers responded "very likely" or "somewhat likely." However, this question does not specify that, in the hypothetical scenario relevant to Makena, the FDA-approved product would be approved for the indication being studied, or approved for one or more other indications. Without this specificity, we do not know which of these two *very* different scenarios the responders considered when answering this question. Further, only 41% of providers surveyed stated that they have ever had a role in advising, guiding, or recommending patient participation in a clinical trial. This finding calls into question whether many providers responding to Q19 have considered the chance their patients could be randomized to receive a placebo. Due to the limitations of Q19, CDER questions the validity of Covis' assertion that a large majority of providers would recommend that a pregnant patient enroll in a trial in which the patient may receive a placebo instead of a product that has been approved by FDA to treat the patient's condition. This assertion is particularly weak given the length of time it took to conduct Trial 003 and how recruitment rates changed after Makena was approved in the U.S. Particularly in light of the difficulties in recruiting U.S. patients to Trial 003 after Makena was approved by FDA, it is doubtful that recruiting would be easier for another randomized, placebo-controlled trial while Makena remains on the market.

Covis' assertions regarding patient willingness to enroll in a randomized, placebo-controlled trial for an FDA-approved product are also not supported by the survey results. The survey asked patients, "How likely would you be to take a prescription drug during pregnancy that is intended to treat preterm birth and that is" (1) "Approved by the FDA and recommended by your health care provider," (2) "Being studied by researchers and not approved by the FDA," or (3) "Being studied by researchers but already approved by the FDA." Although patients indicated that they would be more likely to take a drug being studied and already approved by FDA, as opposed to a drug being studied and not approved by FDA, this question did not gather any information about

patient willingness to *participate* in such study, with the chance that they could receive a placebo. Contrary to the sponsor's claim, the survey results do not provide persuasive evidence that it would be easier to recruit patients for a randomized, placebo-controlled trial of Makena while it remains FDA-approved than if it were removed from the market.

Given that Trial 003 was not completed for nearly a decade, despite careful planning for recruitment challenges in the U.S. after Makena's approval, it is unlikely that another randomized, placebo-controlled trial could be completed in a timely manner even if Makena's approval is withdrawn. Trial 003 enrolled 1,700 subjects from 93 clinical sites in and outside the U.S. Although recruitment took place largely outside of the U.S., in countries where Makena was not on the market, it still took almost 10 years to conduct. The sponsor's proposed RCT has an estimated sample size ranging from approximately 1,200 to 3,200 subjects. It does not follow that another, potentially much larger RCT, would be completed in less time than Trial 003. Also, CDER is concerned that the sample size estimates for a future RCT rely on point estimates that are out of date. Any future trial could have equivocal, if not negative, results, and 10 to 15 years from now we may be in the same position that we are in today, except that Makena would likely have been marketed for more than two decades without demonstration of clinical benefit for its approved use.[176]

Lastly, even if a randomized, placebo-controlled trial were to be conducted, it is unclear if it would be appropriate to focus on Covis' new target population of women with a prior sPTB < 34 weeks. While a prior sPTB < 34 weeks may predict a higher risk of a subsequent PTB, the evidence does not show that this risk factor confers a differential response to Makena.

I. Makena Should Not Remain on the Market While Further Studies Are Conducted

Even if another RCT that could potentially establish Makena's effectiveness in light of the failure of Trial 003 were to be initiated, the prospect of such a study should not delay withdrawal of Makena.

First, as explained above, the legal standard for withdrawal is met. Makena's confirmatory trial failed to verify clinical benefit or even show an effect on the gestational age endpoint that was the basis of accelerated approval. The available evidence fails to show that Makena is effective for its approved use in its indicated population (or any identified subset of that population), and Makena carries risks, some of which are serious. These risks are unacceptable in light of the lack of substantial evidence that Makena is effective. Second, as noted above, a randomized, double-blind, placebo-controlled study is not feasible, at least in the United States, while Makena is the only approved therapeutic agent for prevention of recurrent PTB.

[176] The sponsor suggests that it could use data from observational studies to support or refute the benefit of Makena prior to the conclusion of a randomized, placebo-controlled trial. This is not an appropriate option. As discussed above in section V.G., the many unknowns of preterm birth and the limitations of observational studies preclude CDER's reliance on the findings of observational studies to determine Makena's efficacy.

