Covis Pharma GmbH's Briefing Materials
In Response To The Center for Drug Research and Evaluation's
Notice Of Opportunity For A Hearing
And Proposal To Withdraw Approval Of
MAKENA® (hydroxyprogesterone caproate injection)
250 mg/mL (NDA No.: NDA 21-945)

Docket No. FDA-2020-N-2029

Francesco Tallarico
General Counsel &
Head of Government Affairs & Policy

Aziza Johnson
Executive Vice President
Regulatory & Scientific Affairs

A. Raghav Chari, Ph.D.
Chief Innovation Officer

COVIS PHARMA GMBH

Rebecca K. Wood
Sean C. Griffin
Emily Marden
Kelly Cho
Shamelle Richards

SIDLEY AUSTIN LLP
1501 K Street, NW
Washington, DC 20005

Counsel for Covis Pharma GmbH

September 16, 2022

TABLE OF CONTENTS

ATTACHMENT D TO THE APPENDIX SURVEY OF MFMU NETWORK SITES 41

TABLES

FIGURES

LIST OF ABBREVIATIONS AND TERMS

17-OHPC	17 α-hydroxyprogesterone caproate
ACOG	American College of Obstetricians and Gynecologists
ACRHD	Advisory Committee for Reproductive Health Drugs
AE	Adverse Event
AHR	Adjusted Hazard Ratio
ANDA	Abbreviated New Drug Application
ARR	Absolute Risk Reduction
ASQ	Ages and Stages Questionnaires
BMI	Body Mass Index
BRUDAC	Bone, Reproductive and Urologic Drugs Advisory Committee
CCR	California Cancer Registry
CDER	Center for Drug Evaluation and Research
CFR	Code of Federal Regulations
CGMP	Current Good Manufacturing Practice
CI	Confidence Interval
CRL	Crown-Rump Length
EHR	Electronic Health Record
EPPPIC	Evaluating Progestogens for Preventing Preterm birth International Collaborative
FAERS	FDA Adverse Event Reporting System
FDA	Food and Drug Administration
FDARA	FDA Reauthorization Act of 2017
FDASIA	Food and Drug Administration Safety and Innovation Act
FDCA	Federal Food, Drug, and Cosmetic Act
GA	Gestational Age
GAO	Government Accountability Office
HHS	Department of Health and Human Services
IMM	Irreversible Morbidity or Mortality
IND	Investigational New Drug
IPD	Individual Patient Data
IPINT	Inter-Pregnancy Interval

JAMA	The Journal of the American Medical Association
LMP	Last Menstrual Period
MFM	Maternal Fetal Medicine
MFMU	Maternal Fetal Medicine Units
mGA	Mean Gestational Age
MPPRC	Medical Policy and Program Review Council
mrpGA	Most Recent Prior Spontaneous Delivery at Gestational Age
MTO	More than One Prior SPTB
NCL	National Consumers League
NDA	New Drug Application
NECC	New England Compounding Center
NICHD	National Institute of Child Health and Human Development
NICU	Neonatal Intensive Care Unit
NIH	National Institutes of Health
NISS	Newly Identified Safety Signal
NOOH	Notice of Opportunity for a Hearing
OH	Orthostatic Hypotension
ORUDAC	Obstetrics, Reproductive and Urologic Drugs Advisory Committee
PBPA	Preterm Birth Prevention Alliance
PCAC	Pharmacy Compounding Advisory Committee
PROLONG	Progestin's Role in Optimizing Neonatal Gestation
PTB	Preterm Birth
RADM	Rear Admiral
RCT	Randomized Controlled Trial
REMS	Risk Evaluation and Mitigation Strategies
RR	Relative Risk
RWD	Real-World Data
RWE	Real-World Evidence
SAE	Serious Adverse Event
SMFM	Society for Maternal-Fetal Medicine
SOC	Standard of Care
SOH	Symptomatic Orthostatic Hypotension

SPTB	Spontaneous Preterm Birth
USPI	United States Package Insert
VTE	Venous Thromboembolism
WHO	World Health Organization

I. INTRODUCTION

Covis Pharma GmbH (Covis) submits this briefing book with respect to the Center for Drug Evaluation and Research's (CDER) Notice of Opportunity for a Hearing (NOOH) and proposal to withdraw approval for Makena (hydroxyprogesterone caproate injection)—also called 17-OHPC, 17 α-Hydroxy-progesterone Caproate, 17-HPC, or 17P.[1]

For over a decade, Makena has been the standard of care in the United States, relied on by maternal-fetal medicine (MFM) specialists to help indicated women with a history of preterm birth carry their babies closer to term. In the United States, the impact of preterm birth is disproportionately borne by Black and other minority women as well as by socioeconomically disadvantaged populations and these women would be most impacted by the withdrawal of Makena.

It is not necessary or appropriate for the Food and Drug Administration (FDA) to withdraw the approval of Makena at this time. There is strong evidence from the National Institutes of Health (NIH)-initiated multi-site, double-blind, placebo-controlled clinical trial (Trial 002), known as the "Meis trial," that Makena reduces the risk of preterm birth in the at risk U.S. patient population. Further analyses of the Meis trial detailed below make clear that the Meis trial data show an especially strong effect in reducing preterm birth prior to 35 weeks gestational age in women who are most at risk of preterm birth, namely those with a prior spontaneous preterm birth (SPTB) of <34 weeks gestational age. Notably, the reduction in preterm birth prior to 35 weeks gestational age is an endpoint that has been empirically correlated with a reduction in neonatal morbidity and mortality and is therefore, in CDER's words, an "established surrogate" or an "intermediate clinical endpoint," and likely would not require a confirmatory study. Accordingly, it is entirely possible that CDER could have initially granted a full approval, rather than accelerated approval, to Makena based on the strength of Meis alone.

CDER granted accelerated approval to Makena in 2011 based on the Meis trial and required a confirmatory trial, which the prior sponsors worked with FDA to design and implement. That confirmatory study, PROLONG (Trial 003), while not positive, does not undermine the efficacy demonstrated in the Meis trial. There were flaws in the conduct of PROLONG that ultimately made it incapable of measuring the efficacy of Makena. PROLONG was conducted in a low-risk population, located primarily outside the U.S., which led to

[1] The December 4, 2020 submission identified data, analyses, and information on which the sponsor intended to rely at the hearing. *See* 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) 250 mg/mL, Docket. No. FDA-2020-N-2029-0051 (Dec. 4, 2020), https://www.regulations.gov/comment/FDA-2020-N-2029-0051 [hereinafter December 4, 2020 Hearing Response]. On March 30, 2022, Covis submitted certain additional categories of data and information for inclusion in the hearing for the Presiding Officer's review, at her request. *See* Letter from Covis, to Celia Witten, Ph.D., M.D., Docket. No. FDA-2020-N-2029-0210 (Mar. 30, 2022), https://www.regulations.gov/document/FDA-2020-N-2029-0210. Dr. Witten subsequently granted Covis' request for inclusion. *See* Letter from Celia Witten, Ph.D., M.D., to Rebecca Wood and Christine Hunt, Docket. No. FDA-2020-N-2029-0212 (Apr. 21, 2022), https://www.regulations.gov/document/FDA-2020-N-2029-0212 [hereinafter Witten Apr. 21, 2022]. Covis incorporates by reference the administrative record for Makena, and the contents of previous submissions and correspondence, including those on Docket. No. FDA-2020-N-2029, and all appendices and attachments.

significantly lower event rates in both the placebo and treatment arms, and left the study not appropriately powered to demonstrate efficacy. Further, the ex-U.S. population was subject to different standards of care than the U.S. population—including with respect to the critical assessment of gestational age. Although PROLONG did not confirm efficacy, it did offer additional support for the safety of Makena. As detailed below, PROLONG does not negate the strong efficacy findings of the Meis trial and therefore does not lead to a conclusion that Makena must be withdrawn from the market.

Since it became the sponsor of Makena in 2021, Covis has enlisted recognized experts—including biostatisticians and clinical trialists, many of whom have significant FDA experience—to analyze the available data to evaluate an additional study that could further confirm the efficacy of Makena. As detailed below and in the attached Appendix, Covis is proposing several approaches to a new randomized controlled confirmatory study in the prior SPTB <34 weeks of gestational age high-risk population. Covis also is willing to undertake an observational study.

Given the demonstrated clinical benefit and safety profile of Makena, there is no reason to deprive women of the only available FDA-approved treatment to reduce the risk of preterm birth in high-risk women while an additional study is conducted. To the contrary, withdrawing Makena while an additional study is undertaken would lead to a number of unintended and detrimental consequences, including driving physicians and their patients to riskier options such as compounding, and likely make trial enrollment more difficult. In addition, given recent changes to abortion law in the United States, the need for appropriate reproductive care, especially for the most high-risk patients, is only likely to increase.

FDA has unquestioned authority under the accelerated approval framework to allow a drug to remain on the market while the drug continues to be studied, and FDA has previously made clear that it will use every regulatory option at its disposal, other than withdrawal, to ensure the availability of a drug that benefits patients. Rather than withdraw Makena from the market, FDA should carefully consider all available options for future study—including a randomized controlled study and an observational study. Covis is committed to working with CDER to narrow the labeling to focus the indication on the most high-risk patients while additional study is undertaken. Covis stands ready to work cooperatively with the Agency on any of these options.

II. EXECUTIVE SUMMARY

Preterm birth is a serious medical condition with a considerable impact in the United States, where the rate of premature births is higher than in other industrialized nations. The burden of preterm birth is not equally distributed in the U.S. Instead, the impact of preterm birth is disproportionately borne by Black and other minority women as well as by socioeconomically disadvantaged populations.

Makena was approved in 2011 for the indication of reduction of the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. CDER granted accelerated approval to Makena based on the strength of the Meis trial, named for its principal investigator, Paul J. Meis, M.D., a leading maternal fetal medicine physician. At the

time, CDER stated that the Meis trial was "adequate, well-controlled and very persuasive," and provided "compelling" evidence of clinical benefit.[2] The accelerated approval of Makena was based upon CDER's finding that the Meis trial had demonstrated a compelling effect on reduction in the risk of preterm births <37 weeks gestation. This is an intermediate clinical endpoint—which is *itself* a measurement of a therapeutic effect—rather than a surrogate endpoint, which is more commonly used for accelerated approval. Notably, the Meis trial also demonstrated statistically significant reductions in the risk of preterm birth at <35 weeks and at <32 weeks gestational age, both "established surrogate endpoints" strongly correlated with a reduction in neonatal morbidity and mortality.[3] In a 2014 guidance document, CDER stated that the delay in delivery demonstrated in the Meis trial can be viewed as an *intermediate* clinical endpoint, meaning that this delay in delivery, in itself, provides a therapeutic effect.[4] It follows, therefore, that demonstration of a therapeutic effect on preterm birth prior to 35 weeks gestational age should not require a confirmatory study. In this context, it appears that the strength of the efficacy data in the Meis study could have supported a full approval, with no confirmatory study, rather than an accelerated approval.

The Meis trial was immediately recognized as a major advance in the field of obstetrics and was viewed by peer-review to be highly impactful, as evidence by its publication in the *New England Journal of Medicine*. Shortly thereafter, leading medical societies such as the American College of Obstetricians and Gynecologists (ACOG) and Society for Maternal-Fetal Medicine (SMFM) issued statements endorsing 17-OHPC and establishing it as the standard of care for prevention of preterm birth in women with history of previous spontaneous preterm birth.

As a part of the 2011 accelerated approval, CDER required a confirmatory study, PROLONG, "to verify and describe the clinical benefit" of Makena. This study was initiated prior to the approval of Makena, and designed and completed by Covis' predecessors. As a result of recruitment challenges and discussion with FDA, PROLONG primarily enrolled patients in Ukraine and Russia. Ultimately, patients from the Ukraine and Russia made up 61% of the study participants (79% of the ex-US population), versus 23% of study participants from the U.S., and proved to have a much lower risk profile and were subject to a very different standard of care than the Meis population. There were also fewer Black women in particular: Black women comprised 59% of the Meis population compared to only 6.7% for PROLONG. Moreover, gestational age of the patients' qualifying spontaneous preterm delivery (defined as delivery from 20 to 36 weeks of gestation) was likely determined in Ukraine and Russia using the last menstrual period (LMP)—generally known to be an unreliable estimate used only when ultrasonography facilities are not available—whereas ultrasound examinations are generally used for gestational dating in the U.S. As ACOG acknowledged, given the observed low rates of preterm birth, PROLONG was underpowered to assess efficacy, and there may have been an

---

[2] FDA Briefing Document, "NDA 021945 Hydroxyprogesterone Caproate Injection (trade name Makena)," Bone, Reproductive, and Urologic Drugs Advisory Committee Meeting, 11, 21 (Oct. 29, 2019), https://www.fda.gov/media/132003/download [hereinafter FDA Briefing Document].

[3] CDER, NDA 21945 Clinical Review at 15, https://www.fda.gov/media/80892/download [hereinafter, Clinical Review]; FDA Briefing Document at 20, ("FDA determined that further study was needed to provide confirmatory evidence of the drug's efficacy in terms of direct clinical benefit on neonatal outcomes or through an established surrogate such as the rate of preterm birth prior to 35 and 32 weeks gestation").

[4] FDA, EXPEDITED PROGRAMS FOR SERIOUS CONDITIONS—DRUGS AND BIOLOGICS, GUIDANCE FOR INDUSTRY at 19 (May 2014), https://www.fda.gov/media/86377/download [hereinafter Expedited Programs Guidance].

unintentional selection bias in enrollment in the placebo-controlled trial due to treatment guidelines recommending the use of 17-OHPC. The PROLONG study was completed in 2019 and ultimately did not meet its objective of showing a reduction in preterm birth <35 weeks and neonatal morbidity/mortality. As detailed below, given its shortcomings, PROLONG could not confirm the Meis trial findings because of flaws in the enrollment of its study population. As such, PROLONG does not and cannot invalidate the clinical benefit observed in the Meis study.

FDA has the authority to allow Makena to remain on the market and available to patients during further study under the accelerated approval framework. This approach is supported by significant legal, scientific, and policy rationales. Most importantly, it is in the best interest of the public health to allow Makena to remain on the market as an option for physicians to prescribe to women at high risk for preterm birth. Covis—a leading pharmaceutical company that stands ready to work collaboratively with FDA and that is focused on providing therapeutic solutions for patients with life-threatening conditions and chronic illnesses—remains ready to discuss any and all data-driven next steps with the Agency, including a properly designed, adequate and well-controlled clinical trial, an observational study, and narrowing the labeling to focus the indication on the most high-risk patients while additional study is undertaken. Withdrawing Makena now would not be consistent with the data-driven, patient-focused approach the Agency has taken in similar circumstances. Withdrawal in these circumstances would, as explained below, be contrary to the public health, contrary to law, and an abuse of discretion.

- **Makena has a compelling efficacy profile meriting further study, as demonstrated by the Meis trial**

The Meis trial showed a statistically significant reduction in the risk of delivery prior to 37, 35, and 32 weeks of gestation in patients treated with 17-OHPC compared to those receiving placebo, among women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. The Meis trial so clearly demonstrated the substantial benefit of Makena for women at higher risk of preterm birth that the trial was ended early and compounded 17-OHPC became widely used in the obstetrics community even before accelerated approval of Makena was granted by CDER. CDER based its 2011 accelerated approval of Makena on <37 weeks as the intermediate clinical endpoint, while acknowledging that the Meis trial also demonstrated statistically significant reduction in preterm birth at <35 and <32 weeks gestational age. Reducing the risk of preterm birth at <35 and <32 weeks of gestational age are, according to CDER, "established" surrogates, which are well-known to correlate to meaningful reductions in neonatal morbidity and mortality and do not need confirmatory evidence of clinical benefit.[5] In its 2014 Expedited Programs for Serious Conditions – Drugs and Biologics Guidance, CDER acknowledges the Meis result as demonstrating an *intermediate* clinical endpoint:

> Examples of cases in which FDA has used an intermediate clinical endpoint to support accelerated approval include the following: . . . A treatment for preterm labor was approved based on a demonstration of delay in delivery. Under

[5] Clinical Review at 15; FDA Briefing Document at 20.

accelerated approval, the sponsor was required to conduct postmarketing studies to demonstrate improved long-term postnatal outcomes.[6]

In the decade since Makena's approval in 2011, CDER has used inconsistent terminology (surrogate endpoint vs. intermediate clinical endpoint) in describing the endpoints of the Meis trial and gestational age of delivery. In recent correspondence with Covis and FDA, CDER acknowledged this inconsistency and agreed to use the correct term, "intermediate clinical endpoint," going forward. This clarification is important because unlike a surrogate endpoint—which is not itself a measure of clinical benefit—an intermediate clinical endpoint is explicitly "a measurement of a therapeutic effect."[7] Moreover, as further explained below, CDER's own guidance states that an intermediate clinical endpoint will "usually" support traditional approval in circumstances such as here, in contrast to a surrogate endpoint, which needs to be validated before it can be used for traditional approval.[8]

The results of the Meis trial provided consistent and robust evidence of efficacy across all subgroups evaluated and across all evaluated endpoints.[9] Medical societies such as ACOG have publicly endorsed the use of 17-OHPC since the Meis trial, and consequently, use of Makena has been the standard of care in the U.S. for more than 10 years to prevent preterm birth for women with history of previous spontaneous preterm birth.

To the extent that there previously have been questions about whether the ubiquity of Makena would undermine the ability to undertake a new confirmatory study, as demonstrated below, significant reanalysis supports the feasibility of such an option.

- **Further analysis by experts demonstrates that women with a prior spontaneous preterm birth of <34 weeks gestational age are the highest risk population for subsequent preterm birth**

Covis acquired AMAG, the prior sponsor of Makena, at the end of 2020, and became the sponsor of Makena in 2021. Since that time, Covis has sought the input of key experts to help it further understand both the available data on Makena and the feasibility of a confirmatory study that can be executed in the U.S. in an appropriate population within a reasonable timeframe. The maternal-fetal medicine community has known for some time that prior history of an early spontaneous preterm birth increases the risk of another preterm birth.[10] Other factors that have been previously recognized at a qualitative level are the presence of multiple spontaneous preterm births in the patient's history as well as a short interval between pregnancies. Covis sought to confirm its quantitative understanding of these factors, for which it analyzed both the

[6] Expedited Programs Guidance at 19.

[7] *Id.* at 18.

[8] *Id.*

[9] FDA Briefing Document at 8; Baha Sibai et al., *Re-examining the Meis Trial for Evidence of False-Positive Results*, 136 OBSTET. GYNECOL. 622-627 (2020), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7431135/.

[10] *See* Michael S. Esplin et al., *Estimating recurrence of spontaneous preterm delivery*, 112 OBSTET. GYNECOL. 516-23 (Sept. 2008), https://pubmed.ncbi.nlm.nih.gov/18757647/; S. Katherine Laughon et al., *The NICHD Consecutive Pregnancies Study: Recurrent preterm delivery by subtype*, 210 AM. J. OBSTET. GYNECOL. 131.e1-131.e8 (Feb. 2014), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3934564/.

data and efficacy results from the Meis trial, as well as electronic health records from Dorsata—a leading maternal health technology company with access to records of over 210,000 pregnancies. These data confirm the previously recognized relationships between risk factors and preterm birth outcome while also confirming specific cut-points that would be useful from a clinical trial design standpoint.

The available data shows that earlier prior spontaneous preterm birth is the most significant predictor of subsequent preterm birth. In addition, Covis analyzed the population in Dorsata who had a prior spontaneous preterm birth at various cutoff gestational ages (ranging from 28 through 36 weeks). The data are shown in Table 1 below and demonstrate that there is a clear drop-off in risk past week 34. These data support the selection of <34 weeks for prior SPTB gestational age as an appropriate high-risk subgroup in whom we would propose further study of Makena, and are consistent with existing literature.[11]

Presented below are several risk factor models developed using logistic regression, including the best one, two, and three parameter models within Dorsata. The primary outcome that the analyses tried to predict was the likelihood of the patient experiencing a subsequent preterm birth before week 34. The results of this model are not significantly impacted by varying the predicted outcome to 33 or 35 weeks—the purpose was to pick an outcome that was acknowledged by Ob/Gyns and neonatologists as significantly impacting morbidity/mortality of the neonate. The result of the Dorsata model is shown below in Table 1:

Table 1
Dorsata (Excluding 17-OHPC-Treated Subjects) Best N-variable Models Predicting PTB <34 Weeks

Model/Var#	Variable(s)	Odds Ratio (95% CI)	P-value
Best 1-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.82, 0.91)	<0.0001
Best 2-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.82, 0.91)	<0.0001
Var #2	Smoking during Pregnancy	0.51 (0.18, 1.46)	0.21
Best 3-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.81, 0.90)	<0.0001
Var #2	Smoking during Pregnancy	0.52 (0.18, 1.48)	0.22
Var #3	Alcohol Use during Pregnancy	0.64 (0.28, 1.45)	0.28

As can be seen from the above table, specific measures of prior pregnancy history are a strong predictor of a subsequent preterm birth <34 weeks. Given the above models, Covis also constructed similar one through three parameter models using the Meis data to see if the real-world data in Dorsata (from 2018-2021) were congruent with the Meis data from twenty years earlier. The Meis risk models are shown below in Table 2:

[11] *See supra* note 10.

Table 2
Meis (Vehicle-only) Best N-variable Models Predicting PTB <34 Weeks

Model/Var#	Variable(s)	Odds Ratio (95% CI)	P-value
Best 1-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.89 (0.82, 0.98)	0.013
Best 2-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.89 (0.81, 0.98)	0.013
Var #2	Smoking during Pregnancy	2.84 (1.19, 6.76)	0.019
Best 3-variable model			
Var #1	Smoking during Pregnancy	3.12 (1.27, 7.68)	0.013
Var #2	Mean Gestational Age of Prior Pregnancies (Weeks)	0.90 (0.82, 0.99)	0.030
Var #3	Inter-pregnancy Interval (Years)	0.84 (0.71, 1.00)	0.049

As can be seen in the tables above, the analyses consistently showed the gestational age of prior pregnancies to be a strong predictor of preterm birth across different variable models as well as different populations. We note that while mean gestational age may not be a practical measure for clinical use or labeling, it stands as a proxy for the totality of the patient's prior pregnancy history. The Appendix delves into more detail on the proposed measures that would be utilized in modifying inclusion criterion for a future clinical study.

- **Covis has identified feasible options for a further confirmatory trial of Makena in the high-risk patient population**

Covis has evaluated several study designs that could address Makena's clinical benefit in a high-risk patient population including: a randomized, blinded placebo-controlled study, an observational study longitudinally following risk-matched treated versus non-treated patients in large databases, and hybrid studies involving two or more approaches. The Appendix to this document includes the rationales for and details of proposed study designs, including the inclusion criteria, endpoints, and sample size estimates. These proposals represent the culmination of Covis' work with a multidisciplinary scientific advisory panel, consisting of renowned experts in their fields, and its months-long collaboration with Dorsata.

Among the different proposals presented therein, a randomized controlled trial (RCT) can effectively address confounding variables and is Covis' preferred approach to an additional study. The RCT would be conducted in women with a prior spontaneous preterm birth of <34 weeks gestational age, an inclusion criterion which Covis' analyses described above show to be a particularly strong predictor of preterm birth risk. Covis' preferred endpoint would be a continuous endpoint, such as increase in delivery time from randomization focused on <35 weeks, rather than the conventional categorical endpoints of preterm birth rate at specific cutoffs (e.g., <32, <35, or <37 weeks) used in Meis and PROLONG. This endpoint has many advantages, including higher sensitivity than a categorical endpoint and the ability to provide a clearer understanding of neonatal benefit. Covis' feasibility assessment suggests that a trial using a continuous endpoint can be conducted in 3-6 years, provided that a sufficient number of high-risk patients are enrolled from sites with high incidence rates of spontaneous preterm birth. Notably, 11 of 12 surveyed prospective trial investigators from academic MFMU network sites

that participated in the Meis trial have expressed interest in participating in a new placebo-controlled trial to evaluate the efficacy of 17-OHPC. The remaining provider surveyed declined such participation on grounds that a trial was not necessary at all, as the Meis trial sufficiently settled the question of efficacy.[12]

To date, CDER has declined to meet with Makena's sponsors about options for further study or data generation. For example, although CDER acknowledged during the 2019 BRUDAC meeting that it would be willing to discuss "other ideas [than randomized trials] that can achieve the same objective" for Makena,[13] it has not provided AMAG or Covis with an opportunity for such discussions. Shortly after the 2019 meeting, AMAG sought to meet with CDER to discuss the scope and design of additional confirmatory studies and requested a Type C meeting. CDER denied that request, refusing to enter into a dialogue with AMAG.[14] Covis respectfully requests that its proposal for additional studies of Makena receive proper review and consideration by the Agency and continues to welcome a cooperative path forward in the best interest of patient care.

- **In addition to the Meis trial and PROLONG, more than a decade of real-world use supports the positive safety profile of Makena**

Postmarket surveillance data further supports the safety of Makena. As is required for all FDA-approved drugs, Covis and its predecessors have maintained post-marketing surveillance throughout the life of the drug. Among the more than 350,000 women treated with Makena, no new safety concerns, signals, or risks have been identified in nearly 10 years of use.

CDER has pointed to known potential risks of Makena described in its labeling such as thromboembolic events, depression, allergic reactions, decreased glucose tolerance, fluid retention that may worsen maternal conditions such as pre-eclampsia, and injection site adverse reactions, as additional rationale to withdraw the drug when balanced against Makena's benefit. In practice, only a tiny percentage of the more than 350,000 women treated with Makena have reported experiencing these adverse events. For example, during the past decade of Makena use, only 37 of 356,327 patients (0.01%) have reported thromboembolic events. Comparatively, the background rate of venous thromboembolism (VTE) in pregnant women is 0.12%.[15]

- **Recent studies further support Makena's positive benefit-risk profile**

The favorable benefit-risk profile of 17-OHPC has been well-established not only through decades of real-world use, but also in more recent studies. In Section VII.A.5 below, Covis describes six studies—Bastek, EPPPIC, Manuck, Carter, Schuster and Price—that continue to show 17-OHPC's positive benefit-risk profile in patients. For instance, the first ever

---

[12] Attachment D to the Appendix, Blackwell MFMU Trial Site Survey.

[13] Bone, Reproductive, and Urologic Drugs Advisory Committee Advisory Committee Meeting, 295:08-295:09 (Oct. 29, 2019), https://www.fda.gov/media/136108/download [hereinafter Advisory Committee Meeting Transcript].

[14] *See* Letter from Christine Nguyen, Acting Dir., Div. of Bone, Reproductive and Urologic Products, ODE III, CDER to Helen Milton, Senior Vice President, Regulatory Affairs, AMAG Pharma USA, Inc. (Mar. 11, 2020).

[15] Syed Bukhari et al., *Venous thromboembolism during pregnancy and postpartum period*, 97 EUR. J. INTERN. MED. 8-17 (Mar. 2022), https://pubmed.ncbi.nlm.nih.gov/34949492/.

individual patient data (IPD) meta-analysis of 17-OHPC, which included five randomized trials, revealed that 17-OHPC reduces the risk of preterm birth before 34 weeks in high-risk singleton pregnancies as well as favorable reductions before <28 and <37 weeks.[16] A number of recent studies also have supported Makena's safety profile, with the authors of a recent randomized, double-blind, placebo-controlled trial concluding that their trial "supports the safety of [17-OHPC] administration during pregnancy."[17] These studies directly address the efficacy and/or safety of Makena.

A recent article by Murphy et al. does not undermine the safety profile of Makena.[18] As CDER's Division of Epidemiology II (DEPI II) Team Leader concluded, the study's limitations "preclude this study from contributing definitively to this drug safety issue," as the study "provides insufficient evidence to support regulatory action regarding a long-term cancer risk in offspring who were exposed in utero to 17-OHPC."[19]

As explained in Section VII.A.7, Murphy describes a retrospective analysis of historical use of a different drug, called Delalutin where the clinical use was entirely distinct from modern clinical use with Makena. Murphy acknowledges that although Delalutin and Makena both contain 17-OHPC, the two drugs differ in "the timing, frequency, and pregnancy-related indications."[20] Notably, ACOG recognized the inapplicability of Murphy to Makena and issued an announcement shortly after the article was published, pointing out its "limitations in the design" and stating "the study's findings are not conclusive and should not influence practice" with respect to Makena.[21]

CDER closed its Newly Identified Safety Signal (NISS) process with a determination that the risk or status was "indeterminate," and not that it had found an "identified" risk. As such, CDER itself acknowledged that Murphy does *not* identify a link between 17-OHPC and cancer. This conclusion was reached universally by the CDER offices and divisions reviewing the NISS. As CDER's own internal analysis, and declarations from biostatistical experts show, this study is methodologically flawed, and neither reliable nor relevant to the issues raised in this proceeding.

---

[16] The EPPPIC Group, *Evaluating Progestogens for Preventing Preterm Birth International Collaborative (EPPPIC): Meta-Analysis of Individual Participant Data from Randomised Controlled Trials*, 397 THE LANCET 1183-94 (2021), https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)00217-8/fulltext.

[17] Joan T. Price et al., *Weekly 17 Alpha-Hydroxyprogesterone Caproate to Prevent Preterm Birth Among Women Living With HIV: A Randomised, Double-blind, Placebo-controlled Trial*, 8 LANCET HIV e605-13 (Oct. 2021), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8476342/pdf/main.pdf.

[18] Caitlin C. Murphy et al., *In Utero Exposure to 17α-Hydroxyprogesterone Caproate and Risk of Cancer in Offspring*, 226 AM. J. OBSTET. GYNECOL., 132.e1, 132.e8 (2022), https://pubmed.ncbi.nlm.nih.gov/34767803/ [hereinafter Murphy].

[19] CDER, Division of Epidemiology II (DEPI II), Office of Surveillance and Epidemiology Review (OSE), Office of Pharmacovigilance and Epidemiology (OPE), Team Leader Review, Epidemiology: Review of published paper (Jun. 22, 2022).

[20] *Id*.

[21] ACOG, *ACOG Guidance on 17-OHPC Remains Unchanged*, ROUNDS (Nov. 12, 2021), https://www.magnetmail.net/actions/email_web_version.cfm?ep=ZXoixPhGZdQ3e6Q2dvjdZDwQFTvi0y3E8vmMV8yEYSCen1PqjHurjEQW5OcZEat1bek8Es8Fl1Bc-OK2WWEQeqpXXi6RJogR0lF-bOcAh12TWv9Ju9GGNuZlrp6THaS3.

- **Clinician and patient choice should be respected**

FDA's proposal to withdraw Makena from the market stands in stark contrast with the broad and deep conviction of well-regarded scientists, clinicians, advocacy groups, and patients who highly value preserving Makena as an option for women at high risk for preterm birth.

Major professional societies including ACOG and the SMFM continue to support the use of 17-OHPC to prevent recurrent preterm birth for women with a prior spontaneous preterm birth.[22] Organizations committed to addressing health disparities that disproportionately affect minority women—including the Black Women's Health Imperative, the Preterm Birth Prevention Alliance, In Our Own Voice, and the National Black Nurses Association—have expressed concerns that the withdrawal of Makena would deepen existing health inequities, particularly for Black women. Clinicians and patients should not be denied the ability to make an informed treatment determination in favor of using Makena for appropriate patients.

- **Withdrawal of Makena would deprive an underserved and vulnerable patient population of an extensively-studied, safe, FDA-approved treatment option while exposing them to risks from unsafe and unproven alternatives**

Withdrawal of accelerated approval for Makena and its generic equivalents would have significant negative public health consequences for high-risk pregnant women, including Black and minority women and socially-disadvantaged populations. As demonstrated by Covis' analysis discussed further in Section VII.A.2.b and the Appendix, the data suggest that Black women in the Meis trial experienced a benefit from 17-OHPC treatment earlier in pregnancy and as a subgroup showed significant reduction of preterm birth events for <32 weeks and <35 weeks gestational age. These patterns, at a minimum, indicate that withdrawal of Makena will deprive a patient group of a therapeutic option which has been shown to be powerfully effective, and certainly would suggest that the effects should be more carefully studied before considering any withdrawal.

To be clear, withdrawal of Makena at this time would deprive healthcare providers of what is now the only FDA-approved treatment to reduce the risk of preterm birth—a risk those groups disproportionately bear. Further, it would be fundamentally unfair, unsound, and inappropriate to deprive high-risk pregnant women—many of whom are Black or members of other minority groups—of the use of Makena based on data from PROLONG, a study in which they were not adequately represented.

---

[22] ACOG, *ACOG Statement on 17p Hydroxyprogesterone Caproate* (Oct. 25, 2019), https://www.acog.org/en/news/news-releases/2019/10/acog-statement-on-17p-hydroxyprogesterone-caproate; ACOG, *Practice Advisory: Clinical Guidance for Integration of the Findings of the PROLONG Trial: Progestin's Role in Optimizing Neonatal Gestation* (Oct. 25, 2019), http://web.archive.org/web/20201023110456/https://www.acog.org/clinical/clinical-guidance/practice-advisory/articles/2019/10/clinical-guidance-for-integration-of-the-findings-of-the-prolong-study; SMFM, *SMFM Statement: Use of 17-alpha hydroxyprogesterone caproate for prevention of recurrent preterm birth*, at 3 (Jul. 2020), https://els-jbs-prod-cdn.literatumonline.com/pb/assets/raw/Health%20Advance/journals/ymob/SMFM_Statement_PROLONG-1572023839767.pdf.

No other therapies—including compounded drugs, cervical cerclages, or generic Delalutin (a 17-OHPC product that is neither labeled for use in pregnant women nor available in the same dosing or modes of administration as Makena)—could fill this void. If Makena is withdrawn, the law requires FDA also to bar compounding of 17-OHPC, but FDA's practice suggests that compounding could still be available for some time, thereby posing additional risks to this already high-risk population of pregnant women.

Pharmacy compounding does not require compliance with the rigorous current good manufacturing practice (cGMP) standards that are applicable to approved products such as Makena. Compounding has a troubled history in the U.S. and has led to numerous identified safety risks for patients, as exemplified by the New England Compounding Center (NECC) meningitis outbreak of 2012, in which 64 patients receiving nonsterile compounded therapy died of fungal meningitis.[23] As of 2017, FDA had issued more than 130 warning letters advising compounders of significant violations of federal law and overseen more than 100 recalls involving compounded drugs. For compounded 17-OHPC, during the eight years since the 2012 NECC outbreak, there were at least 26 safety recalls.[24]

These facts are particularly troubling considering that in a recent survey of approximately 400 obstetricians, gynecologists, and maternal-fetal medicine specialists, more than a quarter of physicians answered that they are very likely to recommend compounded medication if there are no approved alternatives.[25] Similarly, a number of 2019 advisory committee members indicated that they would resort to compounded 17-OHPC if Makena were to be withdrawn from the market.[26] CDER reviewers also acknowledged, as recently as 2018, that "from a product quality standpoint, the purity and potency of compounded HPC products cannot always be assured."[27] Relying on compounded versions of Makena to meet the needs of patients and physicians wishing to have access to a 17-OHPC product would pose significant additional safety risks in the treatment of pregnant women.

In addition, in light of recent changes in abortion law within the United States, it is likely that the need for appropriate reproductive care for the most high-risk patients will only increase. Studies have shown unintended pregnancies to be strongly associated with preterm birth and low

---

[23] *See generally* FDA Press Release: New England Compounding Center Pharmacist Sentenced for Role in Nationwide Fungal Meningitis Outbreak (Jan. 31, 2018), https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/press-releases/january-31-2018-new-england-compounding-center-pharmacist-sentenced-role-nationwide-fungal.

[24] David L. Gandell et al., *FDA Approved vs. Pharmacy Compounded 17-OHPC – Current Issues for Obstetricians to Consider in Reducing Recurrent Preterm Birth*, 36 CURR. MED. RES. OPIN. 1393-1401 (Jun. 7, 2020), https://pubmed.ncbi.nlm.nih.gov/32544354/.

[25] *See* Attachment A to the Appendix below.

[26] *See, e.g.*, Advisory Committee Meeting Transcript, *supra* note 13 at 309:15-309:18.

[27] CDER, NDA 021945 Cross-Discipline Team Leader Review (Feb. 14, 2018), https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945Orig1s012SumR.pdf [hereinafter Cross-Discipline Team Leader Review].

birthweight.[28] Accordingly, the need for therapies such as 17-OHPC to support women at risk of preterm birth in delivering healthy babies is likely to increase.

- **FDA's removal of Makena from the market would undermine the Agency's commitment to important public health priorities, including diversifying clinical trials and improving maternal and infant health**

Congress and FDA have long been committed to improving diversity in clinical trials. The Agency has recognized the importance—both as a matter of scientific reliability and sound public health policy—of considering demographic factors when evaluating the therapeutic benefit of a medicinal product. FDA's commitment to diversity in clinical trial participation would be undermined if the Agency nevertheless relied on a study that underrepresented minority groups to determine the safety and efficacy of a drug that is most beneficial for such groups. The burden of preterm birth represents a significant U.S. public health problem, and one that has proved challenging to address. Indeed, CDER has acknowledged that "having effective treatment available to prevent preterm delivery in women who are at risk is of immense societal value."[29] Despite the health policy priority placed on improving maternal-child health, preterm birth persists as an area of unmet need in drug development.

CDER's proposal to withdraw Makena's approval based on PROLONG would be contrary to the Agency's priority of ameliorating health inequity and the recognized need of the high-risk group at issue here. It also would be unwarranted given that CDER has the authority to exercise—and in the past has exercised—regulatory flexibility to allow Makena to remain on the market to help vulnerable populations that bear the greatest impact of preterm birth and see most benefit from Makena treatment.

- **PROLONG supports the safety of Makena and does not undermine the strong efficacy findings of the Meis trial**

The strength of the Meis trial's results made constructing a confirmatory study that mirrored the trial's higher-risk population exceedingly difficult. With 17-OHPC quickly established as the standard of care in the United States, providers were unwilling to enroll patients in a trial of 17-OHPC where they might be given a placebo instead. Given the ubiquity of Makena use in the U.S., based on the strong evidence from the Meis trial and resultant recommendations from professional societies, it became challenging to construct a confirmatory study that mirrored the higher-risk patient population that participated in the Meis trial. Therefore, unlike the Meis trial, PROLONG was conducted largely outside the U.S. because of

[28] *See* Rachel Treisman, *States With the Toughest Abortion Laws Have the Weakest Maternal Supports, Data Shows*, NPR (Aug. 18, 2022), https://www.npr.org/2022/08/18/1111344810/abortion-ban-states-social-safety-net-health-outcomes (citing Amici Curiae Brief of 547 Deans, Chairs, Scholars And Public Health Professionals, The American Public Health Association, The Guttmacher Institute, and The Center for U.S. Policy, in Support of Respondents, *Thomas E. Dobbs, State Health Officer of the Mississippi Department of Health, et al. v. Jackson Women's Health Organization, et al.*, 945 F.3d 265 (No. 19-1392) (2022), https://www.supremecourt.gov/DocketPDF/19/19-1392/193302/20210921172339465_19-1392%20Brief.pdf).

[29] *Id*. at 2.

extensive challenges in enrolling patients in the U.S. where 17-OHPC had already become the widely relied upon standard of treatment, making assignment to placebo challenging.

Thus, even though PROLONG was a larger study than Meis, the PROLONG population was vastly different from Meis and starkly underrepresented individuals from the populations most at risk from preterm birth in the U.S. This is contrary to FDA's express recommendations on ensuring that data generated "reflect the racial and ethnic diversity of the population expected to use the medical product if approved, and … identify effects on safety or efficacy outcomes that may be associated with, or occur more frequently within these populations."[30] PROLONG enrolled far fewer Americans, and fewer American Black women in particular: there were 273 (59%) Black women in the Meis trial compared to only 111 (29%) Black women in the U.S. PROLONG subset. Moreover, the U.S. PROLONG subset had a different risk profile than the Meis subset, in that only about a third of the patients in the subset fell under the classification of high risk defined by prior spontaneous preterm birth <34 weeks—too few to draw any conclusions of efficacy for 17-OHPC in this subpopulation.

There are significant disparities experienced by American Black women in preterm birth rates, that are not explained by geographic or socioeconomic factors. According to the March of Dimes, the preterm birth rate among Black women is 14.0%—which is 51% higher than the rate among all other women in the U.S.[31] Race is one of the many risk factors, along with lower education level, substance use, stress and other factors influencing economic status, that have been linked to preterm birth.

As ACOG acknowledged, PROLONG may have had an unintentional selection bias against enrolling higher-risk patients in the U.S., likely due to desire by physicians to treat their highest risk patients with therapy rather than risk the patients being randomized to placebo.[32] Consequently, the U.S. PROLONG participants had lower risk factors for preterm birth than the participants in the Meis trial, and outside of the U.S., PROLONG primarily enrolled women that had vastly different social determinants of health.

Another important factor in the PROLONG result is the methodology for verifying the prior qualifying gestational age of the singleton spontaneous preterm delivery for patients in the study. One of the inclusion criteria was documented history of a previous singleton spontaneous preterm delivery, defined as delivery from 20 to 36 weeks of gestation following spontaneous preterm labor or preterm premature rupture of membranes. In the U.S., ultrasound has been used consistently to measure crown-rump and femur length in the past decade to determine gestational age, while such practices are not common outside the U.S., and particularly not in Russia and Ukraine.[33] Intake forms for the so-called high-risk patients (prior SPTB <34 weeks of

---

[30] FDA, DIVERSITY PLANS TO IMPROVE ENROLLMENT OF PARTICIPANTS FROM UNDERREPRESENTED RACIAL AND ETHNIC POPULATIONS IN CLINICAL TRIALS GUIDANCE FOR INDUSTRY, Draft Guidance (Apr. 2022), https://www.fda.gov/media/157635/download.

[31] 2021 March of Dimes Report Card (2021), www.marchofdimes.org/materials/March-of-Dimes-2021-Full-Report-Card.pdf.

[32] ACOG Statement on 17p, *supra* note 22.

[33] *See* Marina P. Shuvalova et al., *Maternity Care in Russia: Issues, Achievements, and Potential*, 37 J. OBSTET. GYNAECOL. CAN. 865–871 (2015), https://pubmed.ncbi.nlm.nih.gov/26606698/; Daniel F. O'Keefe and Alfred

gestational age) in PROLONG confirm that prior medical records were used to qualify patients, but no requirement was included regarding the method to identify qualifying gestational age. A review of the individual patient records also shows significant variance in the dates of the qualifying pregnancies, some of which go back into the late 1990s and early 2000s. Based on the PROLONG protocol which instructed that the gestational age be determined "where possible" based on the last menstrual period (LMP) and ultrasound examination,[34] it is highly likely that LMP was utilized to qualify many of the ex-U.S. PROLONG patients. As ACOG has recognized, the LMP method is generally unreliable as it does not account for irregularities in cycle length, variability in the timing of ovulation, or even the fact that women may inaccurately recall their last menstrual period. In fact, it has been estimated that up to one quarter of the preterm births that were classified using LMP may in fact not be preterm.[35] The variability of the LMP method may have given rise to incorrect estimates of the prior qualifying gestational age for ex-U.S. PROLONG participants, making it likely that the inclusion criteria were not actually met for some of these patients. These uncertainties further confound interpretation of the PROLONG data.[36]

The analysis of data conducted by Covis of a cut-point of prior SPTB gestational age <34 offers another way to demonstrate that the Meis population was higher risk than PROLONG-US and helps explain why the event rate in PROLONG-US was markedly lower than that of Meis.

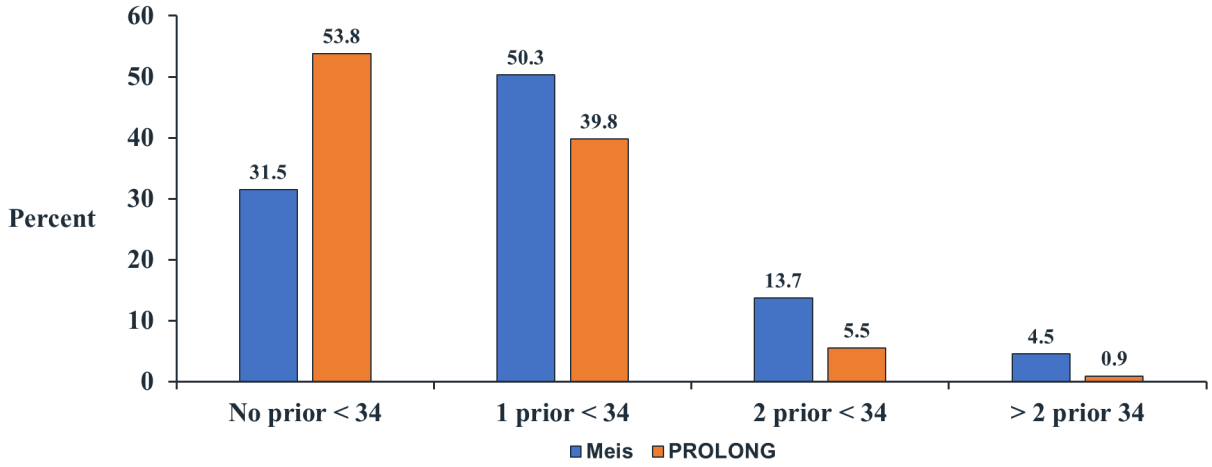
Abuhamad, *Obstetric ultrasound utilization in the United States: Data from various health plans*, 37 SEMIN. PERINATOL. 292-94 (Oct. 2013), https://pubmed.ncbi.nlm.nih.gov/24176148/; *See* S. Arbuzova, *Genetic Services in the Ukraine*, 5 EUR. J. HUM. GENET. 183-87 (1997) (noting that general ultrasound screening twice during the second trimester of pregnancy, rather than the first, is recommended), https://pubmed.ncbi.nlm.nih.gov/9450221/.

[34] Study Protocol, A Phase 3B, Multi-Center, Randomized, Double-Blind Study of Hydroxyprogesterone Caproate Injection, 250 mg/ml, Versus Vehicle for the Prevention of Preterm Birth in Women With a Previous Singleton Spontaneous Preterm Delivery, Protocol Number: 17P-ES-003, Version 5.0 at 24-25 (Nov. 4, 2013).

[35] Michael S. Kramer, et al., *The Validity of Gestational Age Estimation by Menstrual Dating in Term, Preterm, and Postterm Gestations*, 22 JAMA 3306-3308 (Dec. 9, 1998), https://jamanetwork.com/journals/jama/article-abstract/375526.

[36] *See* Sol P. Juárez et al., *Preterm disparities between foreign and Swedish born mothers depend on the method used to estimate gestational age. A Swedish population-based register study*, 16 PLOS ONE e0247138, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7899337/ (finding that the use of last menstrual period to calculate gestational age led to an inaccurate assessment of preterm birth disparities between Swedish and foreign-born women).

Figure 1
Frequency of Prior Qualifying SPTB <34 Weeks (Meis vs. PROLONG-US)



In addition to the above distributions, the differences in the average prior gestational ages for all patients in Meis and PROLONG-US were analyzed. Individual records showed many more patients in PROLONG-US who had early SPTBs followed by two or more full-term births. Many of these patients would not be considered high-risk by treating MFMs, yet were enrolled in the study. The average number of prior preterm births for patients in PROLONG-US was lower compared to patients in Meis. The lower risk of the patients resulted in a lower event rate, rendering the study underpowered to draw any conclusions about the efficacy of 17-OHPC.

All of these factors underscore the limitations of PROLONG, notwithstanding the prior sponsor and CDER's collaboration on its study design and planned sample size before the study was initiated.[37] The Ukrainian and Russian patients and standard of care (SOC) are just not applicable to the U.S. patients and practice.

CDER has argued that a single positive trial such as Meis, even if well-conducted, may have biases or may reflect a chance finding—a position that has been thoroughly refuted in the literature.[38] CDER also has suggested that PROLONG undermines the strength of the Meis findings. Critically, however, CDER does not acknowledge the converse—that a single trial can be a false negative that should not be relied on to undermine a compelling RCT that has been supported by over a decade of real-world experience. This is particularly true in the case of PROLONG where the low event rate markedly impacted the study's power and led to an increased chance of a false negative. Moreover, CDER fails to acknowledge the severe limitations that undermine the PROLONG trial's ability to evaluate the effectiveness of 17-OHPC. In short, PROLONG does not negate Meis and does not disprove the significant benefits of reducing preterm birth demonstrated in the Meis trial in a representative higher risk patient population in the U.S.

---

[37] *See* CDER, NDA 21945 Summary Review 19, 26 (Feb. 3, 2011), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000SumR.pdf [hereinafter NDA 21945 Summary Review].

[38] *See* Sibai et al., *supra* note 9 at 622.

Moreover, PROLONG did not reveal any unexpected or new safety concerns compared to the Meis trial, thereby reaffirming the overall safety profile of Makena. Indeed, the integrated dataset from Meis and PROLONG demonstrates similar rates with respect to thromboembolic disorders, allergic reactions, decreased glucose tolerance, fluid retention that may worsen maternal conditions such as pre-eclampsia, depression, and injection site adverse reactions between Makena and placebo.

- **FDA's accelerated approval framework permits a drug to remain on the market while additional data are generated, and FDA precedent and practice support maintaining Makena's approval status**

In codifying the accelerated approval program, Congress gave FDA considerable discretion to approve drugs through accelerated approval and, once approved, to keep such drugs on the market. For example, in authorizing FDA to approve a drug for serious and life-threatening conditions upon a determination that it has an effect on a surrogate endpoint or intermediate endpoint reasonably likely to predict clinical benefit,[39] Congress gave FDA discretion to determine whether the endpoint meets this standard. Similarly, Congress authorized—but did not require—FDA to condition approval upon a confirmatory trial.[40] Most relevant here, Congress provided that if a required post-approval study fails to verify clinical benefit, FDA "may"—*but need not*—withdraw approval.[41] This contrasts with other sections of the statute (not under the accelerated approval provisions) that state that FDA "shall" withdraw approval in certain circumstances.[42] The accelerated approval framework is designed to allow discretion when the public health benefits from flexibility rather than constraint, and is peppered with permissive ("may") language. By layering discretion atop discretion, Congress sought to give FDA the necessary flexibility to address unmet medical needs for serious or life-threatening diseases or conditions. At the same time, FDA's discretion is not unbounded. FDA must engage in reasoned decision-making, treating like situations alike, considering all relevant evidence and aspects of the problem, and explaining any departures from past policy or practice.

In practice, FDA rarely pursues withdrawal of a drug granted accelerated approval. In the past 30 years, CDER has withdrawn fewer than 10 drugs/indications after the failure of confirmatory study despite having approved more than 270 drugs or indications under accelerated approval.[43] When a confirmatory study fails to verify clinical benefit, FDA has made clear that it will prioritize patient interests and explore a wide range of regulatory options.

The same factors that have led FDA not to withdraw drugs following a failed confirmatory study apply here: Makena has a positive safety profile, the pre-approval data

---

[39] 21 U.S.C. § 356(c)(1)(B) (2020).

[40] *Id.* § 356(c)(2)(A).

[41] *Id*. § 356(c)(3).

[42] *See id*. § 355(e).

[43] *See* FDA, CDER Drug and Biologic Accelerated Approvals Based on a Surrogate Endpoint As of December 31, 2021, https://www.fda.gov/media/151146/download (revised May 13, 2022) [hereinafter CDER Drug and Biologic Surrogate Endpoint Approvals]; Sue Sutter, *Accelerated Approval Withdrawals Through the Years*, THE PINK SHEET: INFORMA PHARMA INTELLIGENCE (Apr. 25, 2019), https://pink.pharmaintelligence.informa.com/PS125183/Accelerated-Approval-Withdrawals-Through-The-Years.

provided compelling evidence of an important clinical benefit for at risk populations, there are identifiable populations that would be placed at risk if a drug were withdrawn, and there are significant safety concerns associated with Makena alternatives, including those of unregulated compounding and potential reliance on off-label use of Delalutin.