Furthermore, withdrawal is needed to maintain the integrity of the accelerated approval pathway. As noted above, based on all of the evidence known to CDER, there is a lack of substantial evidence that the drug is effective for its approved indication. It could take a decade or more to complete and submit results of another randomized, double-blind, placebo-controlled trial that could potentially be sufficient to provide evidence of Makena's effectiveness, particularly in the face of the negative result from Trial 003. Allowing the continued marketing of Makena would mean that patients would be unnecessarily exposed to the risks of this drug without any demonstration of clinical benefit. Further, use of drugs that have not been shown to be effective incurs false hopes and other burdens, such as unnecessary utilization of health care resources. If CDER cannot withdraw drugs under accelerated approval when the grounds for withdrawal are satisfied—and CDER has determined that the drug should be withdrawn because, among other things, the prospect for demonstrating effectiveness is remote, at best, well into the second decade of the drug's approval—this would frustrate Congress' purpose in establishing the accelerated approval pathway, which provides for both earlier approval and expedited withdrawal of certain drugs.

Covis makes several additional arguments in support of retaining Makena on the market while additional studies are conducted. We address those arguments here.

1. That Makena Is Currently the Standard of Care for Prevention of Recurrent Preterm Birth Does Not Weigh in Favor of Retaining Its Approval

Covis states that the medical and patient community continues to support HPC as a treatment option[177] and that withdrawal of Makena would leave an at-risk patient population without a treatment option.[178] Covis further asserts that withdrawing Makena's approval would fail to respect clinician and patient choice and would deprive patients of the option to choose to use Makena while it continues to be studied.

Although CDER carefully considers the views of all stakeholders, it is imperative that its drug approval and withdrawal decisions be driven by the available data on safety and effectiveness. Based on that evidence, Makena's confirmatory trial failed to verify clinical benefit, Makena is no longer shown to be effective for its indicated population or any identified subset of that population, and Makena has an unfavorable benefit-risk profile. For those and other reasons described above, Makena should be withdrawn.

Furthermore, that Makena is considered the standard of care for reducing the risk of recurrent PTB is likely highly influenced by its status as the only therapeutic agent approved for this use. If CDER determines that the available evidence for a drug no longer supports its continued approval, it would be inappropriate to withhold withdrawal because the drug is the standard of care. Moreover, it is especially important for FDA to take action based on evidence that emerged after the drug became the standard of care in circumstances in which, as here, CDER

[177] Sponsor's Response to NOOH at 40.

[178] *Id.* at 44.

has determined that the drug is no longer shown to be effective and the benefit-risk balance is unfavorable. Withdrawal of Makena would make it clear that there is not substantial evidence of effectiveness for its currently approved use, which could help change the current standard of care to align with the best available evidence. Moreover, withdrawal of approval could help facilitate recruitment in a new randomized, double-blind, placebo-controlled trial of Makena or other potential therapies to treat this condition.

Finally, FDA has a duty to ensure FDA-approved drugs are in fact safe and effective. Physicians and patients rely on FDA to ensure FDA-approved drugs are effective. Maintaining Makena's approval in an effort to preserve choices, despite the negative results of Trial 003 would be a disservice to both groups and give the false impression that Makena is shown to be safe and effective for its labeled indication.

2. The Continued Availability of HPC If Makena Were Withdrawn Is Not a Basis to Conclude that Makena Should Remain on the Market

Covis argues that other forms of HPC that would remain on the market should Makena be withdrawn are inadequate alternatives to Makena, and that leaving prescribers and patients with only these options would have "significant negative public health consequences."[179, 180] The sponsor further asserts that "[o]ff-label use of generic Delalutin [(which contains the same active ingredient in the same strength as Makena)] would necessarily be less safe for patients than using Makena according to its labeled uses" because the labeling of Delalutin "does not contain any of the information necessary to enable prescribers to use the drug safely and for the purposes for which it is intended."[181]

CDER acknowledges that Makena is the only therapy currently approved to prevent recurrent PTB. However, CDER disagrees that Makena should remain on the market because other HPC-containing drug products could be prescribed off-label for this indication. This is not a basis under the statute to maintain the approval of a drug that is no longer shown to be effective. Further, any predictions by Covis that there would be widespread off-label prescribing of generic

---

[179] *Id.*

[180] Prior to Makena's approval under the accelerated approval pathway, a drug product containing the same active ingredient in Makena, HPC, 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 solely on safety considerations because it occurred prior to the 1962 Kefauver-Harris Amendments to the FD&C Act, which required substantial evidence of effectiveness, in addition to demonstrated safety, for FDA approval. 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. In the 1990s, the application holder for the Delalutin NDAs discontinued marketing this product and, in September 1999, requested that FDA withdraw its approval. 65 Fed. Reg. 55264 (Sept. 13, 2000). At the time of withdrawal, which was determined not to be for reasons of safety or effectiveness, Delalutin was indicated for several gynecological uses in non-pregnant women but had no approved obstetrical indications. 75 Fed. Reg. 36419 (Jun. 25, 2010).

[181] Sponsor's Response to NOOH at 50 (internal quotations omitted). We note that Delalutin's marketing status would be unaffected by Makena's withdrawal. Although it contains the same amount of the same active ingredient as Makena, its approved indications do not overlap with Makena's indication.

versions of Delalutin for Makena's indication if Makena is withdrawn are speculative. Providers prescribe drugs off-label when they believe an approved drug is effective for the unapproved indication. Maintaining the approval of Makena would suggest that FDA continues to believe it is effective for reducing the risk of recurrent PTB. In contrast, withdrawing Makena due to lack of evidence of effectiveness would send a clear message to providers, which may reduce off-label prescribing of HPC-containing drug products for Makena's currently-approved indication.

With respect to Covis' argument regarding availability of Makena's labeling for those practitioners choosing to prescribe off-label, we note that such practitioners would have access to the same sources of information that they otherwise utilize when making decisions to prescribe off-label, such as practice guidelines and published literature. In sum, irrespective of access to other HPC products, the available evidence does not show that Makena is effective for its approved use, and its benefit-risk profile is unfavorable. Makena satisfies the statutory criteria for withdrawal, and it should be withdrawn.

Covis further argues that "compounding HPC will be unlawful" if Makena is withdrawn, and that compounded HPC would increase safety risks to patients compared to Makena because of compounding-specific risks.[182] CDER does not agree that FDA should refrain from withdrawing approval of Makena in an effort to reduce compounding of HPC. HPC, like thousands of other drugs marketed in the U.S., may be eligible for compounding provided specific conditions that Congress set forth in the FD&C Act are met.[183] But the prospect that a drug with the same active ingredient may be compounded is not an appropriate basis to conclude that a drug that is no longer shown to be effective should remain approved. Such a position would abrogate the statutory standard for FDA approval to be supported by substantial evidence of effectiveness. Moreover, withdrawing Makena's approval due to a lack of evidence of effectiveness may reduce compounding of HPC for Makena's currently-approved indication, as it would send a clear message to providers that Makena's risk-benefit profile is unfavorable.

3. Withdrawing Makena Is Consistent with Precedent

Covis asserts that the Agency has chosen not to withdraw approval of "numerous" drug products where confirmatory studies did not verify clinical benefit.[184] However, Covis' examples of other situations in which CDER did not pursue withdrawal of drugs approved under accelerated approval are inapposite, and they do not change the conclusion that withdrawal is warranted here.

---

[182] Sponsor's Response to NOOH at 45–48.

[183] Under sections 503A and 503B of the FD&C Act, drug products can be compounded only if they meet certain conditions. These conditions include, among others, restrictions on compounding drug products that appear on a list developed by regulation of drug products that have been withdrawn or removed from the market for reasons of safety or effectiveness, or that are essentially copies of an approved or commercially available drug product.

[184] Sponsor's Response to NOOH at 54–59. Covis specifically discusses ProAmatine (midodrine), Iressa (gefitinib), and Iclusig (ponatinib) as examples in its response to the NOOH.

CDER agrees with Covis that the "withdrawal authority is discretionary"[185] and that CDER possesses various regulatory options when a confirmatory trial fails to verify clinical benefit. When considering those options, CDER evaluates the available information about the particular drug at issue and the trial results to determine the appropriate regulatory approach, consistent with the law and the science, to achieve its public health mission. Thus, as stated in the Decision of the Commissioner regarding the Proposal to Withdraw Approval for the Breast Cancer Indication for AVASTIN (Bevacizumab), each decision to withdraw or not withdraw the accelerated approval of a product is made on its own merits.[186] The facts and science underlying FDA's regulatory decisions regarding withdrawal of a specific drug are specific to that drug.