- **FDA should maintain the approval status of Makena while follow-up studies are undertaken and consider whether narrowing the indication to high-risk women is appropriate**

FDA policy and practice support the Agency maintaining the approval of Makena while a follow-up study is undertaken. Since acquiring Makena in late 2020, Covis has worked with experts to assess several study designs for investigations aimed at evaluating Makena's clinical benefit in high-risk patients. Based on the results of this assessment, Covis believes that it will be feasible to conduct a new confirmatory study focused on the patients at highest risk of preterm birth, namely women with a prior spontaneous preterm birth <34 weeks. Moreover, PROLONG suffered significant recruitment challenges as it was initiated shortly after ACOG endorsed 17-OHPC as the standard of care and after FDA approved Makena as the first-ever pharmacotherapy indicated for reduction of preterm birth. Physicians in the U.S. were then reluctant to enroll their highest risk patients in a clinical trial and risk being randomized to placebo, and patients declined to participate in the trial. More than ten years have passed since these developments. Recent developments related to the PROLONG efficacy findings and CDER's proposal to withdraw the Makena approval have raised questions about the efficacy of Makena. Although these developments do not support withdrawal of approval, they have raised questions that need to be addressed by further clinical study.[44] Institutional review boards and clinical investigators that, a decade ago, hesitated to participate in PROLONG would now recognize that equipoise exists to support the enrollment of patients in a randomized clinical study conducted in the United States. Moreover, in the event that any ex-U.S. sites are included in a new confirmatory study, PROLONG has taught the importance of requiring a more rigorous assessment of gestational age as well as assuring that any patient enrolled meet the entry criteria. Covis includes several proposals for such studies as well as the expected timeframe for completion in the attached Appendix.

Covis has taken a number of steps, detailed below in the Appendix, to assess and ensure the feasibility of such a study in the U.S. including outreach to potential sites, consultation with contract research organizations, and surveys of relevant physician and patient populations to understand willingness to participate in a confirmatory study of an FDA approved drug. In Covis' initial survey, a large majority of physicians surveyed (78%) say they are likely to recommend a pregnant patient enroll in a placebo-controlled study comparing the efficacy of a product vs. placebo only when FDA has approved the product.[45] Furthermore, an even larger majority of physicians (88%) say that it is important for treatment options to be approved by FDA before recommending them to their pregnant patients.

[44] A survey of 322 providers showed that more than a third (36%) of doctors say they are prescribing progesterone by injection less often than they did three years ago, despite it being the most prescribed therapy for highest risk patients. *See* Attachment C to the Appendix.

[45] *See* Attachment A to the Appendix.

A follow-up survey affirmed these results. In that survey, 80% of physicians with experience treating patients at risk for preterm birth say they are likely to recommend a pregnant patient enroll in a placebo-controlled study for the prevention of preterm birth when the product is FDA approved. These numbers shift significantly when the product is not FDA approved, with only 39% of physicians reporting they would recommend participation. Strikingly, only 15% would recommend participation if the product's marketing approval has been withdrawn.[46] Additionally, according to a survey of 12 MFMU network sites that participated in the Meis trial, 11 of 12 prospective investigators have indicated an interest in participating in a new study. One of these physicians explicitly stated a belief that Meis settled the question of 17-OHPC's efficacy in the affirmative, and there is no need for a new study.[47]

Therefore, contrary to concerns that continued market availability of Makena would hamper recruitment efforts, *removing* Makena from the market would risk exacerbating recruitment challenges for an RCT. Recruitment in pregnant women and minority groups is known to be challenging because of risk sensitivity and historical mistreatment in clinical research. Knowledge that a clinical trial is being conducted for a drug FDA has withdrawn from the market would lead to even greater hesitancy among potential participants, and potentially thwart any effort to conduct meaningful further study of the drug. Instead, FDA should fully consider whether additional steps are needed, including changes to Makena's labeling. To that end, Covis is committed to working with CDER to narrow the labeling to focus the indication on the most high-risk patients while additional study is undertaken

* * * * *

Withdrawal is unwarranted given the totality of the data and compelling public health considerations, as well as Covis' willingness to conduct an additional confirmatory study in an at-risk patient population. The Meis trial provides strong data supporting the clinical benefit of Makena, including demonstrating a statistically significant reduction in subsequent preterm birth at <37 weeks gestational age ($p<0.001$) as well as <35 and <32 weeks, which are "established" surrogate endpoints or intermediate clinical endpoints strongly correlated with a reduction in neonatal morbidity and mortality. The safety of Makena is supported by the Meis trial, PROLONG, and over a decade of real-world use of the drug. Covis has worked with experts and has identified the high-risk patient population of women with a prior spontaneous preterm birth of <34 weeks as the group most at risk of subsequent preterm birth. With its expert advisors, Covis has developed several proposals for further confirmatory study of Makena based on this high-risk group, including using increase in delivery time (weeks) from randomization as a potential endpoint. For all of these reasons, Covis urges the advisory committee to act in the interest of public health and permit Makena to remain on the market while further study is conducted. Withdrawal of an important treatment option from at risk patients and their obstetricians is not warranted.

[46] *See* Attachment D to the Appendix, *supra* note 12.

[47] *See id.*

III. QUESTIONS PRESENTED

Question 1: *Do the findings from Trial 003 [PROLONG] verify the clinical benefit of Makena on neonatal morbidity and mortality from complications of preterm birth? (for discussion and vote)*

Covis stipulates that the findings from Trial 003 (PROLONG) do not verify the clinical benefit of Makena on neonatal morbidity and mortality from complications of preterm birth. When a confirmatory study fails to provide additional confirmation of clinical benefit, that is the beginning, not the end, of the required analysis.

As CDER has acknowledged, there is nothing in the accelerated approval statute or FDA precedent that mandates that FDA withdraw a drug where a confirmatory study fails. Indeed, in many circumstances it would be inappropriate to withdraw a drug notwithstanding a failed confirmatory study. During its rulemaking process establishing the accelerated approval regulations, for example, CDER recognized that "[a] drug with clear clinical effectiveness in a subset of the population, but not in the population described in labeling, would have its labeling revised to reflect the data. *Withdrawal would be inappropriate under such circumstances*."[48] Moreover, as CDER has elsewhere recognized, "[e]ach decision to withdraw or not withdraw the accelerated approval of a product must be made on its own merits and unique set of facts and considerations for each product."[49] As detailed below, the circumstance presented here do not justify withdrawal.

Question 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? (for discussion and vote).*

Response: *Yes.* Section 506(c) of the FDCA contemplates FDA approval of drugs for serious and life-threatening conditions upon a showing that the drug has an effect on an intermediate clinical endpoint or a surrogate endpoint that is reasonably likely to predict clinical benefit. Trial 002 (Meis) provides substantial evidence of such an effect. More specifically, Trial 002 provides substantial evidence that 17-OHPC reduces the risk of preterm birth at <37 weeks of gestation (the primary endpoint), as well as the risk of preterm birth at <35 weeks and the risk of preterm birth at <32 weeks (both secondary endpoints). The results from Trial 002 are particularly relevant to high-risk U.S. patient populations, including Black women. The results from Trial 002 provide substantial evidence of effectiveness for Makena's intended use and a strong justification for keeping Makena on the market pending further study.

In addition to the previously performed analyses on the Meis study, Covis has reanalyzed the data from a time-to-event perspective, which is further explained below in Section VII.A.2.a below. The data were analyzed both for the overall Meis population and also for a higher-risk population having a prior spontaneous preterm birth <34 weeks. For both of these populations, we show the Kaplan-Meier plots along with hazard ratios. The x-axis shows gestational age,

[48] 57 Fed. Reg. 58,942, 58,956 (Dec. 11, 1992) (emphasis added).

[49] Witten Apr. 21, 2022, *supra* note 1, at 2.

with an "event" including miscarriage, stillbirth, and live birth. The Kaplan-Meier analysis below was censored at week 37 for all patients given that was the primary endpoint for Meis.

Figure 2
Time-to-Event Analysis for Overall Meis Population

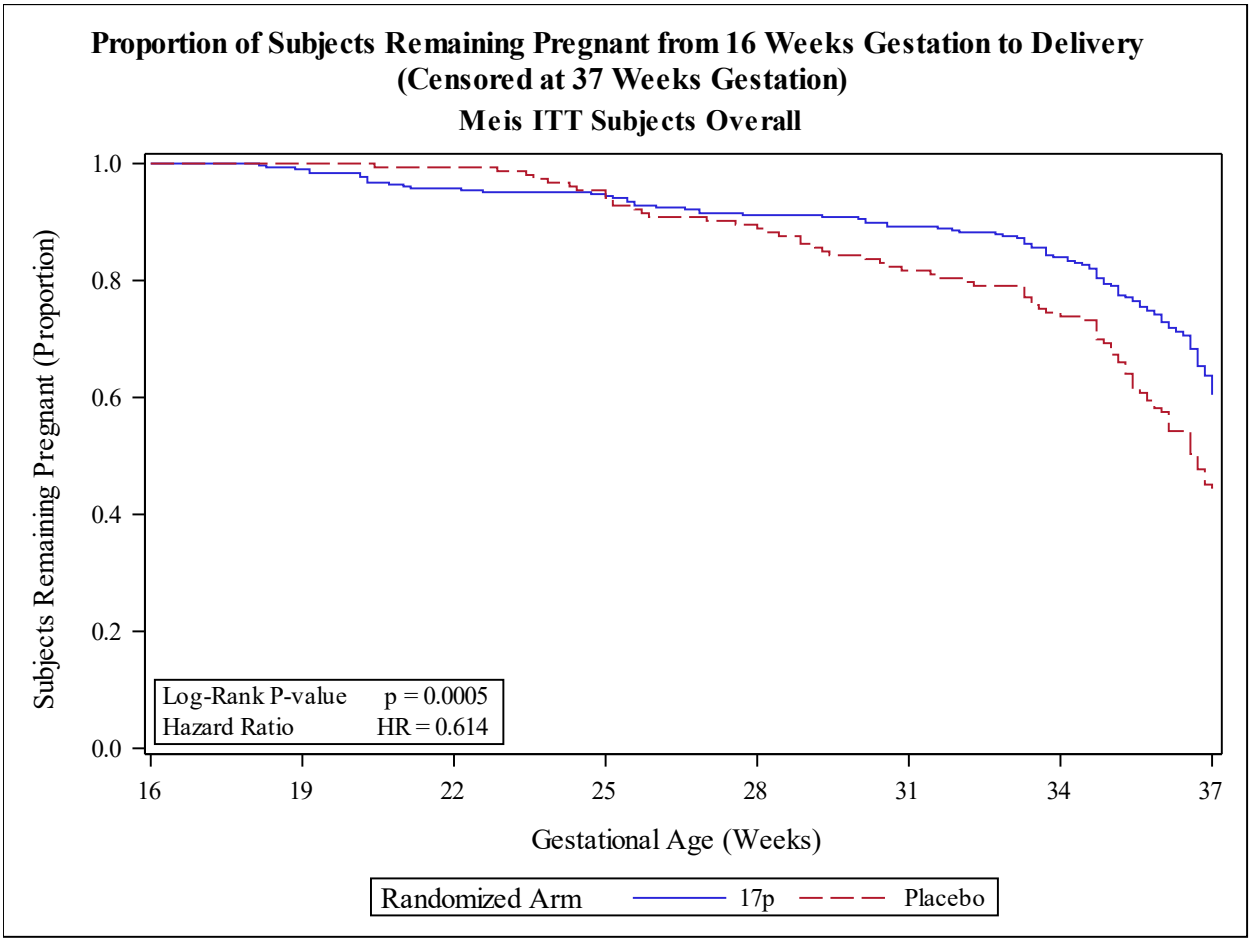
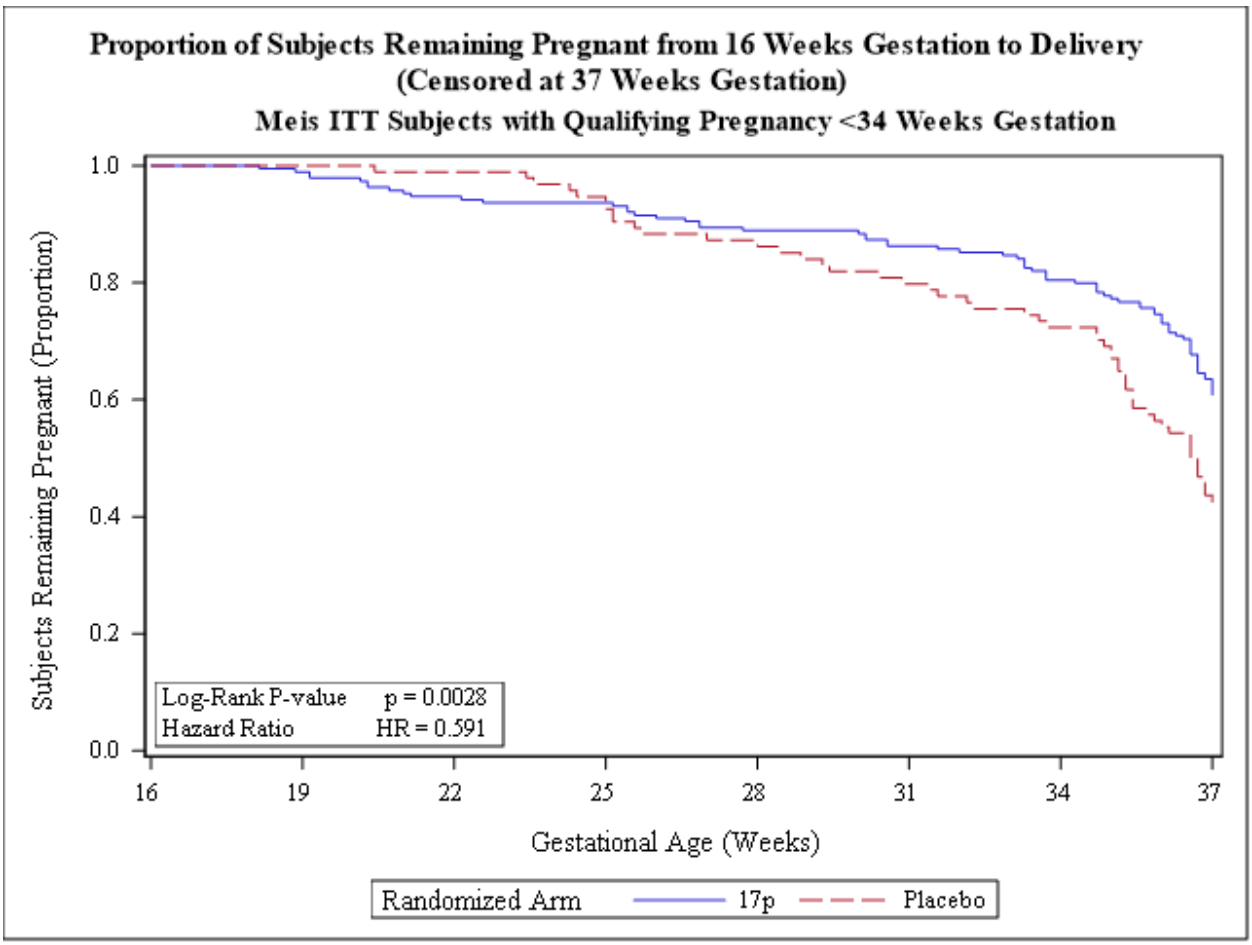


Figure 3
Time-to-Event Analysis for Meis Population With Prior SPTB <34 Weeks



The hazard ratios in both cases show a consistent, strong reduction in event rate for both populations, and the separation of the curves clearly show the benefit from 17-OHPC treatment that accrues in the patient population during the gestational time-course. This analysis demonstrates not only that there is efficacy in the overall population (which has been well-documented), but also that there is efficacy in the high-risk population of prior SPTB <34 weeks, which aligns with Covis' proposed inclusion criterion for a subsequent clinical study.

Although the results from Trial 002 are sufficient to support continued marketing, those data are not the only available evidence of effectiveness. Other studies provide additional support, including studies demonstrating an extension of the gestational period in women treated with 17-OHPC. These studies are further described in Section VII.A.5 of this document.

Even though Trial 003 did not verify the effect seen in Trial 002, Trial 003 did not and cannot invalidate the effectiveness seen in Trial 002. Rather, Trial 003 is best understood as a study that failed to confirm clinical benefit as a result of studying a population with a lower-risk population than that studied in Meis, which led to a much lower-than-anticipated event rate. In addition, a super-majority of patient enrollment outside of the U.S. at trial sites, such as those in

Ukraine and Russia, with standards of care that differ from U.S. standards, including standards that led these sites to inadequately control for gestational age, further reduces the relevance of Trial 003. The efficacy seen in Trial 002 is particularly meaningful for high-risk U.S. patient populations, including Black women, who are especially important to FDA's benefit-risk assessment.

Moreover, an exploratory post-hoc analysis of the PROLONG-US data[50] supports Covis' position that there is efficacy in the high-risk population.[51] Covis analyzed the time from randomization to birth or 35 weeks (whichever came first) for 17-OHPC versus placebo as a measure of "weeks gained on 17-OHPC." The time-from-randomization was capped at 35 weeks gestational age in order to ensure that weeks gained were relevant for neonatal development and would contribute to a reduction of neonatal morbidity and mortality (further details in the Appendix). As Table 3 below shows, an examination of the PROLONG-US data reveals a clear numerical increase in weeks gained for Makena versus placebo for subgroups with a large number of preterm birth risk factors. As shown in Manuck et al. (2016) and Richter et al. (2019), this is clinically significant as the addition of 1-2 weeks of gestational age prior to week 35 is associated with marked reduction in neonatal morbidities.[52]

[50] *See* Appendix Section I.D for an explanation of the statistical methodology used to perform this analysis.

[51] *See* Spong et al., *Progesterone for prevention of recurrent preterm birth:*

Impact of gestational age at previous delivery, 193 A. J. OBSTET. GYNECOL. 1127 (2005); Mercer et al., *Are Women with Recurrent Spontaneous Preterm Births Different from Those Without Such History?*, 194 A. J. OBSTET. GYNECOL. (2006).

[52] *See* Tracy A. Manuck et al., *Preterm Neonatal Morbidity and Mortality by Gestational Age: A Contemporary Cohort*, 215 A. J. OBSTET. GYNECOL. 103.e1–103.e14 (2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4921282/; Lindsay A. Richter et al., *Temporal Trends in Neonatal Mortality and Morbidity Following Spontaneous and Clinician-Initiated Preterm Birth in Washington State, USA: A Population-Based Study*, 9 BMJ OPEN e023004 (2019). Covis has also performed an exhaustive literature search regarding preterm morbidity incidences at various gestational ages and can provide additional information as well as validation of the tables contained herein at CDER's request.

Table 3
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Known Risk Factor Subgroup Among Subjects Randomized Prior to 20 Weeks GA for PROLONG-US

Risk Factor Subgroup	N Total	Estimated treatment effect* (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
Overall	389	0.49	-0.04	1.01	0.0684
Most Recent Prior Spontaneous Delivery at GA<35 (mrpGA<35)	137	1.30	0.30	2.29	0.0113
Black Subjects with mrpGA<35	51	1.57	-0.28	3.42	0.0936
Subjects with Inter-pregnancy Interval <5 Years (IPINT<5) and mrpGA<35	112	1.55	0.34	2.76	0.0126
Subjects with More than One Prior sPTB<37 (MTO37) and mrpGA<35	23	0.99	-0.74	2.72	0.2470
Subjects with IPINT<5 and MTO37 and mrpGA<35	16	2.08	-0.54	4.69	0.1099
Black Subjects with IPINT<5 and mrpGA<35	38	1.75	-0.77	4.26	0.1673
Black Subjects with MTO37 and mrpGA<35	9	-0.10	-0.57	0.37	0.6056

* Within group estimates for the 17P treatment effect (weeks gained from randomization, capped at GA=35) based on model including: Treatment, Mean GA of Prior Spontaneous Deliveries (mGA), and GA at Randomization.

As shown in the Appendix, this analysis was repeated for the Meis population, which showed a similar pattern in terms of the increase in weeks gained with increasing preterm birth risk of the pregnant women. Other analyses on PROLONG-US presented in the Appendix show a systematic improvement of weeks gained for subjects with increasingly worse prior preterm birth history. Contrary to CDER's assertions that the Meis and PROLONG studies altogether do not provide substantial evidence of effectiveness, these analyses strongly suggest a consistent picture that 17-OHPC may be the most effective for the highest-risk patients and highlight the need for further focused studies on high-risk subgroups.

Question 3A: *Should FDA allow Makena to remain on the market? (for discussion). As part of that discussion, you may discuss:*

> *1. whether the benefit-risk profile supports retaining the product on the market;*
>
> *2. 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: *Yes*—Makena should remain on the market pending further study. FDA is not required to withdraw accelerated approval if a post-approval study fails to verify clinical

benefit.[53] The accelerated approval statute, instead, states that FDA "may" withdraw a drug under these circumstances, vesting FDA with the discretion and the obligation to ensure that its decisions regarding accelerated approval are in the best interests of the public health.[54] This permissive approach stands in contrast to other aspects of the Food, Drug, and Cosmetic Act (FDCA) that use the mandatory term, "shall" to explicitly constrain FDA's discretion.[55]

Dr. Richard Pazdur, Director of FDA's Oncology Center of Excellence explained the importance of this approach:

> There are many reasons that a trial fails and that could be the size of the trial, the endpoint they used, the population that they defined. . . . To remove a drug from the market or even an indication is a big deal and not in the public's best interest if you can understand why that trial failed. . . . We have to have that flexibility rather than just a draconian approach.[56]

Similarly, Dr. Billy Dunn, Director of CDER's Office of Neuroscience, recently described FDA's regulatory remine as one that cannot be decoupled from regulatory flexibility: "Our underlying legal authority is clear in not only allowing, but also endorsing and encouraging the application of regulatory flexibility in the setting of serious and life-threatening diseases…."[57] Dr. Dunn highlighted the importance of regulatory flexibility when considering data supporting a drug's effectiveness in areas where there is serious unmet need:

> In appropriate circumstances, such as serious and life-threatening diseases and settings of substantial unmet need, regulatory flexibility applied to assessments of effectiveness means increased tolerance for concluding that a drug is effective when there is residual uncertainty that the drug may not actually be effective, which would be a conclusion at risk of being a false positive and decreased tolerance for concluding that a drug is ineffective when there is residual

---

[53] *See* 21 U.S.C. § 356(c)(3).

[54] *See id*.

[55] *See e.g.* 21 U.S.C. § 355(e). (Discussing drugs that pose an "immediate" threat to public health, the statute states that "The Secretary shall, after due notice and opportunity for hearing to the applicant, withdraw approval of an application with respect to any drug under this section if the Secretary finds (1) that clinical or other experience, tests, or other scientific data show that such drug is unsafe for use under the conditions of use upon the basis of which the application was approved….")

[56] Derrick Gingery, *US FDA Pushes Back Against Critics: Breakthrough Is Not A Drug 'Beauty Contest,'* THE PINK SHEET: INFORMA PHARMA INTELLIGENCE (Dec. 10, 2019) (Pazdur further asserted that "[f]ailure to confirm clinical benefit in a completed trial may reflect the possibility that the drug does not in fact confer clinical benefit, but it also may reflect, for example, unforeseen limitations in trial design, rather than clear evidence of lack of effectiveness. The most appropriate regulatory approach must be governed by the unique factors of the particular case.") ,https://pink.pharmaintelligence.informa.com/PS141340/US-FDA-Pushes-Back-Against-Critics-Breakthrough-Is-Not-A-Drug-Beauty-Contest.

[57] *See* Sarah Karlin-Smith, *Tolerating False Positives: Amylyx, FDA, And The Legal Case For Broad Regulatory Flexibility*, THE PINK SHEET (Sep. 8, 2022), https://pink.pharmaintelligence.informa.com/PS146975/Tolerating-False-Positives-Amylyx-FDA-And-The-Legal-Case-For-Broad-Regulatory-Flexibility.

> uncertainty that the drug may actually be effective, which would be a conclusion at risk of being a false negative.[58]
>
> It is important to note when considering one study plus confirmatory evidence, the single study may be a study of conventional persuasiveness rather than the highly persuasive study we prefer to see when considering a true single study in isolation. The degree of persuasiveness required for approval may be influenced by many things including the seriousness of the disease, whether there is an unmet need, and the character of the confirmatory evidence.[59]

As this view reflects, a decision concerning the withdrawal of a drug approved under accelerated approval must be made with great care. This is particularly true where, as here, the therapy has a proven track record of success recognized by pertinent medical specialists, has been the standard of care recognized by respected medical societies, there is no other FDA-approved therapy for the approved indication, and alternative therapies would pose additional risks for vulnerable patients. Even where a confirmatory study is viewed as a failed study, there is no requirement that a drug or indication be summarily withdrawn; instead, there are a number of important legal, scientific, and public health considerations that govern the proper path forward.

The benefit-risk profile of Makena is positive and supports continued availability. The Meis trial (Trial 002) alone provides substantial evidence of effectiveness; so much so that it could have supported a full approval. The Meis trial is further supported by other studies, and it is not invalidated by the inconclusive results from PROLONG (Trial 003). Nor is the available evidence of effectiveness outweighed by safety concerns. The Meis trial (Trial 002) showed that Makena was not associated with an increased incidence of pregnancy complications, such as preeclampsia, gestational diabetes, or clinical chorioamnionitis. Additionally, the incidence of neonatal death was lower in the 17-OHPC group as compared with placebo, and there was no difference in combined fetal and neonatal mortality between the two treatment groups. PROLONG (Trial 003) further ruled out a doubling of risk of fetal or early infant death, thereby resolving a potential safety signal from the Meis trial.[60] Trial 003 also showed comparable results in the rate of gestational diabetes in Makena vs. placebo groups, and other adverse events were low and comparable between treatment groups. Other studies similarly support the conclusion that Makena is generally safe when used as indicated.

As described more fully below and in the attached Appendix, Covis is prepared to undertake an additional study to verify and describe the efficacy of Makena in reducing preterm birth. We believe that this would be best conducted as a randomized control trial targeting a high-risk population with high incidence rates of spontaneous preterm birth, and also are willing to undertake an observational study.

---

[58] *Id.*

[59] Dr. Billy Dunn, Opening Statement, Meeting of the Peripheral and Central Nervous System Drugs Advisory Committee Meeting (Sept. 7, 2022), https://www.youtube.com/watch?v=6PMOyqd6WfA.

[60] *See* Sibai et al., *supra* note 9 at 622, 624.

Question 3B: *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? (for vote)*

Response: *Yes*. The Agency should not withdraw the only FDA-approved therapy for reducing the risk of preterm birth. In light of Makena's positive benefit-risk profile and the substantial evidence of effectiveness provided by U.S.-focused Trial 002, Makena should remain available to patients who need it while additional study is done. Moreover, recent legal changes with respect to abortion in the United States are likely to increase the need for therapies to support women at risk of preterm birth, which will become an unmet need should the Agency withdraw approval for Makena. Finally, FDA should consider whether it would be appropriate to narrow the indication to high-risk women.

IV. PROCEDURAL HISTORY

Makena is indicated for reducing the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. Makena was designated as an orphan drug in 2007 for the prevention of preterm birth in singleton pregnancies.

CDER granted accelerated approval to Makena or 17-OHPC in 2011 based on the strength of the Meis trial (Trial 002), which demonstrated that the risk of delivery prior to 37 weeks of gestation was significantly reduced in the patients treated with 17-OHPC compared to those receiving placebo, among women with a singleton pregnancy who have a history of singleton spontaneous preterm birth. FDA also acknowledged that, in addition to the demonstrated reduction in preterm birth at <37 weeks, the Meis trial demonstrated a significant reduction in preterm birth at <35 and <32 weeks, in women with a prior spontaneous preterm birth.

CDER's accelerated approval of Makena was based upon its finding that the Meis trial had compellingly demonstrated an effect on what the Center incorrectly labeled as a surrogate endpoint, i.e., reduction in preterm births <37 weeks gestation. Makena was approved for the indication of reduction of the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth.[61] A once weekly treatment is indicated to begin between 16 weeks, 0 days and 20 weeks, 6 days of gestation and to continue until week 37 or delivery, whichever occurs first.[62] In addition to the original Makena intramuscular multi-dose formulation, and a preservative-free single-dose formulation, CDER also has approved five generic versions of Makena,[63] as well as the Makena Auto-Injector for subcutaneous use.[64]

---

[61] *See* Makena Prescribing Information (Feb. 2011), https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021945s000lbl.pdf.

[62] *See id.*

[63] American Regent, Inc.'s ANDAs 210723 (Jun. 21, 2018) and 210724 (Aug. 9, 2019); Slayback Pharma LLC's ANDAs 210618 (Dec. 28, 2018) and 210877 (Mar. 22, 2019); Sun Pharmaceuticals Industries' ANDA 208381 (Apr. 9, 2019); Eugia Pharma Specialties Limited's ANDAs 211070 and 211071 (Apr. 16, 2019); Aspen Pharma USA Inc.'s ANDA 211777 (Aug. 8, 2019).

[64] *See* CDER, NDA 021945/S-012 Supplement Approval Letter (Feb. 14, 2018), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2018/021945Orig1s012ltr.pdf.

As a condition of Makena's accelerated approval, its previous sponsor, AMAG Pharmaceuticals, Inc. (AMAG), completed a confirmatory trial, known as the PROLONG trial (also called Trial 003). The PROLONG trial did not meet its two pre-specified objectives.[65] As detailed below, the conduct of the PROLONG trial was flawed and it, therefore, could not confirm the Meis trial findings. The great majority of patients were enrolled outside the United States, with inadequate controls for differences in measuring gestational age, assessing patients for eligibility criteria related to prior preterm birth, and other criteria related to local standard of care, which was a deficient design that should have been recognized from the outset. CDER, which agreed on the study design and received regular updates on trial enrollment and conduct from Makena's previous sponsors, including through annual reports,[66] also was, or should have been, aware of these deficiencies.[67] Further monitoring the event rate in the overall study would have confirmed the low event rate and highlighted the challenges with the study much sooner. PROLONG studied a significantly different patient population from the Meis trial, with much lower underlying rates of preterm birth and markedly different social and demographic characteristics. In addition, given the different patient population, PROLONG was inadequately powered to assess the efficacy of 17-OHPC.

Shortly after PROLONG was completed, in October 2019, CDER convened its Bone, Reproductive and Urologic Drugs Advisory Committee (BRUDAC) to consider the PROLONG trial results and the status of Makena's approval. BRUDAC was composed of sixteen voting members, only six of whom treated pregnant women in their clinical practice. The remaining ten members consisted of four epidemiologists/statisticians, three physicians who treated infants and children, an endocrinologist, a patient representative, and a consumer representative. After extensive discussion, BRUDAC reached a divided conclusion: nine members recommended that CDER pursue withdrawal of approval for 17-OHPC, and seven members recommended that the CDER leave 17-OHPC on the market and require new confirmatory data.[68] Notably, five of the six experts with actual experience treating pregnant women voted *against* withdrawing Makena from the market and supported further study of the drug.

After the BRUDAC vote, AMAG sought to meet with CDER to discuss a data-driven path forward, including several proposals for initiating additional studies. CDER denied that request by letter dated March 11, 2020, stating that it was, "premature to hold the requested meeting at this time," considering the Center's ongoing deliberations.[69] CDER did not grant or entertain dialogue with AMAG about options for further data generation.

Instead, on October 5, 2020, nearly a year after the divided BRUDAC meeting and eight months after denying AMAG's meeting request—and with no further communication with AMAG—CDER sent a NOOH to AMAG and generic manufacturers of 17-OHPC, initiating the

---

[65] Blackwell, *supra* note 12.

[66] FDA regulations require sponsors of approved drugs to submit an annual report to the Agency that provide a status report of certain postmarketing studies, including postmarket studies required under the accelerated approval program. *See* 21 C.F.R. § 314.81(b)(2)(vii).

[67] *See* NDA 21945 Summary Review, *supra* note 37 at 19, 26.

[68] *See generally* Advisory Committee Meeting Transcript, *supra* note 13 at 08-09.

[69] *See* Letter from Christine Nguyen, Acting Dir., Div. of Bone, Reproductive and Urologic Products, ODE III, CDER to Helen Milton, Senior Vice President, Regulatory Affairs, AMAG Pharma USA, Inc. (Mar. 11, 2020).

process for withdrawing the drug. AMAG responded within the allotted fifteen days to request a hearing. Shortly thereafter, on November 16, 2020, Covis Pharma Group (Covis) announced its acquisition of AMAG. On December 4, 2020, AMAG submitted its response to the NOOH.[70] On March 5, 2021, AMAG transferred all rights and ownership of the Makena NDA to Covis, and duly notified FDA of the change in ownership.[71]

On August 18, 2021, then-Chief Scientist, RADM Denise Hinton—to whom then-Acting Commissioner of Food and Drugs, Janet Woodcock, M.D., had delegated the matter following her recusal—granted the request for a hearing and appointed Celia M. Witten, Ph.D., M.D., Deputy Director of the Center for Biologics Evaluation and Research (CBER), as the presiding officer. In granting Covis' request, RADM Hinton recognized that the sponsor's December 2020 submission "provides specific challenges to the factual and scientific bases underlying CDER's proposal" and "raises genuine and substantial issues of fact appropriate for a hearing."[72] RADM Hinton also concluded that BRUDAC was the appropriate advisory committee to participate in the hearing.[73] Earlier this year, the advisory committee's charter was renewed, and its name changed to Obstetrics, Reproductive and Urologic Drugs Advisory Committee (ORUDAC), with the new name reflecting the transfer of responsibility for indications related to diseases such as osteoporosis and metabolic bone disease.[74]

On April 21, 2022, Dr. Witten determined that the hearing would take place for two and a half days in a virtual format.[75] Dr. Witten granted Covis' request to include within the scope of the hearing Covis' proposals for further study evaluating Makena's effectiveness that it has continued to develop since the sponsor's last submission in December of 2020. In doing so, Dr. Witten "recognize[d] that the availability and feasibility of certain study designs in light of existing data and information regarding Makena's effectiveness might have some relevance in determining whether FDA should withdraw approval while additional confirmatory studies are conducted."[76]

On June 13, 2022, Dr. Witten announced the proposed dates of October 17-19, 2022, for the hearing as well as voting and discussion questions to be posed to ORUDAC.[77] Although Covis proposed to stipulate to voting question one—that the findings from Trial 003 (PROLONG) do not verify the clinical benefit of Makena on neonatal morbidity and mortality

---

[70] 21 C.F.R. § 314.530(c)(3).

[71] *See* 21 C.F.R. § 314.72(a) ("An applicant may transfer ownership of its application. . . .").

[72] *See* Letter from RADM Denise Hinton to Rebecca Wood and Vincent Amatrudo, Docket No. FDA-2020-N-2029-0072, at 5 (Aug. 18, 2021), https://www.regulations.gov/document/FDA-2020-N-2029-0072.

[73] *See id.* at 7-8.

[74] *See* Advisory Committee; Obstetrics, Reproductive and Urologic Drugs Advisory Committee; Renewal, 87 Fed. Reg. 16477 (Mar. 23, 2022).

[75] *See* Witten Apr. 21, 2022, *supra* note 1. Covis requested an in-person hearing. *See* Letter from Rebecca Wood, Sidley Austin LLP, to Celia Witten, Ph.D., M.D., Presiding Officer, Docket No. FDA-2020-N-2029-0204 (Feb. 14, 2022), https://www.regulations.gov/document/FDA-2020-N-2029-0204.

[76] *Id.* at 6.

[77] *See* Witten Apr. 21, 2022, *supra* note 1.

from complications of preterm birth—to allow a greater focus on the matters that are in dispute,[78] CDER declined to agree to a joint stipulation.[79]

V. LEGAL AND REGULATORY FRAMEWORK

In 1992, in response to the HIV/AIDS epidemic, FDA created the accelerated approval program to expedite approval of promising new therapies for serious and life-threatening illnesses.[80] The program, which FDA promulgated via regulation, permitted the Agency to approve drugs that treat serious or life-threatening illnesses based upon evidence that the drug had an effect on a surrogate endpoint reasonably likely to predict clinical benefit.[81] In 1997, Congress codified the accelerated approval program by ratifying FDA's fast-track regulations and accelerated approval regulations together as one "fast track" statutory scheme under Section 506 of the FDCA.[82] Fifteen years later, Congress further expanded the accelerated approval program as part of the Food and Drug Administration Safety and Innovation Act (FDASIA) of 2012.[83] Emphasizing the importance of the "expedited development and review of innovative new medicines to address unmet medical needs for serious or life-threatening diseases or conditions,"[84] Congress encouraged FDA to "utilize innovative and flexible approaches" to grant accelerated approvals.[85]

Pursuant to that policy, Section 506 vests FDA with discretion over multiple aspects of the approval process. Under Section 506, FDA may grant accelerated approval to:

> a product for a serious or life-threatening disease or condition . . . upon a determination that the product has an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit, or on a clinical endpoint that can be measured earlier than irreversible morbidity or mortality, that is reasonably likely to predict an effect on irreversible morbidity or mortality or other clinical benefit, taking into account the severity, rarity, or prevalence of the condition and the availability or lack of alternative treatments.[86]

---

[78] Letter from Rebecca Wood, Sidley Austin LLP, to Celia Witten, Ph.D., M.D., Presiding Officer, Docket No. FDA-2020-N-2029-0220 (Jul. 1, 2022), https://www.regulations.gov/comment/FDA-2020-N-2029-0220.

[79] E-mail from Christine Hunt to Rebecca Wood, Sidley Austin LLP, Docket No. FDA-2020-N-2029-0222 (Jul. 27, 2022), https://www.regulations.gov/document/FDA-2020-N-2029-0222.

[80] *See* New Drug, Antibiotic, and Biological Drug Product Regulations; Accelerated Approval, 57 Fed. Reg. 58942 (Dec. 11, 1992) (promulgating FDA's accelerated approval rule).

[81] *See generally* 21 C.F.R. §§ 314.500 *et seq*.; *see also* 21 C.F.R. part 601 subpart E.

[82] *See* Food and Drug Administration Modernization Act (FDAMA), Pub. L. No. 105-115, § 112, 111 Stat. 2296, 2309-12. (1997).

[83] *See* Food and Drug Administration Safety and Innovation Act (FDASIA), Pub. L. No. 112-144 § 901, 126 Stat. 993, 1083-84 (2012). Whereas FDA's version had been limited to drugs that "provide meaningful therapeutic benefit to patients over existing treatments," 21 C.F.R. § 314.500, Congress loosened this requirement, directing FDA only to "tak[e] into account … the availability or lack of alternative treatments," 126 Stat. at 1084 (codified at 21 U.S.C. § 356(c)(1)(A).

[84] FDASIA § 901(a)(1)(C), 126 Stat. at 1082.

[85] 21 U.S.C. § 356(e).

[86] *Id.* § 356(c)(1)(A).

The statute clearly distinguishes between a surrogate endpoint and a "clinical endpoint that can be measured earlier than irreversible morbidity or mortality," which is commonly referred to as an intermediate clinical endpoint. Both a surrogate endpoint and intermediate clinical endpoint can support accelerated approval: specifically, a surrogate endpoint or intermediate clinical endpoint that is reasonably likely to predict clinical benefit can be granted accelerated approval, but may be subject to certain requirements, including a postmarket study to determine if the endpoint, in fact, predicts the clinical benefit. Significantly however, in contrast to a surrogate endpoint that needs to be "validated" before it can support traditional approval, CDER's guidance states that an intermediate clinical endpoint for products for serious conditions "will *usually* be considered under traditional approval procedures. Approvals based on such clinical endpoints will be considered under the accelerated approval framework only when it is essential to determine effects on IMM or other clinical benefit in order to confirm the predicted clinical benefit that led to approval."[87] This distinction is important because, as CDER has acknowledged, it approved Makena based on an "intermediate clinical endpoint"; the data giving rise to that approval therefore could have supported a full approval.

Importantly, Section 506 explicitly contemplates a benefit-risk judgment taking into account such factors as the severity, rarity, or prevalence of the condition and the availability or lack of alternative treatments.[88] The statute recognizes that in determining whether an endpoint is "reasonably likely to predict clinical benefit," FDA may consider "epidemiological, pathophysiological, therapeutic, pharmacologic, or other evidence developed using biomarkers, for example, or other scientific methods or tools."[89] Additionally, the statute authorizes (but does not require) FDA to impose conditions on accelerated approvals, including requiring the drug's sponsor to "conduct appropriate post-approval studies to verify and describe the predicted effect on irreversible morbidity or mortality or other clinical benefit."[90]

Underscoring the importance it attached to the accelerated approval framework, Congress also afforded FDA discretion in deciding when such approval should be withdrawn. Section 506 defines certain circumstances in which FDA "*may* withdraw approval of a product approved under accelerated approval."[91] Those circumstances include when a required confirmatory study "fails to verify and describe" the "predicted effect on irreversible morbidity or mortality or other clinical benefit," or when "other evidence demonstrates that the product is not safe or effective under the conditions of use."[92] In such circumstances, FDA may—but need not—decide to

---

[87] Expedited Programs Guidance at 18 (emphasis added).

[88] In its recent draft guidance on Benefit-Risk Assessment for New Drug and Biological Products, FDA similarly recognized its ability to "incorporate[] broader public health considerations" into its benefit-risk assessment for new drugs. FDA, DRAFT GUIDANCE FOR INDUSTRY, BENEFIT-RISK ASSESSMENT FOR NEW DRUG AND BIOLOGICAL PRODUCTS, at 5 (Sept. 2021), https://www.fda.gov/media/152544/download. FDA stated that it "recognizes that when a drug is developed to treat serious diseases for which there are few or no approved therapies, greater uncertainty or great risks may be acceptable." *Id.* at 11. FDA also emphasized the role of "patient experience data" in informing benefit-risk assessments, *id.* at 1, noting that "[p]atient experience data can inform nearly every aspect of FDA's benefit-risk assessment throughout the drug lifecycle." *id.* at 11.

[89] 21 U.S.C. § 356(c)(1)(B).

[90] *Id.* § 356(c)(2)(A).

[91] *Id.* § 356(c)(3) (emphasis added).

[92] *Id.* § 356(c)(3)(B)-(C).

withdraw approval.[93] The statute makes clear that FDA may instead permit the drug to stay on the market pending further study.

At the same time, FDA's discretion is not unbounded. Rather, in exercising its discretion, FDA must engage in "reasoned decision-making,"[94] lest its decision be "arbitrary" or "capricious" and therefore unlawful.[95] To satisfy this requirement of reasoned decision-making, FDA must examine "the relevant data" and articulate "a satisfactory explanation" for its decision, "including a rational connection between the facts found and the choice made."[96] FDA also must treat similarly situated parties similarly, and "provide either good reasons[] or a reasoned analysis" justifying any break with past practice or policy.[97] Conversely, FDA may not rely on factors Congress has not intended it to consider, fail to consider an important aspect of the problem, or offer an explanation for its decision that runs counter to the evidence before it.[98] As courts have recognized, FDA also must consider the reliance interests of pregnant women who have already been prescribed the drug.[99] FDA must also consider all relevant aspects of the problem, including the potential to inadvertently encourage off-label use of older, poorly labeled products and the potential to encourage reliance on less safe compounded products.

In addition, the 21st Century Cures Act directs FDA to incorporate patient perspectives into its decision-making. Among other things, Congress has directed FDA to consider "patient preferences with respect to treatment,"[100] and emphasizes the importance of efforts to "take into account women and minorities and … focus[] on reducing health disparities,"[101] including by greater focus on "safe and effective therapies" for pregnant women.[102]

In practice, withdrawal of accelerated approval is rarely pursued by FDA. In the 30 years that accelerated approval has been available, CDER has approved more than 270 drugs or indications under the accelerated approval authorities, and has withdrawn fewer than 10 drugs or indications after the failure of confirmatory studies to verify and describe the predicted clinical benefit of the drug.[103] According to CDER, ~40% of its accelerated approvals (112 of 278) have not yet been converted to full approvals, with a substantial percentage involving delayed or

---

[93] *Id.* § 356(c)(3).

[94] *Dep't of Homeland Security v. Regents of the Univ. of Cal.*, 140 S. Ct. 1891, 1905 (2020).

[95] 5 U.S.C. § 706(2)(A).

[96] *Dep't of Commerce v. New York*, 139 S. Ct. 2551, 2569 (2019).

[97] *Whitman-Walker Clinic, Inc. v. U.S. Dep't of Health & Hum. Servs.*, 485 F. Supp. 3d 1, 38 (D.D.C. 2020) (citing *FCC v. Fox Tele. Stations, Inc.*, 556 U.S. 502, 515 (2009), and *Motor Vehicle Mfrs. Ass'n of U.S., Inc. v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29, 42 (1983)).

[98] *State Farm*, 463 U.S. at 43.

[99] *Regents*, *supra* note 94, 140 S. Ct. at 1913 ("When an agency changes course … it must 'be cognizant that longstanding policies may have engendered serious reliance interests that must be taken into account.'" (quoting *Encino Motorcars, LLC v. Navarro*, 136 S. Ct. 2117, 2126 (2016))); *MediNatura, Inc. v. FDA*, 998 F.3d 931, 940-41 (D.C. Cir. 2021).

[100] 21st Century Cures Act, Pub. L. 114-255 (2016), § 3001.

[101] *Id*. § 2031.

[102] *Id*. § 2041.

[103] *See* CDER Drug and Biologic Surrogate Endpoint Approvals; *see also* Sutter, *supra* note 43. https://pink.pharmaintelligence.informa.com/PS125183/Accelerated-Approval-Withdrawals-Through-The-Years.

pending confirmatory studies.[104] This is further supported by a 2019 article in JAMA Internal Medicine, which concluded that in the context of oncology drugs, only ~20% of accelerated approval drugs have had successful confirmatory trials that reported an improvement in overall survival, confirming clinical benefit.[105] As detailed in this study, ~20% of accelerated approval drugs have had confirmatory studies that used surrogate endpoints that were the same as those used in the preapproval studies, and another ~20% have had confirmatory studies that used surrogate endpoints that were different from those used in the preapproval studies.[106]

FDA has been clear that it will explore the full range of available regulatory options in those cases where the confirmatory study fails to demonstrate clinical benefit, with priority placed on the interests of the patient.[107] FDA explained its approach in detail in a letter sent to the Government Accountability Office (GAO) in 2009, in response to GAO's recommendation that the Agency:

> Clarify the conditions under which the agency would utilize its authority to expedite the withdrawal of drugs approved based on surrogate endpoints under the accelerated approval process if sponsors fail to complete required confirmatory studies with due diligence, or if studies are completed, but fail to demonstrate the clinical effectiveness of the drugs.[108]

FDA further rejected GAO's recommendation, stating that it would be "difficult, if not impossible, to provide further clarification as to when it might utilize its authority to expedite withdrawal of a drug approved on the basis of surrogate endpoints."[109] Instead, FDA outlined a set of key guiding principles for when withdrawal is appropriate.

---

[104] *See* CDER Drug and Biologic Surrogate Endpoint Accelerated Approvals.

[105] *See* Bishal Gyawali, et al., *Assessment of the Clinical Benefit of Cancer Drugs Receiving Accelerated Approval,* 179 JAMA INTERN MED 906-13 (2019), https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2733561.

[106] *See id.* at 910. In at least one case, the drug received full approval even though confirmatory trials did not confirm its clinical benefit. *See id.* at 911. According to another recent article that examined drugs that received accelerated approval from CDER between 1992 and 2021, there are at least three cases in which trial enrollment challenges and/or a change in the treatment landscape rendered the confirmatory study unfeasible. In these instances, CDER indicated that there were alternative options to confirm clinical benefit. *See* Ginny Beakes-Read et al., *Analysis of FDA's Accelerated Approval Program Performance, December 1992-December 2021*, 56 THER. INNOV. REGUL. SCI. 698 (2022), https://pubmed.ncbi.nlm.nih.gov/35900722/.

[107] Indeed, in a variety of contexts, CDER takes a similar approach, including when considering whether to rescind a previously granted breakthrough therapy designation (BTD). According to a recent draft guidance, one potential basis for BTD rescission is "[e]merging data for the designated drug [that] no longer support[s] a finding that 'preliminary clinical evidence indicates that the drug may demonstrate substantial improvement over existing therapies'" or "if a phase 3 trial intended to definitively show the designated drug's effect fails to meet its primary endpoint." FDA, DRAFT GUIDANCE FOR INDUSTRY, CONSIDERATIONS FOR RESCINDING BREAKTHROUGH THERAPY DESIGNATION, 2-3 (Jun. 2022), https://www.fda.gov/media/159359/download. Yet, the draft guidance makes clear that even in such circumstances, CDER has significant discretion to maintain the status quo and that CDER's decision "will depend on the facts specific to that drug development program." *Id.* This is particularly true if "initial data were promising, and there are significant issues with the conduct and design of a subsequent study." *Id.*

[108] U.S. Gov't Accountability Office, New Drug Approval, FDA Needs to Enhance Its Oversight of Drugs Approved on the Basis of Surrogate Endpoints, 36 (Sept. 2009), https://www.gao.gov/new.items/d09866.pdf [hereinafter GAO Report].

[109] *Id.* at 37.

First, FDA explained that it is aware a confirmatory study may fail for reasons that have nothing to do with the clinical efficacy of the product studied:

> Failure to confirm clinical benefit in a completed trial may reflect the possibility that the drug does not in fact confer clinical benefit, but it also may reflect, for example, unforeseen limitations in trial design, rather than clear evidence of lack of effectiveness. The most appropriate regulatory approach must be governed by the unique factors of the particular case.[110]

Dr. Richard Pazdur, Director of FDA's Oncology Center of Excellence, has similarly acknowledged this consideration:

> There are many reasons that a trial fails and that could be the size of the trial, the endpoint they used, the population that they defined. . . . To remove a drug from the market or even an indication is a big deal and not in the public's best interest if you can understand why that trial failed. . . . We have to have that flexibility rather than just a draconian approach.[111]

Dr. Billy Dunn, Director of CDER's Office of Neuroscience, made similar comments during a recent Advisory Committee meeting where he emphasized that regulatory flexibility is a "a well-recognized concept" and a "foundational construct" of FDA's regulatory regime:

> In appropriate circumstances, such as serious and life-threatening diseases and settings of substantial unmet need, regulatory flexibility applied to assessments of effectiveness means increased tolerance for concluding that a drug is effective when there is residual uncertainty that the drug may not actually be effective, which would be a conclusion at risk of being a false positive and decreased tolerance for concluding that a drug is ineffective when there is residual uncertainty that the drug may actually be effective, which would be a conclusion at risk of being a false negative.[112]
>
> It is important to note when considering one study plus confirmatory evidence, the single study may be a study of conventional persuasiveness rather than the

---

[110] *Id.* at 61.

[111] Gingery, *supra* note 56.

[112] *See* Sarah Karlin-Smith, *supra* note 57. Dr. Dunn explicitly cited 21 CFR 314.105 Part C ("FDA will approve an NDA after it determines in the drugs meets the statutory standards for safety and effectiveness. While the standards apply to all drugs, the many kinds of drugs that are subject to the statutory standards and the wide range of uses for those drugs *demand flexibility* in applying the standards. Thus, FDA is required to exercise its scientific judgment. To determine the kind and quantity of data and information an applicant is required to provide for a particular drug to meet the statutory standards….") (emphasis added) and 21 CFR 312.80 ("The Food and Drug Administration has determined that it is appropriate to exercise the *broadest flexibility* in applying the statutory standards while preserving appropriate guarantees for safety and effectiveness. These procedures reflect the recognition that physicians and patients are generally willing to accept greater risks or side effects from products that treat life-threatening and severely-debilitating illnesses than they would accept for products that treat less serious illnesses. These procedures also reflect the recognition that the benefits of the drug need to be evaluated in light of the severity of the disease being treated….") (emphasis added).