In the case of Makena, CDER carefully evaluated the data and science underlying Makena's negative efficacy results and concluded that the standard for withdrawal is met—Trial 003 failed to verify clinical benefit, and Makena is no longer shown to be effective for its approved indication. Moreover, the failure of Trial 003 to show either clinical benefit or an effect on the endpoint that was the basis of accelerated approval makes Makena's situation highly unusual, if not unique. None of the examples cited by Covis of drug products approved under accelerated approval for which CDER did not pursue withdrawal,[187] including that of ProAmatine (midodrine), involved a confirmatory trial that failed to demonstrate an effect on the surrogate or intermediate clinical endpoint that was the basis for the accelerated approval.[188] In addition, as discussed above, as a policy matter, CDER believes there is no public health justification for maintaining Makena's approval, especially in light of the length of time that would likely be needed to generate efficacy data sufficient to overcome the negative result of Trial 003 and the infeasibility of conducting another placebo-controlled trial in U.S. women while Makena remains on the market.

In addition to the failure of Trial 003 to demonstrate an effect on the endpoint that served as the basis for accelerated approval, Makena is distinguishable in other ways from examples Covis describes. For instance, CDER did not initially propose to withdraw approval for Iressa after a confirmatory trial failed to verify clinical benefit in part because the trial produced clear

---

[185] *Id*. at 3.

[186] *See* Decision of the Commissioner regarding the Proposal to Withdraw Approval for the Breast Cancer Indication for AVASTIN (Bevacizumab) citing, *Edison Pharm. Co., Inc. v. Food and Drug Admin*., 600 F.2d 831, 843 (D.C. Cir. 1979) (An applicant's "failure to meet the specific statutory requirements governing [new drug application] approval cannot be excused on the basis of prior [FDA] action with regard to another drug.") at 59.

[187] *Supra* note 18490.

[188] *See, e.g*., Thatcher N, et al., Gefitinib Plus Best Supportive Care In Previously Treated Patients With Refractory Advanced Non-Small-Cell Lung Cancer: Results From A Randomised, Placebo-Controlled, Multicentre Study (Iressa Survival Evaluation in Lung Cancer). The Lancet. 2005; 366 (9496): 1527-1537, https://pubmed.ncbi.nlm.nih.gov/16257339/ (discussing gefitinib); Kim ES, et al. Gefitinib Versus Docetaxel In Previously Treated Non-Small-Cell Lung Cancer (INTEREST): A Randomised Phase III Trial. The Lancet. 2008;372 (9652):1809-1818, https://pubmed.ncbi.nlm.nih.gov/19027483/ (discussing gefitinib); Pulte ED, Chen H, Price LSL, Gudi R, Li H, Okusanya OO, Ma L, Rodriguez L, Vallejo J, Norsworthy KJ, de Claro RA, Theoret MR, Pazdur R. FDA Approval Summary: Revised Indication and Dosing Regimen for Ponatinib Based on the Results of the OPTIC Trial. Oncologist. 2022 Mar 4;27(2):149-157, https://pubmed.ncbi.nlm.nih.gov/35641211/ (discussing ponatinib).

evidence that the drug had an effect in some subjects. Here, there is no compelling evidence that Makena has an effect for any identified subgroup of subjects. When additional studies of Iressa (gefitinib) failed to verify clinical benefit, the sponsor voluntarily withdrew Iressa from the market.[189] Notably, after the product was withdrawn, the sponsor conducted further trials which demonstrated that the subgroup of subjects that benefited from Iressa contained a genetic mutation in their tumor, and thus Iressa was later approved for just this biomarker-selected population.[190] As to Iclusig (ponatinib), CDER did not propose to withdraw approval when the sponsor voluntarily suspended marketing in part because the confirmatory trial was ongoing. Although CDER conducted another benefit-risk assessment after the confirmatory trial raised safety concerns, it concluded that the drug's benefits outweighed the safety concerns if those concerns were addressed through specific safety measures.[191] And when the confirmatory trial was completed, it verified the clinical benefit of Iclusig.[192] For Makena, the evidence does not demonstrate that the drug provides any benefit to subjects. In the absence of benefit, the benefit-risk assessment is unfavorable, because Makena, like all drugs, carries risks.

Finally, Covis' suggestion that withdrawal of a drug approved under accelerated approval is rare[193] ignores that many drugs are voluntarily withdrawn when a confirmatory trial fails to verify clinical benefit. There are numerous examples.[194] And when the sponsor of Avastin

---

[189] 77 Fed. Reg. 24723 (Apr. 25, 2012), https://www.govinfo.gov/content/pkg/FR-2012-04-25/pdf/2012-9944.pdf.

[190] CDER, NDA 206995, Division Director Summary Review (July 13, 2015), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/206995Orig1s000SumR.pdf.

[191] *See* FDA Drug Safety Communication: FDA requires multiple new safety measures for leukemia drug Iclusig; company expected to resume marketing (Dec. 20, 2013), https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-fda-requires-multiple-new-safety-measures-leukemia-drug-iclusig.

[192] Pulte ED, Chen H, Price LSL, Gudi R, Li H, Okusanya OO, Ma L, Rodriguez L, Vallejo J, Norsworthy KJ, de Claro RA, Theoret MR, Pazdur R. FDA Approval Summary: Revised Indication and Dosing Regimen for Ponatinib Based on the Results of the OPTIC Trial. Oncologist. 2022 Mar 4 ;27(2):149–157, https://pubmed.ncbi.nlm.nih.gov/35641211/.

[193] *See* Sponsor's Response to NOOH at 55. Covis asserts that, under the accelerated approval program, "CDER has approved more than 240 drugs/indications, and sought to withdraw fewer than 10 drugs/indications after the failure of confirmatory studies to verify and describe the predicted clinical benefit of the drug." Sponsor's Response to NOOH at 55. As of June 30, 2022, CDER has granted accelerated approval to 282 drug products. Of those 282, clinical benefit has been confirmed for 140 products and those applications have been converted to traditional approval. Of the remaining products, 34 have been withdrawn from the market, and 108 are awaiting confirmation of clinical benefit. Of the products awaiting confirmation of clinical benefit, 73 were approved after June 30, 2019. In other words, 73 of 108 products pending confirmation of clinical benefit were approved within the past three years.

[194] For instance, in addition to the Iressa example described above, Lartruvo (olaratumab) was voluntarily removed from the market after the confirmatory trial failed to verify clinical benefit. 85 Fed. Reg. 43587 (July 17, 2020), https://www.govinfo.gov/content/pkg/FR-2020-07-17/pdf/2020-15516.pdf. The same is true for certain indications of Ethyol (amifostine), Synercid (quinupristin, dalfopristin), Keytruda (pembrolizumab), Tecentriq (atezolizumab), and Opdivo (nivolumab). *See* NDA 20-221/S-020, Ethyol (amifostine) for Injection, Supplement Approval (Mar. 28, 2006), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2006/020221s020ltr.pdf; NDA 50-747/S-011, Synercid I.V. (quinupristin and dalfopristin for injection), Supplement Approval (Nov. 12, 2010), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2010/050747s011,050748s009,050748s010ltr.pdf; BLA 125514/S-106, Keytruda (pembrolizumab) injection, Supplement Approval/Release From Postmarketing Requirement (Mar. 30, 2021),

declined to voluntarily withdraw an indication from the market after the confirmatory trial failed to verify clinical benefit, CDER proposed that it be withdrawn by FDA. The indication was subsequently withdrawn. Here, when the sponsor declined to voluntarily withdraw Makena from the market after the confirmatory trial failed not only to verify clinical benefit, but also to show effect on the surrogate endpoint that served as the basis for accelerated approval, CDER proposed that Makena be withdrawn by FDA. It is appropriate, and in the best interest of public health, for FDA to withdraw Makena.