> highly persuasive study we prefer to see when considering a true single study in isolation. The degree of persuasiveness required for approval may be influenced by many things including the seriousness of the disease, whether there is an unmet need, and the character of the confirmatory evidence.[113]

In other words, observing that a confirmatory trial has failed does not end the inquiry. To the contrary, the Agency is obliged to understand *why* the trial failed and what implications, if any, that failure has for maintaining access to the therapy.

Second, FDA has explained that it will explore the full range of regulatory options available to it in charting a path forward for the drug, with the priority placed on patient well-being, particularly where the approved drug is the only approved therapeutic option:

> By definition, drugs approved under accelerated approval represent significant therapeutic advances for patients with serious and life-threatening illnesses. FDA must carefully evaluate what other options are available to patients at the time it is considering regulatory action for failure to confirm clinical benefit. In some cases a drug for which clinical benefit has not been confirmed may be the only approved therapeutic option for patients with the disease. Removing the drug from the market and leaving patients with no treatment may be unacceptable. In such a case FDA must consider the benefits of continued availability of the drug, which by definition under the accelerated approval program was shown to have an effect on the surrogate endpoint that was the basis for approval, versus the risk that patients may actually be using an ineffective drug and exposing themselves only to its risks. FDA must also consider the possibility that, despite results from confirmatory studies that may appear to indicate that a drug does not provide clinical benefit, there may be a subset of patients for whom the drug may nevertheless be effective.[114]

Third, FDA has stated that even where a confirmatory study *demonstrates* a safety risk to patients, withdrawal may not be appropriate when alternative regulatory options are available:

> Outside of a situation where a confirmatory trial clearly demonstrates harm to the patients (e.g., decreased survival for patients with cancer treated with the accelerated approval drug), FDA believes that [FDA's response to a study that fails to confirm clinical benefit] must be considered on its merits and that the criteria in the existing regulations and statutory provisions . . . provide FDA with sufficient authority and flexibility to make balanced decisions that protect the program from abuse by sponsors and ensure that patients will continue to have access to needed treatments.[115]

---

[113] Billy Dunn, *supra* note 59.

[114] *Id.*

[115] GAO Report at 64-65.

Over the life of the accelerated approval program, FDA has used this approach to chart a path forward for continued marketing of numerous therapies after the failure of confirmatory studies to verify and describe the predicted clinical benefit of the drug.

The case of PROAMATINE® (midodrine HCl) illustrates FDA's patient-focused, flexible approach. Midodrine HCl received an accelerated approval in 1996, for the treatment of symptomatic orthostatic hypotension (OH). Confirmatory studies submitted in 2005, were deemed by FDA to have "failed" to verify clinical benefit.[116] Instead of withdrawing the drug, FDA worked with the sponsor and ANDA holders to design and facilitate two additional studies.[117] It was only after the sponsor and ANDA holders failed to meet the newly set postmarketing trial target dates by the Agency that FDA issued a NOOH proposing to withdraw midodrine HCl. Even then, FDA posted an update clarifying that its NOOH did not represent an actual withdrawal, but rather was taken to instigate the collection of data to establish the efficacy of midodrine HCl for the benefit of the patient population:

> *FDA has two goals with respect to midodrine: (1) to obtain high quality data on the effectiveness of the medication and (2) to maintain access for patients to the medication throughout this process.*
>
> . . .
>
> A key point is that FDA's [NOOH] . . . represented a step in the regulatory process—a step that reflects both the regulatory requirement for manufacturers to verify the clinical benefit of accelerated approval products and *the agency's position that more data about the benefits of midodrine would help doctors and patients understand who can benefit from the drug and how best to use it*.[118]

FDA also noted that since the issuance of the NOOH, it had heard from "several professional organizations that support the use of the product" as well as "[m]any patients and doctors, [who] believe[d] through experience that the medication . . . help[ed] patients substantially."[119]

Even as the NOOH process was proceeding, FDA continued to support additional confirmatory trials that could support continued marketing authorization of midodrine.[120] FDA

---

[116] *See* Letter from Mark S. Robbins, Ph.D., J.D., Executive Vice President and General Counsel, Upsher-Smith Laboratories, Inc., to Division of Docket Management, Food and Drug Administration, Docket No. FDA-2007-N-0475-0014 (Oct. 4, 2007), https://www.regulations.gov/document?D=FDA-2007-N-0475-0014 (citing FDA's March 8, 2007 Joint Meeting Minutes).

[117] *See* Letter from Janet Woodcock, M.D., Director, CDER, Docket No. FDA-2007-N-0475-0018 (Aug. 12, 2009), https://www.regulations.gov/document?D=FDA-2007-N-0475-0018.

[118] FDA, Midodrine Update, https://wayback.archive-it.org/7993/20171115034833/https:/www.fda.gov/Drugs/DrugSafety/ucm225444.htm (Sept. 10, 2010) (emphasis added).

[119] *See id.*

[120] Memorandum from N. Stockbridge, M.D., Ph.D., on Midodrine Hydrochloride (Jan. 11, 2011), https://wayback.archive-

again explained its actions as being in the interest of patients, given the seriousness of the condition and the absence of other available therapies:

> Midodrine is the only drug approved for the treatment of symptomatic orthostatic hypotension (SOH), a rare but serious condition in which patients are unable to maintain blood pressure in an upright position and thereby become dizzy or faint, making it difficult or impossible for patients to carry out their daily life activities. *If marketing approval for midodrine is withdrawn at this time, patients with SOH will be left with no approved therapeutic options*.[121]

Ultimately, the drug sponsor submitted a supplement with the results of additional studies on midodrine HCl in 2015, *nineteen years* after its original approval, and *ten years* after its first set of failed confirmatory studies were submitted to FDA, and in 2017, FDA stated that it was in the process of reviewing the new confirmatory studies. Even though PROAMATINE was subsequently discontinued by its sponsor, FDA determined in 2019 that the withdrawal *was not* based on considerations of safety or efficacy. To date, generic midodrine remains on the market, and FDA has not yet opined publicly on whether the second set of confirmatory studies offer sufficient verification of clinical benefit.[122] The Agency also has recently approved a generic version of midodrine HCl in January of 2021,[123] indicating that it intends to provide patients continued access to this product.

FDA took a similar approach in the case of the oncology drug IRESSA® (gefitnib). There, FDA allowed the product to remain on the market for several years to allow ongoing and new clinical trials to be completed, even after the Agency concluded in December 2004 that the results of the confirmatory trial failed to verify that the drug prolonged survival.[124] Rather than withdrawing accelerated approval, FDA restricted access to certain patient populations, including

---

it.org/7993/20170406045513/https:/www.fda.gov/downloads/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/UCM239484.pdf.

[121] Letter from Abigail Brandel, Counsel to CDER and Carla Cartwright, Counsel to CDER to G. Matthew Warren, Senior Regulatory Counsel, Office of the Commissioner of Food and Drugs, Docket No. 2007-N-0475-0036 (Jan. 13, 2012), https://www.regulations.gov/document?D=FDA-2007-N-0475-0036.

[122] Determination That PROAMATINE (Midodrine Hydrochloride) Tablets, 2.5 Milligrams, 5 Milligrams, and 10 Milligrams, Were Not Withdrawn From Sale for Reasons of Safety or Effectiveness, 84 Fed. Reg. 56459 (Oct. 22, 2019) ("The clinical benefit of PROAMATINE (midodrine hydrochloride) tablets, 2.5 mg, 5 mg, and 10 mg, remains subject to verification."); Letter from Mary Ross Southworth, Pharm.D., Dep. Dir. for Safety, Division of Cardiovascular and Renal Products, Office of Drug Evaluation I, CDER to Mihaela MacNair, Ph.D., M.Sc., Shire Development LLC at 2 (Feb. 7, 2017), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2017/019815Orig1s010ltr.pdf ("We remind you that your supplement submitted March 30, 2015 containing the results of studies intended to address the requirements under 21 CFR 314.510 to verify and describe clinical benefit of midodrine remains under review by the Agency.").

[123] ANDA 214734 was approved on January 21, 2021. The ANDA is therapeutically equivalent to Shire's midodrine HCl. *See* Capital Market, *Alembic Pharmaceuticals Receives USFDA Approval for Midodrine Hydrochloride Tablets,* BUSINESS STANDARD, Jan. 22, 2021, https://www.business-standard.com/article/news-cm/alembic-pharmaceuticals-receives-usfda-approval-for-midodrine-hydrochloride-tablets-121012200455_1.html.

[124] *See* FDA, New Labeling and Distribution Program for Gefitinib (IRESSA) (last updated May 10, 2016), http://web.archive.org/web/20170722191240/www.fda.gov/drugs/drugsafety/postmarketdrugsafetyinformationforpatientsandproviders/ucm163112.htm ("[FDA] is not considering market withdrawal of gefitinib at this time. New clinical trials are being developed, other ongoing trials are being completed, and there will be further analysis of the completed trials described above. These will determine the future role of gefitinib treatment").

those patients who had already taken IRESSA and whose doctors believed it was benefitting them.[125] In justifying its action, FDA reflected that there were factors—other than lack of efficacy—that could have caused the confirmatory trial to fail. For example, FDA noted that there was a suggestion, though not yet proven by controlled clinical trials, that the genetic markers could have caused some patients to respond to the drug and others to not respond.[126] FDA noted that there was "clear evidence in individual patients of significant clinical benefit (e.g., shrinkage of large tumors and prolonged survival in patients with end-stage disease) that could not be ascribed to factors other than drug effect."[127] Ultimately—given that another drug had been approved for the same indication by the time the confirmatory trial reported out—FDA worked with the sponsor to restrict access to IRESSA for new patients while the company undertook additional trials. It was only after further studies failed to confirm clinical benefit that FDA asked the sponsor to voluntarily withdraw IRESSA.[128]

Significantly, after the NDA was withdrawn in 2012, IRESSA returned to the market three years later when the sponsor identified the biomarker (a specific epidermal growth factor receptor (EGFR) mutation) responsible for response to IRESSA. The sponsor submitted a new NDA, and it was granted full approval. On July 13, 2015, FDA approved IRESSA for a narrower indication, namely patients with certain biomarkers: EGFR exon 19 deletions or exon 21 (L858R) substitution mutations as detected by an FDA-approved test.[129] Scholars have suggested that past clinical trials with IRESSA and similar agents may have failed because the actual responders (i.e., patients with the right biomarkers) represented too small a proportion of the patients in the trials.[130]

In addition, FDA has taken a measured response to the question of withdrawal, even where the confirmatory trial raised safety issues. The drug, ICLUSIG® (ponatinib hydrochloride), received accelerated approval in 2012 for several cancer indications. After safety issues arose in the confirmatory trial, FDA requested that the sponsor suspend marketing of the drug pending a benefit/risk re-examination. ICLUSIG was allowed to return to the market less than three months later but with (1) a narrower label, containing a more restricted indication that largely removed the drug from second-line use and relegated its use to the third and fourth-line and later settings, that reduced the targeted patient by half, (2) stronger warnings, i.e., a boxed warning on arterial and thrombotic risks, and (3) a Risk Evaluation and Mitigation Strategy (REMS) consisting of a communication plan to inform healthcare professionals about risks of vascular occlusion and thromboembolism from ICLUSIG. Additional postmarketing requirements were also imposed to better clarify and characterize cardiovascular events as well

---

[125] *See id*.

[126] *See id.*

[127] GAO Report at 63.

[128] *See* 77 Fed. Reg. 24723 (Apr. 25, 2012).

[129] *See* CDER, NDA 206995 Approval Letter (Jul. 13, 2015), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2015/206995Orig1s000ltr.pdf.

[130] *See* Paul Howard, *Why the FDA Rejected A Drug That Helps Cure Lung Cancer – And What We Can Do To Fix It,* FORBES, (Nov. 6, 2015), https://www.forbes.com/sites/theapothecary/2015/11/06/attacking-the-21st-century-cures-act/?sh=4872ed3c1ffc.

as to characterize the safety of a range of ponatinib doses. Ultimately, ICLUSIG was granted full approval, and its postmarketing requirements were deemed fulfilled in November of 2016.

VI. BASIS FOR ACCELERATED APPROVAL OF MAKENA

A. Reducing Preterm Birth Was—and Remains—A Public Health Priority

Preterm birth, defined as birth before the 37th week of gestation,[131] is recognized as the leading cause of neonatal and infant mortality,[132] as well as the cause of short- and long-term complications for those infants who survive. Infants born prematurely are at much higher risk of death within the first 28 days of life. Infants who survive preterm birth are at a significantly higher risk of short- and long-term complications that place pressure and costs on the healthcare system and families, which estimates suggest is as much as $26.2 billion dollars each year.[133] These costs are often related to short-term complications including: respiratory distress syndrome, bronchopulmonary dysplasia and intraventricular hemorrhage; long-term complications include chronic respiratory problems, rehospitalizations, metabolic disorders and neurodevelopmental problems.[134]

Obstetricians have recognized that the greatest impact in neonatal morbidity and mortality was attributable to early preterm birth, rather than birth at a later gestational age. Over time, there has been a shift in consensus regarding the significance of preventing late preterm birth, namely birth at 34-36 weeks of gestation. Around the time of Makena's approval in 2011, CDER began to recognize that infants born late preterm were less healthy than infants born later in pregnancy. Late preterm infants are more likely than term infants to suffer complications, require intensive and prolonged hospitalization, and experience adverse long-term neurodevelopmental outcome.[135] As such, there was a growing concern about the increase in late preterm birth and a focus on its prevention as an important public health priority.[136]

Today, the picture has become even clearer. The medical and scientific community no longer debate the relative significance of early versus late preterm birth. Instead, it is well understood that risks associated with preterm birth lie on a continuum. Neonatal mortality and

---

[131] *See* ACOG, Preterm Labor and Birth Frequently Asked Questions (last updated Jan. 2022), https://www.acog.org/womens-health/faqs/preterm-labor-and-birth.

[132] *See, e.g.*, Hyagriv N. Simhan, *Preterm Birth is the Leading Cause of Neonatal Mortality and is Responsible for Roughly One-Half of Long-Term Neurologic Sequelae*, AM. J. OBSTET. GYNECOL. (2010), https://pubmed.ncbi.nlm.nih.gov/20452479/; Preterm Birth, WORLD HEALTH ORGANIZATION, (Feb. 19. 2018), https://www.who.int/news-room/fact-sheets/detail/preterm-birth.

[133] *See* The Impact of Premature Birth On Society, MARCH OF DIMES, https://www.marchofdimes.org/mission/the-economic-and-societal-costs.aspx (last updated Oct. 2015).

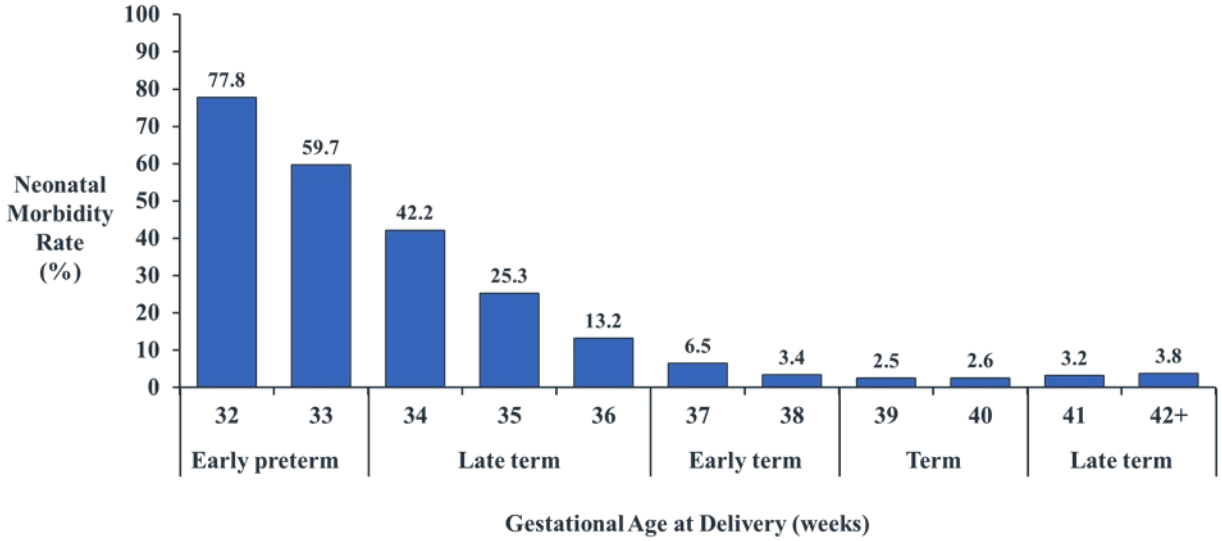
[134] Barbara J. Stoll et al., *Neonatal Outcomes of Extremely Preterm Infants from the NICHD Neonatal Research Network*, 126,3 PEDIATRICS 443-56 (Sept. 2010), https://pubmed.ncbi.nlm.nih.gov/20732945/; Cynthia Gyamfi-Bannerman, et al., *Antenatal Betamethasone for Women at Risk for Late Preterm Delivery*, 374 N. ENGL. J. MED. 1311-1320 (Apr. 2016), https://www.nejm.org/doi/full/10.1056/nejmoa1516783 .

[135] *See* CDER, NDA 21945 Medical Review (Apr. 21, 2014) at 31, https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/125514Orig1s000MedR.pdf [hereinafter NDA 21945 Medical Review].

[136] *See* Joyce A. Martin et al., *Born A Bit Too Early: Recent Trends in Late Preterm Births*, NCHS DATA BRIEF, No. 24, (2009), https://pubmed.ncbi.nlm.nih.gov/19922725/; *see also* Medical Review at 31.

morbidity are known to increase as gestational age decreases (Figure 4).[137] Though late preterm infants have increased mortality and morbidity outcomes when compared with infants born to term, they have a more "benign course" compared with early preterm infants.[138]

Figure 4
Rate of Neonatal Morbidity by Gestational Age at Birth, from 32 Weeks Onwards[139]



In addition, one of the most significant risk factors for singleton spontaneous preterm birth is a patient's history of preterm delivery. Approximately 3.3% of pregnant women, or 130,000 U.S. women annually, have a history of prior singleton spontaneous preterm delivery.[140] It is widely recognized in the obstetrical and gynecological community that a pregnancy after a previous spontaneous preterm birth is generally considered a high-risk pregnancy, subjecting the patient's subsequent pregnancy to at least a 2.5-fold greater risk for preterm birth than women without such prior history.[141] Makena is indicated for treatment of this population. As the data

---

[137] *See* Tracy A. Manuck et al., *Preterm Neonatal Morbidity and Mortality by Gestational Age: A Contemporary Cohort*, 215 A. J. OBSTET. GYNECOL. 103.e1–103.e14 (2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4921282/; Stephanie Todd, et al., *A Composite Neonatal Adverse Outcome Indicator Using Population-Based Data: An Update*, 5 INT. J. POPUL. DATA SCI. 1337 (2020), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7893849/; Sven Cnattingius, et al., *Apgar Score and Risk of Neonatal Death among Preterm Infants*, 383 N. ENGL. J. MED. 49–57 (2020), https://www.nejm.org/doi/10.1056/NEJMoa1915075.

[138] *See* Manuck et al., at 3.

[139] *See* Todd et al., at 9.

[140] Joann R. Petrini et al., *Estimated Effect of 17 Alpha-Hydroxyprogesterone Caproate on Preterm Birth in the United States*, 105 OBSTET. GYNECOL. 267-72 (2005), https://pubmed.ncbi.nlm.nih.gov/15684150/; Cande V. Ananth et al., *Recurrence of Spontaneous Versus Medically Indicated Preterm Birth*, 195 AM. J. OBSTET. GYNECOL. 643-50 (2006), https://pubmed.ncbi.nlm.nih.gov/16949395/.

[141] Brian M. Mercer 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*, 181 AM. J. OBSTET. GYNECOL 1216 (1999), https://pubmed.ncbi.nlm.nih.gov/10561648/#:~:text=Conclusion%3A%20Prior%20spontaneous%20preterm%20deli

excerpted from Laughon et al. (2014) below show (Figure 5 and Table 4), the risk for recurrence increases as the gestational age of the first spontaneous preterm birth decreases.[142] Additionally, a study by Esplin et al. (2008) suggests that the rates of recurrence are highest in women whose prior spontaneous preterm birth occurred before 34 weeks of gestation.[143]

Figure 5
***Incidence of Preterm Delivery by Prior Spontaneous Preterm Delivery Status*[144]**

Table 4
***Risk of Preterm Birth < 37 Weeks of Gestation in Subsequent Delivery by Gestational Age at First Delivery*[145]**

			Preterm birth <37 Weeks in Second Delivery	
Gestational age at first delivery, weeks	**Total n (%)**	**N (%) preterm in 2nd birth**	**Unadjusted RR[a]**	**Adjusted RR[a,b]**
≥ 37	46771 (92.4)	2630 (5.7)	Referent	Referent
34 to < 37	2950 (5.8)	838 (28.9)	5.07 [4.73, 5.42]	4.81 [4.48, 5.15]
28 to < 34	607 (1.2)	226 (37.9)	6.63 [5.95, 7.40]	5.98 [5.37, 6.66]
24 to < 28	152 (0.3)	61 (40.1)	7.03 [5.77, 8.57]	6.42 [5.33, 7.74]
20 to < 24	127 (0.3)	35 (27.8)	4.87 [3.66, 6.47]	4.88 [3.66, 6.50]

Trend for gestational age P<.0001.
[a] Numbers in brackets are 95% confidence intervals.
[b] Models were adjusted for maternal age, race/ethnicity, pre-pregnancy body mass index, insurance, smoke, alcohol, illicit drug use, chronic medical disease.

very,subsequent%20early%20spontaneous%20preterm%20delivery; Jay D. Iams et al., *The Preterm Prediction Study: Recurrence Risk of Spontaneous Preterm Birth. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network*, 178 AM. J. OBSTET. GYNECOL. 1035 (1998), https://pubmed.ncbi.nlm.nih.gov/9609580/.

[142] *See* Laughon et al., *supra* note 10 at 131.e5-131.e6; Tiril Tingleff et al., *Risk of preterm birth in relation to history of preterm birth: a population-based registry study of 213 335 women in Norway,* 129 BR. J. OBSTET. AND GYNECOL. 900-907 (2022), https://pubmed.ncbi.nlm.nih.gov/34775676/.

[143] *See, e.g.*, Esplin et al., *supra* note 10 at 516-23.

[144] Laughon et al., *supra* note 10.

[145] *Id.*

Accordingly, preterm birth remains a significant public health concern, especially for women with a history of early prior spontaneous preterm birth.

B. The Meis Trial Demonstrates Substantial Evidence Of Makena's Efficacy And Was The Basis Of FDA's Accelerated Approval In 2011

At the time the Meis trial was undertaken, preterm birth was "more common in the United States than in many other developed countries and [was] the factor most responsible for the relatively high infant mortality" in the U.S."[146] As such, there was an urgent need for an intervention to reduce recurrent preterm birth among high-risk women, as there was no FDA-approved therapy for this indication.

CDER granted accelerated approval to Makena in 2011, based on the strength of a multi-site, randomized, double-blind, placebo-controlled clinical trial conducted by the Maternal Fetal Medicine Units (MFMU) Network and sponsored by the NICHD. Known as "the Meis trial"—named for its principal investigator, Paul J. Meis, M.D., a leading maternal fetal medicine physician—the trial was immediately recognized as a major advance in the field of obstetrics. When approving Makena, CDER acknowledged that the Meis trial was "adequate, well-controlled and very persuasive," and provides "compelling" evidence of clinical benefit.[147] CDER has also recognized that the Meis trial is "sufficiently persuasive to support drug approval based on the findings of a single adequate and well-controlled trial."[148] Accordingly, this evidence could have provided the basis for full approval at that time.[149]

The Meis trial planned to enroll 500 women with a documented history of singleton spontaneous preterm birth at 19 sites, who were randomly assigned in a 2:1 ratio to receive either 17-OHPC or placebo. Treatment was administered starting between 16 weeks/0 days and 20 weeks/6 days of gestational age and continuing until 37 weeks or delivery, whichever occurred first.

Specifically, the Meis trial found that 17-OHPC:[150]

- Reduced preterm birth prior to 37 weeks gestation from 54.9% to 36.3% with a relative risk (RR) of 0.66 (95% confidence interval (CI): 0.54-0.81; $p<0.001$), translating to a 34% reduction in the primary outcome; and

- Reduced preterm birth at earlier gestational ages, compared to placebo:

---

[146] *See* Paul Meis, *Prevention of Recurrent Preterm Delivery by 17 Alpha-Hydroxyprogesterone Caproate*, 348 N. ENGL. J. MED. 2379 (2003), https://pubmed.ncbi.nlm.nih.gov/12802023/.

[147] FDA Briefing Document at 11, 21. The study was also conducted under good clinical practices (GCP), and pre-approval inspections of the study sites revealed "no concerns regarding the quality and integrity of the data" and "no violations that would impair the acceptability of the clinical data." NDA 21945 Medical Review at 32.

[148] FDA Briefing Document at 8.

[149] *See generally,* FDA, DEMONSTRATING SUBSTANTIAL EVIDENCE OF EFFECTIVENESS FOR HUMAN DRUG AND BIOLOGICAL PRODUCTS DRAFT GUIDANCE FOR INDUSTRY, (Dec. 2019), https://www.fda.gov/media/133660/download.

[150] Meis *supra* note 146 at 2382-2383.

- For delivery <35 weeks gestation: 0.67 (0.48-0.93), p=0.02
- For delivery <32 weeks gestation: 0.58 (0.37-0.91), p=0.02

Significantly, at the second planned interim analysis (conducted when 463 patients had undergone randomization and outcome data were available for 351 patients), an independent data and safety monitoring committee reviewed the study data and determined that the risk of delivery prior to 37 weeks of gestation was significantly reduced in the patients treated with 17-OHPC compared to the placebo arm, with a p-value that was below the pre-specified value in stopping rules (p=0.015).[151] Enrollment was therefore halted based on the pre-specified stopping rules, which established that it would be unethical to continue treating with placebo if robust efficacy was observed.

The results of the Meis trial were so compelling that following publication, the American College of Obstetricians and Gynecologists (ACOG) issued a Committee Opinion recognizing that the trial was stopped early because the results showed a "significant protection against recurrent preterm birth for all races of women who received [17-OHPC]."[152]

Indeed, ACOG issued a Committee Opinion in 2008 establishing 17-OHPC as the de facto standard of care: "*Progesterone supplementation for the prevention of recurrent preterm birth should be offered to women with a singleton pregnancy and a prior spontaneous preterm birth due to spontaneous preterm labor or premature rupture of membranes*."[153]

In understanding Meis, it is important to recognize that CDER has used different terms to describe the endpoints of the Meis trial. This is significant because the type of marketing approval granted (i.e., accelerated approval versus traditional or full approval) can depend on the type of endpoint as well as strength of the evidence supporting the ability of a surrogate or intermediate endpoint to predict clinical benefit.

At times, CDER has characterized the trial's demonstration of Makena's efficacy in reducing preterm births <37 weeks as being an effect on a surrogate endpoint, including in

---

[151] As common in clinical trials, the Meis trial protocol had specified that an external independent data and safety monitoring committee (DSMC) meet periodically to review the trial data during the course of the study. The protocol had also specified what is commonly called "stopping rules," essentially a set of criteria that specify when the trial should be suspended, for example, due to strong evidence of efficacy. In the Meis trial, a Lan-DeMets implementation of the O'Brien-Fleming stopping boundaries was used. A nominal p-value <0.0001 was required to show statistical significance at the first interim analysis, and a nominal p-value <0.015 was required for the second interim analysis. At the second meeting, the DSMC reviewed the interim report and determined that the boundary (p=0.015) was crossed. The committee therefore recommended discontinuation of subject recruitment because 17-OHPC had demonstrated benefit for the primary outcome. *See* NDA 21945 Statistical Review and Evaluation, Clinical Studies at 11 (Oct. 19, 2006), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000StatR.pdf; Meis, *supra* note 146, at 2381.; NDA 21945 Summary Review, *supra* note 37 at 7.

[152] ACOG Committee Opinion, Number 291, *Use of Progesterone to Reduce Preterm Birth*, 102 OBSTET. GYNECOL. 1115 (2003).

[153] ACOG Committee Opinion, Number 419, *Use of Progesterone to Reduce Preterm Birth*, 112 OBSTET. GYNECOL. 963 (2008); *see also* NDA 21945 Medical Review at 16-17 ("This sentence is unambiguous, and has been interpreted as an attempt to create a standard of care.").

CDER's NDA review package from 2011, when Makena was approved.[154] As CDER has recently acknowledged, however, this is incorrect. The Meis trial's demonstration of Makena's efficacy in reducing preterm births <37 weeks is an intermediate clinical endpoint, *not* a surrogate endpoint.[155]

Unlike a surrogate endpoint, which is "not itself a measure of clinical benefit," an intermediate endpoint is "a measurement of a therapeutic effect."[156] For purposes of accelerated approval, an intermediate clinical endpoint is defined as "a measurement of a therapeutic effect that can be measured earlier than an effect on irreversible morbidity or mortality and is considered reasonably likely to predict the drug's effect on [irreversible morbidity or mortality] IMM or other clinical benefit."[157] In contrast, a surrogate endpoint is a marker, such as a laboratory measurement, radiographic image, physical sign, or other measure, that is thought to predict clinical benefit, but is not itself a measure of clinical benefit.[158]

Indeed, in a 2014 guidance document outlining recommendations for the accelerated approval program, CDER states:

> Examples of cases in which FDA has used an *intermediate* clinical endpoint to support accelerated approval include the following: . . . A treatment for preterm labor was approved based on a demonstration of delay in delivery. Under accelerated approval, the sponsor was required to conduct postmarketing studies to demonstrate improved long-term postnatal outcomes.[159]

CDER's characterization of the Meis trial results as an intermediate clinical endpoint is noteworthy because an intermediate endpoint can also be used to support traditional approval (as well as accelerated approval). In fact, in CDER's own words, an intermediate clinical endpoint will "usually" support traditional approval in circumstances such as here, in contrast to a surrogate endpoint, which needs to be validated before it can be used for traditional approval.[160] As CDER has recognized, for products with intermediate clinical endpoints, "[a]n important

---

[154] *See* FDA Briefing Document at 6 ("In 2011, Makena received accelerated approval . . . based on a reduced risk of recurrent preterm birth (PTB) prior to 37 weeks, a surrogate endpoint that FDA considered reasonably likely to predict clinical benefit to the neonate"); NDA 21945 Summary Review, *supra* note 37 at 38, https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000SumR.pdf ("I now believe that a reduction in preterm births <37 weeks gestation is an adequate surrogate endpoint that is reasonably likely to predict clinical benefit in terms of a reduction in neonatal morbidity and/or mortality. In summary, I believe that the efficacy findings from Study 17P-CT-002 are adequate to support approval of HPC under subpart H regulations for accelerated approval").

[155] CDER's Preliminary Briefing Materials at 6 n.4; Letter from Sara Rothman to Celia Witten, Ph.D., M.D., at 3-4 (Aug. 12, 2022).

[156] Expedited Programs Guidance at 18.

[157] *Id*.

[158] *See* Expedited Programs Guidance at 17. Some examples of surrogate endpoints include shorter-term suppression of HIV viral load in plasma (surrogate for reduction of morbidity and mortality associated with HIV disease) and radiographic evidence of tumor shrinkage (surrogate for overall survival in certain cancers). *Id*. at 17-18.

[159] *Id.* at 19 (emphasis added).

[160] *Id.* at 19.

question is whether the demonstrated [effect on an intermediate clinical endpoint] would be a basis for traditional approval."[161] Indeed, CDER has acknowledged that "[a]pprovals for products for serious conditions based on clinical endpoints other than IMM will usually be considered under traditional approval procedures,"[162] which would not require a postmarket confirmatory trial. Approvals based on such clinical endpoints will be considered under the accelerated approval program "only when it is essential" to determine effects on IMM or other clinical benefit, in order to confirm the predicted clinical benefit that led to approval.[163]

CDER has also used the words "well-established surrogate" or "established surrogate" to describe reduction in preterm birth <32 or <35 weeks.[164] CDER has stated that these endpoints would not need confirmatory evidence of clinical benefit.[165] More recently, CDER again signaled that it may consider a certain gestational age at delivery—i.e., <34 weeks—to be a *validated* surrogate endpoint. A surrogate endpoint is "validated" and can be accepted in place of a clinical outcome if sufficient evidence, including evidence from epidemiological studies and clinical trials, show that the surrogate endpoint can be relied upon to predict or correlate with clinical benefit.[166] The significance of a validated surrogate endpoint is that it may be used for traditional approval rather than accelerated approval, which would not require a confirmatory trial to be conducted post-approval. The meeting minutes of CDER's Medical Policy and Program Review Council (MPPRC) meeting in January 2020, which resulted in the CDER's issuance of the NOOH, state the following:

> At this time, the Agency has not determined a certain gestational age at delivery to be a validated surrogate endpoint. Late preterm birth (34-36 weeks gestation) is not considered a validated surrogate endpoint, although early preterm birth (less than 34 weeks) is expected to be more a robust predictor of neonatal outcomes and could be considered for validation in the future.

There are a number of takeaways from these recent characterizations and statements by CDER. First, CDER's characterization of the Meis trial data as an effect on an intermediate clinical endpoint further underscores the strength of the Meis trial findings and the clinical significance of the observed reduction in preterm birth. In addition, because an intermediate endpoint is "a measurement of a therapeutic effect,"[167] CDER's characterization as such suggests that delay in delivery should be considered as a clinical endpoint for future confirmatory trials of Makena and serve as a basis for traditional approval. This is consistent with the evolving view by CDER and experts that earlier preterm birth endpoints could be viewed as validated surrogate

---

[161] *Id.* at 18.

[162] *Id.*

[163] *Id.*

[164] Clinical Review at 15, https://www.fda.gov/media/80892/download; FDA Briefing Document at 20, https://www.fda.gov/media/132003/download ("FDA determined that further study was needed to provide confirmatory evidence of the drug's efficacy in terms of direct clinical benefit on neonatal outcomes or through an established surrogate such as the rate of preterm birth prior to 35 and 32 weeks gestation").

[165] *Id.*

[166] *See* FDA, Surrogate Endpoint Resources for Drug and Biologic Development (current as of Jul. 24, 2018), https://www.fda.gov/drugs/development-resources/surrogate-endpoint-resources-drug-and-biologic-development.

[167] Expedited Programs Guidance at 18.

endpoints, which, like traditional approval, would not require a postmarket confirmatory trial. It would not advance the public health to remove Makena from the market while further studies are conducted to determine whether delay in delivery predicts clinical benefit when, as CDER has acknowledged, an increase in gestational age has been strongly associated with a clinically meaningful reduction in neonatal morbidity and mortality.

CDER agreed in recent correspondence that it would use the term "intermediate clinical endpoint" rather than "surrogate endpoint" going forward, after Covis pointed out CDER's inconsistent use of terminology in describing the endpoints of the Meis trial (surrogate endpoint vs. intermediate clinical endpoint).[168] Previously, CDER had used the term surrogate endpoint to describe the Meis trial data, including in the NDA review package. This prior blurring of the important distinction between the two types of endpoints is significant as the use of a "surrogate endpoint" to describe the Meis trial endpoint minimizes the significance of the data in itself and dismisses the fact that gestational age of delivery is itself a clinical endpoint.

VII. BASIS FOR MAINTAINING APPROVAL OF MAKENA

As detailed below, there is a strong basis for maintaining approval of Makena.

A. Makena Has A Favorable Benefit-Risk Profile Meriting Further Study Rather Than Withdrawal

1. The Meis Trial Provided Compelling Evidence of Makena's Efficacy, Ushering In Makena As A New Standard Of Care In High-Risk Obstetrics Practice

The Meis trial demonstrated the substantial benefit of Makena for women at higher risk of preterm birth—so much so that the trial was ended early and 17-OHPC became widely used in the obstetrics community even before accelerated approval was granted by CDER. Following publication of the Meis trial in the *New England Journal of Medicine*, ACOG released guidelines recognizing the trial and endorsing the use of 17-OHPC.[169] For more than 10 years, use of 17-OHPC has been the standard of care in the U.S. during the second and third trimester for women with history of previous spontaneous birth. Over 350,000 women have been treated with Makena to date.[170]

It is important to note the Meis results in terms of impacts on patients. Based on this data, only 5-6 women would need to be treated with 17-OHPC to prevent one preterm birth <37 weeks of gestation, only 10 women would need to be treated with 17-OHPC to prevent one preterm birth <35 weeks of gestation, and only 12 women would need to be treated to prevent one preterm birth <32 weeks of gestation.[171]

---

[168] CDER Response to Covis' Objections to CDER's Preliminary Briefing Materials at 3-4 (Aug. 12, 2022).

[169] ACOG Committee Opinion Number 291, *supra* note 149; *see also* ACOG Committee Opinion Number 419, *supra* note 152.

[170] Covis Pharma GgmbH, Periodic Safety Update Report (PSUR), Makena® (Hydroxyprogesterone Caproate Injection), 9 (Apr. 1, 2022)

[171] *See* Meis *supra* note 146, at 2382.

The Meis study showed highly statistically significant efficacy results and demonstrated consistent and robust increases in gestational age across all major subgroups of patients in the study. In its review of the Meis trial, CDER acknowledged, "*this treatment benefit appeared independent of race, number of prior preterm deliveries, and gestational age of the prior preterm birth*."[172] This overall relative risk reduction from the Meis trial is discussed in detail in a recent commentary in *Obstetrics & Gynecology.*[173] The authors document the consistent reduction among all subgroups analyzed, as shown in Table 5.

Table 5
Preterm Birth at Less Than 37 Weeks of Gestation by Subgroup (Meis Trial)[174]

Preterm Birth at Less Than 37 Weeks	**17-OHPC** n/N (%)	**Placebo** n/N (%)	**RR (95% CI)**
Overall	111/306 (36.3)	84/153 (54.9)	0.66 (0.54-0.81)[t]
			0.70 (0.57-0.85) [₸]
More than 1 prior preterm birth	41/86 (47.7)	44/63 (69.8)	0.68 (0.52-0.90)
Only 1 prior preterm birth	70/220 (31.8)	40/90 (44.4)	0.72 (0.53-0.97)
Black	64/181 (35.4)	47/90 (52.2)	0.68 (0.51-0.90)
Nonblack	47/125 (37.6)	37/63 (58.7)	0.64 (0.47-0.87)
Unmarried	50/150 (33.3)	43/82 (52.4)	0.64 (0.47-0.86)
Married	61/156 (39.1)	41/71 (57.7)	0.68 (0.51-0.90)
Smoke or substance use	28/85 (32.9)	23/36 (63.9)	0.52 (0.35-0.76)
No smoke or substance use	83/221 (37.6)	61/117 (52.1)	0.72 (0.57-0.92)
Southeast site	23/86 (26.7)	18/40 (45.0)	0.59 (0.36-0.97)
Other sites	92/224 (41.1)	66/113 (58.4)	0.70 (0.56-0.88)

Data are n/N (%), where n=number of patients in the specified category and N= number of patients in the treatment group (overall or in the specified subgroup) with nonmissing delivery data.
t The CI is a 96.5% CI to adjust for the interim analysis.
₸ Adjusted for more than one prior preterm birth using the Cochran-Mantel-Haenszel procedure.

The study also confirmed safety of 17-OHPC in that no safety signals were detected in the study population. The takeaway is clear: Meis stands as clear and compelling evidence of a favorable benefit-risk profile for 17-OHPC.

2. Further Analysis Of The Meis Data Identifies A High-Risk Subgroup That Benefits Most From Makena And Merits Further Study

Covis has worked diligently to reconcile the available data to reach scientifically sound conclusions on the clinical benefit of Makena. As part of this process, Covis collaborated with epidemiological and biostatistical experts, among others, to revisit the data from the Meis and PROLONG trials, and utilize new data from the Dorsata database, with the ultimate goal of understanding how patient risk may be associated with efficacy outcomes.

While there has been a qualitative understanding in the Ob/Gyn/maternal-fetal medicine (MFM) community about the risk factors that predispose a patient for preterm birth, we sought to

[172] FDA Briefing Document at 8.

[173] *See* Sibai et al., *supra* note 9.

[174] The table was created from the Meis data set (and includes previously unpublished data). *Id*.

utilize datasets available to us to formalize (to the extent possible) the relationships between risk factors and the likelihood of preterm birth in the target population.

a. Analysis Of The Dorsata Database Demonstrates That Patients With A Prior Spontaneous Preterm Birth <34 Weeks Gestational Age Are At Highest Risk Of Subsequent Preterm Birth

Covis first modeled the Dorsata database to characterize the risk factors, following which Covis built the same models for PROLONG-US and Meis to compare and better understand the results of each trial. We did not seek to develop a model for the ex-US population in PROLONG given the low incidence rate of preterm birth in this population, as discussed further below.

Dorsata is used by more than 2,300 users including obstetricians and gynecologists and their clinical staff to document care in structured data elements for prenatal and postpartum encounters. These obstetricians and gynecologists care for over 1.3 million unique patients annually. Dorsata's current database comprises over 210,000 pregnancies and continues to grow rapidly. The key to Dorsata's system is that it serves as the primary medical record for obstetrics. Using proprietary software, care plans which incorporate evidence-based clinical guidelines are surfaced for the providers to review within the clinical workflow. The Dorsata system is built on top of the electronic health record (EHR) as an overlay that incorporates the ACOG prenatal flow sheet in the point-of-care system—enabling broad, structured clinical data collection—and allows clinicians to seamlessly manage and track the status of every one of their prenatal patients longitudinally.

Importantly, the data collected is congruent with that collected in registration RCTs. This system captures the majority of the typical demographic and risk factors for the prenatal patient. Covis' analysis of the Dorsata data has confirmed the ability to access significant numbers of pregnancies associated with high-risk patient indicators (for example, history of spontaneous preterm birth <32 or <34 weeks, comorbidities such as hypertension, obesity and smoking, race, socioeconomic status, and other variables known to be associated with high-risk pregnancies) with the ability to characterize all patients comprehensively on their background and demographic characteristics. Additionally, the Dorsata platform tracks medications taken by the patient and captures confirmatory information on whether prescriptions were fulfilled and whether the medication was administered in accordance with the labeled dosing schedule.

The Dorsata dataset analyzed was comprised of approximately 114,000 pregnancies overall from 2018 through 2021, of which 2046 were of patients who had been indicated for treatment with 17-OHPC based on an automatic flag based on their prior history of a spontaneous preterm birth. Covis performed an audit of these 2046 records which resulted in disqualification of 347 of the records for whom the prior spontaneous preterm birth could not be confirmed. The remaining 1699 patients were eligible for analysis. The analysis set in the Dorsata analyses presented in the Appendix attached hereto are from that indicated subset of patients, which is further segmented into patients who were indicated but not prescribed therapy, into patients who were prescribed 17-OHPC therapy but where the prescription was not filled, and finally, patients who were prescribed 17-OHPC therapy. Further details on the Dorsata databases analyzed are available upon request. The risk modeling has been performed on the

population which was indicated for 17-OHPC but not treated, and comprised of 987 subjects after excluding any patients who were missing data elements required for the modeling.

Presented below are several risk factor models based on the analysis of the Dorsata database. These were developed using logistic regression, including the best one, two, and three parameter models within Dorsata. The primary outcome that the analyses tried to predict was the likelihood of the patient experiencing a subsequent preterm birth before week 34. The results of this model are not significantly impacted by varying the predicted outcome to 33 or 35 weeks—the purpose was to pick an outcome that was acknowledged by Ob/Gyns and neonatologists as significantly impacting morbidity/mortality of the neonate. The result of the Dorsata model is shown below in Table 1:

Table 1
Dorsata (Excluding 17-OHPC-treated Subjects) Best N-variable Models Predicting PTB <34 Weeks

Model/Var#	Variable(s)	Odds Ratio (95% CI)	P-value
Best 1-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.82, 0.91)	<0.0001
Best 2-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.82, 0.91)	<0.0001
Var #2	Smoking during Pregnancy	0.51 (0.18, 1.46)	0.21
Best 3-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.81, 0.90)	<0.0001
Var #2	Smoking during Pregnancy	0.52 (0.18, 1.48)	0.22
Var #3	Alcohol Use during Pregnancy	0.64 (0.28, 1.45)	0.28

As can be seen from Table 1, prior pregnancy history and in particular, mean gestational age of prior pregnancies appears to be a strong predictor of a subsequent preterm birth <34 weeks.

Given the above models, Covis also constructed similar one through three parameter models using the Meis data to see if the real-world data in Dorsata (from 2018-2021) were congruent with the Meis data from twenty years earlier. The Meis risk models are shown below in Table 2:

Table 2
Meis (Vehicle-only) Best N-variable Models Predicting PTB <34 Weeks

Model/Var#	Variable(s)	Odds Ratio (95% CI)	P-value
Best 1-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.89 (0.82, 0.98)	0.013
Best 2-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.89 (0.81, 0.98)	0.013
Var #2	Smoking during Pregnancy	2.84 (1.19, 6.76)	0.019
Best 3-variable model			
Var #1	Smoking during Pregnancy	3.12 (1.27, 7.68)	0.013
Var #2	Mean Gestational Age of Prior Pregnancies (Weeks)	0.90 (0.82, 0.99)	0.030
Var #3	Inter-pregnancy Interval (Years)	0.84 (0.71, 1.00)	0.049

Both Table 1 and Table 2 show the importance of prior pregnancy history. While we examined the predictive power of several different measures of previous preterm birth history, it is interesting that both datasets indicated that the mean gestational age (mGA) of previous pregnancies was *the strongest single parameter model*. We also observe that the models do differ in the importance of other factors such as smoking or alcohol use. As the data for these parameters in Dorsata are self-reported by patients and also are not specific with respect to the extent of smoking or alcohol use (occasional vs. habitual), the company would be inclined to utilize the Meis models (which captured these parameters formally as part of the clinical study) to select the additional factors that drive risk. Of note, the PROLONG-US model did not provide any material new conclusions about these risk factors, therefore, we are not presenting those data here. We also point out that the concept of average of gestational ages from the patient's previous birth history is useful from a modeling perspective because it encodes the idea that the patients risk depends on their overall history. However, it is not a practical parameter for defining an inclusion criterion for a future study (and indeed we propose other ways of incorporating this concept as part of the inclusion criterion in a future study).

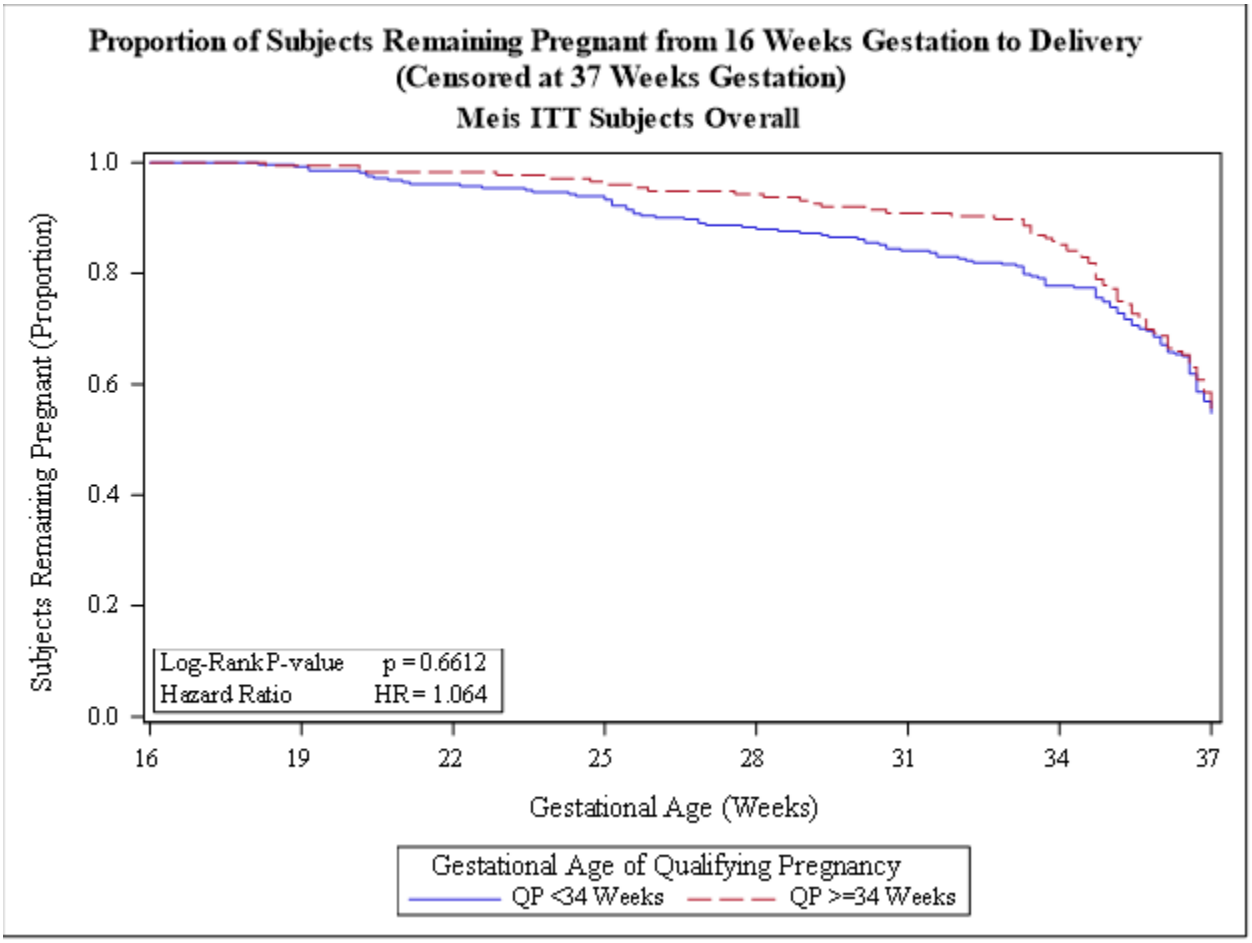
Covis also modeled the Dorsata data to understand what cutoffs in the patients' prior pregnancy history may prove more predictive of a subsequent preterm birth. As shown below in Table 6, the Dorsata data indicate that there is a sharp reduction in the likelihood of a future preterm birth if the patient's history does not include a prior spontaneous preterm birth prior to week 34. This analysis was performed in a population in Dorsata that was indicated for treatment with Makena based on prior history, but did not receive Makena.

Table 6
Count (%) of Subjects with Study Pregnancy Outcome by Gestational Age at Earliest Prior SPTB Dorsata (17-OHPC Untreated)

Gestational Age at Earliest Prior SPTB	Study Pregnancy Outcome	Count	%
20 to <28	Had a SPTB<34	10	11.0
	Not a SPTB	81	89.0
28 to <31	Had a SPTB<34	7	11.1
	Not a SPTB	56	88.9
31 to <33	Had a SPTB<34	14	17.7
	Not a SPTB	65	82.3
33 to <34	Had a SPTB<34	8	18.2
	Not a SPTB	36	81.8
34 to <35	Had a SPTB<34	12	8.2
	Not a SPTB	135	91.8
35 to <36	Had a SPTB<34	11	5.7
	Not a SPTB	182	94.3
36 to <37	Had a SPTB<34	13	4.0
	Not a SPTB	313	96.0
37 to <38	Had a SPTB<34	4	4.2
	Not a SPTB	91	95.8

Covis then examined the Meis data to see whether it was consistent with the finding that there was a sharp reduction in the likelihood of a future preterm birth if the patient's history does not include a prior spontaneous preterm birth prior to week 34. Notably, the Meis data also corroborates the increased risk of preterm birth in the Meis population in the subgroup which had a prior spontaneous preterm birth <34 weeks as seen in the following Time-to-event (Kaplan-Meier) plot (Figure 6) showing the event curves in the (34-36^6 prior spontaneous preterm birth) and <34 week prior preterm birth populations.