4. FDA Cannot Narrow Makena's Labeling to Women with "High Risk" Pregnancies Without Substantial Evidence that Makena Benefits Them

Covis argues that FDA should consider narrowing the drug's labeling to "high risk" pregnancies.[195] Covis, though, fails to provide sufficient support to justify the proposed indication. First, as discussed in Section V.C, there is not substantial evidence of effectiveness to support a narrowed indication in any identified subgroup of Makena's indicated patient population, including "high risk" pregnancies. Further, if such a narrowing is sought, future RCTs would need to provide evidence that clearly demonstrates benefit in a well-defined population and/or pre-defined subpopulation. Covis has not identified a patient population for which Makena shows consistent efficacy, nor has it explained how to define the population that should not use Makena. For example, if women in the U.S. are considered higher risk, the trial data should show a positive trend (i.e., consistent findings of Makena's effect in the U.S. across studies) in such women and in their neonates. Instead, across all the endpoints, no trend exists that favors Makena. Rather, the trend of the data from high-risk groups in clinical investigations published since Trial 002 do not support use of Makena.

5. Leaving Makena on the Market Exacerbates Health Disparities

Covis argues that withdrawal of Makena would deepen existing health inequities, particularly for Black women, given the lack of treatment alternatives for reducing the risk of preterm birth. Covis further suggests that withdrawing Makena could disincentivize research and development in maternal and infant health drug development. In fact, Covis has this backwards.

---

https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2021/125514Orig1s106ltr.pdf; BLA 125514/S-125, Keytruda (pembrolizumab) injection, Supplement Approval/Release From Postmarketing Requirement (Feb. 4, 2022), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2022/125514Orig1s125ltr.pdf; BLA 761034/S-041, Tecentriq (atezolizumab), Supplement Approval (Apr. 13, 2021), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2021/761034Orig1s041ltr.pdf; BLA 761034/S-044, Tecentriq (atezolizumab) injection, Supplement Approval/Fulfillment of Postmarketing Requirement (Oct. 6, 2021), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2021/761034Orig1s044ltr.pdf; BLA 125554/S-095, Opdivo (nivolumab) injection, Supplement Approval/Release From Postmarketing Requirement (Dec. 29, 2020), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2020/125554Orig1s095ltr.pdf; BLA 125554/S-107, Opdivo (nivolumab) injection, Supplement Approval (July 23, 2021), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2021/125554Orig1s107ltr.pdf.

[195] Sponsor's Response to NOOH at 16, 95.

Failing to withdraw Makena from the market, despite the fact that the available evidence does not show Makena to be effective for any group of women, including those at greatest risk of recurrent PTB, would disregard the burdens and risks associated with treatment with Makena. Moreover, as explained above, it would likely thwart the ability of clinical researchers to enroll U.S. women in a placebo-controlled trial, preventing the development of evidence that could potentially overcome the negative results of Trial 003 and potentially show Makena to be effective. Finally, preserving Makena as an option despite the lack of evidence that it works is likely to hinder the development of other drugs that may be shown to be effective for Makena's indication, both because of trial enrollment challenges and due to the uncertainties of how to design, conduct, and interpret trials of new products while Makena remains approved.

Preterm birth is a significant public health concern, and reducing it is a national public health priority. Black women are at most risk for PTB compared to all other racial and ethnic groups in the U.S. According to the CDC, in 2019 through 2021, more than 12% of singleton births to non-Hispanic Black women were preterm compared to 7% of singleton births to non-Hispanic White women.[196] The disparities do not change when one looks at overall preterm birth rates (singleton and multiple gestations).

Unfortunately, there is not substantial evidence that Makena reduces recurrent PTB in Black women or any other identified subgroup. Since Makena was approved in 2011 based on Trial 002, Makena has been studied in several populations with major identified health disparities. Aside from Trial 002, none of these studies show a statistically significant positive treatment effect, or even a trend of positive point estimates across studies, for Makena. Although Trial 002 had a demonstrated treatment effect in favor of Makena when the drug was approved in 2011, this treatment effect has not been corroborated by any of the considerable evidence gathered since accelerated approval was granted. Although there are many social determinants of health and other factors tied to health disparities that impact the risk of PTB, of those that we analyzed, as described in Section V.C., none are associated with a consistent treatment effect of Makena across Trials 002 and 003, nor are they apparent in the additional evidence. The effect seen in Trial 002, a trial in which 60% of subjects were Black, was not seen in the subgroup of Black subjects in Trial 003. Furthermore, the effect was not seen in the Price et al. study, which was conducted in Zambia, or in the Rouse et al. study, in which the trial arms had 22% and 24% Black subjects. Notwithstanding the limitations of observational studies, the effect was also not seen in the Massa et al. study, in which 66.2% of subjects were Black. There was also no treatment effect demonstrated in the publications describing studies from safety net hospital systems or using data from Medicaid-recipient populations.