Figure 6
Time-to-Event Analysis for Meis Population With Prior SPTB <34 and >34 Weeks



The red curve shows the events in the population that had a prior preterm birth from week 34 to week 36[6] (which is the same as the "No prior SPTB <34 weeks" group), while the blue curve shows the events in the prior SPTB <34 week population. We also note that there is support for the <34 week cutoff from previous literature[175], which both show risk tables that are similar to the results we have obtained from Dorsata.

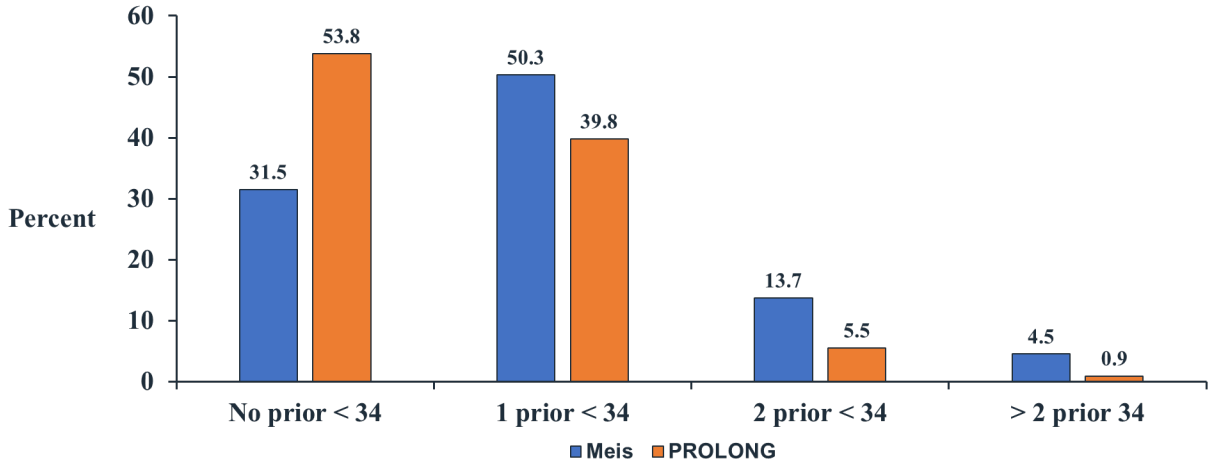
Finally, we note that the above offers an additional explanation for why the ex-US PROLONG data seemed to indicate no benefit from treatment. Not only were the event rates lower due to the low risk of patients (as discussed in Sections VII.A.3.a and VII.A.3.b), but also there is likely to have been an issue with the accuracy of the documented prior pregnancy history, including varying methods of calculating gestational age (as discussed in Section VII.A.3.c).

[175] *See* Esplin et al. and Laughon et al., *supra* note 10.

b. PROLONG-US And Meis Differed Fundamentally With Respect to Risk Factors

Given the above models, Covis examined the differences in the PROLONG-US and Meis populations more closely. The following graphic (Figure 1) compares the frequency of prior spontaneous preterm birth <34 weeks in PROLONG-US compared to Meis.

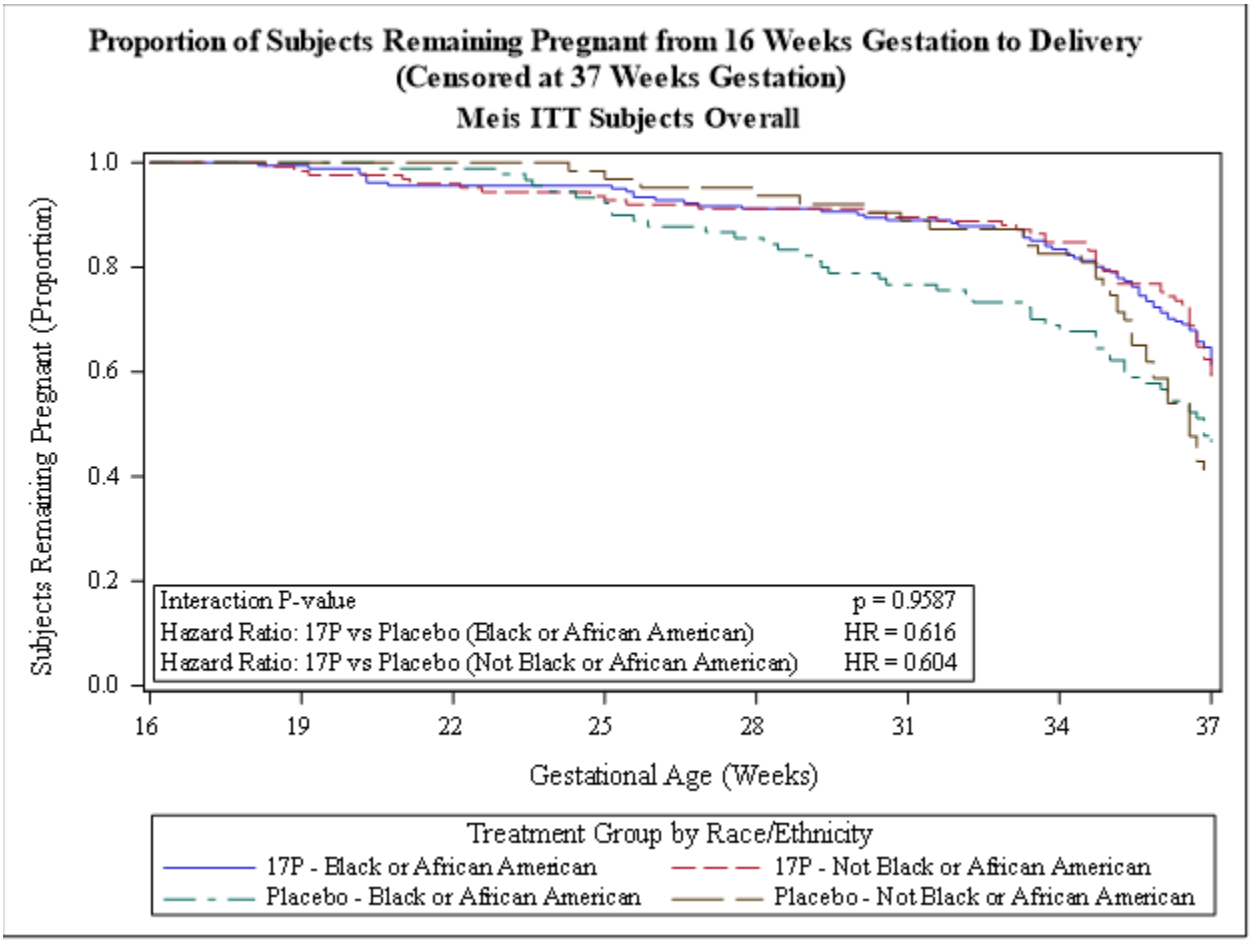
Figure 1
Frequency of Prior Qualifying SPTB <34 Weeks (Meis vs. PROLONG-US)



The graphic shows that nearly 70% of the patients in Meis had a prior spontaneous preterm birth before week 34 compared to less than half of PROLONG-US. The percentage of subjects who had 2 or more prior SPTB <34 weeks is also substantially higher in Meis than PROLONG-US. Additionally, analyzing the pregnancy history of Meis and PROLONG-US patients show a shift in the average gestational age of prior pregnancies, another parameter that Covis' risk modeling above showed to be predictive of a future preterm birth. The importance of prior pregnancy history underscores the need to have accurate and documented prior birth history using first trimester ultrasonography given the noise associated with LMP.

Finally, there are potentially important differences in the racial and socio-economic makeup of the two populations that may influence the event rate in a meaningful manner. One example of this is seen in the Meis data, when analyzed in a time-to-event basis. The figure below (Figure 7) shows, interestingly, that while Meis overall showed similar therapeutic effect of 17-OHPC on preterm birth as measured against a <37 week endpoint, that Black patients in Meis experienced a reduction in events in an earlier gestational timeframe than the non-Black patients.

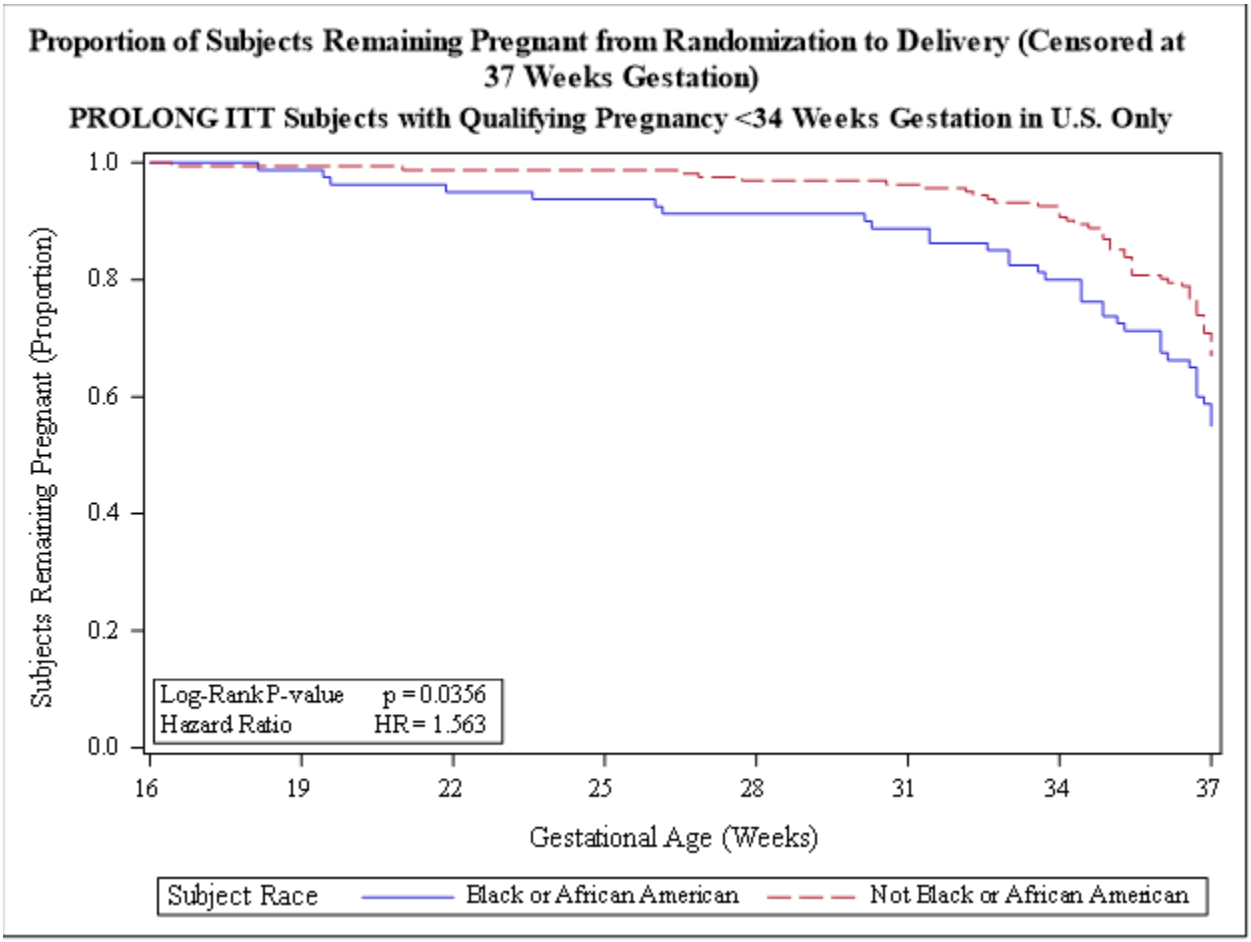
Figure 7
Time-to-Event Analysis for Meis Population (Black vs. Non-Black)



Other ways of looking at the data confirm this effect—whether examined in terms of efficacy with a gestational endpoint <32 or <35 weeks (where the majority of the reduction in event rates is in the Black population), examining these data in an ordinal analysis where the events are binned by early, mid or late-preterm birth outcome, or when looking at a "change from baseline" approach where the measured parameter is the delta between a recent qualifying pregnancy in the individual and the gestational age achieved in the study.

The pattern of earlier events in Black patients is not only seen in the Meis data. The same pattern is seen in high-risk patients (prior SPTB <34 weeks) in PROLONG-US, as seen in the following time-to-event plot comparing Black and non-Black patients.

Figure 8
Proportion of Subjects Remaining Pregnant From Randomization To Delivery (Censored At 37 Weeks Gestation)



Despite the post hoc nature of the above analyses, it is important to acknowledge the differences in the data for Black and non-Black patients. The smaller proportion of Black patients in US-PROLONG coupled with the less severe prior history of preterm birth indicates that the two studies were in different populations, and that PROLONG (or PROLONG-US) does not represent an appropriate confirmatory study for Meis.

Given the generally higher risk for preterm birth seen in the US Black population, the evidence within Meis (if analyzed with efficacy for PTB <35 weeks or earlier) of a race interaction, and the low risk of PROLONG-US based on previous birth history, Covis believes it is important to fully understand the therapeutic benefits of 17-OHPC treatment in this indication. The only way to gain this understanding is to conduct another well-designed study, as is discussed further in Section VII.D, below.

3. PROLONG Does Not Invalidate The Results And Conclusions From The Meis Trial

The PROLONG trial was conducted as a confirmatory trial for 17-OHPC, but did not demonstrate a significant reduction in risk for the two pre-specified co-primary endpoints of reducing neonatal morbidity/mortality[176] and preterm birth <35 weeks. PROLONG does not, however, invalidate the results and conclusions from the Meis trial. As explained in greater detail below, PROLONG studied a "vastly different" patient population, at significantly lower underlying risk of preterm birth and with markedly different social and demographic characteristics from the Meis trial.[177] PROLONG's patient population was also primarily ex-U.S. in a markedly different patient risk profile unlike the Meis trial which was exclusively conducted within the U.S. Challenges in the makeup of the trial ex-U.S., given well-documented differences in preterm birth patterns in the target countries, should have been foreseen. Ultimately, PROLONG is therefore not a *negative* study but instead, because of key differences in enrollment of the study population, was inherently incapable of confirming the Meis trial findings.

a. The PROLONG Population Was A Much Lower Risk Population Compared To The Meis Population

The previous Advisory Committee meeting materials, as well as the published literature demonstrated in detail how the data from PROLONG (both overall and the US subgroup) differed significantly from the Meis data. Table 7 compares the demographics between Meis and PROLONG/PROLONG-US, while Figure 9 compares the placebo rates for preterm birth across the three groups.

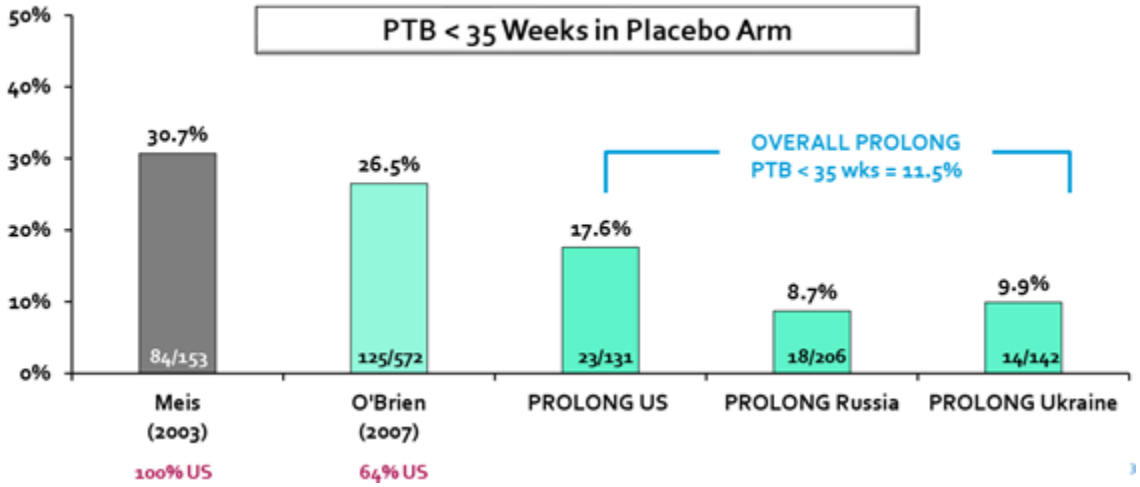
Table 7
Different Social and Demographic Characteristics Across PROLONG and Meis Trials

Demographics/Baseline Characteristics	Meis (N=463) %	U.S. PROLONG (N=391) %	PROLONG (N=1708) %
Age (years), mean ± SD	26.2 ± 5.6	27.6 ± 5.1	30.0 ± 5.2
>1 previous SPTB	28.9	27.4	14.5
GA of prior SPTB (median)	32 wks	34 wks	33 wks
Black/African American	59.0	28.9	6.7
Hispanic or Latino	14.9	13.8	9.1
Unmarried with no partner	50.3	30.7	10.1
Educational status (≤ 12 years)	71.3	50.5	43.7
Any substance use during pregnancy	26.1	28.4	9.3

[176] The composite included any of the following: neonatal death, grade 3 or 4 intraventricular hemorrhage, respiratory distress syndrome, bronchopulmonary dysplasia, necrotizing enterocolitis, or proven sepsis.

[177] Sibai et al., *supra* note 9, at 625.

Figure 9
Comparison of Placebo Rates Across Studies



The differences in patient populations between the two trials arose, in part, due to the extensive recruitment challenges experienced by PROLONG. Even prior to granting accelerated approval to Makena, CDER also predicted the difficulties of conducting a placebo-controlled confirmatory trial of 17-OHPC in light of the drug's widespread use in the U.S. for preterm birth prevention. In fact, in 2009, CDER initially determined that it would not approve the Makena marketing application in its present form, in part, due to lack of sufficient documentation demonstrating the confirmatory trial's feasibility.[178] With the 2008 ACOG Committee Opinion having virtually established 17-OHPC as the standard of care, the Center was concerned that "*Institutional Review Boards (IRBs) and patients may interpret the ACOG committee opinion as indicating that any remaining questions regarding the efficacy and safety of hydroxyprogesterone caproate are not sufficient to justify conducting a placebo-controlled study.*"[179] As such, CDER requested certain documentation of trial feasibility for the Makena application to be approved, including documentation of IRB approval and enrollment of subjects.[180] Makena's sponsor provided enrollment updates and estimated timeline for PROLONG completion based on the observed rate of enrollment, in order to address the Agency's concerns.[181]

Indeed, following Makena's approval in 2011, PROLONG's rate of enrollment from U.S. sites fell significantly as physicians and patients were reluctant to be involved in a placebo-

[178] *See* Letter from Scott Monroe, M.D., Dir., Division Reproductive and Urologic Products, Office of Drug Evaluation III, CDER, to Robb Hesley, Vice President, Business Development, Cytyc Corporation, 2 (Jan. 23, 2009), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000SumR.pdf..

[179] NDA 21945 Medical Review at 72.

[180] *See* NDA 21945 Summary Review, *supra* note 37, at 18-19.

[181] As detailed below, the current landscape has evolved and eased these enrollment concerns. For example, announcement of the PROLONG results (though based on a study that suffered from enrollment flaws) and FDA's initiation of this action to withdraw approval have raised questions about the drug's efficacy, leading to increased physician openness to considering enrolling their patients in a clinical trial.

controlled trial of an FDA approved drug.[182] Major medical centers in the U.S. also largely declined to participate because of existing patient access to 17-OHPC.[183] After Makena's approval, PROLONG's rate of enrollment from U.S. and Canadian sites fell from 11 subjects per month to four subjects.[184] Between July 2014 and June 2015, the rate sank further to two subjects per month.[185] That the population being studied was an orphan population, representing only ~3% of all pregnancies, further added to the enrollment challenge.[186]

Ultimately, PROLONG needed to be conducted primarily outside of the U.S. In the end, women were enrolled at 93 clinical centers in 9 countries with Russia and Ukraine accounting for 61% of the study patients (and 79% of the ex-US study population), while participants from the U.S. accounted for only 23% of the trial population.[187]

b. The Patient Population Enrolled In PROLONG Was At Significantly Lower Risk Of Preterm Birth

As illustrated in Figure 1, above, comparing the PROLONG US and Meis populations, showed significant differences in the number and timing of prior preterm births. Indeed, nearly 70% of the patients in Meis had a prior spontaneous preterm birth before week 34 compared to about half of PROLONG-US. Additionally, analyzing the pregnancy history of Meis and PROLONG-US patients show a shift in the average gestational age of prior pregnancies, a parameter that, as explained above, is predictive of a future preterm birth.

The lower event rate in PROLONG was likely due to the fact that the trial enrolled from a substantially different patient population, primarily outside of the U.S., with markedly different social and demographic characteristics from that of the Meis trial population. As Sibai et al., state in their review of PROLONG:

> When comparing demographics and baseline characteristics, the differences among socioeconomic status surrogates linked to higher rates of preterm birth (e.g., substance use, education level, race) stand out, with most differences driven by patients enrolled in PROLONG outside the United States. Given the health disparities that exist in obstetric care and preterm birth rates in the United States, these differences are noteworthy.[188]

As ACOG has recognized, social determinants of health, or the "conditions in the environment in which people are born, live, work, and age, play equally as important a role [to

---

[182] *See* Sean C. Blackwell et al., *supra* note 4, at 132.

[183] *See id.*

[184] *See* AMAG Pharmaceuticals, Inc., FDA Type B Meeting Request Briefing Package, NDA 021945 at 8 (Aug. 13, 2015).

[185] *See id.*

[186] *See* Petrini et al., *supra* note 140.

[187] The remaining 16% of patients were enrolled in Hungary, Spain, Bulgaria, Canada, Czech Republic, and Italy, each enrolling less than 100 patients.

[188] Sibai et al., *supra* note 9.

biologic and genetic factors] in shaping health outcomes." This is particularly the case with respect to preterm birth. Relevant social determinants of health in this context include:

- Race/ethnicity
- Socioeconomic status
- Education
- Marital status
- Population-level determinants, which include factors such as:
    - Access to healthy foods/good nutrition
    - Availability of affordable housing
    - Access to health services
    - Investments in public safety and social services

Clinicians are well aware that a combination of these factors places women at higher risk of recurrent preterm birth. The Meis trial had enrolled exclusively at 19 university-affiliated Network centers in the U.S. These academic institutions often represent "safety net" hospitals that provide care for the most under-served populations, and were selected to be in the Network partially based on the patient population and high rate of preterm birth. Thus, many of these negative social determinants of health were de facto represented in patients. In contrast, only 18% of the U.S. population in the PROLONG trial were from university-affiliated centers, and the PROLONG trial enrolled primarily in communities where patients had positive social determinants of health.

This unintentional selection bias can be seen when comparing demographics and baseline characteristics between the Meis trial population and the PROLONG population. The differences among socioeconomic status surrogates linked to higher rates of preterm birth stand out, particularly in the PROLONG ex-U.S. population. Different social and demographic characteristics between the PROLONG and Meis populations are provided in Table 7.

Table 7
Different Social and Demographic Characteristics Across PROLONG and Meis Trials

Demographics/Baseline Characteristics	**Meis (N=463) %**	**U.S. PROLONG (N=391) %**	**PROLONG (N=1708) %**
Age (years), mean ± SD	26.2 ± 5.6	27.6 ± 5.1	30.0 ± 5.2
>1 previous SPTB	28.9	27.4	14.5
GA of prior SPTB (median)	32 wks	34 wks	33 wks
Black/African American	59.0	28.9	6.7
Hispanic or Latino	14.9	13.8	9.1
Unmarried with no partner	50.3	30.7	10.1
Educational status (≤ 12 years)	71.3	50.5	43.7
Any substance use during pregnancy	26.1	28.4	9.3

Fewer higher risk patients were enrolled in PROLONG. For example, in the U.S. PROLONG subgroup, the median gestational age of the prior spontaneous preterm birth was a full two weeks later than in the Meis trial (34 weeks versus 32 weeks). Further, the U.S. subgroup had half the proportion of Black women as compared to the Meis trial. There were 273 (59%) Black women in the Meis trial, and only 111 (29%) Black women in the U.S. PROLONG subset.

This is particularly worth noting because there are significant disparities in preterm birth rates in the U.S. between Black and non-Hispanic white women that are not explained by geographic or socioeconomic factors.[189] As demonstrated graphically in Figure 10 below, the March of Dimes 2021 Report Card states that in the U.S., the preterm birth rate among Black women is 14.0%—which is 51% higher than the rate among all other women in the U.S. Native American/Alaskan women experience the next highest rates of preterm birth (11.7%), followed by Hispanic women (9.8%).[190] Figure 11 depicts the proportion of women who have an early preterm delivery (defined as less than 34 weeks of gestation) versus a late preterm delivery, by race. The disparity among Black women is even more stark, with nearly double the rate of early preterm birth compared to all other races. As is well-recognized, early preterm births are more likely to be associated with neonatal complications, some of which may persist, including developmental delays.

[189] *See* Heather Burris et al., *Racial Disparities in Preterm Birth in the US; A Biosensor of Physical and Social Environmental Exposures,* 104 ARCH. DIS. CHILD 931-35 (2019), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6732250/; Brittany Chambers et al., *Using Index of Concentration at the Extremes as Indicators of Structural Racism to Evaluate the Association with Preterm Birth and Infant Mortality—California, 2011–2012,* 96 J. URBAN HEALTH 159-70 (2019), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6458187/; Suzan L. Carmichael et al., *Population-level Correlates of Preterm Delivery Among Black and White Women in the U.S.*, 9 PLOS ONE e94153 (Apr. 16, 2014), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3989227/pdf/pone.0094153.pdf; Britt McKinnon et al., *Comparison of Black-White Disparities in Preterm Births Between Canada and the United States,* 188 CMAJ E19-E26 (Jan. 5, 2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4695373/.

[190] March of Dimes, *supra* note 31, at 12.

Figure 10
Percentage of Live Births in 2017-2019 (Average) Born Preterm[191]

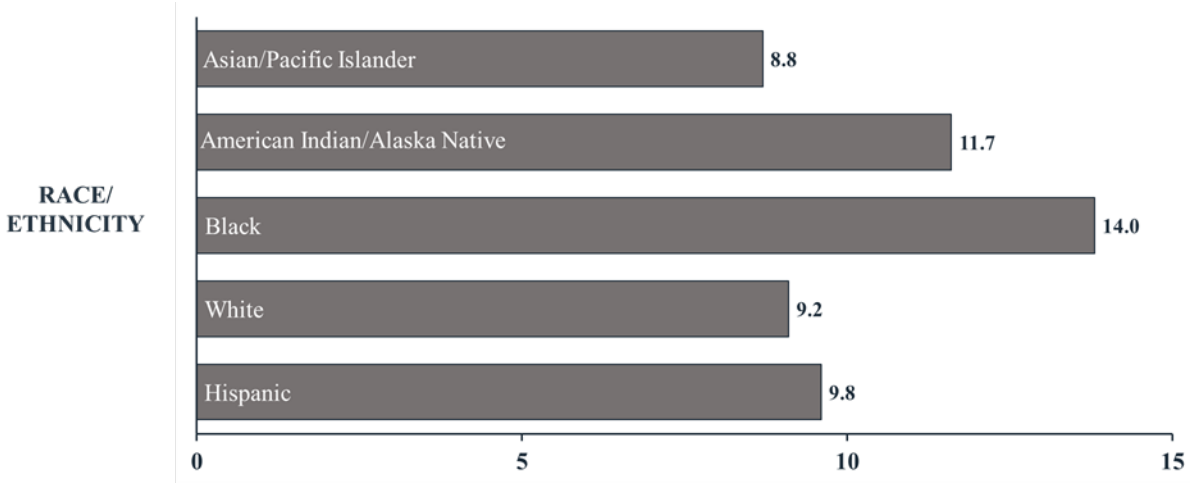
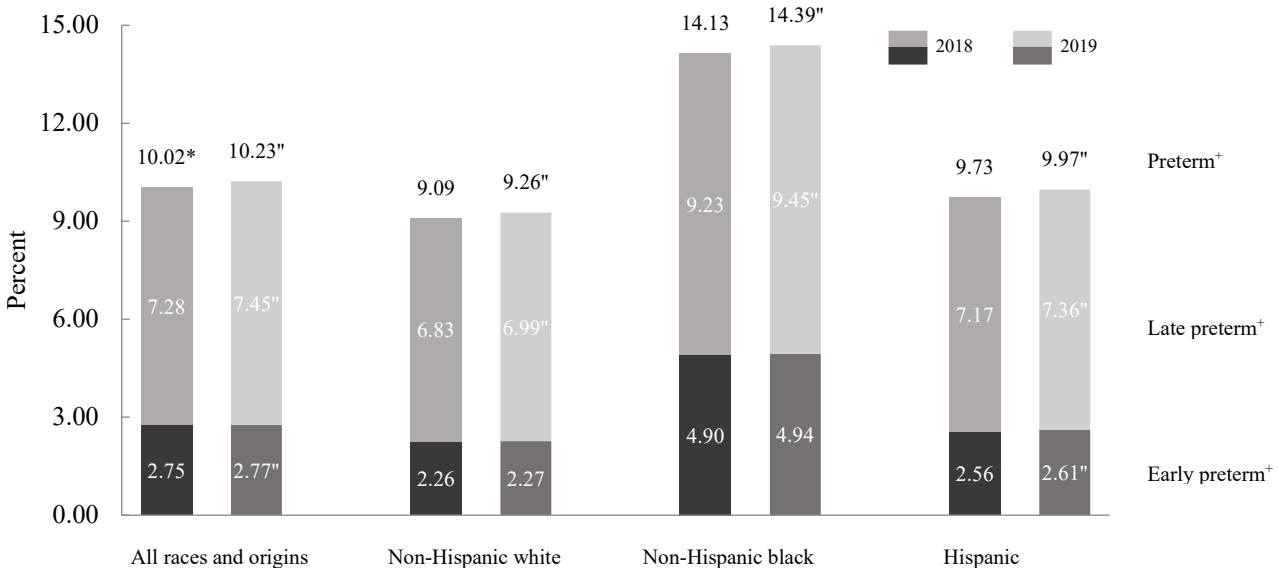


Figure 11
Preterm Birth Rates in the U.S. by Race and Ethnicity (2018 and 2019)[192]



* Data do not add to totals due to rounding.

" Significant increase from 2018 ($p<0.5$).

+ Significant difference between all race and Hispanic-origin groups ($p<0.05$).

PROLONG is therefore flawed in its lack of inclusion of patients at higher risk of preterm birth, including American Black and other minority women and consequently is not

[191] *Id.*

[192] Martin et al., *supra* note 136.

generalizable to the U.S. population that may benefit from treatment. It is well recognized that conditions and efficacy of treatments can vary significantly by minority subpopulation. These populations may be ill-served by studies that are inadequately representative, such as PROLONG, particularly where there are trends that suggest increased efficacy in these groups. The Center's proposal to withdraw Makena is misplaced particularly because PROLONG's patient population was fundamentally different and not representative of the U.S. high-risk population, and the potential outsized impact it will have on populations that were underrepresented in the clinical research.

c. PROLONG Relied On Inconsistent And Unreliable Methods To Verify The Gestational Age Of The Qualifying Delivery

In the last decade, U.S. obstetricians have followed WHO guidelines and generally used ultrasound to measure the crown-rump length (CRL) or gestational sac mean diameter to determine gestational age.[193] An ACOG recommendation explains that first semester ultrasound measurement is accurate in establishing gestational age, while use of ultrasound in the second and third trimester are not similarly reliable for calculations of gestational age. Notably, these obstetric practices were not common in ex-U.S. countries during the time that PROLONG was recruiting (2009 to 2018). In particular, in Russia and Ukraine, where the majority of the PROLONG study population was enrolled, the practice history of gestational age determination has not been as clearly defined. What data exists suggests that the use of ultrasound during the first trimester of pregnancy to determine and document gestational age was less prevalent and/or not standardized when compared with rates of ultrasound utilization in the United States.

Covis' review of the intake forms for PROLONG participants confirms that there was no uniform method used to confirm the gestational age of the patients' qualifying delivery. The study protocol also did not require that a certain method be used. Instead, the protocol instructed that "where possible," the gestational age of the qualifying delivery should be determined by a combination of the last menstrual period method and ultrasound examination, and the screening criteria did not require any verification of the gestational age of the prior birth in the mother's medical record. It is therefore likely that the gestational age of the qualifying delivery was based on the last menstrual period for many of the ex-U.S. PROLONG patients.

The last menstrual period method is generally known to be unreliable and is recommended only when ultrasonography facilities are not available. This method relies on a number of assumptions, namely, a regular menstrual cycle of 28 days, with ovulation occurring on the 14th day after the beginning of the menstrual cycle. As ACOG has recognized, the last menstrual period method therefore does not account for irregularities in cycle length, variability in the timing of ovulation, or even the fact that women may inaccurately recall their last menstrual period. Indeed studies have shown that approximately one half of women inaccurately recall their last menstrual period and as many as 40% of the women experienced more than 5-day discrepancies in the estimated due date between ultrasound dating and last

[193] *See* Scott N. MacGregor and Rudy E. Sabbagha, *Assessment of Gestational Age by Ultrasound*, GLOB. LIB. WOMEN'S MED. (2008), https://www.glowm.com/section-view/heading/Assessment%20of%20Gestational%20Age%20by%20Ultrasound/item/206#.

menstrual period dating. Indeed, it has been estimated that up to one quarter of the preterm births that were classified using last menstrual period may in fact not be preterm.[194]

Further, there has been a marked difference in the use of obstetric ultrasonography in Russia as compared with the United States. By 2015, 90% of women in the United States received ultrasounds during the first trimester compared with just 77% of women in Russia.[195] Moreover, prior to 2009, the use of ultrasonography specifically with respect to gestational measurement was, at best, sporadic. The use of ultrasound along with last menstrual period has been documented in Northwest Russia in the early 1990s, however, only a small portion of these births refer to ultrasound with last menstrual period being the primary method by which gestational age determination was made.[196] Additionally, Covis conducted a comprehensive literature search and was unable to identify any literature pointing to ultrasonography as the standard method of gestational age determination in Russia. Conversely, several published studies on various aspects of maternal and pediatric health which cite last menstrual period as the method for gestational age determination.[197]

The use of ultrasound for gestational age determination in Russia remains limited today. The authors of a recent review of the fetal growth calculation effort in Russia note that "[t]here is no consensus on fetal growth monitoring in modern Russia. Neither the Russian Society of Obstetricians and Gynecologists, nor the Russian Association of Specialists in Ultrasound Diagnostic in Medicine has ever published any clinical recommendations concerning the application of fetal growth charts."[198] Russian clinicians use varying fetal growth charts, with no consistent quantitative methodology or underlying clinical or biological hypothesis.[199] As a result, fetal growth gestational age measurements suffer from several methodological errors

---

[194] Michael S. Kramer, et al., *The Validity of Gestational Age Estimation by Menstrual Dating in Term, Preterm, and Postterm Gestations*, 22 JAMA 3306-3308 (Dec. 9, 1998), https://jamanetwork.com/journals/jama/article-abstract/375526.

[195] *See* Shuvalova et al., *supra* note 33; O'Keefe, et al. *supra* note 33.

[196] *See* Postoev, et al., *Changes in detection of birth defects and perinatal mortality after introduction of prenatal ultrasound screening in the Kola Peninsula (North-West Russia): combination of two birth registries*, 15 BMC Pregnancy Childbirth 308 (Nov. 23, 2015), https://pubmed.ncbi.nlm.nih.gov/26596677/; Anna A. Usynina, *Risk factors for perinatal mortality in Murmansk County, Russia: a registry-based study,* 1 Glob Health Action (2017), https://pubmed.ncbi.nlm.nih.gov/28156197/.

[197] *See* Tatiana B. Makukhina and Viktoriia V. Makukhina, *Ectopic low-lying implantation pregnancy: analysis of outcomes depending on gestation age*, 8 RUSSIAN OPEN MEDICAL JOURNAL 208 (2019); Anna A. Usynina, et al., *Gestation-specific live-born singleton newborns birth weight, length and head circumference percentiles and curves (Arkhangelsk County birth registry data)*, 24 HUMAN ECOLOGY 56-64 (2017); Anton A. Kovalenko et al., *Under-reporting of major birth defects in Northwest Russia: a registry-based study*, 76 INT. J. CIRCUMPOLAR HEALTH (Aug. 30, 2017); Ekaterina E. Sharashova, Erik E. Anda, and Andrej M. Grjibovski, *Early pregnancy body mass index and spontaneous preterm birth in Northwest Russia: a registry-based study*, 14 BMC PREGNANCY AND CHILDBIRTH 1-8 (2014); Tatiana Sherkunova, Changing trends in caesarean section births in Murmansk County, Russia, (Mar. 2014) (unpublished M.P.H. thesis, The Arctic University of Norway); Erik E. Anda et al., *Implementation, Quality Control and Selected Pregnancy Outcomes of the Murmansk County Birth Registry in Russia*, 67 INT. J. CIRCUMPOLAR HEALTH 318-334 (2008).

[198] A. M. Kholin, et al, *Ways to standardise of fetometry in Russia: INTERGROWTH-21st project and its implementation*, 9 Obstetrics and Gynecology (2018), https://en.aig-journal.ru/articles/Podhody-k-standartizacii-fetometrii-v-Rossii-proekt-INTERGROWTH-21-i-ego-vnedrenie.html.

[199] *See id*.

including inaccurate gestational age measurements, inaccuracy in population differentiation and inclusion/exclusion criteria, and deficiencies in imaging standardization protocols.[200] Therefore, although ultrasound screening was mandated for all outpatient clinics in Russia in the year 2000, there is limited evidence of documented ultrasound-estimated gestational age prior to 2009 and after 2009, even where implemented, the methodologies used are inconsistent and not aligned with WHO standards.

In the case of Ukraine, it is even more clear that last menstrual period was the prevailing method for first trimester gestational age determination in PROLONG. Though the vast majority of pregnant women in Ukraine routinely received ultrasounds as far back as 1995, the standard of care in Ukraine recommended second trimester ultrasounds, not first trimester ultrasounds.[201] A comprehensive study of the outcomes of 17,137 pregnancies published in 1999 cited the use of last menstrual period to estimate gestational age.[202] More recently, authors of a 2019 study on maternal alcohol use among Ukrainian mothers noted difficulty in obtaining gestational age when assessing pre-term birth due to the lack of ultrasound screenings and poorly documented last menstrual period records.[203] Similar to the Russian case, use of first trimester ultrasound in Ukraine has been poorly documented and last menstrual period appears to be the primary method used to determine gestational age in early pregnancy.

In sum, these data show that the prior gestational age history of PROLONG subjects in Russia and Ukraine was unreliable both at a qualitative level (i.e., whether or not the subject had a prior spontaneous preterm birth), particularly if the documented SPTB was at week 35 or later due inaccuracies as a result of use of the last menstrual period method, and at a quantitative level, where risk modeling based on prior gestational age history necessarily suffers from reliance on data of poor quality.

d. Leading Medical Organizations ACOG And SMFM Agree That PROLONG Does Not Negate the Strong Findings Of The Meis Trial

The leading medical societies, ACOG and SMFM, recognized the disconnect between the Meis patient population and the PROLONG patient population and what it meant for interpretation of the PROLONG results. Following publication of the PROLONG results, ACOG stated that it would maintain its current recommendation to use progesterone supplementation in women with a prior spontaneous preterm birth.[204] As ACOG stated: *"It is well known that infants born prematurely have increased risks of poor outcomes, including death, and that the risk decreases as gestational age increases.* In fact, *preterm birth is the*

---

[200] *See id.*

[201] *See* Arbuzova, *supra* note 233, at 184 (noting that "General ultrasound screening twice, at 16-18 and at 24-27 weeks of pregnancy, is recommended").

[202] Ruth E. Little et al., *Outcomes of 17 137 Pregnancies in 2 Urban Areas of Ukraine*, 12 AM. J. PUBLIC HEALTH 1832-36 (1999).

[203] Claire D. Coles et al., *Gestational age and socioeconomic status as mediators for the impact of prenatal alcohol exposure on development at 6 months*, 12 BIRTH DEFECTS RES. 789-96 (Jul. 15, 2019), https://pubmed.ncbi.nlm.nih.gov/30378744/.

[204] ACOG Statement on 17p, *supra* note 22; ACOG Practice Advisory, *supra* note 22.

leading cause of neonatal mortality in the United States. Preventing preterm birth can help give babies a better chance at a healthy life."[205] ACOG further acknowledged that PROLONG may have been underpowered to assess efficacy and that, due to treatment guidelines recommending use of 17-OHPC, there may have been an "unintentional selection bias" in enrollment in a placebo-controlled trial. This unintentional selection bias against enrolling higher risk patients in the U.S. in PROLONG likely resulted from the desire by physicians to treat their highest risk patients with therapy rather than risk the patients being randomized to placebo, or patients, after being educated, declining to participate in the trial in favor of receiving treatment.

For its part, the SMFM stated "[b]ased on the evidence of effectiveness in the Meis study, which is the trial with the largest number of US patients, and given the lack of demonstrated safety concerns, SMFM believes that it is reasonable for providers to use 17-OHPC in women with a profile more representative of the very-high-risk population reported in the Meis trial."[206] Notably, SMFM specifically acknowledged that "substantial differences in the [study populations] likely account for the different baseline rates of recurrent [preterm birth] and potentially explain some of the contrasting results observed in the Meis and PROLONG trials."[207]

e. An Exploratory Analysis Of PROLONG-US Data Suggests Efficacy In The High-Risk Population

Covis Pharma has re-analyzed both PROLONG and Meis to construct models that seek to predict the likelihood of a preterm birth in an individual with a prior spontaneous preterm birth. Notably, Covis has rigorously examined several variations of the prior pregnancy history, one of the most significant risk factors for subsequent preterm birth as described above, to further characterize and model the probability of a subject having a preterm birth. Specifically, Covis has explored the average of prior gestational ages of live birth pregnancies (referred to as mean gestational and or mGA), by various cutoffs (<35 weeks, <34 weeks, <32 weeks, <28 weeks etc.), the number of preterm births whose gestational age was less than a given cutoff (<37 weeks, <35 weeks, etc.), and the gestational age of most recent pregnancy preceding the study (referred to as the mrpGA). This last factor was included because there were examples of patient histories in all datasets where a patient had multiple pregnancies but where the more recent pregnancies had continued to full term. In general, clinicians would view such patients as being of lower risk.

In addition to the conventional categorical endpoints of PTB rate at specific cutoffs (e.g., <37, <35. <32 weeks), Covis examined a variety of continuous endpoints that were designed to probe whether 17-OHPC was extending the pregnancy and adding any additional time in utero relative to placebo. Our hypothesis was that these continuous endpoints would be more sensitive than the categorical endpoints used in Meis or PROLONG and may tease out a signal where the categorical endpoints did not show an effect.

---

[205] ACOG Statement on 17p, *supra* note 22.

[206] SMFM *supra,* note 22.

[207] *Id.* at 3.

We performed analyses on the change from baseline to the study pregnancy where baseline was the mrpGA, as well as time from randomization to birth. For both of these continuous endpoints, we saw signals of efficacy in terms of weeks gained by 17-OHPC relative to placebo in PROLONG-US, particularly in subgroups known to be higher risk such as those with a more severe history of preterm births as well as the Black subpopulation. As shown in the Appendix, similar effects were seen in the Meis data when analyzed in this manner. For ex-US PROLONG, however, there was no difference in weeks gained for any subgroups regardless of risk factors.

The endpoint for all analyses was time (weeks) from randomization until the earlier of (1) delivery or (2) 35 weeks gestation (i.e., time capped at 35 weeks gestation). The analysis population included only women randomized up to 19 weeks and 6 days gestation as women randomized at 20 weeks gestation or later were excluded (on account of previous CDER statistical reviews for the Meis study that noted that the treatment effect of 17-OHPC was present only when the subjects were randomized before week 20[208]). All analyses were performed using linear regression with time from randomization (capped at 35 weeks gestation) as the dependent variable. All analyses included treatment and GA at randomization as independent variables. Analyses of high-risk subgroups (Table 3) also included adjustment for mean gestational age of prior spontaneous deliveries. Analyses among subsets defined by the mrpGA among spontaneous deliveries and mGA were also adjusted for mGA (Table 8) and mrpGA (Table 9), respectively. Only spontaneous births at a minimum of 20 weeks GA were included in the calculation of mrpGA and mGA. The reason to cap at 35 weeks gestation was to focus on the period of gestation viewed as most beneficial to the fetus from the perspective of increased time in utero.

Of note, the table shows a clear numerical increase in weeks gained by 17-OHPC versus placebo as we analyze subgroups with a larger number of risk factors.[209] As shown in Manuck et al. (2016) and Richter et al. (2019), this is clinically significant as the addition of 1-2 weeks of gestational age prior to week 35 is associated with marked reduction in neonatal morbidities.[210]

---

[208] NDA 21945, Statistical Reviews at 16, https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000StatR.pdf.

[209] *See* Spong et al., *Progesterone for prevention of recurrent preterm birth: Impact of gestational age at previous delivery*, 193 A. J. OBSTET. GYNECOL. 1127 (2005); Mercer et al., *Are Women with Recurrent Spontaneous Preterm Births Different from Those Without Such History?*, 194 A. J. OBSTET. GYNECOL. (2006).

[210] *See* Tracy A. Manuck et al., *Preterm Neonatal Morbidity and Mortality by Gestational Age: A Contemporary Cohort*, 215 A. J. OBSTET. GYNECOL. 103.e1–103.e14 (2016), Lindsay A. Richter et al., *Temporal Trends in Neonatal Mortality and Morbidity Following Spontaneous and Clinician-Initiated Preterm Birth in Washington State, USA: A Population-Based Study*, 9 BMJ OPEN e023004 (2019).

Table 3
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Known Risk Factor Subgroup Among Subjects Randomized Prior to 20 Weeks GA for PROLONG-US

Risk Factor Subgroup	N Total	Estimated treatment effect* (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
Overall	389	0.49	-0.04	1.01	0.0684
Most Recent Prior Spontaneous Delivery at GA<35 (mrpGA<35)	137	1.30	0.30	2.29	0.0113
Black Subjects with mrpGA<35	51	1.57	-0.28	3.42	0.0936
Subjects with Inter-pregnancy Interval <5 Years (IPINT<5) and mrpGA<35	112	1.55	0.34	2.76	0.0126
Subjects with More than One Prior sPTB<37 (MTO37) and mrpGA<35	23	0.99	-0.74	2.72	0.2470
Subjects with IPINT<5 and MTO37 and mrpGA<35	16	2.08	-0.54	4.69	0.1099
Black Subjects with IPINT<5 and mrpGA<35	38	1.75	-0.77	4.26	0.1673
Black Subjects with MTO37 and mrpGA<35	9	-0.10	-0.57	0.37	0.6056

* Within group estimates for the 17P treatment effect (weeks gained from randomization, capped at GA=35) based on model including: Treatment, Mean GA of Prior Spontaneous Deliveries (mGA), and GA at Randomization.

In order to refine this analysis, with an eye towards defining the inclusion criteria for a future confirmatory study in higher risk patients, we also examined the relationship between prior pregnancy history cut points and weeks gained on 17-OHPC relative to placebo. For PROLONG-US, the two prior pregnancy history measures we analyzed were the mean prior GA and the mrpGA. These results are presented in Tables 8 and 9 below:

Table 8
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Mean Gestational Age (mGA) of Prior Deliveries Among Subjects Randomized at <20 Weeks GA for PROLONG-US

mGA Subgroup	N Total	Estimated treatment effect (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
mGA<28	28	3.48	0.60	6.36	0.0198
mGA<29	34	2.56	0.20	4.93	0.0348
mGA<30	41	2.20	0.23	4.17	0.0295
mGA<31	54	0.96	-0.87	2.79	0.2970
mGA<32	56	0.97	-0.82	2.75	0.2817
mGA<33	81	1.01	-0.27	2.29	0.1214
mGA<34	101	0.89	-0.54	2.32	0.2186
mGA<35	142	0.42	-0.65	1.48	0.4399
mGA<36	191	0.47	-0.36	1.30	0.2688
mGA<37	254	0.54	-0.17	1.24	0.1351

Table 9
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Most Recent Prior Gestational Age (mrpGA) of Previous Deliveries Among Subjects Randomized at <20 Weeks GA for PROLONG-US

mGA Subgroup	N Total	Estimated treatment effect (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
mrpGA<28	37	3.33	0.93	5.73	0.0081
mrpGA<29	45	2.48	0.39	4.56	0.0210
mrpGA<30	51	2.20	0.39	4.00	0.0183
mrpGA<31	57	1.82	0.22	3.42	0.0262
mrpGA<32	64	1.60	0.16	3.04	0.0297
mrpGA<33	84	1.26	0.08	2.45	0.0371
mrpGA<34	101	1.44	0.29	2.59	0.0149
mrpGA<35	137	1.43	0.42	2.44	0.0058
mrpGA<36	195	0.98	0.22	1.74	0.0118
mrpGA<37	248	0.96	0.24	1.67	0.0090

The above two tables demonstrate a monotonic relationship between the weeks gained on 17-OHPC and the risk of the subject as defined by either the GA from their pregnancy immediately prior to enrolling in the study or the mean of their prior GAs.

We have also repeated this analysis for the ex-US PROLONG subgroup but did not see a benefit from 17-OHPC treatment in terms of the weeks gained since randomization. This may be due to:

1. The overall risk level of the ex-US subjects is low. While the preterm birth rate in ex-US PROLONG is enriched relative to the general population, the PTB rate for the placebo group in ex-US PROLONG, Russia and Ukraine was 20%, 17% and 21% respectively (compared to 28% in PROLONG-US).

2. Classifying patients according to their mrpGA or mean prior GA has issues particularly with respect to Russia and Ukraine, where standards for determination of gestational age have been inconsistently applied. As we note above, prior births likely involved determination of gestational age using LMP, particularly given the dates of the qualifying births in Russia and Ukraine and the indirect evidence of the use of LMP in various registries and databases in those countries. The uncertainty in the use of LMP in these countries renders the validity of subjects' mrpGA suspect, and indeed it is surprising that there is even a remnant of any trend remaining from this analysis.

3. Neither measure of prior birth history (mean GA or mrpGA) correlate with birth outcome in ex-US PROLONG. In other words, they are not good models for predicting risk for ex-US PROLONG. This may be expected on account of the uncertainties in the prior birth history. Given the overall lack of signal seen in ex-US PROLONG, it is not surprising that subgroup analyses by these prior history measures also showed no signal.

For the same reason, we believe that any subgroup analysis on the ex-US PROLONG subjects using risk factors are unlikely to provide insight because prior birth history – the single most important predictor of preterm birth risk – has not been measured in a comparable fashion to the US population.

Tables 3, 8 and 9 above show that there is a treatment effect from 17-OHPC when assessed based on a more sensitive endpoint such as weeks gained from randomization. In addition, due to capping of the weeks gained, the incremental weeks gained from 17-OHPC treatment focused on <35 weeks for which there is general consensus on benefit for the neonate with respect to morbidity/mortality as shown in Manuck et al. (2016) and Richter et al. (2019). Covis believes that the concept of added weeks of gestation has a clearer clinical interpretation in comparison to a categorical endpoint such as the rate of preterm birth at a given cutoff such as 35 weeks. Further, by picking the cut point at which we cap weeks gained, we ensure that any difference between the treatment arms is focused in a time window that is clinically relevant for neonatal development.

In sum, these analyses give rise to a strong suggestion that 17-OHPC may be effective for the highest-risk patients and highlight the need for further focused studies in this cohort.

4. The Meis Trial And PROLONG Establish Makena's Favorable Safety Profile

The safety of 17-OHPC for pregnant women and their babies has been demonstrated by the Meis trial and by PROLONG.

The Meis trial demonstrated the positive safety profile of Makena, and in CDER's own words, "[t]here were no safety findings," as noted in the Center's review of the trial at the time.[211] The most common type of adverse event (AE) reported during the Meis study was injection site reactions, which was expected as patients received weekly intramuscular injections. The incidence of pregnancy complications, such as preeclampsia, gestational diabetes, or clinical chorioamnionitis, as well as the incidence of serious adverse events (SAEs), was not different between the 17-OHPC and placebo arms.[212] There was a non-statistically significant trend toward an increase in the second trimester miscarriage rate and stillbirth rate in the 17-OHPC arm. Conversely, however, the incidence of neonatal deaths was reduced in the 17-OHPC group, and the overall incidence of combined fetal and neonatal mortality from treatment onset to delivery was similar in both groups.[213]

The follow-up study (Study 17P-FU) which examined outcome data at two years of age or greater on the children born to women treated in the Meis study also revealed no differences in developmental delays, safety concerns related to overall health or physical development, or genital or reproductive anomalies between children with *in utero* exposure to placebo versus 17-OHPC. The authors of the follow-up study therefore concluded, "this study provides reassurance

[211] NDA 21945 Medical Review at 63.

[212] *See* AMAG Pharmaceuticals, Advisory Committee Briefing Materials – Bone, Reproductive, and Urologic Drugs Advisory Committee Meeting, 26 (Oct. 29, 2019), https://www.fda.gov/media/132004/download.

[213] NDA 21945 Medical Review at 63.

that 17 α- hydroxyprogesterone caproate is safe for the fetus when administered in the second and third trimesters."[214] Although PROLONG did not confirm 17-OHPC's efficacy, it did reaffirm the favorable maternal and fetal safety profile of 17-OHPC by meeting the key safety objective, which was to rule out a doubling in the risk of fetal and early infant death in the 17-OHPC group compared to placebo. A "doubling of risk" was selected and agreed upon with FDA based on sample size consideration as well as clinical relevance given the expected low rate of the outcome. Assuming 4% fetal or early infant death rate in both treatment groups, a sample size of 1707 provided 83% power to rule out doubling in risk of fetal or early infant death. For context, to rule out a 1.5-fold increase in risk, the sample size needed would be ~4300.

In PROLONG, the rate of fetal/early infant death was low in both treatments groups; 1.7% in the 17-OHPC and 1.9% in placebo groups (RR = 0.87, 95% CI 0.42–1.81). Given that the upper bound of the 95% CI was <2.0, a doubling in the risk of fetal/early infant death was excluded. *Thus, the primary safety objective was achieved for PROLONG*. Of note, in the PROLONG trial, the rate of miscarriage for 17-OHPC was lower than placebo: RR 0.28 (95% CI 0.08–0.94).

With regard to stillbirth, 1.1% and 0.5% of patients in the 17-OHPC and placebo groups experienced a stillbirth with a RR of 2.07 (95% CI: 0.59-7.29). There is no known biological hypothesis indicating that 17-OHPC would increase the risk of stillbirth. Moreover, Dr. Baha Sibai conducted a blinded review of the clinical study report narratives for each of the 12 stillbirth cases; 11 of these 12 cases had identified underlying contributing factors distinct from 17-OHPC (e.g., infection, abruption, placental infarcts).[215]

Indeed, CDER itself acknowledged in the October 2019 Briefing materials: "Although the number of fetal and neonatal deaths are too low to draw definitive conclusions, *the findings of this safety outcome appear to be similar between placebo and Makena*."[216]

Moreover, a recent publication by Sibai *et al*. in the *Journal of Perinatology*, explains that the integrated safety data from the Meis and PROLONG trials demonstrate a favorable safety profile, comparable to placebo, for maternal and fetal risks.[217] Excerpted below from this publication, Table 10 contains the relevant obstetrical outcomes/events in 2% or more of women in the 17-OHPC group, or at a higher rate in the 17-OHPC group versus Placebo. As can be readily observed by examining the data in the table, the Adverse Events in the integrated dataset were low and comparable between 17-OHPC and placebo. Notably, discussion of safety at the FDA Advisory Committee was limited, presumably because both the Committee and FDA recognize that: 1) PROLONG successfully achieved its primary safety objective 2) no new safety concerns were raised with PROLONG and 3) the integrated dataset is reassuring in regard to both maternal and fetal safety.

---

[214] Northen et al., *Follow-Up of Children Exposed In Utero to 17 alpha-Hydroxyprogesterone Caproate Compared With Placebo*, 110 OBSTET. GYNECOL. 865-72 (2007), https://pubmed.ncbi.nlm.nih.gov/17906021/.

[215] Baha Sibai et al., *Safety Review of Hydroxyprogesterone Caproate in Women With a History of Spontaneous Preterm Birth*, 41 J. PERINATOL. 718-25, 722 (2021), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8049867/.

[216] FDA Briefing Document at 45 (emphasis added).

[217] Sibai et al., *supra* note 215, at 723-24.

Table 10
Relevant Obstetrical Outcomes and Events Occurring in ≥2% of the Women in the 17-OHPC Group or at a Higher Rate in the 17-OHPC Versus Placebo Group[218]

	Meis		PROLONG		Integrated	
	17-OHPC	Placebo	17-OHPC	Placebo	17-OHPC	Placebo
	N=310 *n* (%)	*N*=153 *n* (%)	*N*=1128 *n* (%)	*N*=578 *n* (%)	*N*=1438 (%)	*N*=731 (%)
Admission for preterm labor (other than delivery admission)	49 (16.0)[c]	21 (13.8)	187(16.5)[c]	84 (14.5)	16.4[c]	14.4
Preeclampsia or gestational hypertension	27 (8.8)[c]	7 (4.6)	47 (4.2)[c]	30 (5.2)	5.2[c]	5.1
Nausea	18 (5.8)	7 (4.6)	55 (4.9)	26 (4.5)	5.1	4.5
Gestational diabetes	17 (5.6)[c]	7 (4.6)	35 (3.1)[c]	21 (3.6)	3.6[c]	3.8
Headache	4 (1.3)	0	68 (6.0)	28 (4.8)	5.0	3.8
Injection site pruritus	18 (5.8)	5 (3.3)	42 (3.7)	23 (4.0)	4.2	3.8
Injection site swelling	53 (17.1)	12 (7.8)	5 (0.4)	2 (0.3)	4.0	1.9
Back pain	4 (1.3)	1 (0.7)	50 (4.4)	20 (3.5)	3.8	2.9
Vomiting	10 (3.2)	5 (3.3)	42 (3.7)	19 (3.3)	3.6	3.3
Urticaria	38 (12.3)	17 (11.1)	5 (0.4)	0	3.0	2.3
Constipation	2 (0.6)	1 (0.7)	38 (3.4)	17 (2.9)	2.8	2.5
Insomnia	2 (0.6)	1 (0.7)	36 (3.2)	13 (2.2)	2.6	1.9
Cervical incompetence/cerclage	5 (1.6)	2 (1.3)	34 (3.0)	16 (2.8)	2.4	2.2
Injection site nodule	14 (4.5)	3 (2.0)	18 (1.6)	9 (1.6)	2.2	1.6
Diarrhea	7 (2.3)	1 (0.7)	23 (2.0)	13 (2.2)	2.1	1.9
Oligohydramnios[d]	11 (3.6)[c]	2 (1.3)	9 (0.8)[c]	12 (2.1)	1.4[c]	1.9
Chorioamninitis[d]	11 (3.6)[c]	5 (3.3)	9 (0.8)[c]	2 (0.3)	1.4[c]	1.0
Cholestasis[d]	0	0	3 (0.3)	5 (0.9)	0.2	0.7
Maternal depression[e]	0	0	5 (0.4)	5 (0.9)	0.3	0.7
VTE[d]	1 (0.3)	0	0	1(0.2)	0.07	0.1

VTE, venous thromboembolism
c N=306 as denominator for these AEs for Meis; N=1130 as denominator for these AEs for PROLONG.
d Included in table given medical relevance for this therapeutic class.
e Included in table given reference in the product insert. Maternal depression was based on AE reporting.

Other potential risks associated with 17-OHPC have been investigated and addressed. For example, though neither ACOG nor individual clinicians raised concerns regarding miscarriage or stillbirth in the Meis trial, the 2006 FDA Advisory Committee focused on a potential "safety signal" for miscarriage and stillbirth, given the observed rates in the 17-OHPC group (3.6%) compared to placebo (1.3%).[219] Therefore, an important consideration in PROLONG was the systematic evaluation of fetal loss to exclude increased risk with 17-OHPC.

[218] Sibai et al., *supra* note 215.

[219] Transcript, FDA Reproductive Health Advisory Committee Meeting on Gestiva, 148:3-5 (Aug. 29, 2006), https://downloads.regulations.gov/FDA-2020-N-2029-0188/attachment_1.pdf.

There are additional NICHD MFMU trials, such as Harper (n= 852), Rouse (n=661), and Caritis (n=134),[220] that did not evaluate the efficacy of 17-OHPC for the prescribed indication (i.e., singleton pregnancy who have a history of singleton spontaneous preterm birth) and therefore are of limited value in evaluating efficacy.[221] Nevertheless, these studies support the positive maternal-fetal safety profile of 17-OHPC.

As a condition of FDA approval of 17-OHPC, in addition to the PROLONG trial, Study 004 (a follow-up study of infants born to mothers in the PROLONG study) was required. This study has been completed and in children born to women treated with 17-OHPC during second (at least 16 weeks of gestation) and third trimester of pregnancy, exposure of 17-OHPC was not associated with behavioral or neurodevelopmental sequelae in children up to ~25 months (the age of assessment in this study).[222] These data confirm the earlier follow-up study from the Meis trial by Northen et al., which reassured the safety of 17-OHPC for the fetus when administered in the second and third trimesters.[223]

Study 004 collected 245 completed Ages and Stages Questionnaires (ASQ), which evaluates five areas of development: communication, gross motor skills, fine motor skills, problem solving and personal-social behavior. If a child scored below the threshold for any domain, he or she was referred for a Bayley-III and neurological exam. The ASQ screen positive subjects were equally distributed between the two groups: 22.8% in the 17-OHPC group vs. 21.8% in the placebo group. The RR of a positive ASQ screen result for subjects in the 17-OHPC group relative to placebo was 1.04 with a 95% CI for the true RR of 0.64 to 1.70. In the Bayley-III Analysis Population, there were 12.8% in the 17-OHPC group and 13.4% in the placebo group that had a score of borderline or worse (score <80), RR=0.96 (95% CI 0.48, 1.91). Similar findings were seen on the neurological exam, with abnormal exam found in 22.8% in the 17-OHPC group versus 21.8% in the placebo group.[224] Thus, Study 004 confirms the findings of

---

[220] Margaret Harper et al., *Omega-3 Fatty Acid Supplementation to Prevent Recurrent Preterm Birth: A Randomized Controlled Trial*, 115 OBSTET. GYNECOL. 234-42 (Feb. 2010), https://pubmed.ncbi.nlm.nih.gov/20093894/; Dwight J. Rouse et al., *A Trial of 17 Alpha-Hydroxyprogesterone Caproate to Prevent Prematurity in Twins*, 357 N. ENGL. J. MED. 454-61 (2007), https://pubmed.ncbi.nlm.nih.gov/17671253/; Steve N. Caritis et al., *Prevention of Preterm Birth in Triplets Using 17 Alpha Hydroxyprogesterone Caproate A Randomized Controlled Trial*, 113 OBSTET. GYNECOL. (Feb. 2009), https://pubmed.ncbi.nlm.nih.gov/19155896/.

[221] Harper et al. assessed whether the addition of an omega-3 fatty acid supplement would reduce preterm birth in women with at least one prior spontaneous preterm birth receiving 17-OHPC. Rouse et al. conducted an RCT of 17-OHPC in 655 women carrying twins, to evaluate 17-OHPC's effect on preterm birth rate and fetal/neonatal outcome. In doing so, the authors specifically acknowledged, "[i]n singleton gestations, 17 alpha-hydroxyprogesterone caproate (17P) has been shown to reduce the rate of recurrent preterm birth. This study was undertaken to evaluate whether 17P would reduce the rate of preterm birth in twin gestations." Rouse et al., at 454. Caritis et al., similarly conducted an RCT of 17-OHPC in 134 women carrying triplets. Makena's labeling specifically states that "Makena is not intended for use in women with multiple gestations or other risk factors for preterm birth." Makena, Prescribing Information at 1 (Feb. 2018), https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s012lbl.pdf.

[222] *See* 17P-FU-004, A Prospective Noninterventional Follow-Up Study of Children Aged 23 to 24 Months, Born to Mothers Who Received Hydroxyprogesterone Caproate Injection, 250 mg/mL, or Vehicle for the Prevention of Preterm Birth (Follow-Up to Study 17P-ES-003 [PROLONG]), Clinical Study Report, https://clinicaltrials.gov/ct2/show/NCT01146990 [hereinafter 17P-FU-004 CSR].

[223] Northen et al., *supra* note 214.

[224] *See* 17P-FU-004 CSR.

the Northen follow-up study to the Meis trial, and provides reassurance of safety when 17-OHPC is given during the time period approved for initiation (i.e., between 16 weeks and 20 weeks/6 days).

In short, there should be no question that the maternal and fetal safety of Makena was confirmed in PROLONG.

5. Recent Studies Further Support The Positive Benefit-Risk Profile Of 17-OHPC

There are a number of studies that further support the efficacy and safety of 17-OHPC. Contrary to CDER's characterization,[225] an observational study by Bastek et al. provides support for efficacy of 17-OHPC, as the authors concluded in their own words, "[t]he significant shift in [gestational age distribution at delivery] towards the late preterm period in TP2 may be due to the introduction of 17P use at our institution."[226] Bastek et al. compared the preterm birth rate and gestational age distribution at delivery among women who delivered at an urban medical center over two time periods, pre-17-OHPC (TP1: January 2004 to December 2005) and post-17-OHPC (TP2: January 2008 to December 2009), to determine the public health impact of 17-OHPC treatment.[227] The time periods were chosen so as the institution adopted a policy in 2006 that 17-OHPC be prescribed to eligible women, defined as women with a singleton pregnancy and a history of prior SPTB (consistent with Makena labeling).

Significantly, Bastek et al. concluded that among women eligible to receive 17-OHPC that delivered preterm within the study period, the gestational age distribution at delivery was 10 days later in TP2 than in TP1 (33.13 vs. 31.64 weeks, $p<0.01$). Moreover, from TP1 to TP2, there were significantly fewer preterm births between 21 and 33 6/7 weeks (54.37 vs. 35.00%) and more preterm births during the late preterm period (45.63 vs. 65.00%, $p<0.01$). Though the authors concluded that there was no difference in the overall rate of preterm birth <37 weeks between TP1 and TP2, this shift toward *late* PTB is clinically relevant, as outcomes of late preterm infants are generally improved compared to that of earlier gestational ages. They noted that there was a significantly greater percentage of Black women, a subgroup at increased risk for preterm birth, in TP2 than TP1. As the authors explained, the data provide "evidence that 17P may have brought us closer towards mitigating the adversity associated with prematurity, which is of great public health significance." [228]

Unadjusted and adjusted OR for having a PTB during each pre-specified gestational age period were calculated, as shown in Table 11 below. Women were 2.3-fold more likely to deliver a preterm infant during the late preterm period in TP2 than TP1 (OR 2.30, 95% CI 1.49-3.54) after adjusting for biologically plausible confounders such as race, insurance, and age.

---

[225] *See, e.g.,* Division Decision Memo, NDA 021945 Makena (hydroxyprogesterone caproate), Withdrawal of Accelerated Approval, at 18-19 (Oct. 5, 2020).

[226] *See* Jamie A. Bastek et al., *Trends in prematurity: What do changes at an urban institution suggest about the public health impact of 17-alpha hydroxyprogesterone caproate?*, 16 MATERN. CHILD HEALTH J. 564–68 (Apr. 12, 2011), https://link.springer.com/article/10.1007/s10995-011-0783-z.

[227] *Id.* at 564.

[228] Bastek et al., *supra* note 226, at 568.

Table 11
Odds of Women Eligible for 17-OHPC Delivering a Preterm Infant During Each Gestational Age Range in TP2 Compared to TP1[229]

Gestational Age Range (Weeks)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)[a]
$21^0 – 23^6$	0.38 (0.16-0.90)	0.39 (0.16-0.93)
$24^0 – 27^6$	0.67 (0.31-1.46)	0.67 (0.31-1.47)
$28^0 – 32^6$	0.56 (0.31-1.01)	0.53 (0.29-0.96)
$32^0 – 34^6$	0.77 (0.43-1.39)	0.75 (0.42-1.36)
$34^0 – 36^6$	2.21 (1.45-3.39)	2.30 (1.49-3.54)

a Adjusted for confounders including race, age, and insurance.

To further test their hypothesis that the trends observed were due to 17-OHPC use, the authors performed analyses on two groups of women to whom 17-OHPC would not be recommended in a current pregnancy: women with a history of prior preterm birth who are currently pregnant with twins and women with no history of prior preterm birth. Neither of these subgroups demonstrated a similar shift to a later gestational age, supporting the authors' hypothesis. Despite these findings, CDER has mischaracterized the conclusions of this study in its Division Decision Memo recommending withdrawal of accelerated approval and in its preliminary briefing book. The Division Decision Memo, for instance, focuses only on statistics that seemingly support CDER's position (e.g., that there was no difference in the institution's rate of preterm birth <37 weeks), while completely omitting any discussion of the authors' conclusions noted above.[230]

A meta-analysis recently published by the Evaluating Progestogens for Preventing Preterm birth International Collaborative (EPPPIC) and funded by the Patient-Centered Outcomes Research Institute (PCORI), also supports the favorable efficacy profile of 17-OHPC. The study is the largest existing individual patient data (IPD) meta-analysis of progestogens—vaginal progesterone, intra-muscular 17-OHPC, and oral progesterone—used to prevent preterm birth, including participant-level data from 31 trials and over 11,000 women and 16,000 offspring. The study included five randomized trials for intramuscular 17-OHPC, two of which were the Meis and PROLONG studies. The EPPPIC study is the first IPD meta-analysis of 17-OHPC in singleton gestation pregnancies.

In conclusion, the EPPPIC authors determined that "[v]aginal progesterone and 17-OHPC both reduced birth before 34 weeks' gestation in high-risk singleton pregnancies."[231] Specifically, 17-OHPC and vaginal progesterone reduced the relative risk of early preterm birth in high-risk singleton pregnancies before 34 weeks, RR of 0.83, 95% CI 0.68-1.01 and RR of 0.78, 95% CI 0.68-0.90, respectively, although the authors acknowledged that the CI for 17-OHPC "just crossed the line of no effect."[232] The authors, however, noted that the two-stage forest plots for preterm birth earlier than 34 weeks showed all but two trials (one vaginal progesterone and PROLONG) lay to the left of equivalence. According to the authors, while some heterogeneity existed between vaginal progesterone trials, there was less variation for 17-

[229] *Id.* at 567.

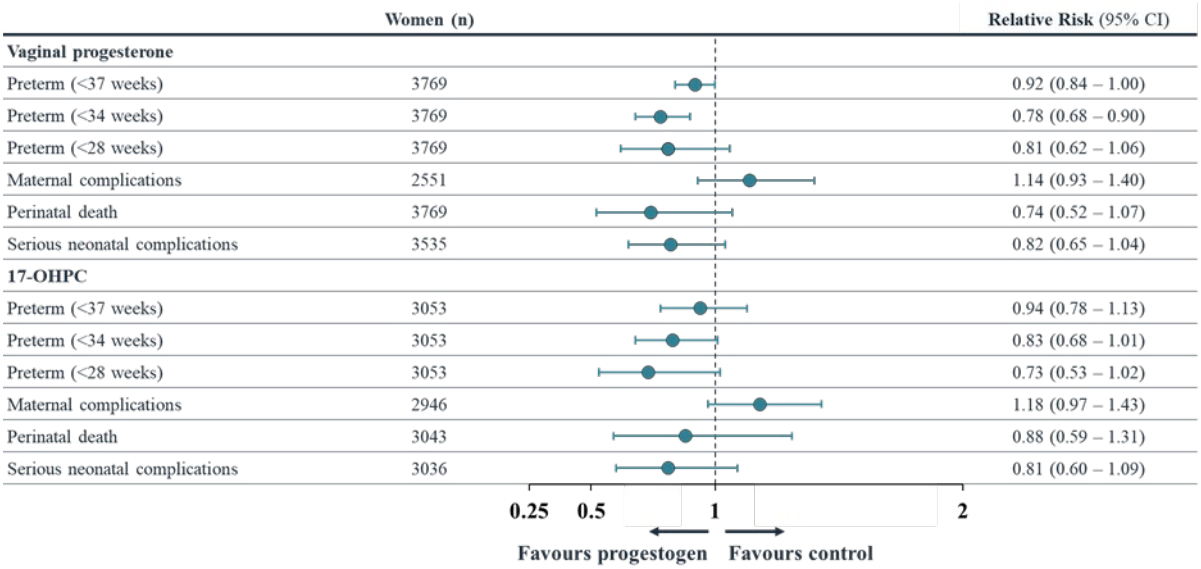
[230] Division Decision Memo *supra* note 225, at 18-19.

[231] The EPPPIC Group *supra* note 16, at 1183.

[232] *Id.* at 1186.

OHPC.[233] The two drugs also showed favorable reductions at before 28 weeks and 37 weeks and indicated potential reductions in serious neonatal complications and incidence of low birthweight infants, as provided in Figure 12 below. It should be recognized that the vaginal progesterone studies were predominantly conducted in women with a short cervix, which is a different risk factor than the prior SPTB. Notably, and in contrast to 17-OHPC, there is *no* RCT of vaginal progesterone in women with a prior SPTB that has demonstrated a reduction in preterm birth compared to placebo.

Figure 12
Preterm Birth and Neonatal Outcomes For Vaginal Progesterone and 17-OHPC



The same MFMU trial by Harper et al., also was used for another analysis by Manuck et al., to identify clinical factors that may predict 17-OHPC response. In order to do so, the authors classified 852 women enrolled (all of whom received 17-OHPC) as a responder or non-responder, with "responder" defined as an individual woman extending her gestational period by three weeks or more compared to the gestational age of her earliest prior preterm birth.[234] The researchers identified several risk factors for non-response to 17-OHPC (gestational age of the previous earliest preterm birth, placental abruption in the current pregnancy, gonorrhea or chlamydia infection, and male fetus) suggesting that "women who experience a recurrent PTB or recurrent PTB at a similar gestational age while receiving 17-OHPC have distinctly different clinical and biologic characteristics than those who receive 17-OHPC and deliver at term."[235] Notably, as many as 78.9% of the 852 women enrolled were considered responders, i.e., had

[233] *Id.* at 1186-87.

[234] Tracy A. Manuck et al., *Nonresponse to 17-alpha hydroxyprogesterone caproate for recurrent spontaneous preterm birth prevention: clinical prediction and generation of a risk scoring system*, 215 AM. J. OBSTET. GYNECOL. 622.e1–622.e8 (2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5086280/.

[235] *Id.* at 7.

pregnancies treated with 17-OHPC deliver at least three weeks later compared to the gestational age of the earliest PTB without 17-OHPC treatment. As illustrated in Figure 13:

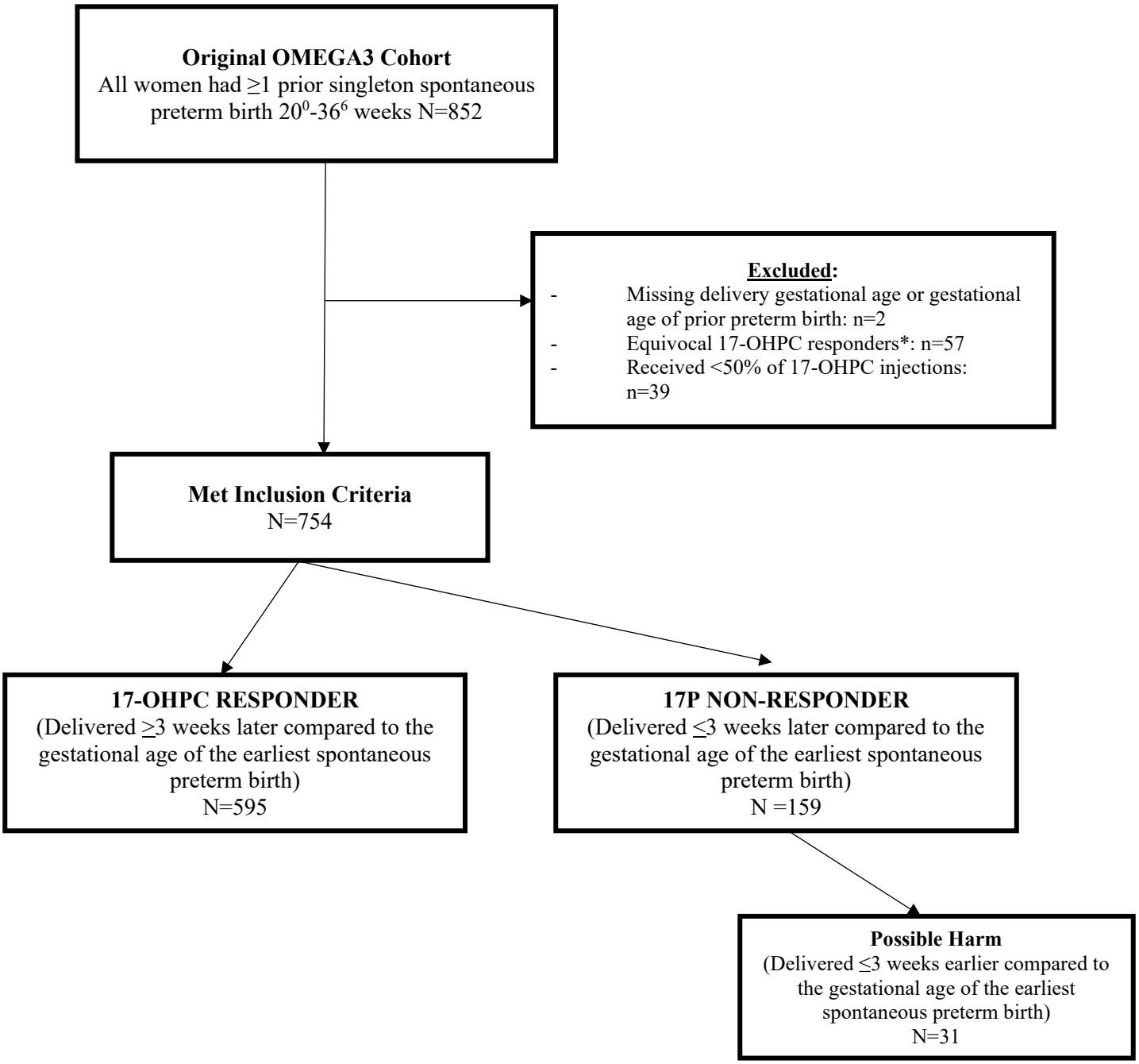
- 595 women (78.9%) were deemed "responders" who delivered 3 weeks or more later than their earliest spontaneous preterm birth, as the flow diagram below depicts

- 159 women (21.1%) were deemed "non-responders" as they delivered <3 weeks compared to their earliest spontaneous preterm birth

- 39 women were excluded from the analysis due to non-compliance (received <50% of doses), missing data, and/or if a woman delivered <3 weeks from her prior GA, but the delivery was 37 weeks or later (i.e., a term pregnancy). For example, if a prior spontaneous preterm birth gestational age was 35 weeks and 17-OHPC therapy delivery was at 37 weeks, Manuck deemed these cases "equivocal."

Although there was not placebo comparison here, these data from MFMU Network Centers suggest that if removal of 17-OHPC occurred, a significant number of women would not have the benefit of extending pregnancy by three weeks or more, compared to their prior earliest delivery.

Figure 13
Prolongation of Pregnancy with 17-OHPC[236]



[236] *Id.*

* Equivocal 17-OHPC responders: Earliest preterm birth was 35-36 weeks, delivered during early term period (37-38 weeks).

Additional cross-sectional data also confirms the real-world benefit linked to treatment with Makena. Carter et al. published a retrospective cohort study using MarketScan® data in 2019 that reviewed 3,374 pregnancies.[237] The researchers found that early 17-OHPC initiation and higher 17-OHPC treatment compliance were associated with lower preterm birth rates. Specifically:

- Women with an early 17-OHPC start (16 weeks to <21 weeks) were less likely to deliver preterm than those with a late start (21 weeks to < 29 weeks) (absolute risk reduction (aRR) 0.88; 95% CI 0.79–0.97; p = 0.02).

- Less compliant patients (receiving <25% of recommended doses) had a higher preterm birth rate than those receiving >85% of recommended doses (aRR 1.5; 95%CI 1.2–1.7; p <0.01).

This further supports the pharmacological activity of 17-OHPC; i.e., if a medication was not effective, one would not expect that receiving fewer doses (such as starting late) or being non-compliant would impact rates of preterm birth.

A recent study by Schuster et al. also used MarketScan® data to conduct a retrospective cohort study of deliveries between April of 2008 and January 2015 and concluded no significantly increased risk of venous thromboembolism (VTE) from 17-OHPC. The study included as many as 4,775,667 delivery hospitalizations, with 18,745 of these women having received 17-OHPC during pregnancy. Among women who did not receive 17-OHPC, 0.52% of women had a VTE diagnosis compared to 0.61% receiving 17-OHPC (RR 1.18, 95% CI 0.98-1.42). The study also revealed that there was no increased risk of VTE for patients receiving versus not receiving 17-OHPC for both during the antenatal period as well as during delivery hospitalizations.[238]

In addition, the safety profile of 17-OHPC is further supported by a study by Price et al., which was a randomized, double-blind, placebo-controlled trial in pregnant women with HIV in Zambia.[239] 800 women were enrolled and randomized 1:1 to receive either 17-OHPC or placebo. According to the authors, "[t]his trial supports the safety of [17-OHPC] administration during pregnancy." Adverse reactions occurred equally between the two arms, and overall incidence was consistent with previous trials of 17-OHPC, as provided below in Tables 12 and 13. While the trial did not show a reduction in preterm birth (<37 weeks of gestation) or stillbirth, the data cannot be considered to accurately reflect 17-OHPC's effectiveness, as the trial studied a very distinct population from that indicated for Makena. Not only did the women in the study have HIV, but women with history of previous spontaneous preterm delivery were also *excluded* from the study under the exclusion criteria. In contrast, Makena is intended for use *only* in women who have a history of singleton spontaneous preterm birth. As a result, there was

---

[237] Ebony B. Carter et al., *Practical considerations with 17-Hydroxyprogesterone caproate for preterm birth prevention: does timing of initiation and compliance matter?*, 39 J. PERINATOL 1182-89 (2019), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6890226/.

[238] Meike Schuster et al., *17-Alpha Hydroxyprogesterone Caproate and Risk for Venous Thromboembolism During Pregnancy,* J. MATERN. FETAL NEONATAL MED. 1-2 (2021), https://www.tandfonline.com/doi/abs/10.1080/14767058.2021.1911997?journalCode=ijmf20.

[239] Price et al., *supra* note 17.

a high proportion of women with previous term deliveries participating in the trial, and the overall frequency of preterm birth or stillbirth in the study population was substantially low—9%, which was even lower than the rate in the local population.

Table 12
Maternal and Neonatal Adverse Events For 17-OHPC and Placebo[240]

Adverse Event	17-OHPC N=399	Placebo N=401	Risk Difference (95% CI)	Relative Risk (95% CI)
Maternal Outcomes				
Pregnancy-induced hypertension	13 (3%)	6 (2%)	1.8% (-0.3 to 3.9)	2.2 (0.8 to 5.7)
Pre-eclampsia	6 (2%)	9 (2%)	-0.7% (-2.6 to 1.1)	0.7 (0.2 to 1.9)
Eclampsia	0 (0%)	1 (<1%)	-	-
Maternal death	0 (0%)	1 (<1%)	-	-
Antepartum hemorrhage	6 (2%)	7 (2%)	-0.2% (-2.0 to 1.5)	0.9 (0.3 to 2.5)
Preterm prelabor rupture of membranes	5 (1%)	7 (2%)	-0.5% (-2.2 to 1.2)	0.7 (0.2 to 2.2)
Oligohydramnios	2 (1%)	2 (1%)	-	-
Polyhydramnios	22 (6%)	19 (5%)	0.8% (-2.3 to 3.8)	1.2 (0.6 to 2.1)
Chorioamnionitis	1 (<1%)	0 (0%)	-	-
Caesarean delivery	73 (18%)	77 (19%)	-0.9% (-6.3 to 4.5)	1.0 (0.7 to 1.3)
Neonatal Outcomes				
Birthweight <10th percentile for gestational age[a]	95/392 (24%)	93/394 (24%)	0.6% (-5.3 to 6.6)	1.0 (0.8 to 1.3)
Birthweight <3rd percentile for gestational age[a]	28/392 (7%)	47/394 (12%)	-4.8% (-8.9 to -0.7)	0.6 (0.4 to 0.9)
Birthweight <2500g[b]	41/395 (10%)	46/395 (12%)	-1.3% (-5.6 to 3.1)	0.9 (0.6 to 1.3)
Birthweight <1500g[b]	7/395 (2%)	7/395 (2%)	0.0% (-1.8 to 1.8)	1.0 (0.4 to 2.8)
1-min Apgar score <7 among liveborn infants[c]	12/336 (4%)	9/338 (3%)	0.9% (-1.7 to 3.5)	1.3 (0.6 to 3.1)
5-min Apgar score <7 among liveborn infants[d]	7/137 (5%)	2/144 (1%)	-	-

[240] *Id.* at e610.

Adverse Event	17-OHPC N=399	Placebo N=401	Risk Difference (95% CI)	Relative Risk (95% CI)
Neonatal intensive care unit admission among liveborn infants[e]	29/389 (7%)	24/390 (6%)	1.3% (-2.2 to 4.8)	1.2 (0.7 to 2.0)
Neonatal supplemental oxygen requirement among liveborn infants[e]	13/389 (3%)	10/390 (3%)	0.8% (-1.6 to 3.2)	1.3 (0.6 to 2.9)
Neonatal assisted ventilation among liveborn infants[e]	1/389 (<1%)	1/390 (<1%)	-	-
Neonatal death[f]	14/388 (4%)	7/386 (2%)	1.8% (-0.5 to 4.1)	2.0 (0.8 to 4.9)

- Statistical tests were not done when there were fewer than five events in either group.
a Birthweight not recorded for 10 neonates and birthweight centile not able to be calculated for another 4 neonates
b Birthweight not recorded for 10 neonates
c 21 stillborn infants excluded and 1-minute Apgar not recorded for 105 liveborn infants
d 21 stillborn infants excluded and 5-minute Apgar not recorded for 498 liveborn infants
e 21 stillborn infants excluded
f 21 stillborn infants excluded and 5 infants with undocumented vital status at 28 days of life

Table 13
Adverse Events For 17-OHPC and Placebo[241]

Adverse Event	17-OHPC N=399	Placebo N=401	Risk Difference (95% CI)	Relative Risk (95% CI)
Headache	25 (6%)	17 (4%)	2.0% (-1.1 to 5.1)	1.5 (0.8 to 2.7)
Diarrhea	18 (5%)	10 (2%)	2.0% (-0.5 to 4.6)	1.8 (0.8 to 3.9)
Cough	14 (4%)	8 (2%)	1.5% (-0.8 to 3.8)	1.8 (0.7 to 4.1)
Itching	9 (2%)	12 (3%)	-0.7% (-3.0 to 1.5)	0.8 (0.3 to 1.8)
Rash	12 (3%)	5 (1%)	1.8% (-0.2 to 3.8)	2.4 (0.9 to 6.8)
Pain	5 (1%)	9 (2%)	-1.0% (-2.8 to 0.8)	0.6 (0.2 to 1.7)
Swelling	8 (2%)	6 (2%)	0.5% (-1.3 to 2.3)	1.3 (0.5 to 3.8)
Nausea or vomiting	4 (1%)	7 (2%)	-	-
Dizziness	0 (0%)	2 (<1%)	-	-
Urticaria	1 (<1%)	0 (0%)	-	-
Erythema	0 (0%)	1 (<1%)	-	-
Nodule	0 (0%)	0 (0%)	-	-
Anaphylaxis	0 (0%)	0 (0%)	-	-

- Statistical tests were not done when there were fewer than five events in either group.

CDER also relies on other observational studies such as Nelson et al., Hakim et al., Wang et al., and Massa et al. in support of its position.[242] These studies have significant limitations,

[241] *Id.* at e610.

[242] David B. Nelson, et al., *17-alpha hydroxyprogesterone caproate did not reduce the rate of recurrent preterm birth in a prospective cohort study*, 216. AM J OBSTET. GYNECOL. 600.e1–600.e9. (2017),

including the use of irregular claims data, inconsistent patient compliance in light of 17-OHPC's approved indication, and suspect inclusion criteria, rendering them of limited use in informing considerations of Makena's efficacy. For example, Nelson et al. conducted a prospective cohort study, comparing 430 women with a prior spontaneous preterm birth ≤35 weeks that were given compounded 17-OHPC at Parkland Hospital of the University of Texas Southwestern Medical Center, to the hospital's historical cohort. According to the authors, the comparison did not support the efficacy of 17-OHPC.

The comparison to the historical cohort is, however, inherently flawed because of significant uncertainties about the study population that undermine any conclusions that can be drawn. According to the authors, "the historical rate of recurrent birth ≤35 weeks" in the Parkland Hospital general obstetric population was 16.8% based on 5,787 women during 1988-2011, when 17-OHPC was allegedly not in use.[243] This explanation is ambiguous, and critically, it is inappropriate to compare a specific group such as those with a history of spontaneous preterm birth and known to be one of the highest risk groups, to a general population with a "recurrent preterm birth." In addition, there are questions as to whether 17-OHPC was truly not being used at Parkland prior to 2011, given that the hospital was one of the original 19 MFMU trial sites for the Meis trial[244] and that the use of 17-OHPC was recommended by ACOG as early as 2003 and further established as the standard of care in 2008. These ambiguities render Nelson et al. inconclusive, and the study should not be relied upon.

Massa et al. is similarly flawed. Therein, the authors published findings from a retrospective study that compared treatment with 17-OHPC to no treatment between January 1, 2008 and January 29, 2019. Makena was approved by FDA in February 2011; however, utilization of the FDA-approved formulation was slow to change from pharmacist-compounded formulations, particularly in Medicaid populations, which comprised more than 80% of the population the Massa authors studied. Unlike Makena, these compounded formulations were not made under cGMPs, and could have a sub-potent dose (among other potential problems, including sterility). Despite these significant distinctions, there is no identification by the Massa et al. authors of the proportion of women who received the compounded formulations versus 17-OHPC.

Further, documentation of the number of doses received was incomplete for more than half of the population.[245] The study therefore does not reflect the indicated use of 17-OHPC, which should be administered on weekly intervals, beginning between weeks 16 and 20/6 and continuing until Week 37, or delivery, whichever occurs first. Additionally, though the authors adjusted for certain risk factors associated with preterm birth, no adjustment was made for the

---

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5449222/; Joe B. Hakim, et al., *Effectiveness of 17-OHP for Prevention of Recurrent Preterm Birth: A Retrospective Cohort Study*, AM J PERINATOL, (2021), https://pubmed.ncbi.nlm.nih.gov/34972229/; Katherine Massa, et al. *Pregnancy duration with use of 17-a-ydroxyprogesterone caproate in a retrospective cohort at high risk of recurrent preterm birth.,* AM J OBSTET GYNECOL MFM. (2020); Xi Wang, et al., *Utilization, and Effectiveness of 17-Alpha Hydroxyprogesterone Caproate (17OHPC) in a Statewide Population-Based Cohort of Medicaid Enrollees*, AM J PERINATOL, (2021), https://pubmed.ncbi.nlm.nih.gov/34784617/.

[243] Nelson, et al., at 5 and 21.

[244] NDA 21-945, Clinical Review, at 22 (Feb. 3, 2011), https://www.fda.gov/media/80892/download.

[245] Massa et al, *supra* note 242, at 7.

number of full-term deliveries. A prior full-term delivery is well-recognized to decrease the risk of a future recurrent preterm delivery. The authors report that women who were not treated had significantly ($p<0.001$) fewer full-term deliveries than those treated with 17-OHPC: 35% of those not treated had full-term deliveries, as compared to 57.9% of those treated, and 16.9% of treated had ≥2 full-term deliveries, as compared to 35.0% of those not treated. Taken together, these undermine any conclusions that can be drawn from Massa et al.

Like Massa et al., the Wang et al. study fails to account for the use of compounded 17-OHPC and lack of compliance with 17-OHPC's indicated dosing regimen. Wang et al. examined Medicaid-covered 17-OHPC-eligible pregnancies from 2014-2016 in Pennsylvania, including those using compounded formulations, without considering potential impacts of such formulations. The authors acknowledge that compliance and the dosing regimen in study subjects may have differed significantly from the labeled indication.[246] Based on the label for Makena, a woman who has a full term pregnancy should have between 16-21 doses of 17-OHPC over the course of the pregnancy. However, 50.2% – just one-half the studied population – received 16 or more doses based upon pharmacy claims, despite the majority of those subjects (70%) having a full-term delivery. In addition, the authors note that 62.2% received 17-OHPC between 16-26 weeks, with 32% of the subjects started treatment at less than 16 weeks, timing that is inconsistent with the approved labeling, and an unknown number of women starting therapy after 20 weeks and 6 days, the latest that 17-OHPC is indicated to begin. Finally, there is ambiguity as to whether the weekly injections were completed, given the likelihood of at-home treatment.

Ultimately, the authors acknowledge that the study did not show that 17-OHPC was associated with a reduced risk of recurrent preterm birth. Notably, however, they attribute this lack of efficacy in part to the low utilization and adherence to the drug in the study population and state that the findings "suggest there may be barriers to implementing recommended preventive clinical therapies among pregnant persons enrolled in public insurance" and that further research is needed "with a lens toward remedying disparities in receipt of care."[247]

Finally, CDER incorrectly asserts that Hakim et al. supports its position with regards to Makena, but fails to mention the study's significant limitations as well as the fact that the study revealed a high-risk cohort that seemed to benefit from 17-OHPC. Specifically, the authors acknowledge that the highest risk group, i.e., those with a prior SPTB <28 weeks, showed a decrease in PTB with 17-OHPC treatment:

> Note that in the highest risk group (patients whose first preterm birth was at or before 28 weeks' gestation) the estimated RR was<1, but the small sample size (n=62) was inadequate to declare this effect statistically significant. This would be consistent with (but does not confirm) the hypothesis that the drug might be most effective for those at the highest risk, potentially demonstrating the utility of further focused studies on high-risk subgroups.[248]

[246] Wang et al., *supra* note 242, at 10.

[247] *Id.*

[248] Hakim et al., *supra* note 242.

Notably, the authors relied on a commercial database, and the resulting study population was extremely low risk, with only 18.1% of untreated women having a SPTB <37 weeks, a rate even lower than the PROLONG population. The authors assert that their study mirrored the same exclusion and inclusion criteria as Meis, but this appears incorrect as the study appears limited to the first and second recorded pregnancy, rather than properly selecting the last pregnancy and then looking back in time to see if the woman had a prior spontaneous PTB. An approach relying on the last pregnancy, rather than the first and second recorded pregnancy, would ensure that the full pregnancy history was included in the analysis, in the same manner as the Meis study. Ultimately, this lack of attention to methodology undermines the value of the study.

6. Real-World Makena Use For Over A Decade Supports A Favorable Safety Profile

Postmarket surveillance data further supports the safety profile of Makena. CDER has pointed to known potential risks of Makena described in its labeling, such as thromboembolic events, depression, allergic reactions, decreased glucose tolerance, fluid retention that may worsen maternal conditions such as pre-eclampsia, and injection site adverse reactions.[249] Table 14 below provides the number of reported adverse events for each of these known risks within the last decade.

Significantly, of the more than 350,000 women have been exposed to Makena as of August 31, 2022, only a tiny percentage of these women have reported experiencing labeled potential adverse events. For example, during the past decade of Makena use, only 37 of 356,327 patients (0.01%) have reported thromboembolic events (a labeled potential risk). Notably, the background rate of venous thromboembolism (VTE) among pregnant women is much higher: around 1.2 per 1000 deliveries (0.12%).[250]

Injection site reactions is the most frequently reported adverse event. This is consistent with Makena's labeling, which relies on the Meis trial data for characterization of these adverse reactions. According to the labeling, the most common adverse reactions reported from the Meis trial were injection site reactions.[251] It is worth noting that the labeling describes that the placebo group in the trial also experienced injection site reactions, including pain (32.7%), swelling (7.8%) and pruritus (3.3%).[252]

---

[249] *See, e.g.,* CDER's NOOH at 7, https://www.regulations.gov/document/FDA-2020-N-2029-0001.

[250] Bukhari et al., *supra*, note 15, at 8.

[251] Makena Prescribing Information at 1 and 7 (Feb. 2018), https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s012lbl.pdf.

[252] *Id*. at 7.

Table 14
Annual Estimated Patient Exposure and Certain Adverse Events

Year	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**Grand Total**
Estimated Patient Exposure	**7001**	**18636**	**27449**	**39776**	**54233**	**67522**	**56524**	**35382**	**24647**	**19669**	**5488**	**356327**
Thromboembolic events	1	1	2	2	2	5	2	5	6	8	2	36
Depression	1	5	3	17	30	31	46	33	27	16	14	223
Decreased glucose tolerance	3	2	3	9	46	63	43	30	40	19	10	268
Fluid retention with pre-eclampsia	1	5	7	10	29	49	52	42	52	32	16	295
Allergic reactions	18	19	45	62	114	115	187	133	114	112	39	958
Injection site reactions	112	177	399	710	1787	2407	6294	6285	3800	2698	1149	25818
Grand Total	136	209	459	810	2008	2670	6624	6528	4039	2885	1230	27603

In addition, among the more than 350,000 women treated, no new well-founded safety concerns, signals, or risks have been identified in more than 10 years of use.[253] The rates of any adverse events reported since Makena's approval in 2011 are consistent with risks identified in Makena's approved labeling.[254]

As noted above, CDER recently closed a NISS for Makena with respect to the risk of cancer in offspring of women who took hydroxyprogesterone caproate during pregnancy.[255] CDER has acknowledged that this NISS was based solely on Murphy.[256] As detailed below, CDER's own evaluation showed that Murphy provides no basis for a conclusion that Makena is associated with any risk of subsequent cancer in the children of treated women. In addition, FDA's own database, FDA Adverse Event Reporting System (FAERS), contains no reports of cancer or tumor from 17-OHPC.

7. CDER's Closure Of A NISS, Based On A Scientifically Flawed Article, Further Supports Makena's Positive Safety Profile

CDER previously suggested that it intends to rely on Murphy to question the safety profile of Makena. Any such reliance would be misplaced in light of the inherent limitations in that article, expert biostatistician declarations, physician guidance, and CDER's own analysis.

a. Murphy Is Neither Reliable Nor Relevant To This Proceeding

Murphy did not analyze data from use of Makena. Instead, Murphy describes a retrospective analysis of mother-child dyads receiving prenatal care in the Oakland, California area between 1959 and 1966 from the Kaiser Foundation Health Plan (Kaiser).[257] The investigators obtained medical records beginning six months before pregnancy through delivery to try to determine the number and timing of injections of Delalutin that the mothers received during pregnancy.[258] The investigators then attempted to identify through a California registry the incidence of any cancer in the offspring through 2019.[259] Based on a comparison of those datasets, the authors purported to find that *in utero* exposure to Delalutin was associated with "a higher risk of any cancer" and to claim "particularly striking associations with exposure in the first trimester and three or more injections."[260]

---

[253] There is a question in the comments to the docket regarding whether there could be an association between 17-OHPC and autism or attention-deficit/hyperactivity disorder (ADHD). *See* Simon Ball Emails re Makena (Aug. 12, 2022), https://www.regulations.gov/document/FDA-2020-N-2029-0223; Comment from Jill Escher (Sept. 13, 2022), https://www.regulations.gov/comment/FDA-2020-N-2029-0240. To our knowledge, neither FDA nor any published literature has suggested such a connection. Postmarket surveillance data show that among the more than 350,000 women exposed to Makena to date, there have been only five adverse events alleging ADHD and/or autism.

[254] Covis Pharma, *supra* note 170.

[255] Letter from Meredith Hillig, M.S., Office of Rare Diseases, Pediatrics, Urologic and Reproductive Medicine, Division of Urology, Obstetrics, and Gynecology, CDER to Covis Pharma GmbH (Jun. 8, 2022).

[256] Response from Meredith Hillig, M.S. to Lavonne M. Patton, Ph. D (Jun. 10, 2022).

[257] *See* Murphy et al., *supra* note 18, at e2.

[258] *See id.*

[259] *See id.*

[260] *Id.* at e7.

There are several methodologic shortcomings that affect the interpretation and undermine the reliability of the study's results. Notably, ACOG has recognized limitations in the study design as well as the peer review and publication process, and concluded that Murphy does not change the risk-benefit profile of Makena. Shortly after Murphy was published, ACOG issued the following announcement:

> **ACOG Guidance on 17-OHPC Remains Unchanged**
>
> ACOG is aware of the November 8 study published in the American Journal of Obstetrics and Gynecology linking the risk of cancer in offspring to people administered 17-hydroxyprogesterone caproate early during pregnancy. Due to limitations in the design, the study's findings are not conclusive and should not influence practice. ACOG guidance remains unchanged: 17-OHPC can be offered at 16 weeks of gestation or later to pregnant individuals with prior preterm birth in the context of a shared decision-making process incorporating the available evidence for progesterone supplementation, the values and preferences of the pregnant person, and the resources available. The peer review and publication process, additional information regarding the study population, exposures, methods and absolute risks observed in this study, as well as replication in other populations, will help to further understand any potential risks and balance them against the benefits. As with all guidance, ACOG will continue to monitor the literature and revise clinical recommendations as appropriate.[261]

Further, Murphy is not relevant to considerations of the safety and efficacy of Makena because it describes a retrospective analysis of Delalutin, not Makena. As the authors acknowledge, the historical use of Delalutin differs in "the timing, frequency, and pregnancy-related indications" as compared to modern clinical practice with Makena and its generic equivalents.[262] Differences in timing are crucial. Makena is not approved for administration prior to at least the 16th week of pregnancy—well into the second trimester. A significant number of women Murphy studied were likely exposed to Delalutin early in the first trimester:

> [T]he median gestational age at administration across all women receiving Delalutin was 10 weeks with a 25th percentile of 7 weeks. This means that in the first trimester group there is likely a large number of women who received Delalutin very early in their pregnancy, although these data are not presented in the manuscript.[263]

Indeed, the authors claim that the most "striking" results were seen in offspring who were exposed to Delalutin during the first trimester.[264] As discussed further below, there was

---

[261] ACOG, *supra* note 22.