Treatment with Makena adds burden to pregnant women, and those burdens are heaviest for those with historic and current unequal distribution of social, political, economic, and environmental resources. Treatment with Makena is not without risk. The current evidence does

[196] Martin JA, Osterman MJK. Exploring the decline in the singleton preterm birth rate in the United States, 2019–2020. NCHS Data Brief, no 430, Figure 2. Hyattsville, MD: National Center for Health Statistics. 2022, available at https://www.cdc.gov/nchs/data/databriefs/db430.pdf (last visited Sept. 14, 2022)

not show that Makena leads to the prevention of recurrent PTB or improved health for the neonate. It is incorrect to conclude that withdrawal of Makena would deepen existing health inequities, particularly for Black women. And narrowing the indication to "high-risk" pregnancies would amount to a false representation of Makena's effectiveness for women at greatest risk of PTB. It would also make it much more challenging, if not infeasible, to study either Makena or other promising treatments in the populations that most need an effective drug.

VI. CONCLUSION

For the reasons discussed above, Makena should be withdrawn from the market. CDER recognizes the enormous public health challenges we face with PTB, the hardship faced by the affected children and their families, and the need for effective treatment. When CDER approved Makena under accelerated approval, we took a risk to make it available earlier in the hope that it would ultimately be confirmed to be effective in reducing neonatal morbidity and mortality. An important counterbalance to the earlier approval that this pathway affords is the availability of expedited withdrawal when the statutory and regulatory criteria are satisfied. Trial 003 did not verify Makena's previously expected benefit, and considering all of the evidence available today, the clear scientific conclusion is that Makena is no longer shown to be effective for its currently approved use for either its indicated population or any subset thereof. Thus, two independent statutory and regulatory criteria for withdrawal are met. Makena's benefit-risk profile is unfavorable. Allowing it to remain on the market in the hope that it may eventually be shown to be effective would undermine the integrity of the accelerated approval pathway and unnecessarily expose patients to the risks associated with a drug that is not shown to be effective.

Appendices

1. Advisory Committee on Reproductive Health Drugs (ACRHD) Meeting on Gestiva[197] (August 26, 2006)

- Federal Register Notice
- Agenda
- Advisory Committee for Reproductive Health Drugs Roster
- Adeza Briefing Document
- Adeza Addendum
- Adeza Bibliography of Appendices
- Adeza Slides
- FDA Backgrounder
- FDA Appendix
- FDA Bibliography
- FDA Slides Gillen
- FDA Slides Monroe
- FDA Slides Romero
- FDA Slides Wesley
- Transcript
- Transcript Index
- Summary of Meeting Minutes

2. BRUDAC Meeting on Makena (October 29, 2019)

- AMAG Backgrounder
- FDA Backgrounder
- BRUDAC Committee Roster
- BRUDAC Final Agenda
- BRUDAC Final Meeting Roster

[197] The 2006 ACRHD meeting used the then-proposed name "Gestiva" for the drug product that was ultimately approved and marketed as Makena.

- BRUDAC Final Questions
- Minutes
- AMAG Slides
- FDA Slides
- Transcript
- Webcast Recording

3. CDER's Review of Trial 003, CDER's Statistical Review and Evaluation, NDA 021945/S-023, Makena (June 30, 2020)

4. CDER's Decisional Memorandum, NDA 021945 Makena (hydroxyprogesterone caproate) (October 5, 2020)

5. Addendum to CDER Decisional Memorandum, NDA 021945 Makena (hydroxyprogesterone caproate), EPPPIC (January 14, 2022)

6. Subgroup Figures Including Shrinkage Using 6,000,000 and 60,000 Iterations

7. Notification of Newly Identified Safety Signal to Covis Pharma GmbH (June 9, 2022)

8. Decision of the Commissioner, Proposal to Withdrawal Approval for the Breast Cancer Indication for AVASTIN (Bevacizumab) (November 18, 2011)

9. CDER's Review of Trial 003, CDER's Clinical Review, NDA 021945/S-023, Makena (October 5, 2020)

10. FDA-approved Prescribing Information for Makena

Selected References

Adams MM, Elam-Evans LD, Wilson HG, Gilbertz DA. Rates of and factors associated with recurrence of preterm delivery. JAMA. 2000;283(12):1591–1596.