[262] *See* Murphy et al., *supra* note 18, at e8.

[263] Declaration of Anita F. Das, Ph.D., (Jun. 8, 2022), https://www.regulations.gov/search?filter=FDA-2020-N-2029-0233, at ¶ 3 [hereinafter Das Declaration]; *accord* Declaration of Brent A. Blumenstein, Ph.D. (Jul. 22, 2022), at¶ 8 [hereinafter Blumenstein Declaration].

[264] Murphy et al., *supra* note 18, at e7.

not a statistically significant increased risk of any cancer when Delalutin was first administered in the second or third trimesters.[265]

Differences in indication also are critically important. Delalutin was administered to pregnant women for a wide variety of reasons.[266] Murphy states that 41% of women in the study were treated with Delalutin for a threatened miscarriage[267] (one of the Delalutin indications[268]). The authors do not specify, however, the indication for the remaining 59% of women treated. Given that the publication also reports that, "most (n=165 of 234, 70.5%) of the offspring were first exposed in the first trimester," one would expect that the majority of this patient population was not treated for recurrent preterm birth, as treatment for this indication is initiated no earlier than 16 weeks of gestation.[269] In contrast, Makena and its equivalents are intended for use only to reduce the risk of preterm birth and only in women (i) with a singleton pregnancy, (ii) who have a history of singleton spontaneous preterm birth, and (iii) who do not have multiple gestations or other risk factors for preterm birth.[270] Accordingly, the conditions of use of Delalutin in the 1960s bears no resemblance to modern clinical practice with respect to Makena.[271]

Moreover, the article is not sufficiently reliable for multiple reasons. As an initial matter, the primary outcome measure (all cancer risk) is based on an analysis of pregnant women exposed to Delalutin versus those not exposed.[272] The authors' trimester-delimited exposure analysis, however, suggests that any risk from Delalutin is limited to exposure during the first trimester of pregnancy.[273]

---

[265] *See* Das Declaration, *supra* note 263, at ¶ 3; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[266] *See* 36 Fed. Reg. 18115 (Sept. 9, 1971); 38 Fed. Reg. 27947, 27948 (Oct. 10, 1973); K.H. Meyerhoff et al., *The Use of 17-Alpha-Hydroxyprogesterone Caproate to Maintain Pregnancy*, 4 CURR THER RES CLIN EXP. 499-505 (1962), https://pubmed.ncbi.nlm.nih.gov/13935332/.

[267] Threatened miscarriage (also called threatened abortion) refers to vaginal bleeding before 20 weeks of pregnancy in the presence of an undilated cervix, suggesting a risk for pregnancy loss. *See* Robert E. Nesbitt, *The Outcome of Pregnancy Complicated By Threatened Abortion*, 2 CLIN. OBSTET. GYNECOL. 97-109 (Mar. 1959), https://pubmed.ncbi.nlm.nih.gov/13639320/; Michelle Mouri, et al., *Threatened Abortion*, In StatPearls (National Library of Medicine Bookshelf: NBK430747) (Sept. 9, 2021), https://www.ncbi.nlm.nih.gov/books/NBK430747/. *See also* Mount Sinai Health Library, *Miscarriage - threatened*, https://www.mountsinai.org/health-library/diseases-conditions/miscarriage-threatened.

[268] Delalutin (hydroxyprogesterone caproate) injection 125 mg/mL and 250 mg/mL was originally approved based on a finding of safety under NDAs 10-347 and 16-911. The approved indications reference the treatment of several gynecologic and obstetrical conditions, including the treatment of habitual and threatened miscarriage (also called habitual and threatened abortion). *See* NDA 21-945, Office Director Memo at 2 (Feb. 3, 2011), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000ODMemo.pdf.

[269] *See* Murphy et al., *supra* note 18, at e3.

[270] *See* Makena USPI §2.1 (Revised Feb. 2018), https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s013lbl.pdf.

[271] *See generally Humana of Aurora, Inc. v. Heckler*, 753 F.2d 1579, 1583 (10th Cir. 1985) ("When an agency [acts] on a study not designed for the purpose and which is limited and criticized by its authors on points essential to the use sought to be made of it, the administrative action is arbitrary and capricious").

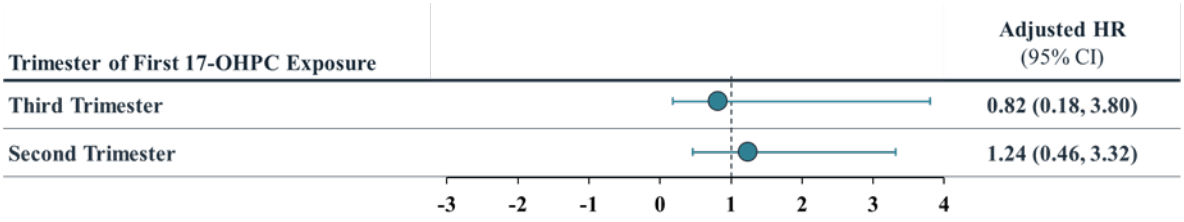
[272] We assume for purposes of this discussion, without conceding, that "all cancer risk" is a meaningful statistic.

[273] *See* Murphy et al., *supra* note 18, at 137.e7, tbl 3.

As shown in Figure 14 below and as the attached declaration in Appendix A from Anita Das, Ph.D.—a biostatistics expert with more than 25 years of experience in designing and conducting statistical analyses of clinical trial data and the lead biostatistician for both the Meis and PROLONG trials—explains, the hazard ratios and confidence intervals presented by the Murphy article authors make clear that 17-OHPC, when as administered as indicated on Makena's FDA-approved label, is not associated with the risks the authors claim to find:

> The study showed a significant increased risk of any cancer only when Delalutin was administered in the first trimester…. There was not a statistically significant increase[d] risk of any cancer when Delalutin was first administered in the second or third trimesters, with adjusted hazard ratios (aHR) of 1.24 (95% CI 0.46 – 3.32) and 0.82 (95% CI 0.18 – 3.80), respectively.[274]
>
> A hazard ratio (HR) provides a measure of the risk of an event (in this case, cancer) and is the ratio of patients experiencing the event over a period of time in the exposed group versus the unexposed group. An HR of one means that the two groups are experiencing an equal risk of the event (i.e., cancer). An HR of greater than one suggests that women in the exposed group had an increased chance of experiencing cancer as compared to women in the unexposed group. The 95% confidence interval (CI) provides a range of values that the HR is expected to fall between 95% of the time. If 1.0 is included in the 95% CI, it can be concluded that the hazard ratio is not statistically significant, and there is no difference in the risk of the event in the exposed and unexposed patients. Given that 1.0 is included in the 95% CIs…, this study provides no evidence that Delalutin when administered in the second or third trimester is associated with an increase in all cancer risk.[275]

Figure 14
95% CI and aHR for Trimester of First 17-OHPC Exposure



Further, the study suffers from a lack of sufficient events to support the authors' conclusions. This shortcoming seems particularly pronounced with respect to the authors' claim that "male offspring had an additional risk of cancer associated with exposure in late pregnancy."[276] That claim is based on a small number of person-years of follow up (1436.5 for

[274] Das Declaration, *supra* note 263, at¶ 3; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[275] Das Declaration, *supra* note 263, at¶ 4; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[276] *See* Murphy et al., *supra* note 18, at 132.e7.

late pregnancy compared to 376,390.5 for not exposed) and only four cancer cases.[277] This results in an imprecise estimate of the risk (aHR = 2.59) and a relatively wide 95% CI, ranging from an aHR of 1.07 (very close to 1.0) to 6.28.[278] Accordingly, the authors' conclusion is not supportable.

It is also questionable whether the authors utilized appropriate methods to minimize bias. Though the authors assert that they minimized "the possibility of bias [or systematic error] due to measurement error,"[279] it is not possible to verify this claim without access to the Child Health and Development Studies (CHDS) data used to identify mother-child dyads, the medical records used to identify the administration of Delalutin, or the California Cancer Registry (CCR) data used to identify cancers. Based on the available data, there is at least one identifiable problem with their data sourcing, which necessarily impacts the validity of their analysis. The authors purport to be evaluating lifetime cancer risk between 1959 and 2019, yet they rely exclusively on CCR data to identify cancers.[280] According to its website, the California Registry was not created until 1985 and did not start collecting data until 1988.[281] Thus, it is not clear whether *any* cancer that occurred prior to 1988 is reflected in the CCR data. Even after that date, reporting compliance was incomplete.[282] The information bias associated with no reporting (prior to 1988) and underreporting (after 1988) would be particularly problematic if it were differential (i.e., occurring to a different extent in the exposed and unexposed groups and could dramatically alter the results of this retrospective analysis.). Covis notes that CDER has attempted to rely on a poorly designed retrospective study to support a safety signal, yet CDER has been unwilling to discuss retrospective RWE-based study designs (e.g. a historical control study) Covis has proposed as part of an effort to mine and follow the data on Makena's clinical benefit.

The authors' approach to confounders also is insufficient. A confounder refers to a variable or factor that is not included in an analysis or study design, but which may have an effect on the risk of the disease.[283] The authors identify as "measured" confounders only basic

---

[277] Das Declaration, *supra* note 263, at¶ 7; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[278] *See* Das Declaration, *supra* note 263, at ¶ 7; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8; Murphy et al., *supra* note 18, at 132.e5.

[279] *See* Murphy et al., *supra* note 7 at e9.

[280] *See id*. at e2.

[281] *See* California Cancer Registry, *Policies and Procedures for Access to and Disclosure of Confidential Data from the California Cancer Registry,* at 6 (Jan. 2014), https://www.ccrcal.org/download/82/site-pdf-links/7430/ccrpoliciesprocedures_v05-2_3-19-2.pdf; American Cancer Society, *California Cancer Facts & Figures 2017*, at 41 (2017), https://www.ccrcal.org/download/68/special-reports-for-all-cancers/3334/california-cancer-facts-and-figures-2017.pdf.

[282] *See* Jennifer L. Malin et al., *Validity of Cancer Registry Data for Measuring the Quality of Breast Cancer Care* 94 J. NATL. CANCER INST. 835 (Jun. 5, 2002), https://pubmed.ncbi.nlm.nih.gov/12048271/; Robert Martinsen et al., *Collecting Comorbidities from Statewide Administrative Data*, Poster, North American Association of Central Cancer Registries (NAACCR) Annual Conference, UC Davis Institute for Population Health Improvement (2015), https://www.naaccr.org/wp-content/uploads/2016/11/P-13-3294.pdf; Audrey H. Choi et al., *Underreporting of Gastrointestinal Stromal Tumors: Is the True Incidence Being Captured*?, 19 J. GASTROINTEST. SURG. 1699 (May 22, 2015), https://pubmed.ncbi.nlm.nih.gov/26001370/.

[283] Das Declaration, *supra* note 263, at ¶ 5; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

demographic and birth characteristics for the offspring in the mother-child dyads.[284] They identify similarly few confounders for the mothers.[285] They, however, ignore numerous confounders, such as the mothers' histories of cancer, broader family histories of cancer, maternal use of tobacco or alcohol during pregnancy, offspring use of tobacco or alcohol, offspring environmental exposures, offspring BMI and other health factors, and offspring education and income, which are complex and overlapping risk factors that contribute to cancer causation that cannot be ignored based on a probabilistic analysis.[286] Indeed, the authors only controlled for two confounders – birth year and maternal body mass index. Their failure to properly control for confounders "could have produced a spurious association between Delalutin and cancer risk."[287]

Even as to measured confounders, the authors' analysis is deficient. Missingness (i.e., data missing from a population sample) was substantial for many measured confounding variables. High levels of missingness generally indicate the potential for bias and suggest that imputation—i.e., replacing missing data with a value that has been statistically determined—was inappropriate. In fact, the authors' own sensitivity analysis showed that imputation systematically reduced the aHR and the CIs (see Table 15 below comparing these numbers with and without imputation). The authors use multiple imputation to assess the impact of missing data and this analysis generally resulted in an aHR closer to 1 with a smaller lower bound of the 95% CI. For example, for second trimester exposure, the aHR was 1.24 (95% CI 0.46 – 3.32) and using multiple imputation was 1.03 (0.33 -3.20), a reduction of 0.21 in the aHR and 0.13 in the lower bound of the CI. This analysis shows that missing data has an important impact on the results.[288] Moreover, as explained by Dr. Das, though the authors fail to provide a multiple imputation analysis for analysis by sex, "[G]iven the trend presented… it can be postulated that the aHR and lower bound of the CI for males/late pregnancy would be reduced such that the lower bound of the CI could easily be below 1, indicating no increased cancer risk for male patients."[289]

[284] *See* Murphy et al., *supra* note 18, at 132.e3-132.e4.

[285] *See id.* at 132.e3, 132.e5.

[286] *See id.* at 132.e9 and Online Supplement. Moreover, the authors have not actually disclosed how they conducted such analysis. They did not, for example, disclose the "bias parameters" that they chose to evaluate or the modes nor was a study protocol in place.

[287] Das Declaration, *supra* note 263, at ¶ 5; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[288] Das Declaration, *supra* note 263, at ¶ 6; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[289] Das Declaration, *supra* note 263, at ¶ 7; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

Table 15
Comparison of Adjusted Hazard Ratios and Confidence Intervals With and Without Imputation

(as reported in Table 3 and Supplemental Table 2 in Murphy et al. Publication)

Variables	aHR (95% CI)	aHR (95% CI) using multiple imputation for missing values
In Utero Exposure to 17-OHPC		
Not Exposed	1.00	1.00
Any Exposure	1.99 (1.31-3.02)	1.85 (1.23-2.77)
Trimester of First 17-OHPC Exposure		
Not Exposed	1.00	1.00
First Trimester	2.57 (1.59-4.15)	2.28 (1.45-3.60)
Second Trimester	1.24 (0.46-3.32)	1.03 (0.33-3.20)
Third Trimester	0.82 (0.18-3.80)	0.82 (0.11-5.81)
Number of 17-OHPC Injections		
Not Exposed	1.00	1.00
1-2 Injections	1.80 (1.12-2.90)	1.68 (1.06-2.66)
≥ 3 injections	3.07 (1.34-7.05)	2.86 (1.25-6.56)

Murphy also suffers from inadequate statistical modeling, namely model misspecification, that fails to properly account for death as a competing event. It is well-understood that competing events can preclude the event under investigation (in this case cancer) from occurring.[290] In order to address a competing event, such as death, the cumulative incidence function for the event of interest must be calculated using appropriate survival analysis methods.[291] The absence of such methods can lead to biased estimates, for example, an overestimation of cumulative incidence.[292] The attached declaration in Appendix A from Brent Blumenstein, Ph.D.—a biostatistician and clinical trialist who has served on numerous FDA advisory committees—explains that the authors' approach to modeling leads to incorrect estimates that cannot support their conclusions:

> The main statistical model used to analyze the data was the proportional hazards regression (PHR) model (also known as Cox regression) with time to first cancer incidence as the event. Children without cancer were censored at time last known to be cancer free or death. The misspecification issue is that death is a competing event because death before cancer prevents the observation of cancer. Censoring

[290] *See e.g.*, Jaya M Satagopan, et al., *A Note on Competing Risks in Survival Data Analysis*, 91 BR. J. CANCER 1229 (Oct. 4, 2004), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2410013/; Bryan Lau, Stephen Cole, and Stephen J Gange, *Competing Risk Regression Models for Epidemiologic Data*, 170 AM. J. EPIDEMIOL. 244 (Jul. 15, 2009), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2732996/.

[291] *See id.*; Peter C. Austin, Douglas S. Lee, and Jason P. Fine, *Introduction to the Analysis of Survival Data in the Presence of Competing Risks*, 133 CIRCULATION 601 (Feb. 9, 2016), https://www.ahajournals.org/doi/10.1161/circulationaha.115.017719.

[292] *See* Noah A. Schuster et al*., Ignoring Competing Events in the Analysis of Survival Data May Lead to Biased Results: A Nonmathematical Illustration of Competing Risk Analysis*, 122 J. CLIN. EPIDEMIOL. 42 (Jun. 2020), https://www.jclinepi.com/article/S0895-4356(19)31061-3/fulltext.

> for time last known to be cancer free is characterized by the possibility of further observation, whereas death precludes the possibility of further observation.[293]
>
> Therefore, the validity of the conclusions presented in the article are based on incorrect estimates because an incorrect statistical model was used. Without having the data set the impact on conclusions of using a statistical analysis model treating death as a competing event rather than as a censored observation cannot be assessed.[294]

Valid scientific conclusions cannot be drawn from a study that contains such major limitations. As observed by Dr. Das, "the results are therefore difficult to interpret and cannot be considered to accurately present the association of 17-OHPC and cancer."[295]

In addition, to the extent new safety information about Makena were to come to light—and it has not—the Federal Food, Drug, and Cosmetics Act (FDCA) requires FDA in the first instance to follow the Risk Evaluation and Mitigation Strategies (REMS) provisions to assess and address any new safety signals when the Agency "becomes aware of new safety information and makes a determination that such a [REMS] is necessary to ensure that the benefits of the drug outweigh the risks of the drug."[296] FDA has not made such a determination, and has initiated no action to require further assessment. FDA has tools to address new safety signals but, to date, has not invoked them here.

b. CDER's Own Multidisciplinary Review Acknowledged Numerous Flaws In Murphy That Preclude It From Being Used To Draw Conclusions About Makena's Safety

Following a thorough investigation into its opened NISS, including epidemiological, pharmacological, and statistical evaluations, CDER concluded that "there are significant issues with attempting to apply the results of the Murphy study to the current regulatory and clinical environment" for Makena[297] CDER's own analysis of flaws in the Murphy study's design and analysis mirrors that of Covis' biostatistical experts, concluding that Murphy is a considerably flawed study that relied on a weak study design, unsupported underlying assumptions, and unsuitable statistical methods to find an association between 17-OHPC and cancer.

- **CDER's Epidemiology Review Undermines Murphy**

Several members of CDER's Division of Epidemiology II (DEPI II), Office of Surveillance and Epidemiology Review (OSE), Office of Pharmacovigilance and Epidemiology (OPE)—including an epidemiologist, a team leader, and the Division Director—evaluated

---

[293] Blumenstein Declaration, *supra* note 263, at ¶ 5 (Jul. 22, 2022).

[294] *Id*. at ¶ 7.

[295] Das Declaration, *supra* note 263, at ¶ 8; *accord* Blumenstein Declaration, *supra* note 263, at ¶ 8.

[296] 21 U.S.C. § 355–1(a)(2)(A) (2010).

[297] CDER, Division of Epidemiology II (DEPI II), Office of Surveillance and Epidemiology Review (OSE), Office of Pharmacovigilance and Epidemiology (OPE), Director's memo on the assessment of the publication: "In utero exposure to 17-a hydroxyprogesterone caproate and risk of cancer in offspring" by Murphy et al. NISS #1004783, at 3 (Ju1. 15, 2022) [hereinafter DEPII Division Director Memo].

Murphy and found the study design and authors' analysis to be lacking in the rigor necessary to draw valid conclusions about any link between 17-OHPC and cancer risk.

CDER's epidemiology team members acknowledged that the Murphy study suffered from numerous flaws including:

(1) the lack of a study protocol and a clear *a priori* hypothesis;

(2) the absence of a conceptual framework;

(3) a high likelihood of residual confounding;

(4) an inordinately high number of conducted statistical analyses along with insufficient event numbers to support the authors' conclusions;

(5) no clearly communicated information on sample size or power;

(6) failure to use a multivariate adjusted statistical model;

(7) inconsistency in the timing of 17-OHPC exposure between study subjects and Makena's FDA-approved indication;

(8) questionable generalizability; and

(9) unreliable effect estimates.

Accordingly, the CDER Team Leader concluded that the study's limitations "preclude this study from contributing definitively to this drug safety issue" and stating that Murphy "provides insufficient evidence to support regulatory action regarding a long-term cancer risk in offspring who were exposed in utero to 17-OHPC."[298]

This CDER team's findings corroborate and supplement the conclusions of Covis' expert witnesses. CDER's DEPI II Team Leader and Division Director both recognized Murphy's failure to use a pre-specified study protocol,[299] which is inconsistent with FDA best practices for a study of this type.[300] Murphy did not have an *a priori* hypothesis for its analysis, reducing the study's "ability to provide valuable insight into the association between in utero exposure to 17-OHPC and risk of cancer in the offspring."[301] Murphy was designed without a conceptual framework, which is particularly problematic in the context of carcinogenesis, a multistage process that is "one of the most complex processes in biology." As explained by CDER's reviewer, when designing a study, investigators must properly address assumptions and identify variables necessary to properly control for confounding prior to defining variables and choosing

[298] CDER, Division of Epidemiology II (DEPI II), Office of Surveillance and Epidemiology Review (OSE), Office of Pharmacovigilance and Epidemiology (OPE), Team Leader Review, Epidemiology: Review of published paper, at 4, 7, 8 (Jun. 22, 2022)

[299] *Id*. at 3; DEPI II Division Director Memo at 2.

[300] *See* DEPI II Division Director Memo at 2.

[301] DEPI II Team Leader Review at 4, 7.

statistical analysis methods. Murphy lacked clarity about treatment effect and causal pathways needed to ensure that study design and data analysis avoid "ambiguity about the treatment effect estimated . . . ."[302]

CDER's analysis also pointed to multiple sources of residual confounding. Murphy failed to address whether "women who took 17-OHPC during pregnancy may be more likely to be exposed to other risk factors for developing cancer in the offspring that are unmeasured or unadjusted in this study."[303] Additionally, the investigators failed to use propensity score matching to adjust for multiple variables and routinely adjusted for only one or two variables.[304] Further, as acknowledged by the Murphy authors "[t]he association between in utero exposure to 17-OHPC and cancer in offspring may be confounded by factors shared between mother and offspring, which were not measured in the CHDS."[305] Using the example of maternal smoking, CDER's reviewer explained how shared factors may lead to confounding. Maternal smoking is linked to preterm birth and it can also influence the likelihood of an offspring's uptake of smoking. Smoking in adulthood would directly affect cancer risk independent of any direct link between maternal smoking and offspring cancer risk. The Murphy investigators failed to use "a more advanced design and/or analytic methods" to properly account for such confounding.[306] Moreover, the authors did not collect data on cancer risk factors that may have occurred in adulthood and "if these cancer risk factors were more present among the exposed compared to the unexposed offspring, an increased association could possibly be observed."[307] Finally, the Murphy investigators' PBA analysis, which was intended to address confounding by unmeasured covariates, evaluated only unmeasured individual confounders and "residual confounding from multiple confounders could synergically affect the effect estimates more than individual unmeasured confounder[s] in either direction."[308]

CDER's review also concluded that given the small number of exposed cases and the large number of statistical analyses run, positive associations between 17-OHPC exposure and offspring cancer risk were likely to have been found based on chance alone. CDER's review described the increased risk of finding spurious associations when conducting analyses of multiple outcomes with few cases:

> Stratifying on outcomes with as few as one exposed case in the analyses, while bolstering study power by retaining all study controls, *is highly likely to identify at least a few spurious false positive associations*. In the context of so few exposed cases,

[302] *Id*. at 5.

[303] *Id*. at 6.

[304] *Id*.; *See also* CDER, Division of Epidemiology II (DEPI II), Office of Surveillance and Epidemiology Review (OSE), Office of Pharmacovigilance and Epidemiology (OPE), DEPI II Epidemiology: Literature Review, at 10-11 (Jul. 15, 2022) [hereinafter DEPI II Epidemiologist Review].

[305] Murphy et al., *supra* note 18 at 132e3.; *See also* DEPI II Team Leader Review at 7; DEPI II Epidemiologist Review at 11.

[306] DEPI II Team Leader Review, *supra* note 19, at 7.

[307] DEPI II Epidemiologist Review, *supra* note 304, at 11.

[308] *Id*.

misclassification [of] even a single in utero 17-OHPC exposure could lead to large differences in resultant hazard ratios.[309]

Additionally, Murphy's "conclusions regarding the risk of cancer in male offspring are unsupportable given the "increased variability of effect estimates in this subgroup due to the small event number."[310]

CDER identified additional flaws in Murphy:

- Murphy failed to provide any information on sample size or power in the published article.[311]

- Murphy's study investigators failed to use a fully multivariate adjusted statistical model, with many of the final adjusted hazard ratio results including just one or two covariates. CDER's reviewer explained that "[m]ethods for covariate inclusion, such as requiring the covariate to have a 10% change on the resultant hazards ratio, are likely inferior to fully multivariate adjusted models."[312]

- CDER's epidemiologist acknowledged the inconsistency between the timing of exposure to 17-OHPC among those studied by Murphy as compared with women treated with Makena based on the drug's FDA-approved indication and usage. CDER's reviewer noted that the Murphy study is limited in "its capability to directly examine the effects of the first exposure at gestational weeks 16-20 as Makena is currently administered in clinical practice, because most exposed during these weeks in the study by Murphy et al. were also exposed during the first trimester . . . ."[313]

CDER's reviewers also questioned the generalizability of the Murphy study's results. In addition to the inconsistency between timing of study population exposure and modern clinical practice,[314] the study's results may not be broadly generalizable given that all of the women studied were covered under a health maintenance organization (HMO), leaving women covered by a preferred provider organization (PPO) or uninsured women unrepresented.[315]

CDER pointed out that Murphy failed to address other issues that may have changed the effect estimates observed:

- Murphy's use of "year of last contact" to create a *date* of last contact may have affected effect estimates.[316]

---

[309] DEPI II Team Leader Review, *supra* note 19, at 6 (emphasis added).

[310] DEPI II Epidemiologist Review at 13.

[311] DEPI II Team Leader Review, *supra* note 19, at 3.

[312] *Id*. at 6.

[313] DEPI II Epidemiologist Review at 8, 12-13.

[314] DEPI II Division Director's Memo at 2.

[315] DEPI II Epidemiologist Review at 9.

[316] *Id*. at 12.

- It is unclear whether the number (%) of loss-to-follow up is comparable between the unexposed and exposed group given that the use of the California Cancer Registry leads to a loss of cancer status data for any individual who moved out of the state. This could lead to either an under- or over-estimate of the effect size.[317]

- **CDER's Pharmacology Review Undermines Murphy**

CDER's Pharmacology Review also found numerous flaws in Murphy. Valid conclusions about a safety risk cannot be drawn from Murphy in the absence of a plausible biological pathway through which 17-OHPC can give rise to cancer. CDER's pharmacological reviewer identified no clear mechanism for how cancer could occur based on maternal exposure to 17-OHPC. Indeed, CDER's evaluation referenced an in-depth literature review that found that "17-HPC was generally without adverse outcomes following prenatal exposure in both nonclinical and clinical studies."[318]

CDER's pharmacology reviewer explained that there are two pathways by which a molecule may cause cancer: direct DNA damage leading to mutations, or engagement with a normal physiological pathway leading to continued proliferation. Considering both pathways, CDER's reviewer concluded that "[i]n the case of 17-HPC the first route is less likely."[319] With respect to the second pathway, existing data does not clearly show that 17-OHPC binds with any receptor that would, in turn, lead to cancer-causing proliferation.[320] Ultimately, CDER's "integrated analysis suggests that it is unlikely that 17-HPC is increasing cancer risk via the cytoplasmic progesterone receptors" and establishes that "[t]here are insufficient data to evaluate the potential cancer risk for 17-HPC acting through any of the membrane progesterone receptors."[321]

With respect to risk differences between the first trimester and later trimesters, CDER's reviewer further concluded that "17-HPC acting at PR-A or PR-B appears to have insufficient biological activity to cause any super-physiological actions."[322] A similar finding was observed when data on dose-response activity was analyzed. CDER's reviewer explained that "there are insufficient published data to draw any reliable inferences about dose or exposure-response as they relate to adverse events in response to in utero exposure to 17-HPC."[323] Accordingly, the available data does not support a hypothesis linking 17-OHPC to cancer.

- **CDER's Statistical Review Undermines Murphy**

Echoing the analyses of CDER's epidemiology reviewers and Covis' biostatistical experts, CDER's statistical review identified multiple limitations in Murphy's statistical

---

[317] *Id*.

[318] CDER, Division of Applied Regulatory Science, Office of Clinical Pharmacology (DARS/OCP) Review of 17-α Hydroxyprogesterone Caproate and Trans-Generational Carcinogenesis (NISS 1004783), at 8 (Jul. 26, 2022).

[319] *Id.* at 9.

[320] *Id*.

[321] *Id*. at 1.

[322] *Id*. at 10.

[323] *Id*. at 10 (emphasis omitted).

methodology. Not only did the Murphy study investigators fail to pre-specify a protocol or statistical analysis plan, making it "difficult to interpret the statistical significance of the results,"[324] their main analysis utilized unclear methods for handling missing data and a PBA analysis that risked model misspecification. Moreover, according to CDER's reviewer, the authors' faulty analytical approach is subject to systematic error, or bias, resulting from a failure to leverage appropriate methods to achieve covariate balance, and to potential confounding due to inadequate variable adjustment. In addition, CDER's statistical review affirms that the Murphy study investigators' data interpretation did not properly take into account acknowledged differences in timing of 17-OHPC treatment between historical uses and current clinical practice, and relied on small event numbers in reaching unreliable conclusions.

CDER statistical reviewers noted numerous problems with how Murphy was conducted:

- DBVII's reviewer noted that "DBVII requested the study protocol from the corresponding author. However, the author stated that the study was conducted without a protocol and provided the user's manual for the Child Health and Development Studies (CDHS), which was the source of the data for the Murphy study. *Due to the absence of study protocol, it is unclear what analyses were planned, executed, or considered*."[325]

- CDER's reviewer noted that although Murphy used multiple imputations for missing values in their sensitivity analysis, "it is not clear how missing data were handled in the main analysis."[326]

- CDER's statistical reviewer noted that there was no pre-specified protocol or statistical analysis plan (SAP), and "[w]ithout prespecification, it is not possible to interpret the statistical significance of the findings."[327]

- CDER noted that the authors' claim that there was no difference in follow-up between these groups is questionable given that "covariate balance was not assessed using methods such as standardized mean difference." Further, no information was provided on the ability of the investigators to achieve covariate balance after adjustment and "[p]ropensity score methods, which allows for covariate balance assessment after adjusting, were not used." CDER's reviewer concluded that it is unclear whether the authors appropriately decreased residual bias.[328]

[324] CDER, Office of Translational Sciences (OTS), Office of Biostatistics (OB), Division of Biometrics VII (DBVII), Statistical Review of Murphy et al. (2022), "In utero exposure to 17α-hydroxyprogesterone caproate and risk of cancer in offspring" at 6 (Jun. 27, 2022).

[325] *Id*. at 1-2 (emphasis added).

[326] *Id*. at 3.

[327] *Id*. at 5.

[328] *Id*.

- Since year of birth and maternal BMI were the only variables included in the Murphy Article authors' Cox PH model, "[o]ther potential confounders that could have been clinically meaningful were not adjusted based on the variable selection criteria."[329]

- The authors' used PBA analysis, which "can be susceptible to design or model misspecification of the main analysis." Additionally, the Murphy investigators' analysis relied on unmeasured confounding rendering their analysis "not conclusive."[330]

- CDER's reviewer observed that the timing of 17-OHPC exposure (first trimester) and the condition treated (threatened abortion) for the women in the Murphy study is not comparable to clinical practice or the FDA-approved indication and usage of Makena.[331]

- Although the Murphy authors claim that the results for male offspring are based on data comparable to current clinical practice, this was based on a small number of cancer events (n=4), CDER's reviewer noted that "[w]omen exposed in the second trimester had no statistically significant difference compared to women with no exposure."[332]

Thus, as CDER's review itself concluded, "major limitations in the [Murphy] study design and analysis methods hinder the interpretability and validity of the study results. . . . From a statistical perspective, because of the major limitations of the study, the evidence of the reported increased cancer risk in 17-OHPC exposed offspring is inconclusive."[333]

In short, these statements from CDER's own NISS assessment reveal the extent to which the Center's own reviewers questioned whether Murphy can be relied on as evidence of a link between 17-OHPC treatment and cancer risk in offspring. CDER's reviews all lead to the same unmistakable conclusion: significant limitations in the design of Murphy preclude using the study to draw any valid inferences about the safety of Makena.

c. CDER's NISS Concluded That Murphy Raises No Identified Risk For Makena And Supplies No Support For Adverse Regulatory Action

As just detailed, CDER's multidisciplinary epidemiological, pharmacological, and statistical review of the Murphy article during the NISS illustrate the deficiencies in the Murphy Article. In light of these analyses, CDER classified the NISS as an "indeterminate risk." CDER's Manual of Policies & Procedures (MAPP), defines an "intermediate risk" as "[a]n untoward occurrence for which, following a comprehensive assessment, the findings are

[329] *Id*.

[330] *Id.*

[331] *Id.*

[332] *Id*.

[333] *Id*. at 6.

inconclusive with regard to the association with the medicinal product of interest."[334] According to CDER's normal approach, "CDER refers to indeterminate risks because in some instances, due to uncertainties about or inconsistences in the data available about a safety signal, it is not possible to ascertain whether there is adequate evidence of an association between a drug and an adverse event."[335]

Moreover, by classifying the NISS as "indeterminate," CDER has made clear that there is no "identified" risk, which is defined as "[a]n untoward occurrence for which there is adequate evidence of an association with the medicinal product of interest."[336] Therefore, by definition, there is no "adequate evidence of an association" between Delalutin (the drug studied in Murphy et al.) and all cancer risk. It follows *a fortiori* that there is no "adequate evidence of an association" between Makena (which was not studied) and all cancer risk.

In addition, it is unclear whether CDER initiated the NISS process for Delalutin, the actual drug at issue in Murphy. We are aware of at least two generic versions of Delalutin that have received FDA approval: Aspen Global Inc.'s ANDA 200271 (approved Aug. 24, 2015) and Eugia Pharma Specialties Ltd's ANDA 211142 (approved May 9, 2019). The Orange Book suggests that at least one of these (the first) continues to be available on the market. If CDER has not initiated a NISS process for Delalutin, it would raise questions about whether the NISS here was conducted under CDER's normal process in order to assess a reported safety signal for 17-OHPC, or whether the NISS was instituted to provide support for CDER's position in this hearing.

B. The Medical And Patient Community Continue To Support 17-OHPC As An Important Treatment Option

Medical and patient groups have expressed sustained support for maintaining access to Makena despite CDER's proposal to withdraw Makena from the market. On October 25, 2019, the same day the PROLONG trial results were published, ACOG issued a Practice Advisory maintaining its clinical recommendation to use progesterone supplementation in women with prior spontaneous preterm birth.[337] Following CDER's issuance of its NOOH, ACOG issued a statement in a similar vein, stating that its recommendations remained unchanged from the above mentioned Practice Advisory.[338] SMFM also released a statement following CDER's NOOH, reaffirming its support for the use of 17-OHPC and maintaining its previous recommendations as outlined in an October 2019 statement.[339]

---

[334] CDER, Manual of Policies and Procedures (MAPP) 4121.3, at 15 (Apr. 30, 2020).

[335] *Id.* at 15.

[336] *Id.* at 14.

[337] *See* ACOG, Practice Advisory*; supra* note 22.

[338] ACOG, *ACOG Statement on FDA Proposal to Withdraw 17p Hydroxyprogesterone Caproate* (Oct. 7, 2020), https://www.acog.org/news/news-releases/2020/10/acog-statement-on-fda-proposal-to-withdraw-17p-hydroxyprogesterone-caproate.

[339] SMFM, *SMFM Responds to the FDA's Proposal that Makena and Generic Equivalents be Withdrawn from the Market* (Oct. 5, 2020), https://s3.amazonaws.com/cdn.smfm.org/media/2543/Makena,_10.5.pdf.

Notably, both ACOG and SMFM have suggested that differences in rates of efficacy outcomes between the Meis trial and PROLONG were likely the result of significant variation in the composition of the trials' study populations. ACOG remarked that the patient populations in PROLONG had "divergent sociodemographic characteristics and a substantially lower preterm birth rate" compared to the populations in the Meis trial.[340] ACOG further noted that an "unintentional selection bias" may have occurred in the U.S. PROLONG sites, resulting in higher-risk women avoiding participating in the study out of fear that they would be randomized to placebo treatment.[341] SMFM similarly attributed the discordant results of the Meis and PROLONG studies, in part, to the "substantial differences" in the study populations and "significantly different" baseline rates of preterm birth.[342]

Physicians have also expressed concern about the negative implications that withdrawal of Makena would have on the population of women at greatest risk for preterm birth. A group of MFM physicians opposed to CDER's proposal to withdraw the approval of Makena explained that when a:

> [M]inority demographic group at greatest risk for a serious medical problem appears to obtain significant benefit [from Makena], any decision that will ultimately make it impossible to obtain the drug should be undertaken cautiously. This issue is particularly pressing when that minority group may be the least able to find and financially afford work-arounds to obtain the needed mediation in our complex medical system that has a history of failing to serve them well.[343]

These considerations have even greater force in the current landscape where recent changes to abortion law in the United States are likely to increase the need for appropriate reproductive care, especially for the most high-risk patients. Studies have shown unintended pregnancies to be associated with increased risk of preterm birth.[344] States with the strongest prohibitions on abortion also tend to have weaker maternal support and poorer maternal and child health outcomes, including low infant birth weight, infant mortality, and adverse childhood experiences.[345] For instance, according to the CDC, five of the bottom six states with the lowest birthweight (defined as babies born weighing less than 2,500 grams or 5 lbs 8 oz)—Mississippi, Alabama, Georgia, South Carolina and Louisiana—already have abortion bans in place, while the sixth state, Wyoming, has a near-total abortion ban that has recently been delayed by court order.[346] As a result, women living in States with restrictive abortion laws instituted as a result

---

[340] ACOG, *Practice Advisory*, *supra* note 22.

[341] *Id*.

[342] *See* SMFM Statement, *supra* note 22.

[343] Michael F. Greene et al., *Preterm Birth and 17OHP —Why the FDA Should Not Withdraw Approval*, 383 N. ENGL. J. MED. e130(1)-e130(3) (Nov. 3, 2020), https://pubmed.ncbi.nlm.nih.gov/33140924/.

[344] *See supra* note 28 (citing Amici Curiae Brief of 547 Deans, Chairs, Scholars And Public Health Professionals, The American Public Health Association, The Guttmacher Institute, and The Center for U.S. Policy, in Support of Respondents, *Thomas E. Dobbs, State Health Officer of the Mississippi Department of Health, et al. v. Jackson Women's Health Organization, et al.*, 945 F.3d 265 (No. 19-1392) (2022)).

[345] *See supra* note 28.

[346] *See id.* (citing CDC, Percentage of Babies Born Low Birthweight by State, https://www.cdc.gov/nchs/pressroom/sosmap/lbw_births/lbw.htm).

of recent developments are likely to have the highest need for therapeutics to reduce the risk of preterm birth and to be disproportionately impacted by the withdrawal of Makena's approval.

Organizations dedicated to improving the health of minority women, including the Black Women's Health Imperative, In Our Own Voice, National Minority Quality Forum, and the National Black Nurses Association, have been particularly active in expressing concerns about the impact that the withdrawal of Makena would have on Black women. These organizations note that Black women remain twice as likely to have preterm deliveries as other women and express a hope that Makena would "remain available to all eligible pregnant people, so that no person experiencing pre-term birth, is left without any access to a safe treatment option."[347] Given the disparity in incidence rates of preterm birth due to race, removing access to the only approved therapy for prevention of preterm birth could further deepen existing maternal and infant health inequities.[348]

In addition, a number of leading consumer organizations, including the National Consumers League (NCL) and HealthyWomen, have expressed support in continued use of Makena since the publication of the PROLONG trial results and the October 2019 Advisory Committee meeting. In its comment to the Center's NOOH, the NCL referred to its unsuccessful outreach efforts to urge CDER to maintain patient access to Makena and explore other methods of confirming Makena's clinical benefit.[349] The NCL outreach letter was signed by 15 leading women's health and consumer organizations and physicians, including:

- American Society for Reproductive Medicine
- Association of Women's Health
- Obstetric & Neonatal Nurses
- March of Dimes
- National Birth Equity Collaborative
- National Coalition for Infant Health
- National Medical Association

These organizations emphasize that CDER should consider alternative methods to define patient populations that most benefit from 17-OHPC rather than withdraw marketing approval and create a void in approved treatments for prevention of recurrent preterm birth.[350] High-risk pregnancy support organizations such as Sidelines and the Preterm Birth Prevention Alliance have also submitted comments, arguing that removal of 17-OHPC from the market "is not in the

---

[347] *See* Comment from Black Women's Health Imperative, Docket No. FDA-2020-N-2029-0006 (Oct. 19, 2020), https://www.regulations.gov/document?D=FDA-2020-N-2029-0006.

[348] *Id*.

[349] *See* Letter from American Society for Reproductive Medicine et al. to Patrizia Cavazzoni, Acting Director, CDER (Jun. 18, 2020), https://web.archive.org/web/20201130151719/https://d3n8a8pro7vhmx.cloudfront.net/ncl/pages/5160/Attachments/original/1593477202/Maternal_and_Infant_Health_Stakeholder_Letter_6-18-2020.pdf?1593477202.

[350] *Id*.; *see also* Comment from HealthyWomen, Docket No. FDA-2020-N-2029-0003 (posted on Oct. 14, 2020), https://www.regulations.gov/document?D=FDA-2020-N-2029-0003.

best interest of patients, their babies, or medical providers."[351] Sidelines, the PBPA, and others recognize that 17-OHPC has helped thousands of families and that withdrawing 17-OHPC without leaving behind a comparable alternative would be "ill-advised, irresponsible and unnecessary."[352] Notably, seven members of the October 2019 Advisory Committee, six of whom are obstetricians, voted in favor of keeping the product on the market while further studies are undertaken. Furthermore, 15 members of Congress responding to a full and fair consideration stakeholder letter recognized the importance of sensitivity to the disproportionate burden of preterm birth and other pregnancy-related complications borne by minority women.[353] They urged FDA to "work to ensure maternal health is prioritized…" and "consider stakeholder voices… as well as options for meeting the scientific evidence for efficacy."[354]

Patients who have relied on Makena to carry their pregnancies to term echo fears about the consequences of the removal of Makena from the market. Illustrative of numerous comments is the following patient experience:

> I lost one child to preterm birth and I believe the use of Makena directly contributed to the healthy birth of my living daughter. The thought of Makena or generic no longer being approved and available to women at high risk for preterm labor is shocking and terrifying, especially because I am currently pregnant and was fully expecting the Makena shot to be part of my medical care. Please do not remove this option for women in need.[355]

> After experiencing a spontaneous preterm delivery, you carry that fear throughout all future pregnancies, and it is beyond unnerving. Makena gives you hope and provides a sort of comfort throughout subsequent pregnancies. Knowing that there was a drug made specifically for my situation and a team of caring people behind that drug that would see me through my entire course of therapy was what gave me the confidence and ability to get through what I thought wasn't an option for me. It's scary to think that if it weren't for Makena I may not have been given the chance to be a mother. I truly feel that Makena was a major contributing factor to delivering my healthy full-term babies.[356]

These voices are critically important in making any decision that will affect access to healthcare. The 21st Century Cures Act directs FDA to consider "patient preferences with

---

[351] *See* Comment from the Preterm Birth Prevention Alliance, Docket No. FDA-2020-N-2029-0070 (posted on Jun. 21, 2021), https://www.regulations.gov/comment/FDA-2020-N-2029-0070.

[352] *See* Comment from Sidelines National Support Network, Docket No. FDA-2020-N-2029-0005 (Oct. 19, 2020), https://www.regulations.gov/document?D=FDA-2020-N-2029-0005.

[353] *See* Letter from Madeleine Dean, et al., Members of Congress to Stephen M. Hahn, M.D., Commissioner of FDA, Re: Full and Fair Consideration of 17P Stakeholder Letter, Docket No. FDA-2020-N-2029-0151 (Nov. 25, 2020), https://www.regulations.gov/document/FDA-2020-N-2029-0151.

[354] *See id.*

[355] *See* Comment from Nida Bajwa, Docket No. FDA-2020-N-2029-0064 (Mar. 8, 2021), https://www.regulations.gov/comment/FDA-2020-N-2029-0064.

[356] *See* Comment from Jamila Almonte, Docket No. FDA-2020-N-2029-0058 (Dec. 10, 2020), https://www.regulations.gov/comment/FDA-2020-N-2029-0058.

respect to treatment"[357] and emphasizes the importance of efforts to "take into account women and minorities and … focus[] on reducing health disparities,"[358] including by greater focus on "safe and effective therapies" for pregnant women.[359] Preterm birth is a very serious condition impacting pregnant women and neonates, and causes disproportionate harm to minority communities, thereby increasing health disparities.

It is clear from professional organizations, individual practicing physicians, and patients themselves that 17-OHPC remains an essential tool in treating preterm birth, and that the unavailability of this drug would leave this patient population at risk. OB/GYNs and other clinicians should be permitted to exercise their independent clinical judgment on the benefits of 17-OHPC for their patients. Removing the product from the market would prevent clinicians from prescribing what is, to date, the only safe, FDA-approved treatment indicated for women at high risk for recurrent preterm birth. The decision to use—or not use—17-OHPC should remain an individualized, informed choice based on sound, shared-decision making between health care providers and patients, as echoed by the leading medical societies ACOG and SMFM who oversee physicians that treat these high-risk patients.

C. Makena Should Be Kept On The Market To Avoid Negative Public Health Outcomes

At present, the U.S. has a significant incidence of preterm birth and has seen the rate generally rising in recent years up to 10.1 percent of all live births in 2020.[360] This represents approximately 400,000 preterm deliveries out of 4 million overall deliveries.[361] In addition, according to a 2012 World Health Organization (WHO) report, the U.S. lags significantly behind other industrialized, high-income nations with respect to preterm birth rate.[362] The U.S. average preterm birth rate in 2010 was 12.0%—markedly higher than both the average rate of high-income countries (9.3%) as well as the global average rate of 11.1%. In fact, of the 1.2 million preterm births estimated to occur in high-income regions, almost half of them (42% or 0.5 million) occurred in the U.S.; the U.S. also was included in the top ten countries with the highest number of preterm births, along with other nations including Brazil, India, and Nigeria.

The burden of preterm birth is not, however, shared equitably. Rates of preterm birth among Black and other minority women, as well as socioeconomically disadvantaged women, exceed those of other demographic groups.[363] Studies suggest that Black women have a two-

---

[357] 21st Century Cures Act, Pub. L. 114-255 (2016), § 3001.

[358] *Id*. § 2031.

[359] *Id*. § 2041.

[360] *See generally* March of Dimes, 2021 Report Card, *supra* note 31, at 11.

[361] *See* March of Dimes, 2020 Report Card; March of Dimes, Fact Sheet, The PREEMIE Reauthorization Act of 2018 (Jun. 2018), https://www.marchofdimes.org/materials/March-of-Dimes-PREEMIE-Reauthorization-Act-Fact-Sheet-June-2018.pdf.

[362] *See* WHO, Born Too Soon, The Global Action Report on Preterm Birth, at 26 (2012), https://apps.who.int/iris/rest/bitstreams/53412/retrieve; Hannah Blencowe et al., *National, Regional, and Worldwide Estimates of Preterm Birth Rates in the Year 2010 with Time Trends Since 1990 for Selected Countries: A Systematic Analysis and Implications*, 379 LANCET 2162-72 (Jun. 9, 2012), https://pubmed.ncbi.nlm.nih.gov/22682464/.

[363] March of Dimes, 2020 Report Card, *supra* note 361.

fold increased risk and women from the most disadvantaged neighborhoods have a 27% increased risk for preterm birth.[364] There remains an urgent need for Makena—the only FDA-approved therapeutic option for prevention of preterm birth—particularly in communities disproportionately affected by the preterm birth public health challenges.

Makena has given high-risk women and their healthcare providers hope of avoiding the adverse pregnancy complications associated with preterm birth. As explained by Dr. Washington Hill, a MFM specialist who has urged FDA to consider permitting additional research on the benefits of 17-OHPC rather than removing Makena from the market, if FDA withdraws Makena's approval, these patients would have no viable alternative for preventing preterm birth:

> Despite the high physical, emotional, and financial toll that preterm birth continues to take on our country – and disproportionately on women and families of color – not enough therapeutic tools currently exist to prevent it.[365]
>
> There are important points to consider before removing the only FDA-approved treatment option for use when counseling high-risk pregnant women with a history of spontaneous preterm birth. This includes the possibility of conducting additional research, particularly within high-risk populations, which could help the Agency make the most informed decision possible and align with its historic, overarching emphasis on advancing patient well-being and health equity.[366]

The absence of Makena from the market would lead to harmful public health consequences beyond the loss of the accepted standard of care, based on uncertainty created by CDER's hasty regulatory action, for a condition that is life-threatening for neonates and that has no other FDA-approved treatment options. Makena's favorable safety profile has been well established by multiple clinical trials, ongoing safety monitoring, and subsequent safety analyses. In contrast, potential treatment alternatives—compounded 17-OHPC, off-label use of generic Delalutin, and cervical cerclages—raise distinct safety concerns, including the adequacy of safety-related labeling, and could pose significant risks to patients—many of whom are Black and minority women.[367]

---

[364] *See* Collette N. Ncube et al., *Association of Neighborhood Context with Offspring Risk of Preterm Birth and Low Birthweight: A Systematic Review and Meta-Analysis of Population-Based Studies*, 153 SOC. SCI. MED. 156 (Mar. 2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7302006/; Tracy Manuck, *Racial and Ethnic Differences in Preterm Birth: A Complex, Multifactorial problem*, 41 SEMIN. PERINATOL. 511 (Dec. 2017), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6381592/.

[365] Letter from Washington Hill, M.D. to Patrizia Cavazzoni, M.D., Acting Director, CDER, Docket No. FDA-2020-N-2029-0060, at 1 (Dec. 8, 2020), https://www.regulations.gov/comment/FDA-2020-N-2029-0060.

[366] *Id*. at 2.

[367] Moreover, to the extent CDER believes that Murphy et al. provides a basis to withdraw Makena (and it does not), that same logic would counsel in favor of withdrawing Delalutin, the actual product that is the subject of the publication.

1. Compounding And Other Unproven Treatments Pose Additional Safety Risks

Purported treatment alternatives that might be more widely utilized if Makena were withdrawn from the market—including compounded 17-OHPC, off-label Delalutin, and cervical cerclages—would risk jeopardizing the health of pregnant women.

If FDA withdraws Makena from the market, the law requires that the Agency also add its active ingredient to the list of drugs that have been withdrawn or removed from the market for reasons of safety or effectiveness.[368] Once a drug is added to list of withdrawn or removed drugs, compounding of that drug is illegal.[369] FDA convenes and consults with the Pharmacy Compounding Advisory Committee (PCAC) prior to issuing regulations amending the list of withdrawn or removed drugs except in instances where protection of the public health demands more immediate action.[370] That process can leave a considerable delay between FDA withdrawal of a drug from the market and the addition of the drug to FDA's list of drugs that have been withdrawn or removed from the market for reasons of safety or effectiveness. For example, on July 2, 2014, FDA proposed to revise the list of withdrawn or removed drugs to add 25 new drug products.[371] FDA did not issue regulations updating the list of withdrawn or removed drugs to include these 25 drugs until October 7, 2016, a gap of over two years.[372] Notably, FDA last updated its list of withdrawn or removed drugs on December 11, 2018, almost four years ago.[373]

Thus, though withdrawal of Makena would prohibit compounding of 17-OHPC, it may be possible for compounding to continue for a significant period of time following withdrawal. Moreover, though the Agency's position is incorrect under the Agency's own regulations[374], FDA has stated that it has broad discretion over whether compounding would be prohibited for a withdrawn drug:

---

[368] *See* 21 C.F.R. § 216.24.