Bastek JA, Adamczak JE, Hoffman S, Elovitz MA, Srinivas SK. Trends in prematurity: what do changes at an urban institution suggest about the public health impact of 17-alpha hydroxyprogesterone caproate? Matern. Child Health J. 2012;16:564–568.

Blackwell SC, et al. 17-OHPC to prevent recurrent preterm birth in singleton gestations (PROLONG study): a multicenter, international, randomized double-blind trial. Am. J. Perinatol. 2020;37(02):127–136.

Caritis SN, et al. Prevention of preterm birth in triplets using 17 alpha hydroxyprogesterone caproate a randomized controlled trial. Am. J. Obstet. Gynecol. 2009;113(2):285–292.

Caritis SN, et al. Relationship between 17-alpha hydroxyprogesterone caproate concentration and spontaneous preterm birth, Am. J. Obstet. Gyencol. 2014; 210(2):128.e1–6.

Carter EB, Cahill AG, Olsen MA, Macones GA, Tuuli MG, Stout MJ. Practical considerations with 17-hydroxyprogesterone caproate for preterm birth prevention: does timing of initiation and compliance matter? J. Perinatol. 2019;39(9):1182–1189.

Chang C, et al., Withdrawing approval of Makena—A proposal from the FDA Center for Drug Evaluation and Research, N. Engl. J. Med. 2020.

Goldenberg RL, Mercer BM, Meis PJ, Copper RL, Das A, McNellis D. The preterm prediction study: fetal fibronectin testing and spontaneous preterm birth. NICHD Maternal Fetal Medicine Units Network. Obstet. Gynecol. 1996;87:643–48.

Hakim J, Zhou A, Hernandez-Diaz S, Hart J, Blair J. Wylie B, Beam A, Effectiveness of 17-OHP for prevention of recurrent preterm birth: a retrospective cohort study, Am. J. Perinatol. 2021.

Manuck TA, Stoddard GJ, Fry RC, et al. Nonresponse to 17-OHPC for recurrent spontaneous preterm birth prevention: clinical prediction and generation of a risk scoring system. Am. J. Obstet. Gynecol. 2016;215:622.e1–8.

Massa K, Childress K, Vricella L, et al. Pregnancy duration with use of 17-a-hydroxyprogesterone caproate in a retrospective cohort at high risk of recurrent preterm birth. Am. J. Obstet. Gynecol. MFM. 2020; 2(4):100219.

Meis PJ, Klebanoff M, Thom E, Dombrowski MP, Sibai B, et al. Prevention of recurrent preterm delivery by 17-alpha-hydroxyprogesterone caproate. N. Engl. J. Med. 2003;348:2379–85.

Mercer BM, 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. 1999;181(5 Pt 1):1216.

Murphy CC, et al. In utero exposure to 17α-hydroxyprogesterone caproate and risk of cancer in offspring. Am. J. Obstet. Gynecol. 2022 Jan;226(1):132.e1–132.e14.

Nelson DB, McIntire DD, McDonald J, Gard J, Turrichi P, Leveno KJ. 17-alpha hydroxyprogesterone caproate did not reduce the rate of recurrent preterm birth in a prospective cohort study. Am. J. Obstet. Gynecol. 2017;216:600.e1–9.

Price JT, et al. Weekly 17 alpha-hydroxyprogesterone caproate to prevent preterm birth among women living with HIV: a randomised, double-blind, placebo-controlled trial, The Lancet HIV. 2021;8(10):e605–e613.

Rouse DJ, et al. A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. N. Engl. J. Med. 2007;357:454–461.

Wang X, Garcia S, Kellom K, Boelig R, Matone M, Eligibility, utilization, and effectiveness of 17-alpha hydroxyprogesterone caproate (17OHPC) in a statewide population-based cohort of medicaid enrollees. Am. J. Perinatol. 2021.