[369] 21 U.S.C. § 353a(b)(1)(C); § 353b(a)(4). FDA maintains a single list to implement both Section 503A(b)(1)(C) and Section 503B(a)(4). The two statutory provisions are distinct, however, and the latter does not include any additional procedural requirements that must be met before a drug that has been withdrawn by FDA for reasons of safety or effectiveness is added to the list of drugs that may not be manufactured by outsourcing facilities. Further, the provisions of Section 503A(c)(1) allow FDA to automatically update the list of drugs that may not be compounded in 503A pharmacies where doing so advances the public health.

[370] *See* Letter from Patrizia Cavazzoni, Director, CDER to Public Citizen's Health Research Group, 4, Docket No. FDA-2021-P-0378-0004 (May 6, 2022), https://www.regulations.gov/document/FDA-2021-P-0378-0004.

[371] *See* 79 Fed. Reg. 37,687 (Jul. 2, 2014).

[372] 81 Fed. Reg. 69,668 (Oct. 7, 2016).

[373] 83 Fed. Reg. 63,569 (Dec. 11, 2018). The PCAC convened on June 8, 2022, recommending that lorcaserin hydrochloride be added to the list of withdrawn or removed drugs. The Agency has not yet acted on the Committee's recommendation. *See* FDA, June 8, 2022: Meeting of the Pharmacy Compounding Advisory Committee Meeting Announcement (last updated Jun. 23, 2022), https://www.fda.gov/advisory-committees/advisory-committee-calendar/june-8-2022-meeting-pharmacy-compounding-advisory-committee-meeting-announcement-06082022; Alliance for Pharmacy Compounding, PCAC Recommends Glutathione for Bulks List (Jun. 10, 2022), https://a4pc.org/2022-06/pcac-recommends-glutathione-for-bulks-list/.

[374] 21 C.F.R. § 216.24.

> The inclusion of an entry on the Withdrawn or Removed List prevents all compounding of human drug products that fall within that entry. Such a bar on compounding human drug products might not be appropriate in every instance in which FDA determines that an FDA-approved drug product was withdrawn from sale for reasons of safety or effectiveness. Compounded drug products are not FDA approved, but they can serve an important role for patients whose medical needs cannot be met by an FDA-approved drug product. FDA's determination that an FDA-approved drug product was withdrawn from the market for reasons of safety or effectiveness does not consider compounded drug products, including those that contain components of the drug whose approval was withdrawn. Compounded drug products may raise different considerations that necessitate separate and thorough analysis through FDA's process for considering potential entries for the Withdrawn or Removed List.

As a consequence of either FDA's delayed action or an Agency decision to permit compounding of 17-OHPC, clinicians, who commonly prescribed compounded 17-OHPC before Makena was approved by FDA in 2011,[375] may resort to compounded 17-OHPC should they want to continue administering this treatment. Such compounded 17-OHPC would present risks not associated with FDA-approved Makena.

Compounded drugs are not subject to the rigorous safety and quality controls that accompany regulatory oversight of approved drugs.[376] Unlike the approved drug safety labeling for Makena, the labeling of compounded products remains virtually unregulated. Additionally, compounding pharmacies are exempt from the robust current good manufacturing practice (cGMP) standards that are applicable to approved drug products. The absence of such requirements presents great danger to patients, who may be placed at significant risk for serious injury and death as the result of poor drug quality and unsanitary conditions in compounding facilities.[377] Indeed, in the 2018 Cross Discipline Team Leader Review supporting approval of a preservative-free subcutaneous formulation of Makena, the reviewer acknowledged that compounded 17-OHPC lacks the quality associated with FDA-approved Makena, "Specifically,

---

[375] A national survey of board certified maternal-fetal medicine specialists conducted in 2005 showed that 67% of those physicians prescribed progesterone to help prevent preterm birth. Amen Ness et al., *Impact of the Recent Randomized Trials on the Use of Progesterone to Prevent Preterm Birth: A 2005 Follow Up Survey*, 195 AM. J. OBSTET. GYNECOL. 1174, (2006), https://www.ajog.org/article/S0002-9378(06)00758-7/fulltext. A comparable survey of obstetrician-gynecologists (OB/GYNs) administered two years later found that 74% of responding physicians offered or recommended progesterone for prevention of preterm birth. Zsakeba Henderson et al., *Attitudes and Practices Regarding Use of Progesterone to Prevent Preterm Births*, 26 AM. J. PERINATOL. 529 (Aug. 2009) https://pubmed.ncbi.nlm.nih.gov/19301227/. Data collected between 2004 and 2011 suggests that patients treated with progestogens prior to Makena's approval were likely treated with compounded 17-OHPC. *See* Baha Sibai et al., *Pregnancy Outcomes of Women Receiving Compounded 17 α-Hydroxyprogesterone Caproate for Prophylactic Prevention of Preterm Birth 2004 to 2011*, 29 AM. J. PERINATOL. 635 (Sept. 2012), https://pubmed.ncbi.nlm.nih.gov/22576126/; Advisory Committee Meeting Transcript, *supra* note 13, 276:09-277:02 (Dr. Sibai).

[376] *See* FDA, Compounding and the FDA: Questions and Answers (last updated Jun. 21, 2018), https://www.fda.gov/drugs/human-drug-compounding/compounding-and-fda-questions-and-answers.

[377] *See id*.

from a product quality standpoint, the purity and potency of compounded HPC products cannot always be assured."[378]

FDA safety data on compounded 17-OHPC further highlights the dangers associated with unreliable and unproven alternatives. According to Gandell et al., between 2012 and 2020, there were at least 26 safety recalls of compounded 17-OHPC, including several recalls for lack of sterility assurance, and recalls related to product contamination, adverse events associated with the presence of bacteria and fungi in product suspension fluid, and a fungal meningitis outbreak.[379] According to FDA, the 2012 meningitis outbreak at the NECC led to the death of 64 patients—with hundreds more gravely sickened.[380] As Gandell and colleagues note, such tragedies are particularly troubling given that FDA engages with only a small fraction of all existing compounding pharmacies.[381] Accordingly, the true public health toll from compounded 17-OHPC is unknown and could represent a catastrophic risk to high-risk patients already burdened by systemic inequities, including disparities in access to critical care. In comparison, as an FDA reviewer recently observed, "published literature using doses [of FDA-approved Makena] up to 3.6-fold above that used in the SC auto-injector have not identified any safety concerns."[382] These concerns further illustrate why FDA should allow Makena to remain on the market while additional study is being undertaken, given the time period it would likely take for FDA to remove any compounded 17-OHPC from the market and the heightened safety risks associated with compounded products which does not exist with Makena.

Nevertheless, FDA officials have suggested that off-label use of generic Delalutin may be substituted for treatment with Makena.[383] This suggestion is completely at odds with FDA's longstanding policy against the promotion of the off-label use of approved drugs.[384] As made clear in the drug's labeling, Delalutin, unlike Makena, is not indicated for preterm birth or even for use in pregnant women.[385] FDA encouraging provider use of generic Delalutin as a replacement for Makena would amount to the promotion of an off-label use—a position that cannot be reconciled with FDA's policy with respect to off-label promotion and acknowledgment of "examples of significant adverse consequences that have resulted from off-

---

[378] FDA, Makena (Hydroxyprogesterone caproate) Solution for Injection Cross-Discipline Team Leader Review (Feb. 14, 2018), https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945Orig1s012SumR.pdf.

[379] Gandell et al., *supra* note 24.

[380] *See* FDA, New England Compounding Center Pharmacist Sentenced for Role in Nationwide Fungal Meningitis Outbreak (Jan. 31, 2018), https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/press-releases/january-31-2018-new-england-compounding-center-pharmacist-sentenced-role-nationwide-fungal.

[381] *Id.*

[382] Cross-Discipline Team Leader Review, *supra* note 27.

[383] *See* 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) 250 mg/mL, Docket No. FDA-2020-N-2029-0051, at 16 and Section III.B (Dec. 4, 2020), https://www.regulations.gov/comment/FDA-2020-N-2029-0051.

[384] *See generally* Declaration of Rachel E. Sherman, M.D., Associate Director of Medical Policy, CDER, ¶ 17, *Par Pharm., Inc. v. United States*, No. 1:11-cv-1820 (D.D.C. Jan. 11, 2012).

[385] McGuff Pharmaceuticals, Inc., Hydroxyprogesterone Caproate Injection USP (Aug. 2015), https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/200271lbl.pdf.

label uses of approved drugs" even when such use was supported by the medical community.[386] The suggestion also fails to recognize that Delalutin is not an appropriate replacement for Makena.

Generic Delalutin is not equivalent to important versions of Makena, including a subcutaneous, intramuscular, preservative-free version approved in 2018.[387] In addition, since Delalutin is not approved for preterm birth, the Delalutin labeling lacks instructions for safe use for this therapeutic use and contains no warnings cautioning patients or providers about relevant risks.[388] In fact, generic Delalutin's labeling uses an outdated, less accessible format and does not include any of the detailed patient information that is available on Makena's approved labeling.[389] Additionally, since there is no currently marketed branded Delalutin, the drug's labeling will not be updated to reflect adverse event findings or any other emergent safety information. Furthermore, Delalutin has not been subject to study in pregnant patient populations in the same way that Makena has. For all of these reasons, generic Delalutin could not fill the void left by the loss of Makena, either in terms of safety or the adequacy of how safety information is conveyed to patients.

Physicians testifying during the October 2019 Bone Reproductive and Urologic Drugs Advisory Committee (BRUDAC) meeting held to discuss continued FDA approval of Makena explained the harm to patients that would be caused from compounding and other unsafe alternatives to treatment with Makena:

> [T]he clinical response out there in the field is going to be that our brethren will start prescribing other versions of progesterone, whether it's vaginal, or oral. or some other compounded injectable, and they may all at once; that that could happen or they could put in more cerclages that were unnecessary. So in that regard, I think we're also looking at other ethical implications here, where we're doing harm where we shouldn't be.[390]

> If there is not a 17P FDA-approved version available, many [physicians] will turn to a compounded 17P. Others will advise off-label, unproven medical therapies or choose a surgical option with cervical cerclage, which has not been proven to work and has a greater risk for patient harm.[391]

---

[386] *See* Sherman Declaration, *supra* note 384, at ¶ 17, *Par Pharm., Inc. v. United States.*

[387] *See* Letter from Hylton Joffe, M.D., M.M. Sc. to David Knauss, Sr. Manager, AMAG Pharmaceuticals, Inc., (Feb. 14, 2018), https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2018/021945Orig1s012ltr.pdf.

[388] *See id*.

[389] *See id*; AMAG Pharmaceuticals, Inc., MAKENA® (hydroxyprogesterone caproate injection) for intramuscular or subcutaneous use (February 2018), https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021945s012lbl.pdf; *see also* 71 Fed. Reg. 3921 (Jan. 24, 2006) (In 2006, FDA updated the format for prescription drug product labels to create labeling that is less technical, highlights the most important patient information, and is more accessible to the average reader.).

[390] Advisory Committee Meeting Transcript, *supra* note 13 at 254:22-255:08.

[391] *Id*. at 78:02-08.

Two advisory committee members, who voted to leave Makena on the market while further study was conducted, cited safety concerns related to compounding and cervical cerclages in reaching their decision:

> I think it will be hard to look at someone who had a preterm delivery that had a term delivery on Makena, and then tell her, but it doesn't work, because we can all agree, and we all have, that the data's conflicting, and we don't like things about each trial. But to just toss it out and say we're going to go back to ground zero and put people at risk from potential compounded 17P, I don't think is worth it.[392]

> If we look at what we have, this is the only pharmacotherapy we have for preterm birth that has been shown to work in some populations. The next thing, if we withdraw totally, people will be placed in cerclages, which studies have shown increases preterm birth in this population, and there are no other pharmacotherapies out there, so we'll see patients scrambling to get this. And I just worry about what that will be.[393]

The detrimental impact of the market withdrawal of Makena on health equity is, by itself, concerning, but equally unsettling is that it conflicts with FDA's commitment to diversity in clinical trial and stated health equity goals. FDA's mission is, above all, to "promote the public health by promptly and efficiently reviewing clinical research and taking appropriate action on the marketing of regulated products[.]"[394] FDA action is inappropriate when it undermines the public health.[395] Withdrawal of accelerated approval for Makena and its generic equivalents will have significant negative public health consequences for women, minorities, and socioeconomically disadvantaged populations.

2. Withdrawal Of Makena Would Represent A Departure From FDA's Commitment To Diversity In Clinical Trials And The Government's Health Equity Priorities

Clinical trial diversity and health equity considerations further counsel against withdrawal here. As numerous authors have recognized, pregnant women have traditionally been excluded from clinical research due to a perception that they belong to a vulnerable group, while various manifestations of systemic inequality have served to leave Black Americans and other racial and ethnic minorities out of medical research studies.[396] FDA has instituted policies

---

[392] *Id*. at 310:15-22 (Dr. Hickey).

[393] *Id.* at 311:10-19 (Dr. Eke).

[394] 21 U.S.C. § 393(b)(1).

[395] *See Beaty v. FDA*, 853 F. Supp. 2d 30, 42 (D.D.C. 2012) (FDA acted arbitrarily and capriciously where "[it] acted in a manner contrary to the public health"); *see also FEC v. Democratic Senatorial Campaign Comm*., 454 U.S. 27, 32 (1981) (courts must set aside Agency actions that "frustrate the policy that Congress sought to implement").

[396] *E.g.*, George Sheba et al., *A Systematic Review of Barriers and Facilitators to Minority Research Participation Among African Americans, Latinos, Asian Americans, and Pacific Islanders*, 104 AM. J. PUB. HEALTH e16 (Feb. 2014), https://pubmed.ncbi.nlm.nih.gov/24328648/.

to enhance equity in clinical trials.[397] In accordance with requirements of the Food and Drug Administration Safety and Innovation Act (FDASIA) of 2012, FDA directs study sponsors to "enroll participants who reflect the characteristics of clinically relevant populations with regard to age, sex, race, and ethnicity" and acknowledges that "[i]nadequate participation and/or data analyses from clinically relevant populations can lead to insufficient information pertaining to medical product safety and effectiveness for product labeling."[398]

FDA's withdrawal of Makena based on PROLONG—a failed study that was not fully inclusive of Black and other minority women[399]—would be contrary to the Agency's mandate to improve clinical trial diversity and the prioritization of ameliorating clinical trial inequity as a barrier to health care access.[400] Congress and FDA have recognized the underrepresentation of women and minority groups in clinical trials as a critical problem for the approval of safe and effective therapies. Specifically, the FDA Reauthorization Act of 2017 (FDARA) required that FDA, in coordination with other stakeholders, convene a public meeting to discuss topics pertaining to clinical trial inclusion and exclusion criteria, including "barriers to participation in clinical trials, including . . . regulatory, geographical, and socioeconomic barriers" and "clinical trial design and methods . . . that increase enrollment of more diverse patient populations," and make a report on the topics discussed at the meeting available.[401]

More recently, FDA released guidance aimed at providing sponsors with recommendations on the creation of a Race and Ethnicity Diversity Plan during the drug development process.[402] Race and Ethnicity Diversity Plans facilitate the enrollment of representative numbers of participants from historically underrepresented groups and, importantly, are intended to be submitted at the very beginning of the IND process.[403] FDA explains that these plans "should begin with an assessment of any data that may indicate the potential for a medical product to have differential safety or effectiveness associated with race or ethnicity," and that "[w]hen there are data that indicate that the medical product may perform differentially across the population based on factors associated with race or ethnicity, the Plan

---

[397] *See generally* FDA, Clinical Trial Diversity (last updated Dec. 9, 2021), https://www.fda.gov/consumers/minority-health-and-health-equity/clinical-trial-diversity.

[398] FDA, ENHANCING THE DIVERSITY OF CLINICAL TRIAL POPULATIONS — ELIGIBILITY CRITERIA, ENROLLMENT PRACTICES, AND TRIAL DESIGNSGUIDANCE FOR INDUSTRY, at 5 (Nov. 2020), https://www.fda.gov/media/127712/download.

[399] Blackwell et al. *supra* note 4, at 132-133.

[400] *See* White House, Fact Sheet: President Biden Reignites Cancer Moonshot to End Cancer as We Know It (Feb. 2, 2022), https://www.whitehouse.gov/briefing-room/statements-releases/2022/02/02/fact-sheet-president-biden-reignites-cancer-moonshot-to-end-cancer-as-we-know-it/; White House, The Biden-Harris Administration FY 2023 Budget Makes Historic Investments in Science and Technology (Apr. 5, 2022), https://www.whitehouse.gov/ostp/news-updates/2022/04/05/the-biden-harris-administration-fy-2023-budget-makes-historic-investments-in-science-and-technology/; FDA, FDA Takes Important Steps to Increase Racial and Ethnic Diversity in Clinical Trials (Apr. 13, 2022), https://www.fda.gov/news-events/press-announcements/fda-takes-important-steps-increase-racial-and-ethnic-diversity-clinical-trials.

[401] *See* Pub. L. No. 115-52 § 610(a), 131 Stat. 1051 (codified at 21 U.S.C. § 360bbb note).

[402] *See* FDA, DIVERSITY PLANS TO IMPROVE ENROLLMENT OF PARTICIPANTS FROM UNDERREPRESENTED RACIAL AND ETHNIC POPULATIONS IN CLINICAL TRIALS GUIDANCE FOR INDUSTRY, DRAFT GUIDANCE (Apr. 2022), https://www.fda.gov/media/157635/download.

[403] *See id*. at 2.

should specify the study design features that will support analyses that will inform the safety and effectiveness of the medical product in the relevant racial and ethnic populations."[404]

In addition, Health and Human Services (HHS) Secretary Xavier Becerra has committed to improving maternal health and reduce maternal and infant mortality and morbidity.[405] Regulatory actions that support maternal and infant health drug development encourage further research and development (R&D) efforts.[406] When R&D efforts are aimed at preterm birth prevention, they address an area identified as one of significant unmet need in women's health.[407] FDA's withdrawal of Makena would signal a lack of regulatory sensitivity to the real harms caused by disincentives to R&D. The potential long-term impact of inadequate maternal and fetal R&D would leave untold numbers of women and children to bear the burden of preterm birth risk.

Thus, Congress and FDA have acknowledged that minority subpopulations may respond differently to therapeutic products. A differential response to therapy is particularly important to explore where there are disparities in incidence and a glaring unmet need in drug development, as there are with preterm birth. Failure to consider such factors is not justified, particularly where the health and viability of the most vulnerable, being infants, are at stake.

D. Follow-Up Confirmatory Study Is Appropriate And Should Be Undertaken While Makena Remains Available To Patients

Consistent with FDA policy and practice, the Agency should maintain the approval of Makena while additional study is undertaken to further confirm the benefit of Makena. While this further study is undertaken, Covis is committed to working with CDER to narrow the indication for Makena to high-risk patients. As discussed in Section V above, FDA has stated explicitly that in the absence of a safety risk, it will explore all available regulatory options to keep a drug on the market where there is patient benefit, even though a confirmatory study has failed. This patient-centered approach is appropriate with respect to Makena, and Covis remains committed to working with CDER to find a data-driven approach to confirming the benefit of Makena.

---

[404] *See id*. at 6-7.

[405] *See* Transcript for Finance Committee Hearing for Xavier Becerra (Feb. 24, 2021) ("If confirmed, I will work to ensure that people have access to quality and affordable care, and to ensure that Medicaid makes progress on addressing the major health care challenges facing our country, including maternal health. . . . I will work tirelessly to reduce maternal and infant mortality and morbidity, using the expertise and resources across the many HHS agencies whose missions include ensuring child health.").

[406] *See* Congressional Budget Office, Research and Development in the Pharmaceutical Industry, (Apr. 2021), https://www.cbo.gov/publication/57126; Eric Budish et al., *Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials*, 105 AM. ECON. REV. 2044 (Jul. 2015), https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.20131176.

[407] Sandra S. Retzky and Timothy D. Baker, *Pharmaceutical Drug Development for Women's Health Care: Triumphs, Disappointments, and Market Needs*, 12 MD MED. 17 (2011), https://pubmed.ncbi.nlm.nih.gov/21657173/.

As explained in detail in the Appendix, Covis has spent months in collaboration with healthcare technology company, Dorsata[408] and a multidisciplinary scientific advisory panel, convened with the purpose of proposing frameworks for and evaluating the feasibility of future studies that could confirm Makena's clinical benefit in high-risk women. The scientific advisory panel included leaders in the fields of obstetrics, gynecology, biostatistics, epidemiology, clinical trials, oncology and drug development, who collectively have decades of experience at FDA and on FDA advisory committees.

With these outside experts, Covis has conducted extensive analysis of the available data to identify the patient population most at risk of a subsequent preterm birth. This analysis has included an evaluation of the PROLONG and Meis trial data to explore the extent to which the risk profile of the patients differed between the trials. As discussed in Section VII.A.3.b, PROLONG failed to enroll a high-risk population as reflected by the much lower number of events than planned by protocol. In addition to this analysis, Covis conducted a review of data from the Dorsata database to identify the appropriate high-risk population that could benefit from Makena, as well as the gestational endpoint at which neonatal morbidity and mortality shows the sharpest improvement.

Given these insights, Covis and its expert panel have developed proposals for feasible confirmatory RCTs that can be conducted completely or largely in the identified high-risk population in the U.S. Any sites outside the U.S. would be rigorously monitored to assure compliance with a protocol designed to minimize the impact of differences in standard of care and assure compliance with entry criteria. This data-driven approach will avoid the flaws in the conduct of PROLONG, which enrolled a much lower risk population, and is more likely to confirm the result of the Meis trial. Covis is also proposing that it could conduct observational studies, utilizing the Dorsata database, to generate data on efficacy within a shorter term timeframe. Critically, in developing these proposals, Covis has taken a number of steps, detailed below, to assess and ensure the feasibility of such a study in the U.S. including outreach to potential sites, consultation with contract research organizations, and surveys of relevant physician and patient populations to understand willingness to participate in a confirmatory study of an FDA approved drug.

Taken together, all of these steps make clear that there is ample opportunity to further explore the benefit of Makena in high-risk patients, while the product remains available on the market.

1. Prior Spontaneous Preterm Birth At <34 Weeks Gestational Age Is The Patient Population At Greatest Risk Of Subsequent Preterm Birth

Given the range seen for the placebo incidence rate of spontaneous preterm birth between Meis, PROLONG and other studies, Covis has conducted extensive analyses to identify a high-

[408] As explained in greater detail in Section VII.A.2.a, Dorsata is a healthcare technology company focused on improving women's health. Dorsata provides a maternity care management software platform that is used for decision support, documentation, obstetrical care plans, order entry, and clinical data reporting, among other things. Most relevantly, Dorsata's current database comprises of as many as over 210,000 pregnancies, enabling Covis to perform deep-dive analyses of preterm birth and insights to inform clinical trial development.

risk subpopulation for whom a stronger correlation exists between the inclusion criteria for the subpopulation and spontaneous preterm birth. As discussed in Section VII.A.2.a, above, Covis performed logistic regression analyses on the Meis, PROLONG and Dorsata datasets to examine the relationship among all available patient factors including gestational age of prior qualifying spontaneous preterm births, number of qualifying spontaneous preterm births, validated proxies for social and economic status, demographics and baseline information (e.g., race, age, body weight/BMI, certain comorbidities such as hypertension), smoking, and alcohol usage. This analysis showed that women who have had a prior spontaneous preterm birth <34 week are at highest risk for a subsequent preterm birth. This risk factor is therefore proposed as the inclusion criteria for a follow up RCT.

a. These Analyses Support Inclusion Criteria To Ensure Consistency Of Risk In A Further RCT

The previous analysis highlighted the importance of selecting patients with a prior SPTB <34 weeks. Covis further proposes to refine the inclusion criterion to ensure consistency of risk in the selected population. As discussed in Section VII.A.2.a, above, women who have had a prior spontaneous preterm birth <34 week are at highest risk for a subsequent preterm birth. This risk factor is therefore proposed as the inclusion criteria for a follow up RCT, with two potential additional modifications: (1) the previous singleton qualifying spontaneous preterm birth <34 weeks occurred within the last 5 years of randomization, and (2) documented medical history of first trimester ultrasonography to calculate the gestational age of the qualifying delivery, consistent with ACOG's recommendation that first trimester ultrasound measurement is accurate in establishing gestational age, while ultrasound in the second and third trimester are not accurate.[409] The first modification is important given the existence of many patient records in PROLONG where the qualifying pregnancies occurred early on in a patient's life and the same patient had several full-term births preceding the enrollment in the study. Such patients may not benefit from 17-OHPC treatment if they have had several successful term pregnancies in their recent history. The second criterion addresses the need for accurate determination of prior pregnancy history due to the reliance on this parameter as a surrogate of future PTB risk.

These additional criteria address many of the inadequacies observed in PROLONG, where much of the study population was at a low risk of subsequent preterm birth, and inconsistent and unreliable methods were used to verify the gestational age of the qualifying delivery.[410] In the U.S., obstetricians have generally used ultrasound to measure the CRL or femur length to determine gestational age in the last decade.[411] In contrast, these practices were not common in the PROLONG population ex-U.S. In particular, in Russia and Ukraine, where

---

[409] *See* ACOG Committee Opinion, No. 700, Methods for Estimating the Due Date (Replaces Committee Opinion Number 611, Oct. 2014. Reaffirmed 2022), https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2017/05/methods-for-estimating-the-due-date.

[410] As discussed in Section VII.A.3.a, the PROLONG population consisted primarily of women with lower risk factors for a subsequent preterm birth. This assessment has been further confirmed by Covis' analysis of the Meis, PROLONG, and Dorsata data, which demonstrate that the high-risk population consists of those women with a prior spontaneous preterm birth at <34 weeks gestational age. As discussed in Section VII.A.3.c, the assessment of gestational age for the ex-U.S. patients in PROLONG was based on inconsistent and unreliable methods, including second or third trimester ultrasound or last menstrual period.

[411] *See, e.g.*, ACOG Committee Opinion, No. 700, *supra* note 409.

the majority of the PROLONG study population was enrolled, the use of ultrasound during the first trimester of pregnancy to determine and document gestational age was less prevalent and/or not standardized when compared with ultrasound utilization in the United States.

Data suggest that the use of ultrasound during the first trimester of pregnancy to determine and document gestational age was less prevalent when compared with rates in the United States. By 2015, 90% of women in the United States received ultrasounds during the first trimester compared with just 77% of women in Russia.[412] Further, prior to 2009, the use of ultrasonography specifically with respect to gestational measurement was, at best, sporadic. The use of ultrasound along with LMP has been documented in Northwest Russia in the early 1990s, however, only a small portion of these births refer to ultrasound with LMP being the primary method by which gestational age determination was made.[413] Additionally, Covis conducted a comprehensive literature search and was unable to identify any literature pointing to ultrasonography as the standard method of gestational age determination in Russia. Conversely, several published studies on various aspects of maternal and pediatric health which cite LMP as the method for gestational age determination.[414]

In the case of Ukraine, it is even more clear that LMP was the prevailing method for first trimester gestational age determination. Though the vast majority of pregnant women in Ukraine routinely received ultrasounds as far back as 1995, the standard of care in Ukraine recommended second trimester ultrasounds, not first trimester ultrasounds.[415] A comprehensive study of the outcomes of 17,137 pregnancies published in 1999 cited the use of LMP to estimate gestational age.[416] More recently, authors of a 2019 study on maternal alcohol use among Ukrainian mothers noted difficulty in obtaining gestational age when assessing pre-term birth due to the lack of ultrasound screenings and poorly documented LMP records.[417] Similar to the Russian case, use of first trimester ultrasound in Ukraine has been poorly documented and LMP appears to be the primary method used to determine gestational age in early pregnancy.

2. Published Literature And Prior CDER Evaluations Support Clinical Endpoint of <35 Weeks Gestational Age As Highly Associated With Neonatal Outcomes

For the preterm birth endpoint, Covis proposes that the RCT would use an efficacy endpoint of gestational age of prolonging a pregnancy beyond 35 weeks. CDER and past

---

[412] *See* Shuvalova et al., *supra* note 33; O'Keefe, et al. *supra* note 33.

[413] *See* Postoev, et al., *Changes in detection of birth defects and perinatal mortality after introduction of prenatal ultrasound screening in the Kola Peninsula (North-West Russia): combination of two birth registries*, 15 BMC Pregnancy Childbirth 308 (Nov. 23, 2015), https://pubmed.ncbi.nlm.nih.gov/26596677/; Anna A. Usynina, *Risk factors for perinatal mortality in Murmansk County, Russia: a registry-based study,* 1 Glob Health Action (2017), https://pubmed.ncbi.nlm.nih.gov/28156197/.

[414] *See supra* note 197.

[415] *See* Arbuzova, *supra* note 33, at 184 (noting that "General ultrasound screening twice, at 16-18 and at 24-27 weeks of pregnancy, is recommended").

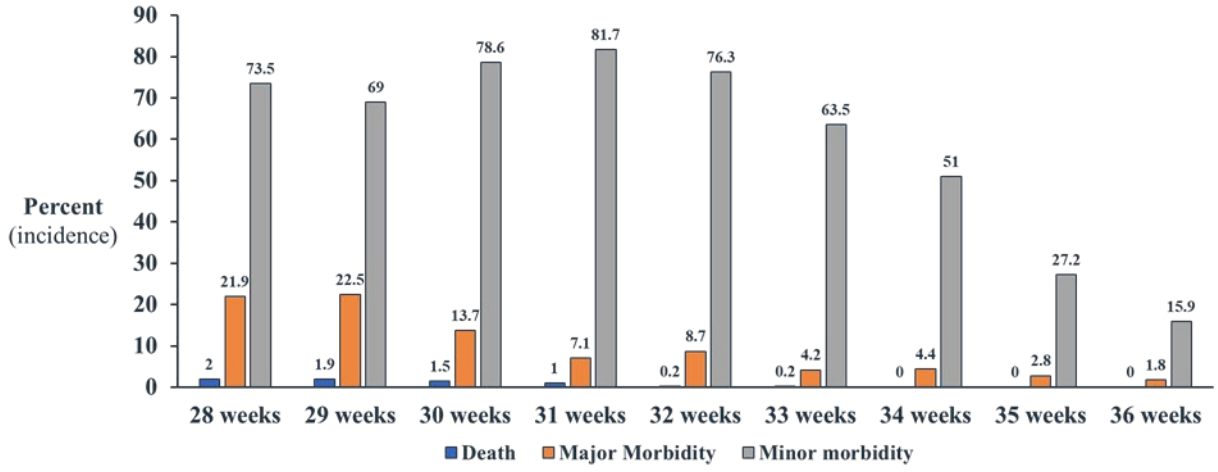
[416] Little et al., *supra* note 202.

[417] Claire D. Coles et al., *Gestational age and socioeconomic status as mediators for the impact of prenatal alcohol exposure on development at 6 months*, 12 BIRTH DEFECTS RES. 789-96 (Jul. 15, 2019), https://pubmed.ncbi.nlm.nih.gov/30378744/.

Advisory Committees (2006) have indicated in the past that they view such an endpoint as meaningful. For example, CDER has describes reduction in preterm birth <32 or <35 weeks as a "well-established surrogate" or "established surrogate".[418] CDER has also stated that these endpoints—in contrast to later preterm birth endpoints such as reduction in preterm birth <37 weeks—would not need confirmatory evidence of clinical benefit.[419] The choice of preterm birth <35 weeks was also selected as the primary endpoint for PROLONG in consultation with FDA.

This is based on the extensive literature supporting the association between increase in gestational age at birth and reduction in neonatal morbidity/mortality. For example, Manuck et al.'s analysis of an obstetric cohort of 115,502 women and their neonates publication in 2016 demonstrates that incidence rates of death, major neonatal morbidity, and minor neonatal morbidity decline significantly with each advancing week of gestation, as shown below.

Figure 15
***Incidence Rates by Gestational Age*[420]**



In particular, the relative reduction in the rates of the morbidities from week 32 to week 36 demonstrates the clinical benefit in extending gestational age of the neonate into week 35 or later. As will be seen below, Covis proposes designing the gestational age endpoint to focus on the benefit of extending gestation through this critical period for which the morbidity/mortality benefits of extended time in utero have been established. Further support for the association

[418] NDA 21945 Clinical Review at 15, https://www.fda.gov/media/80892/download; FDA Briefing Document at 20, https://www.fda.gov/media/132003/download ("FDA determined that further study was needed to provide confirmatory evidence of the drug's efficacy in terms of direct clinical benefit on neonatal outcomes or through an established surrogate such as the rate of preterm birth prior to 35 and 32 weeks gestation").

[419] *Id.* CDER has also described the delay in delivery observed in the Meis trial as an "intermediate clinical endpoint," indicating that the delay in delivery, in itself, provides a therapeutic effect. Expedited Programs Guidance at 19.

[420] Tracy A. Manuck et al., *Preterm Neonatal Morbidity and Mortality by Gestational Age: A Contemporary Cohort*, 215 A. J. OBSTET. GYNECOL. 103.e1–103.e14 (2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4921282/.

between gestational age and morbidity/mortality in the 32-35 week period can also be found in other recent analyses including a large meta-analysis by Teune et al.[421]

For these reasons, Covis proposes that a further RCT have either of two primary endpoints: (1) a measure of the improvement of gestational age in a time window that is strongly correlated with morbidity/mortality benefits for the neonate, with a preference for weeks gained from 17P treatment focused on <35 weeks or (2) a study primarily focusing on a neonatal morbidity/mortality endpoint.

In case the recommendation is to conduct an outcome study on neonatal morbidity and mortality, Covis proposes to revisit the previously utilized definition of the morbidity/mortality endpoint with the Agency. For example, the Manuck article describes both major and minor morbidities, which are all associated with significant healthcare burden to the patient as well as the system. Covis proposes to have an in-depth discussion with the Agency on the selection of an appropriate subset of clinically relevant morbidities to include that allow for a clinical study that may be enrolled and completed within 3-6 years (based on incidence of events, an assumed effect size, and the implied sample size).

Table 16
Classification of Morbidities

Meis/PROLONG	Manuck et al. (2016)	
	Major Morbidities	**Minor Morbidities**
Neonatal death	Persistent pulmonary hypertension	Grade 1 or 2 intraventricular hemorrhage
Grade 3 or 4 intraventricular hemorrhage	Grade 3 or 4 intraventricular hemorrhage	Necrotizing enterocolitis stage I
Respiratory distress syndrome	Seizures	Respiratory distress syndrome
Bronchopulmonary dysplasia	Hypoxic-ischemic encephalopathy	Hyperbilirubinemia requiring treatment
Necrotizing enterocolitis	Necrotizing enterocolitis stage II/III	Hypotension requiring treatment
Sepsis	Bronchopulmonary dysplasia	

Finally, in the case of the gestational age-based primary endpoints we would be looking for a trend that supports improvement in mortality and morbidity outcomes.

3. Based On Extensive Feasibility Analyses, Covis Proposes To Undertake An Additional RCT To Confirm The Benefit Of Makena In The High-Risk Population

Covis is confident that it is feasible to enroll patients at high risk of preterm birth, as defined in terms of the new inclusion criteria (i.e., <34 weeks prior SPTB with a validated recent prior birth gestational ages), particularly if Covis' preferred endpoint, increase in delivery time from randomization for 17-OHPC vs. placebo focused on <35 weeks, is utilized. Given the previously described differences in preterm birth by race, Covis believes it will be important to recruit and enrich for high-risk Black women in this study. Our preliminary assumption is that, depending on the chosen endpoint, we may seek to enroll about a third of the total subjects from the Black population in order to allow for a reasonable sample size and with stratification

[421] Margreet Teune et al., *A Systematic Review of Severe Morbidity in Infants Born Late Preterm*, 205 AM. J. OBSTET. GYNECOL. 374.e1 (2011), https://pubmed.ncbi.nlm.nih.gov/21864824/.

assuring balance for unconfounded exploration of the effects in Black subjects, The proposed study designs are based on the totality of understanding the association of neonatal morbidity/mortality by gestational age and would provide confirmation of Makena's clinical benefit in women at high risk for a future preterm birth. Enrollment will turn on the number of sites and enrollment rate at each individual site, which is the subject of ongoing feasibility studies. A sample of the results from the survey are described below. This enrollment estimate is based on several independent workstreams:

1) A formal RCT feasibility assessment that has identified 19 potential non-academic sites in the US and an (at this point) additional 10 sites in European countries (e.g., France, Germany, Italy, Spain, U.K.) where a similar high-risk subpopulation may be available to recruit. The ex-U.S. feasibility assessment is still ongoing and may identify additional sites in the near term. The estimate is that these sites may be able to contribute 0.5-1 patient with prior SPTB <34 weeks per month per site, therefore accounting for approximately 100 patients per year, assuming some reduction in the number of actual sites that participate after they review the full protocol.

2) A survey conducted within the Dorsata practice network indicating a willingness to participate in a placebo-controlled RCT with 17-OHPC. In the consenting practices, there are about 1200 patients per year who have pregnancies and who have had a prior spontaneous preterm birth <34 weeks. We may estimate between 5-15% of these patients may consent to enroll and screen successfully into the study, suggesting the addition of another 60-180 patients per year from this network

3) In recent outreach to prospective investigators at the 12 academic MFMU network sites that participated in Meis, prospective investigators expressed support for site participation in a new RCT of 17-OHPC. Of 12 investigators, 11 indicated interest in participating in a new trial, and one suggested that a new trial is not warranted because Meis affirmatively settled the question of 17-OHPC's efficacy.[422] This survey also provided some additional insights on the views of these prospective investigators on key features of the protocol. In particular, they overwhelmingly felt that any protocol today would need to include some aspect of rescue for patients presenting with shortening cervix post-randomization, e.g., the option to perform a cerclage on eligible patients. When queried about their views on the endpoints, half the respondents indicated interest in the type of "weeks gained" analysis that we have presented in this supplement. The details of the survey are in Attachment D.

4) In addition to the above, Covis conducted a systematic survey in 400 providers and a similar number of pregnant patients with prior spontaneous preterm birth

[422] Blackwell et al. *supra* note 4.

history to understand attitudes for enrolling in a placebo-controlled RCT. Covis subsequently conducted a follow-up survey including providers from the initial survey, which garnered 150 new respondents. These surveys confirm the willingness of physicians to contribute patients to the RCT and for patients to consent to such a study at significant rates. Importantly we also probed the recruiting effect having an approved product in the RCT has relative to an unapproved product. The willingness of both patients and providers to participate in a study with an approved product was higher than for an unapproved product. Importantly, providers willingness to participate in a study dramatically reduces when a product's approval is withdrawn (which situation we will be in if Makena's Accelerated Approval were withdrawn).

5) The sample size is discussed below, in Attachment B to the Appendix.

Details of the proposed RCTs and feasibility assessments for each are included in the Appendix, and in Table 17:

Table 17
Proposed Randomized Controlled Trials

Potential study	Rationale	Estimated sample size
Impact of treatment with Makena vs. placebo on time from randomization to delivery (weeks)	Established marked neonatal benefit on measures of morbidity/mortality with increase in weeks of gestation	▪ 1:1 randomization – 382 subjects ▪ 2:1 randomization – 429 subjects
Impact of treatment with Makena vs. placebo in achieving full term birth (comparison of percentage of patients achieving 37 weeks)	Moving a patient from a baseline of <34 weeks to full term has clear benefit to the neonate based on the reduction in morbidity event rates	▪ 2:1 randomization – 1152 subjects ▪ 3:1 randomization – 1364 subjects
Impact of treatment with Makena vs. placebo on preterm birth rate < 35 weeks (comparison of percentage of patients)	Established marked increase in morbidity/mortality in infants born prior to week 35	▪ 2:1 randomization – 1458 subjects ▪ 3:1 randomization – 1728 subjects
Impact of treatment with Makena vs. placebo on preterm birth rate < 35 weeks (measured as time-to-event censored at 35 weeks)	Established marked increase in morbidity/mortality in infants born prior to week 35	▪ 2:1 randomization – 1860 subjects ▪ 3:1 randomization – 2280 subjects
Impact of treatment with Makena vs. placebo on preterm birth rate <35 weeks (measured as weeks of gestational age less than 35 weeks)[423]	Established marked increase in morbidity/mortality in infants born prior to week 35 Method weighs delay of very early preterm birth more than mid/late preterm birth	▪ 2:1 randomization – 1413 subjects ▪ 3:1 randomization – 1660 subjects
Impact of treatment with Makena vs. placebo on a composite neonatal morbidity/mortality index	The definition of composite neonatal index subject to discussion with CDER, including potential expansion to include "minor" morbidities	Based on 15%-20% incidence rate of neonatal events ▪ 2:1 randomization – 2676 subjects ▪ 3:1 randomization – 3172 subjects

The first several studies all have identical inclusion criterion, and all feature a gestational age/Preterm Birth endpoint. The sole difference between the options—which leads to the difference in sample size estimates—is the analysis methodology. These, in concept range from the idea of taking patients with prior early SPTB history and assessing the weeks of gestational age gained <35 weeks, to gestational age of prolonging a pregnancy beyond 35 weeks, the difference between the arms of getting patients to full term, the rate of events <35 weeks, and a time-to-event-based methodology. Based on the analyses described in Section VII.A.3.e, Covis proposes conducting a new study that has a single primary endpoint of increase in delivery time from randomization for 17-OHPC vs. placebo focused on <35 weeks. This endpoint has several features that make it the preferred endpoint:

- It is more sensitive than a categorical endpoint such as PTB rate <35, allowing for the ability to test the hypothesis with a smaller sample size—this is critical given the historical challenges with recruiting subjects into a placebo-controlled study for this indication.

- It provides a clearer understanding of where the benefit to the neonate is accruing, i.e., the endpoint measures the difference in gestational age up to <35 weeks.

- The capping mechanism to 35 weeks gestation ensures that any increase in GA is associated with clinical benefit to the neonate.

This option is designed such that the analysis methodology does not give either arm any credit for births after week 35, counting them all the same. Therefore, any differences in the treatment arms would require an improvement of preterm birth outcomes at gestational ages prior to week 35. The numerical difference between arms can then be interpreted directly in terms of the likely benefit in morbidity/mortality of the neonate. Covis welcomes an opportunity to discuss these analysis approaches with the Agency as they present potentially distinct insights on the therapeutic benefit of the intervention.

If, alternatively, there is a desire to obtain data supporting/refuting the benefit of the product in a shorter timeframe (e.g., within 24 months) Covis stands ready to augment the above RCTs with observational studies. As described further in the Appendix, such observational studies could include either a retrospective or prospective cohort study stratified by history of SPTB within the past 5 years prior to the pregnancy being studied, and study the effects of Makena versus no treatment in the prior SPTB <34 week subgroup and separately in the prior SPTB >34 week subgroup. This observational study approach may offer additional insight on the relationship between risk level of the patient and their likelihood to see benefit with Makena. In conducting these studies, Covis may be able to leverage the Dorsata clinical network given the high-quality data that is available from their record system.

Covis may also explore similar observational designs in case there is a need to augment the currently available understanding of the association of preterm birth gestational age and morbidities/mortalities in the infant either in general or to probe certain questions that may not be adequately addressed in the previously conducted studies on neonatal outcomes by gestational age.

4. Withdrawal Of Makena Is Not Justified By Concerns Over Clinical Trial Recruitment While Makena Remains An Approved Drug

Concerns about clinical trial enrollment in a patient population that is underserved and potentially reluctant to participate in an RCT or about clinical equipoise are not sufficient to justify withdrawal. Despite efforts to diversify clinical trial participation in light of the historical exclusion of pregnant women and minorities from such trials, recruitment of both groups remains a challenge.[424] Pregnant women often are uncomfortable with clinical research or influenced by family members to decline participation.[425] Moreover, the long history of exploitation of Black Americans by medical researchers has created a distrust of clinical research within the Black community.[426] Makena's removal from the market would not eliminate these barriers to clinical

[424] Paula M. Frew et al., *Recruitment and Retention of Pregnant Women Into Clinical Research Trials: An Overview of Challenges, Facilitators, and Best Practices*, 59 CLIN. INFECT. DIS. S400 (Dec. 15, 2014), https://pubmed.ncbi.nlm.nih.gov/25425718/; Richard D. Branson et al., *African Americans' Participation in Clinical Research: Importance, Barriers, and Solutions*, 193 AM. J. SURG. 32 (2007), https://pubmed.ncbi.nlm.nih.gov/17188084/.

[425] Frew et al.

[426] *See* Branson et al., at 35.

trial participation. In fact, if FDA were to withdraw approval for Makena, barriers to participation in an RCT would likely be exacerbated.

At the time that PROLONG was recruiting, Makena not only had become the standard of care, but it also was the first drug ever approved for the prevention of preterm birth. The approval of Makena, supported by the strong results produced in the Meis trial, provided a clear barrier to trial participation given physician and patient demand for treatment with Makena, and the reluctance to risk randomization to placebo in an RCT. In the years since, the PROLONG trial's results (though based on a study with flaws in how it was conducted) and FDA's initiation of this action to withdraw approval have raised questions about the drug's efficacy which unequivocally re-establish equipoise, despite Makena's wide use, to support conducting a new study. Put another way, in light of these questions regarding the scope Makena's efficacy, further study is warranted, particularly in a high-risk population. Simultaneously, the lack of any valid safety signals eliminates concerns that leaving Makena on the market places patients at risk, particularly when clinicians would be forced to turn to unproven and riskier treatments such as compounded drugs or surgical procedures as treatment for their anxious patients. Thus, Makena should remain available to patients who need it while additional study is undertaken.

Finally, there is a significant risk that, in addition to making Makena unavailable for U.S. patients who are not in a position to enroll in a clinical trial due to geography and other factors, withdrawal of approval could make a placebo-controlled trial in high-risk patients even more difficult to enroll. Under human subjects protection regulations, investigators must provide a prospective participant any "information that a reasonable person would want to have in order to make an informed decision about whether to participate,"[427] which would necessarily include the FDA approval status of a trial drug. Knowledge of FDA's withdrawal of approval for a therapy would compound these background challenges and considerations and result in further reluctance to participate in a clinical trial.

* * * * *

For more than a decade, Makena has been the virtual standard of care and the only FDA-approved therapy for indicated patients suffering from the risks of preterm labor—a burden that falls hard on U.S. Black women and other minorities. There is no sound public health reason to deprive physicians and their patients from having access to this important therapy while additional confirmatory study is undertaken. Accordingly, Covis is committed to working with the Agency to consider appropriate options for further study and for narrowing the labeling to high-risk women, rather than depriving women of this needed therapy.

VIII. CONCLUSION

For all of these reasons, Makena should be kept on the market while additional study is undertaken.

[427] 45 C.F.R. § 46.116(a)(4).

Respectfully submitted,

Francesco Tallarico
General Counsel &
Head of Government Affairs & Policy

Aziza Johnson
Executive Vice President
Regulatory & Scientific Affairs

A. Raghav Chari, Ph.D.
Chief Innovation Officer

COVIS PHARMA GMBH

Rebecca K. Wood
Sean C. Griffin
Emily Marden
Kelly Cho
Shamelle Richards
SIDLEY AUSTIN LLP
1501 K Street, NW
Washington, DC 20005

Counsel for Covis Pharma GmbH

September 16, 2022

APPENDIX

COVIS' PROPOSAL OF RANDOMIZED CONTROLLED TRIAL (RCT) AND OBSERVATIONAL STUDY OPTIONS TO FURTHER CONFIRM THE CLINICAL BENEFIT OF MAKENA

As explained in detail in Covis' Briefing Book, the benefit of Makena in high-risk patients should be confirmed while the product remains available to the affected patient population. Since becoming the Makena sponsor in 2021, Covis collaborated with healthcare technology company, Dorsata,[1] and has worked with a multidisciplinary scientific advisory panel to evaluate frameworks for and the feasibility of future studies that could confirm Makena's clinical benefit in high-risk women. Covis welcomes the opportunity to further engage with the Agency on a data-driven path forward for this important study.

The scientific advisory panel included leaders in the fields of obstetrics, gynecology, biostatistics, epidemiology, clinical trials, oncology and drug development, who collectively have decades of experience at FDA and on FDA advisory committees. Scientific advisory panelists and advisors include:

- a Ph.D. biostatistician and clinical trialist who has served on multiple FDA advisory committees
- an M.D. clinical trialist, and former assistant commissioner at FDA, who has focused on diversity in clinical trials
- an M.D. clinical drug development expert and former supervisory physician at CDER
- an M.D. clinical trialist, infectious disease specialist, and pharmacoepidemiologist, who has served on multiple FDA advisory committees
- an OB/GYN who served as a medical officer in CDER's Division of Bone, Reproductive and Urologic Products
- an M.D. preventive medicine specialist and former FDA epidemiologist and team leader
- an OB/GYN who previously served as the health policy lead at a major women's health physician organization.

[1] As explained in greater detail in Section I.B below, Dorsata is a healthcare technology company focused on improving women's health. Dorsata provides a maternity care management software platform that is used for decision support, documentation, obstetrical care plans, order entry, and clinical data reporting, among other things. Most relevantly, Dorsata's current database comprises of as many as over 210,000 pregnancies, enabling Covis to perform deep-dive analyses of preterm birth and insights to inform clinical trial development.

With these outside experts, Covis has conducted extensive analysis of the available data to identify the patient population most at risk of a subsequent preterm birth. This analysis has included evaluation of the PROLONG and Meis trial data to explore the extent to which the risk profile of the patients differed between the trials. As discussed in Sections I.A and I.C, PROLONG failed to enroll a high-risk population as reflected by the much lower number of events than planned by protocol. In addition to this analysis, Covis conducted a review of data from the Dorsata database to identify the appropriate high-risk population that could benefit from Makena, as well as the gestational endpoint at which neonatal morbidity and mortality shows the sharpest improvement.

Given these insights, Covis and its expert panel have developed proposals for feasible confirmatory randomized controlled trials (RCTs) that can be conducted in the identified high-risk population in the U.S. This data-driven approach will avoid the flaws in the conduct of PROLONG, which enrolled a much lower risk population, and is more likely to confirm the result of the Meis trial. The remainder of this Appendix presents several proposals for additional studies, which represent a range of designs and also vary in terms of the feasibility to conduct such studies in a reasonable timeframe. The proposals presented here do not foreclose other possible approaches, but reflect extensive data analysis and the thinking of a highly qualified panel of experts well-positioned to critically examine a range of possible study designs and offer finely tuned recommendations. Covis is prepared to work with CDER to discuss these proposals and any other feasible methods of further investigating the value of Makena as an important treatment option for patients.

In addition, the Appendix includes a description of potential observational studies, which could also be utilized to further explore the clinical benefit of Makena.

I. THE RATIONALE FOR CONDUCTING ANOTHER STUDY IN A HIGH-RISK PATIENT POPULATION OF PRIOR SPONTANEOUS PRETERM BIRTH <34 WEEKS GESTATIONAL AGE

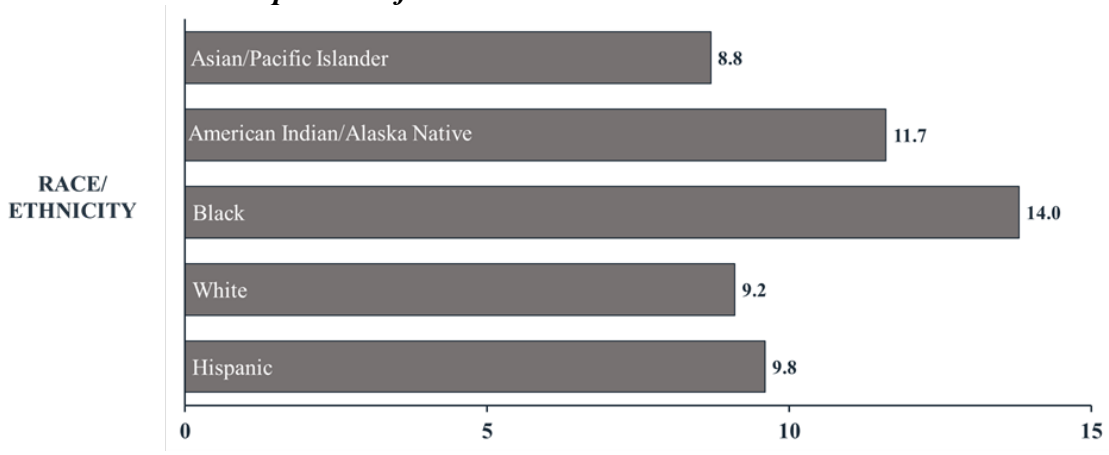
A. The Patient Population In PROLONG Had Much Lower Risk Profile Than That In In The Meis Trial

The previous Advisory Committee meeting materials, as well as the published literature demonstrated in detail how the data from PROLONG (both overall and the US subgroup) differed significantly from the Meis data. Table 7 in the Briefing Materials compares the demographics between Meis and PROLONG/PROLONG-US, while Figure 9 in the Briefing Materials compares the placebo rates for preterm birth across the three groups.

Table 7
Different Social and Demographic Characteristics Across PROLONG Meis Trials

Demographics/Baseline Characteristics	Meis (N=463) %	U.S. PROLONG (N=391) %	PROLONG (N=1708) %
Age (years), mean ± SD	26.2 ± 5.6	27.6 ± 5.1	30.0 ± 5.2
>1 previous SPTB	28.9	27.4	14.5
GA of prior SPTB (median)	32 wks	34 wks	33 wks
Black/African American	59.0	28.9	6.7
Hispanic or Latino	14.9	13.8	9.1
Unmarried with no partner	50.3	30.7	10.1
Educational status (≤ 12 years)	71.3	50.5	43.7
Any substance use during pregnancy	26.1	28.4	9.3

Figure 9
Comparison of Placebo Rates Across Studies



Although it is clear that the overall incidence of events is markedly lower both in the overall and U.S. subgroup for PROLONG, the reasons for the differences in outcomes has not been described in as much detail. With a view to both understanding the reasons for the differences in pregnancy outcome as well as to define the appropriate inclusion criterion for another study, Covis embarked on a detailed modeling exercise to identify the factors that appear to most strongly influence preterm birth rate. While there has been a qualitative understanding in the Ob/Gyn/maternal-fetal medicine (MFM) community about the risk factors that predispose a patient for preterm birth, we sought to utilize datasets available to us to formalize (to the extent possible) the relationships between risk factors and the likelihood of preterm birth in the target population.

B. Modeling From Three Datasets—Dorsata, Meis and PROLONG—Makes Clear That Women With Prior Spontaneous Preterm Birth <34 Weeks Gestational Age Are The Relevant High-Risk Patient Population

Dorsata is used by more than 2,300 users including obstetricians and gynecologists and their clinical staff to document care in structured data elements for prenatal and postpartum

encounters. These obstetricians and gynecologists care for over 1.3 million unique patients annually. Dorsata's current database comprises over 210,000 pregnancies and continues to grow rapidly. The key to Dorsata's system is that it serves as the primary medical record for obstetrics. Using proprietary software, care plans which incorporate evidence-based clinical guidelines are surfaced for the providers to review within the clinical workflow. The Dorsata interface is built on top of the electronic health record (EHR) as an overlay that incorporates the ACOG prenatal flow sheet in the point-of-care system—enabling broad, structured clinical data collection—and allows clinicians to seamlessly manage and track the status of every one of their prenatal patients longitudinally.

Importantly, the data collected is congruent with that collected in registration RCTs. This system captures the majority of the typical demographic and risk factors for the prenatal patient. Covis' analysis of the Dorsata data has confirmed the ability to access significant numbers of pregnancies associated with high-risk patient indicators (for example, history of spontaneous preterm birth <32 or 34 weeks, comorbidities such as hypertension, obesity and smoking, race, and other variables known to be associated with high-risk pregnancies) with the ability to characterize all patients comprehensively on their background and demographic characteristics. Additionally, the Dorsata platform tracks medications taken by the patient and captures confirmatory information on whether prescriptions were fulfilled and whether the medication was administered in accordance with the labeled dosing schedule.

The Dorsata dataset analyzed comprised of approximately 114,000 pregnancies overall from 2018 through 2021, of which 2046 were patients who had been indicated for treatment with 17-OHPC based on an automatic flag based on their prior history of a spontaneous preterm birth. Covis performed an audit of these 2046 records which resulted in disqualification of 347 of the records for whom the prior spontaneous preterm birth could not be confirmed. The remaining 1699 patients were eligible for analysis. The analysis set in the Dorsata analyses presented in this Appendix are from that indicated subset of patients, which is further segmented into patients who were indicated but not prescribed therapy, into patients who were prescribed 17-OHPC therapy but where the prescription was not filled, and finally, patients who were prescribed 17-OHPC therapy. Further details on the Dorsata databases analyzed are available upon request. The risk modeling has been performed on the population which was indicated for 17-OHPC but not treated, and comprised of 987 subjects after excluding any patients who were missing data elements required for the modeling.

Covis first modeled the Dorsata database to characterize the risk factors, following which Covis built the same models for PROLONG-US and Meis to compare the findings. We did not seek to develop a model for the ex-US population in PROLONG given the low incidence rate of preterm birth in this population.

Presented below are several risk factor models developed using logistic regression, including the best one, two, and three parameter models within Dorsata. The primary outcome that the analyses tried to predict was the likelihood of the patient experiencing a subsequent preterm birth before week 34. The results of this model are not significantly impacted by varying the predicted outcome to 33 or 35 weeks—the purpose was to pick an outcome that was acknowledged by Ob/Gyns and neonatologists as significantly impacting morbidity/mortality of the neonate.

Methodologically, the "best" N-variable logistic regression model was selected based on the score test using PROC LOGISTIC in SAS. Each model selected the "best" N-variables from a larger set of potential risk factors, including but not limited to maternal age, race, pre-pregnancy BMI and weight, education level, smoking, alcohol, and drug use, inter-pregnancy interval, and gestational age of prior pregnancies. Gestational age parameters were based on prior spontaneous deliveries, and included mean gestational age, any deliveries <32 or <34 weeks, total numbers of deliveries <32 or <34 weeks, and having more than one prior delivery <32 or <34 weeks

The result of the Dorsata model is shown below in Table 1:

Table 1
Dorsata (Excluding 17-OHPC -treated Subjects) Best N-variable Models Predicting PTB <34 Weeks

Model/Var#	Variable(s)	Odds Ratio (95% CI)	P-value
Best 1-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.82, 0.91)	<0.0001
Best 2-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.82, 0.91)	<0.0001
Var #2	Smoking during Pregnancy	0.51 (0.18, 1.46)	0.21
Best 3-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.86 (0.81, 0.90)	<0.0001
Var #2	Smoking during Pregnancy	0.52 (0.18, 1.48)	0.22
Var #3	Alcohol Use during Pregnancy	0.64 (0.28, 1.45)	0.28

As can be seen from the above, prior pregnancy history and in particular, mean gestational age of prior pregnancies appears to be a strong predictor of a subsequent preterm birth <34 weeks. Given the above models, Covis also constructed similar one through three parameter models using the Meis data to see if the real-world data in Dorsata (from 2018-2021) were congruent with the Meis data from twenty years earlier. The Meis risk models are shown below in Table 2:

Table 2
Meis (Vehicle-only) Best N-variable Models Predicting PTB <34 Weeks

Model/Var#	Variable(s)	Odds Ratio (95% CI)	P-value
Best 1-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.89 (0.82, 0.98)	0.013
Best 2-variable model			
Var #1	Mean Gestational Age of Prior Pregnancies (Weeks)	0.89 (0.81, 0.98)	0.013
Var #2	Smoking during Pregnancy	2.84 (1.19, 6.76)	0.019
Best 3-variable model			
Var #1	Smoking during Pregnancy	3.12 (1.27, 7.68)	0.013
Var #2	Mean Gestational Age of Prior Pregnancies (Weeks)	0.90 (0.82, 0.99)	0.030
Var #3	Inter-pregnancy Interval (Years)	0.84 (0.71, 1.00)	0.049

Both datasets show the importance of prior pregnancy history. While we examined the predictive power of several different measures of previous preterm birth history it is interesting that both datasets indicated that the mean gestational age (GA) of previous pregnancies was the strongest single parameter model. We also observe that the models do differ in the importance of other factors such as smoking or alcohol use. As the data for these parameters in Dorsata are self-reported by patients and also not specific with respect to the extent of smoking or alcohol use (occasional vs. habitual), we would be inclined to utilize the Meis models (which captured these parameters formally as part of the clinical study) to select the additional factors that drive risk. Of note, the PROLONG-US model did not provide any material new conclusions about these risk factors, due to which we are not presenting those data here. We also point out that the concept of average of gestational ages from the patient's previous birth history is useful from a modeling perspective because it encodes the idea that the patients risk depends on their overall history. However, it is not a practical parameter for defining an inclusion criterion for a future study (and indeed we may need other ways of incorporating this concept, whether for inclusion criterion in a future study or for labeling purposes).

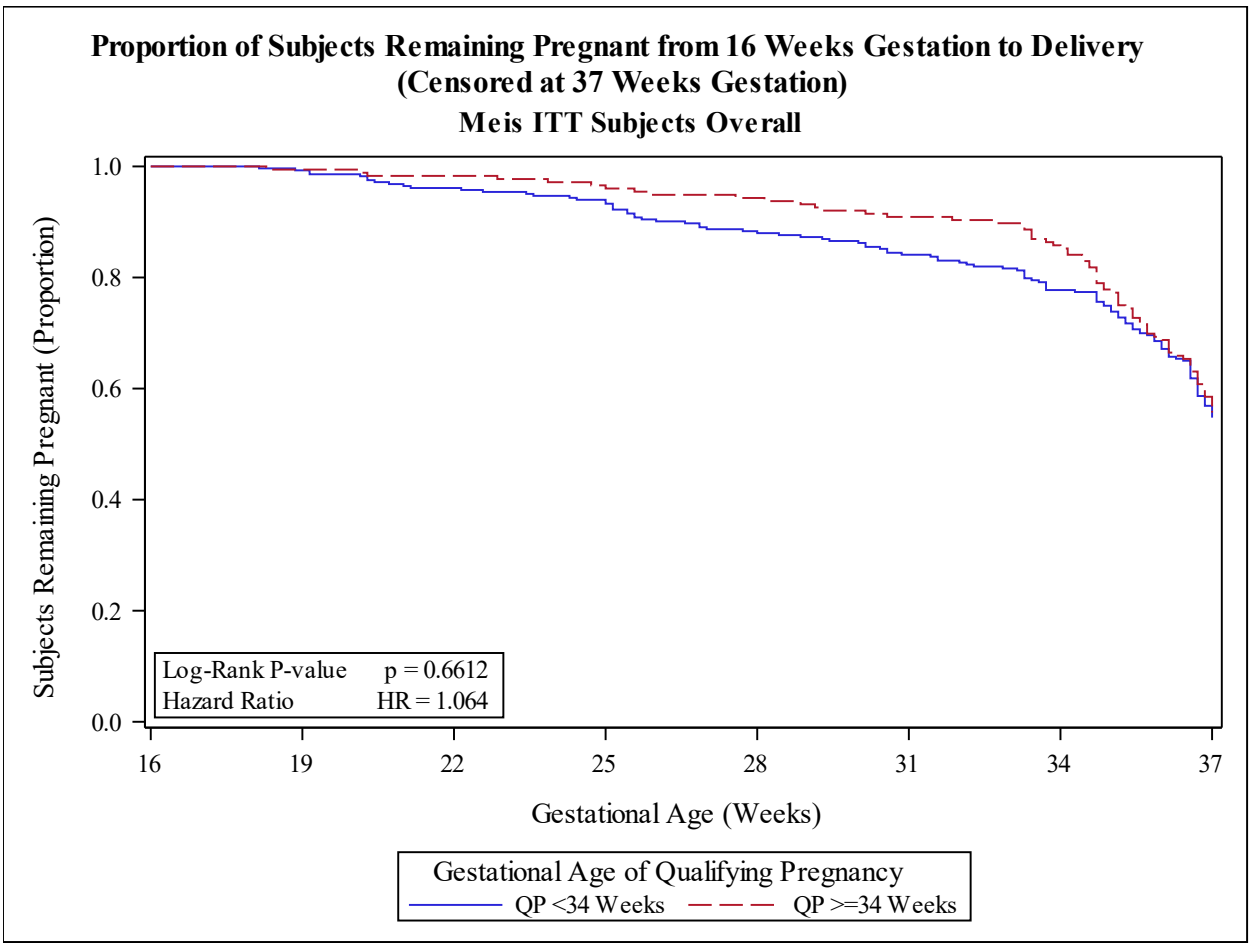
Covis also modeled the Dorsata data to understand what cutoffs in the patients' prior pregnancy history may prove more predictive of a subsequent preterm birth. The Dorsata data indicate that there is a sharp reduction in the likelihood of a future preterm birth if the patient's history does not include a prior spontaneous preterm birth prior to week 34 as shown in Table 6, which analysis was performed in a population in Dorsata that was indicated for 17-OHPC treatment based on prior history but did not receive 17-OHPC medication.

Table 6
Count (%) of Subjects with Study Pregnancy Outcome by Gestational Age at Earliest Prior Preterm Birth Dorsata (17-OHPC Untreated)

Gestational Age at Earliest Prior SPTB	Study Pregnancy Outcome	Count	%
20 to <28	Had a SPTB<34	10	11.0
	Not a SPTB	81	89.0
28 to <31	Had a SPTB<34	7	11.1
	Not a SPTB	56	88.9
31 to <33	Had a SPTB<34	14	17.7
	Not a SPTB	65	82.3
33 to <34	Had a SPTB<34	8	18.2
	Not a SPTB	36	81.8
34 to <35	Had a SPTB<34	12	8.2
	Not a SPTB	135	91.8
35 to <36	Had a SPTB<34	11	5.7
	Not a SPTB	182	94.3
36 to <37	Had a SPTB<34	13	4.0
	Not a SPTB	313	96.0
37 to <38	Had a SPTB<34	4	4.2
	Not a SPTB	91	95.8

Examining the Meis data also corroborates the increased risk of preterm birth in the Meis population in the subgroup which had a prior spontaneous preterm birth <34 weeks as seen in the following Time-to-event (Kaplan-Meier) plot (Figure 6) showing the event curves in the (34-36^6 prior preterm birth) and <34 week prior preterm birth populations.

Figure 6
Time-to-Event Analysis for Meis Population With Prior SPTB <34 Weeks



The red curve shows the events in the population that had a prior spontaneous preterm birth from week 34 to week 36^6 (which is the same as the "No prior SPTB <34 weeks" group), while the blue curve shows the events in the <34 week population. We also note that there is support for the <34 week cutoff from previous literature (Laughon 2014 and Esplin 2008), which both show risk tables that are similar to the results we have obtained from Dorsata.

Finally, we may note that the above offers an additional explanation for why the ex-US PROLONG data seemed to indicate no benefit from treatment. Not only were the event rates lower due to the low risk of patients, but also there is likely to have been an issue with the accuracy of the documented prior pregnancy history.

These additional criteria address many of the inadequacies observed in the recruitment and conduct of PROLONG, where much of the study population was at a low risk of subsequent

preterm birth, and inconsistent and unreliable methods were used to verify the gestational age of the qualifying delivery.[2]

In the last decade, U.S. obstetricians have followed WHO guidelines and generally used ultrasound to measure the crown-rump length (CRL) or gestational sac mean diameter to determine gestational age.[3] An ACOG recommendation explains that first semester ultrasound measurement is accurate in establishing gestational age, while ultrasound in the second and third trimester are not similarly reliable for calculating gestational age. In contrast, these obstetric practices were not common in ex-U.S. countries during the time that PROLONG was recruiting. In particular, in Russia and Ukraine, where the majority of the PROLONG study population was enrolled, the practice history of gestational age determination has not been as clearly defined. What data exists suggests that the use of ultrasound during the first trimester of pregnancy to determine and document gestational age was less prevalent and/or not standardized when compared with ultrasound utilization in the United States.

Covis' review of the intake forms for PROLONG participants confirms that there was no uniform method used to confirm the gestational age of the patients' qualifying delivery. The study protocol also did not require that a certain method be used, rather instructing that "where possible," the gestational age of the qualifying delivery be determined by a combination of the last menstrual period method and ultrasound examination, and the screening criteria did not require any verification of the gestational age of the prior birth in the mother's medical record. It is therefore likely that the gestational age of the qualifying delivery was based on the last menstrual period for many of the ex-U.S. PROLONG patients.

The last menstrual period method is generally known to be unreliable and is recommended only when ultrasonography facilities are not available. This is because the method involves a number of assumptions, namely, a regular menstrual cycle of 28 days, with ovulation occurring on the 14th day after the beginning of the menstrual cycle. As ACOG has recognized, the LMP method therefore does not account for irregularities in cycle length, variability in the timing of ovulation, or even the fact that women may inaccurately recall their last menstrual period. Indeed studies have shown that approximately one half of women inaccurately recall their last menstrual period and as many as 40% of the women experienced more than 5-day discrepancies in the estimated due date between ultrasound dating and LMP dating. Indeed, it has

[2] As discussed in the Covis' Briefing Book, Section VII.A.3.a-b, the PROLONG population consisted primarily of women with lower risk factors for a subsequent preterm birth. This assessment has been further confirmed by Covis' analysis of the Meis, PROLONG, and Dorsata data, which demonstrate that the high-risk population consists of those women with a prior spontaneous preterm birth at <34 weeks gestational age. As discussed in the Briefing Book, Section VII.A.3.c, the assessment of gestational age for the ex-U.S. patients in PROLONG was based on inconsistent and unreliable methods, including second or third trimester ultrasound or last menstrual period.

[3] *See* Scott N. MacGregor and Rudy E. Sabbagha, *Assessment of Gestational Age by Ultrasound*, GLOB. LIB. WOMEN'S MED. (2008), https://www.glowm.com/section-view/heading/Assessment%20of%20Gestational%20Age%20by%20Ultrasound/item/206#.

been estimated that up to one quarter of the preterm births that were classified using LMP may in fact not be preterm.[4]

There has been a marked difference in the use of obstetric ultrasonography in Russia as compared with the United States. By 2015, 90% of women in the United States received ultrasounds during the first trimester compared with just 77% of women in Russia.[5] Further, prior to 2009, the use of ultrasonography specifically with respect to gestational measurement was, at best, sporadic. The use of ultrasound along with LMP has been documented in Northwest Russia in the early 1990s, however, only a small portion of these births refer to ultrasound with LMP being the primary method by which gestational age determination was made.[6] Additionally, Covis conducted a comprehensive literature search and was unable to identify any literature pointing to ultrasonography as the standard method of gestational age determination in Russia. Conversely, several published studies on various aspects of maternal and pediatric health which cite LMP as the method for gestational age determination.[7]

The use of ultrasound for gestational age determination in Russia remains limited today. The authors of a recent review of the fetal growth calculation effort in Russia note that "[t]here is no consensus on fetal growth monitoring in modern Russia. Neither the Russian Society of Obstetricians and Gynecologists, nor the Russian Association of Specialists in Ultrasound Diagnostic in Medicine has ever published any clinical recommendations concerning the application of fetal growth charts."[8] Russian clinicians use varying fetal growth charts, with no consistent quantitative methodology or underlying clinical or biological hypothesis.[9] As a result, fetal growth gestational age measurements suffer from several methodological errors including inaccurate gestational age measurements, inaccuracy in population differentiation and inclusion/exclusion criteria, and deficiencies in imaging standardization protocols.[10] Therefore, although ultrasound screening was mandated for all outpatient clinics in Russia in the year 2000, there is limited evidence of documented ultrasound-estimated gestational age prior to 2009 and after 2009, even where implemented, the methodologies used are inconsistent and not aligned with WHO standards.

In the case of Ukraine, it is even more clear that LMP was the prevailing method for first trimester gestational age determination. Though the vast majority of pregnant women in Ukraine

---

[4] Michael S. Kramer, et al., *The Validity of Gestational Age Estimation by Menstrual Dating in Term, Preterm, and Postterm Gestations*, 22 JAMA 3306-3308 (Dec. 9, 1998), https://jamanetwork.com/journals/jama/article-abstract/375526.

[5] *See* Shuvalova et al., *supra* note 33; O'Keefe, et al. *supra* note 33.

[6] *See* Postoev, et al., *Changes in detection of birth defects and perinatal mortality after introduction of prenatal ultrasound screening in the Kola Peninsula (North-West Russia): combination of two birth registries*, 15 BMC Pregnancy Childbirth 308 (Nov. 23, 2015), https://pubmed.ncbi.nlm.nih.gov/26596677/; Anna A. Usynina, *Risk factors for perinatal mortality in Murmansk County, Russia: a registry-based study,* 1 Glob Health Action (2017), https://pubmed.ncbi.nlm.nih.gov/28156197/.

[7] *See supra* note 197.

[8] A. M. Kholin, et al, *Ways to standardise of fetometry in Russia: INTERGROWTH-21st project and its implementation*, 9 Obstetrics and Gynecology (2018), https://en.aig-journal.ru/articles/Podhody-k-standartizacii-fetometrii-v-Rossii-proekt-INTERGROWTH-21-i-ego-vnedrenie.html.

[9] *See id*.

[10] *See id.*

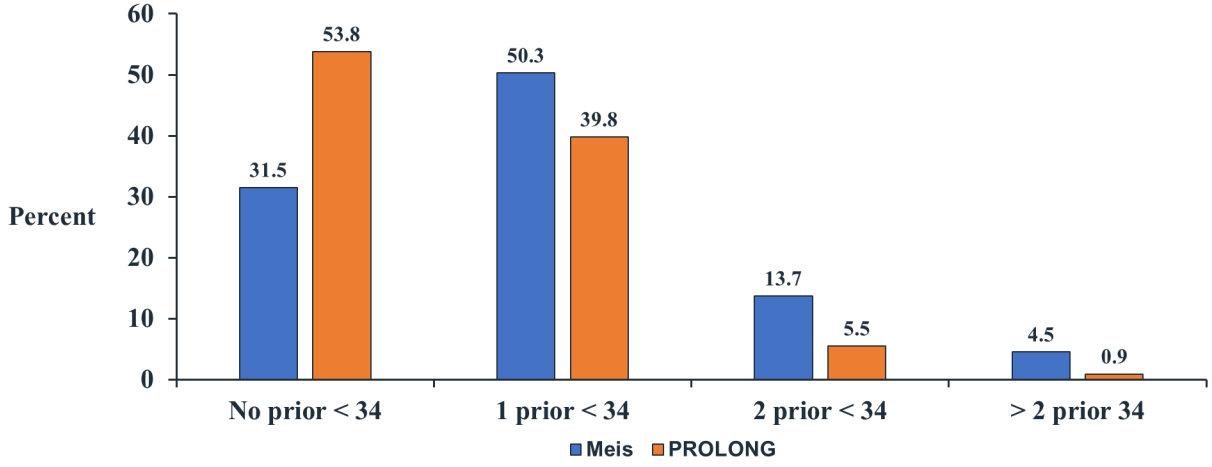
routinely received ultrasounds as far back as 1995, the standard of care in Ukraine recommended second trimester ultrasounds, not first trimester ultrasounds.[11] A comprehensive study of the outcomes of 17,137 pregnancies published in 1999 cited the use of LMP to estimate gestational age.[12] More recently, authors of a 2019 study on maternal alcohol use among Ukrainian mothers noted difficulty in obtaining gestational age when assessing pre-term birth due to the lack of ultrasound screenings and poorly documented LMP records.[13] Similar to the Russian case, use of first trimester ultrasound in Ukraine has been poorly documented and LMP appears to be the primary method used to determine gestational age in early pregnancy.

These data show that the prior gestational age history of PROLONG subjects in Russia and Ukraine was unreliable both at a qualitative level (i.e., whether or not the subject had a prior spontaneous preterm birth), particularly if the documented sPTB was at week 35 or later due inaccuracies as a result of use of the LMP method, and at a quantitative level, where risk modeling based on prior gestational age history necessarily suffers from reliance on data of poor quality.

C. PROLONG-US And Meis Differed Fundamentally With Respect To Risk Factors

Given the above models, we can also now examine the differences in the PROLONG-US and Meis populations more closely. The following graphic (Figure 1) compares the frequency of prior spontaneous preterm birth <34 weeks in PROLONG-US compared to Meis.

Figure 1
Frequency of Prior Qualifying SPTB <34 Weeks (Meis vs. PROLONG-US)



[11] *See* Arbuzova, *supra* note 233, at 184 (noting that "General ultrasound screening twice, at 16-18 and at 24-27 weeks of pregnancy, is recommended").

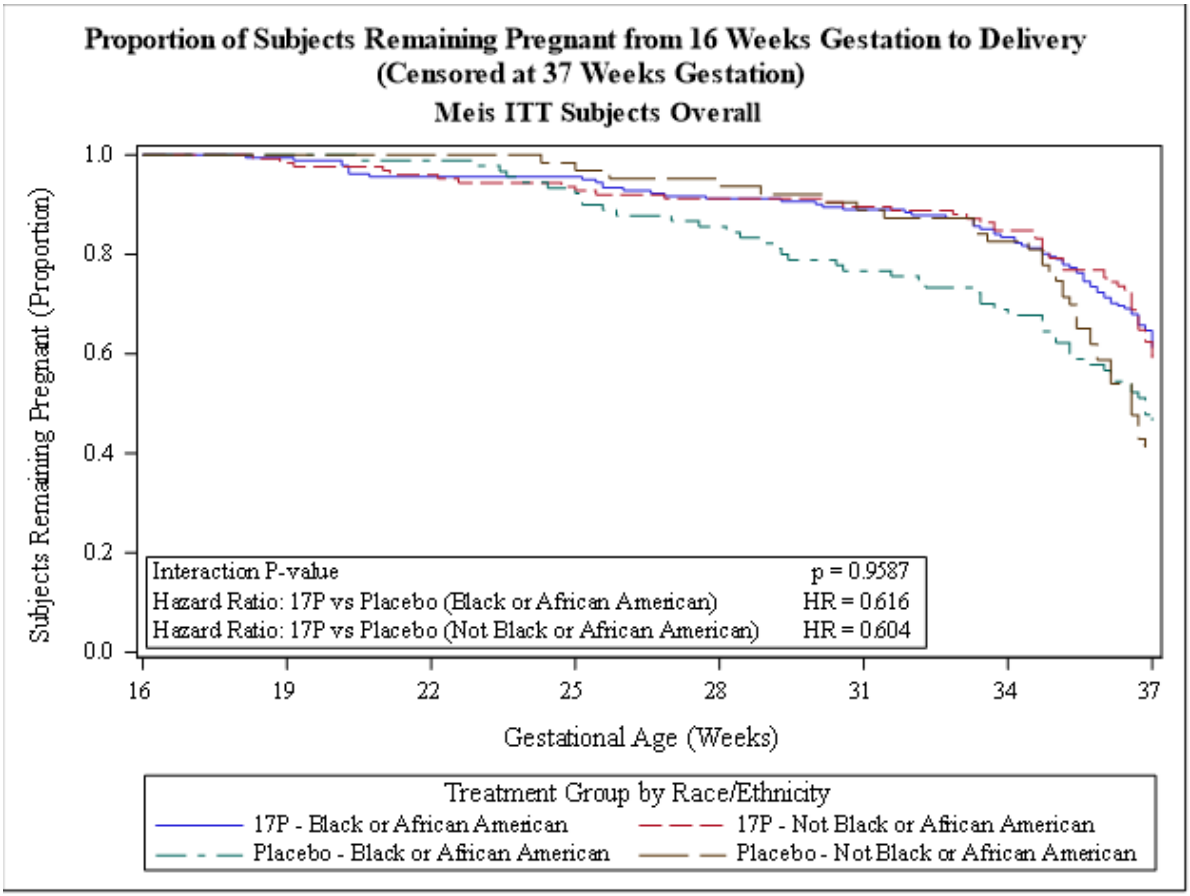
[12] Little et al., *supra* note 202.

[13] Claire D. Coles et al., *Gestational age and socioeconomic status as mediators for the impact of prenatal alcohol exposure on development at 6 months*, 12 BIRTH DEFECTS RES. 789-96 (Jul. 15, 2019), https://pubmed.ncbi.nlm.nih.gov/30378744/.

The graphic shows that nearly 70% of the patients in Meis had a prior spontaneous preterm birth before week 34 compared to about half of PROLONG-US. The percentage of subjects who had 2 or more prior SPTBs <34 weeks is also substantially higher in Meis than PROLONG-US. Additionally, analyzing the pregnancy history of Meis and PROLONG-US patients show a shift in the average gestational age of prior pregnancies, another parameter that Covis' risk modeling above showed to be predictive of a future preterm birth. The importance of prior pregnancy history underscores the need to have accurate and documented prior birth history using first trimester ultrasonography given the noise associated with LMP.

Finally, there are potentially important differences in the racial and socio-economic makeup of the two populations that may influence the event rate in a meaningful manner. One example of this is seen in the Meis data, when analyzed in a time-to-event basis. The figure below shows, interestingly, that while Meis overall showed similar therapeutic effect of 17-OHPC on preterm birth as measured against a <37 week endpoint, that Black patients in Meis experienced a reduction in events in an earlier gestational timeframe than the non-Black patients.

Figure 7
Time-to-Event Analysis for Meis Population (Black vs. Non-Black)

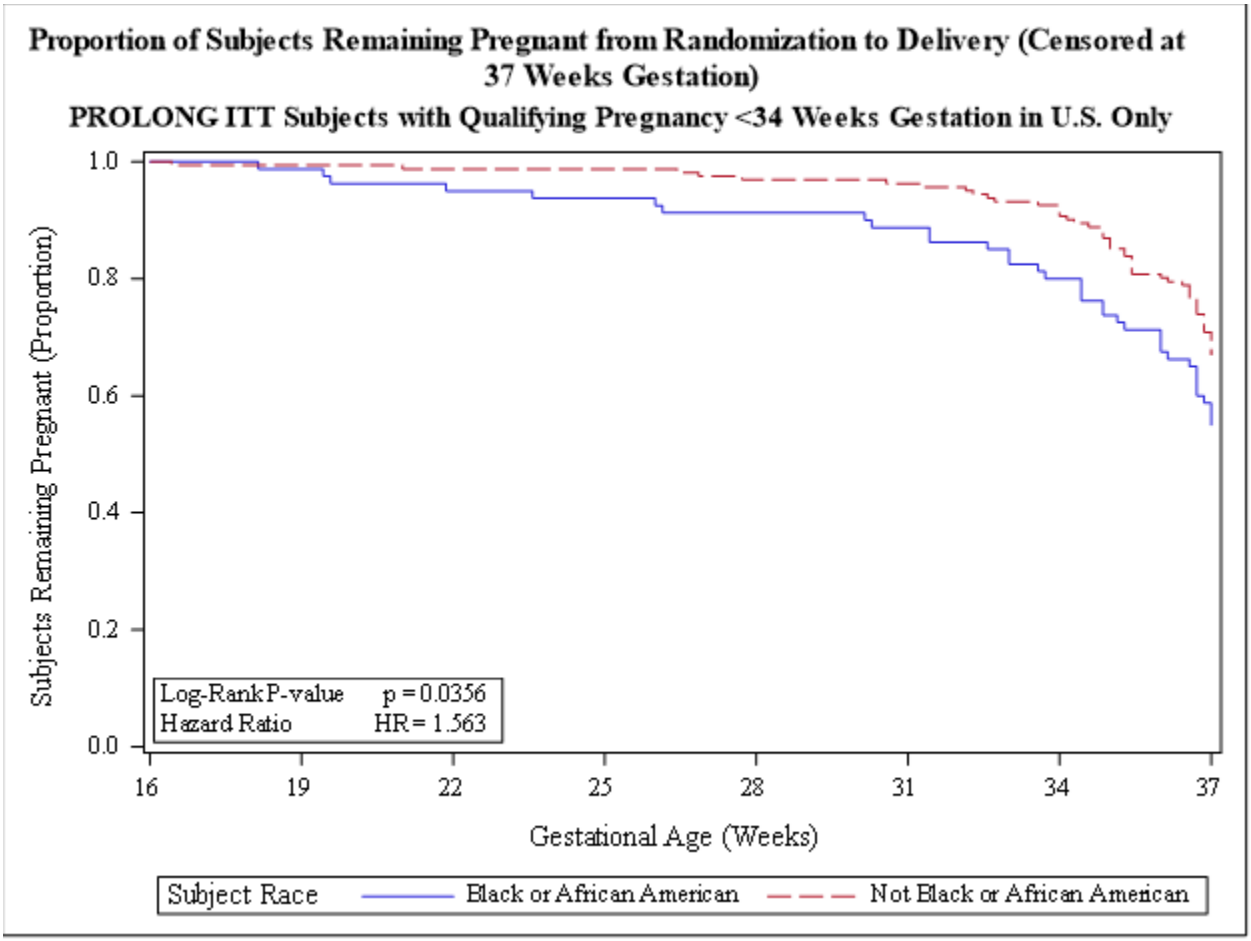


Other ways of looking at the data confirm this effect—whether examined in terms of efficacy with a gestational endpoint <32 or <35 weeks (where the majority of the reduction in

event rates is in the Black population), examining these data in an ordinal analysis where the events are binned by early, mid or late-preterm birth outcome, or when looking at a "change from baseline" approach where the measured parameter is the delta between a recent qualifying pregnancy in the individual and the gestational age achieved in the study.

The pattern of earlier events in Black patients is not only seen in the Meis data. The same pattern is seen in high-risk patients (prior SPTB <34 weeks) in PROLONG-US, as seen in the following time-to-event plot comparing Black and non-Black patients.

Figure 8
Proportion of Subjects Remaining Pregnant from Randomization to Delivery (Censored at 37 Weeks Gestation)



Despite the post hoc nature of the above analyses, it is important to acknowledge the differences in the data for Black and non-Black patients. The smaller proportion of Black patients in US-PROLONG coupled with the less severe prior history of preterm birth indicates that the two studies were in different populations, and that PROLONG (or PROLONG-US) does not represent an appropriate confirmatory study for Meis.

Given the generally higher risk for preterm birth seen in the US Black population, the evidence within Meis (if analyzed with efficacy for PTB <35 weeks or earlier) of a race

interaction, and the low risk of PROLONG-US based on previous birth history, Covis believes it is important to fully understand the therapeutic benefits of 17-OHPC treatment in this indication. The only way to gain this understanding is to conduct another well-designed study.

D. An Exploratory Analysis Of PROLONG-US Data Suggests Efficacy In The High-Risk Population And Value In Using Continuous Endpoints For Future Confirmatory Study

Covis Pharma has re-analyzed both PROLONG and Meis to construct models that seek to predict the likelihood of a preterm birth in an individual with a prior spontaneous preterm birth. Notably, Covis has rigorously examined several variations of the prior pregnancy history, one of the most significant risk factors for subsequent preterm birth as described above, to further characterize and model the probability of a subject having a preterm birth. Specifically, Covis has explored the average of prior gestational ages of live birth pregnancies (referred to as mean gestational age or mGA) with various cutoffs (<35 weeks, <34 weeks, <32 weeks, <28 weeks etc.), the number of preterm births whose gestational age was less than a given cutoff (<37 weeks, <35 weeks, etc.), and the gestational age of most recent pregnancy preceding the study (referred to as the mrpGA). This last factor was included because there were examples of patient histories in all datasets where a patient had multiple pregnancies but where the more recent pregnancies had continued to full term. In general, clinicians would view such patients as being of lower risk.

In addition to the conventional categorical endpoints of PTB rate at specific cutoffs (e.g., <37, <35. <32 weeks), Covis examined a variety of continuous endpoints that were designed to probe whether 17-OHPC was extending the pregnancy and adding any additional time in utero relative to placebo. Our hypothesis was that these continuous endpoints would be more sensitive than the categorical endpoints used in Meis or PROLONG and may tease out a signal where the categorical endpoints did not show an effect.

We performed analyses on the change from baseline to the study pregnancy where baseline was the mrpGA, as well as time from randomization to birth. For both of these continuous endpoints, we saw signals of efficacy in terms of weeks gained on 17-OHPC relative to placebo in PROLONG-US, particularly in subgroups known to be higher risk such as those with a more severe history of preterm births as well as the Black subpopulation. As shown below, similar effects were seen in the Meis data when analyzed in this manner. For ex-US PROLONG, however, there was no difference in weeks gained for any subgroups regardless of risk factors.

The endpoint for all analyses was time (weeks) from randomization until the earlier of (1) delivery or (2) 35 weeks gestation (i.e., time capped at 35 weeks gestation). The analysis population included women randomized up to 19 weeks and 6 days gestation; women randomized at 20 weeks gestation or later were excluded (on account of previous CDER statistical reviews for the Meis study noted that the treatment effect of 17-OHPC was present only when the subjects were randomized before week 20[14]). All analyses were performed using

[14] NDA 21945, Statistical Reviews at 16, https://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/021945Orig1s000StatR.pdf.

linear regression with time from randomization (capped at 35 weeks gestation) as the dependent variable. All analyses included treatment and GA at randomization as independent variables. Analyses of high-risk subgroups (Table 3) also included adjustment for mean gestational age of prior spontaneous deliveries. Analyses among subsets defined by the mrpGA among spontaneous deliveries and mGA were also adjusted for mGA (Table 8) and mrpGA (Table 9), respectively. Only spontaneous births at a minimum of 20 weeks GA were included in the calculation of mrpGA and mGA. The reason to cap at 35 weeks gestation was to focus on the period of gestation viewed as most beneficial to the fetus from the perspective of increased time in utero.

Of note, the table shows a clear numerical increase in weeks gained by 17-OHPC versus placebo as we analyze subgroups with a larger number of risk factors. As shown in Manuck et al. (2016) and Richter et al. (2019), this is clinically significant as the addition of 1-2 weeks of gestational age prior to week 35 is associated with marked reduction in neonatal morbidities.[15]

Table 3
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Known Risk Factor Subgroup Among Subjects Randomized Prior to 20 Weeks GA for PROLONG-US

Risk Factor Subgroup	N Total	Estimated treatment effect* (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
Overall	389	0.49	-0.04	1.01	0.0684
Most Recent Prior Spontaneous Delivery at GA<35 (mrpGA<35)	137	1.30	0.30	2.29	0.0113
Black Subjects with mrpGA<35	51	1.57	-0.28	3.42	0.0936
Subjects with Inter-pregnancy Interval <5 Years (IPINT<5) and mrpGA<35	112	1.55	0.34	2.76	0.0126
Subjects with More than One Prior sPTB<37 (MTO37) and mrpGA<35	23	0.99	-0.74	2.72	0.2470
Subjects with IPINT<5 and MTO37 and mrpGA<35	16	2.08	-0.54	4.69	0.1099
Black Subjects with IPINT<5 and mrpGA<35	38	1.75	-0.77	4.26	0.1673
Black Subjects with MTO37 and mrpGA<35	9	-0.10	-0.57	0.37	0.6056

* Within group estimates for the 17P treatment effect (weeks gained from randomization, capped at GA=35) based on model including: Treatment, Mean GA of Prior Spontaneous Deliveries (mGA), and GA at Randomization.

The same analysis has been performed with the Meis data (Table 18). Qualitatively, the analyses for Meis and PROLONG-US exhibit a similar pattern in terms of the increase in weeks gained with increasing risk of the subjects. To put in context the numbers for weeks gained, it is helpful to refer to the body of literature that clearly establishes the incidence rates of neonatal

[15] *See* Manuck et al., *supra* note 52; Lindsay A. Richter et al., *Temporal Trends in Neonatal Mortality and Morbidity Following Spontaneous and Clinician-Initiated Preterm Birth in Washington State, USA: A Population-Based Study*, 9 BMJ OPEN e023004 (2019). Covis has also performed an exhaustive literature search regarding preterm morbidity incidences at various gestational ages and can provide additional information as well as validation of the tables contained herein at CDER's request.

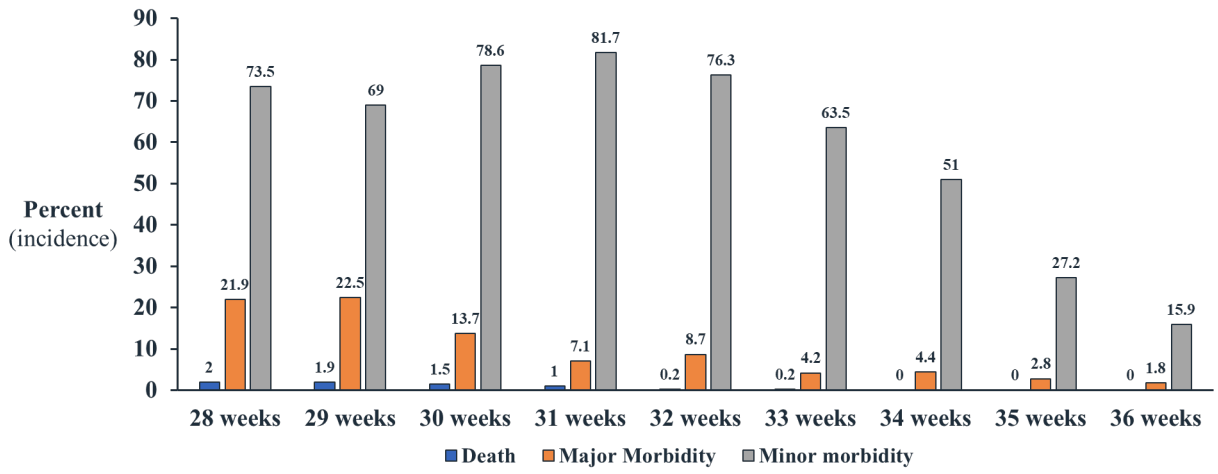
morbidity and mortality. In the cited articles (Manuck 2016 and Richter 2019) we note a week-on-week reduction of the rate of morbidities from weeks 28 onwards through to week 36. By capping weeks gained to 35 weeks of gestation, even a week of gestation has a significant impact on the rate of neonatal morbidities as shown in the mortality/morbidity incidence tables.

Table 18
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Known Risk Factor Subgroup Among Subjects Randomized Prior to 20 Weeks GA for Meis

Risk Factor Subgroup	N Total	Estimated treatment effect* (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
Overall	313	0.82	-0.09	1.73	0.0788
Most Recent Prior Spontaneous Delivery at GA<35 (mrpGA<35)	193	0.81	-0.46	2.09	0.2112
Black Subjects with mrpGA<35	115	1.45	-0.26	3.17	0.0966
Subjects with Inter-pregnancy Interval <5 Years (IPINT<5) and mrpGA<35	165	1.26	-0.02	2.54	0.0544
Subjects with More than One Prior sPTB<37 (MTO37) and mrpGA<35	62	2.67	0.65	4.70	0.0107
Subjects with IPINT<5 and MTO37 and mrpGA<35	56	3.22	1.26	5.18	0.0017
Black Subjects with IPINT<5 and mrpGA<35	96	1.94	0.16	3.71	0.0326
Black Subjects with MTO37 and mrpGA<35	44	1.88	-0.63	4.38	0.1384

*Within group estimates for the 17P treatment effect (weeks gained from randomization, capped at GA=35) based on model including: Treatment, Mean GA of Prior Spontaneous Deliveries (mGA), and GA at Randomization.

Figure 15
Incidence Rates by Gestational Age[16]



In order to refine this analysis, with an eye towards defining the inclusion criteria for a future confirmatory study in higher risk patients, we also examined the relationship between prior pregnancy history cut points and weeks gained on 17-OHPC relative to placebo. For PROLONG-US, the two prior pregnancy history measures we analyzed were the mean prior GA and the mrpGA. These results are presented in Tables 8 and 9 below:

Table 8
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Mean Gestational Age (mGA) of Prior Deliveries Among Subjects Randomized at <20 Weeks GA for PROLONG-US

mGA Subgroup	N Total	Estimated treatment effect (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
mGA<28	28	3.48	0.60	6.36	0.0198
mGA<29	34	2.56	0.20	4.93	0.0348
mGA<30	41	2.20	0.23	4.17	0.0295
mGA<31	54	0.96	-0.87	2.79	0.2970
mGA<32	56	0.97	-0.82	2.75	0.2817
mGA<33	81	1.01	-0.27	2.29	0.1214
mGA<34	101	0.89	-0.54	2.32	0.2186
mGA<35	142	0.42	-0.65	1.48	0.4399
mGA<36	191	0.47	-0.36	1.30	0.2688
mGA<37	254	0.54	-0.17	1.24	0.1351

[16] *See* Manuck et al., *supra* note 52 (Figure 15 shows week-on-week reduction of the rate of morbidities from weeks 28 onwards through to week 36. By capping weeks gained to 35 weeks of gestation, even a week of gestation has a significant impact on the rate of neonatal morbidities. Richter et al. (Tables 22 and 23) provide further validation of the Manuck data as well as incidence of specific morbidities by week of gestational age that confirm the benefit seen for additional weeks *in utero*.

Table 9
Estimated Treatment Effect (Weeks Gained) for 17-OHPC in Subgroups Defined by Most Recent Prior Gestational Age (mrpGA) of Previous Deliveries Among Subjects Randomized at <20 Weeks GA for PROLONG-US

mGA Subgroup	N Total	Estimated treatment effect (weeks gained)	Lower 95% CL	Upper 95% CL	P-value
mrpGA<28	37	3.33	0.93	5.73	0.0081
mrpGA<29	45	2.48	0.39	4.56	0.0210
mrpGA<30	51	2.20	0.39	4.00	0.0183
mrpGA<31	57	1.82	0.22	3.42	0.0262
mrpGA<32	64	1.60	0.16	3.04	0.0297
mrpGA<33	84	1.26	0.08	2.45	0.0371
mrpGA<34	101	1.44	0.29	2.59	0.0149
mrpGA<35	137	1.43	0.42	2.44	0.0058
mrpGA<36	195	0.98	0.22	1.74	0.0118
mrpGA<37	248	0.96	0.24	1.67	0.0090

The above two tables demonstrate a monotonic relationship between the weeks gained on 17-OHPC and the risk of the subject as defined by either the GA from their pregnancy immediately prior to enrolling in the study or the mean of their prior GAs.

We have also repeated this analysis for the ex-US PROLONG subgroup but did not see a benefit from 17-OHPC treatment in terms of the weeks gained since randomization. This may be due to:

1. The overall risk level of the ex-US subjects is low. While the preterm birth rate in ex-US PROLONG is enriched relative to the general population, the PTB rate for the placebo group in ex-US PROLONG, Russia and Ukraine was 20%, 17% and 21% respectively (compared to 28% in PROLONG-US).

2. Classifying patients according to their mrpGA or mean prior GA has issues particularly with respect to Russia and Ukraine, where standards for determination of gestational age have been inconsistently applied. As we note above, prior births likely involved determination of gestational age using LMP, particularly given the dates of the qualifying births in Russia and Ukraine and the indirect evidence of the use of LMP in various registries and databases in those countries. The uncertainty in the use of LMP in these countries renders the validity of subjects' mrpGA suspect, and indeed it is surprising that there is even a remnant of any trend remaining from this analysis.

3. Neither measure of prior birth history (mean GA or mrpGA) correlate with birth outcome in ex-US PROLONG. In other words, they are not good models for predicting risk for ex-US PROLONG. This may be expected on account of the uncertainties in the prior birth history. Given the overall lack of signal seen in ex-US PROLONG, it is not surprising that subgroup analyses by these prior history measures also showed no signal.

For the same reason, we believe that any subgroup analysis on the ex-US PROLONG subjects using risk factors are unlikely to provide insight because prior birth history – the single most important predictor of preterm birth risk – has not been measured in a comparable fashion to the US population.

Tables 3, 8 and 9 above show that there is a treatment effect from 17-OHPC when assessed based on a more sensitive endpoint such as weeks gained from randomization. In addition, due to capping of the weeks gained, the incremental weeks gained from 17-OHPC treatment focused on <35 weeks for which there is general consensus on benefit for the neonate with respect to morbidity/mortality as shown in Manuck et al. (2016) and Richter et al. (2019). Covis believes that the concept of added weeks of gestation has a clearer clinical interpretation in comparison to a categorical endpoint such as the rate of preterm birth at a given cutoff such as 35 weeks. Further, by picking the cut point at which we cap weeks gained, we ensure that any difference between the treatment arms is focused in a time window that is clinically relevant for neonatal development.

In sum, these analyses give rise to a strong suggestion that 17-OHPC may be effective for the highest-risk patients and highlight the need for further focused studies in this cohort.

II. PROPOSED RANDOMIZED CONTROLLED TRIAL (RCT) OPTIONS

The scientific advisory panel agreed that among different types of studies available to Covis, an RCT would most effectively address confounding variables and represent the most scientifically rigorous method of confirming Makena's benefit.

A. RCT Study Inclusion Criteria And Endpoint Definition

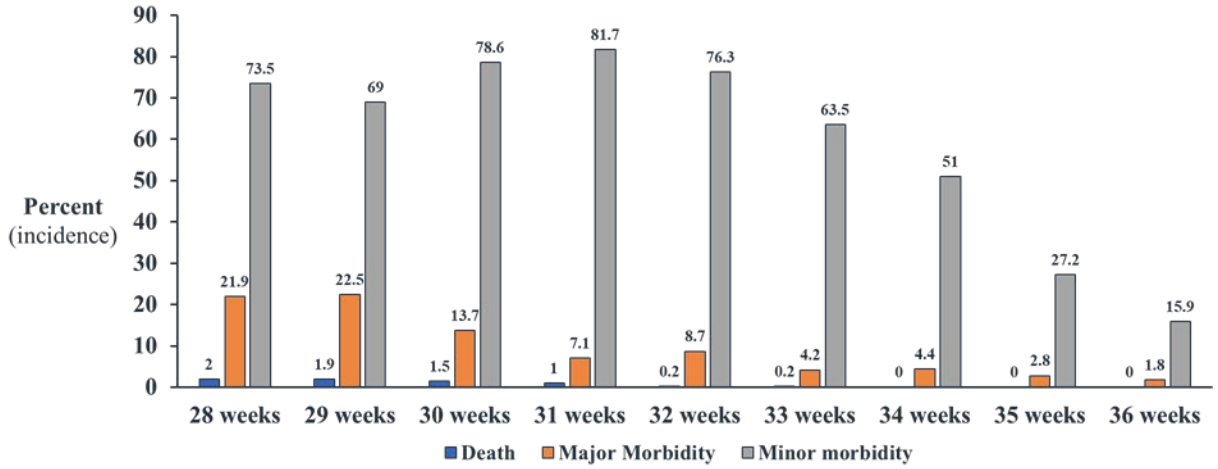
The previous analysis highlighted the importance of selecting patients with a prior SPTB <34 weeks. Covis further proposes to refine the inclusion criterion to ensure consistency of risk in the selected population. These two modifications are: (1) the previous singleton qualifying SPTB <34 weeks occurred within the last 5 years of randomization, and (2) documented medical history of first semester ultrasonography to calculate the gestational age of the qualifying delivery, consistent with ACOG's recommendation that first semester ultrasound measurement is accurate in establishing gestational age, while ultrasound in the second and third trimester are not accurate.[17] The first modification is important given the existence of many patient records in PROLONG where the qualifying pregnancies occurred early on in a patient's life and the same patient had several full-term births preceding the enrollment in the study. Such patients may not benefit from treatment with Makena if they have had several successful term pregnancies in their recent history. The second criterion addresses the need for accurate determination of prior pregnancy history due to the reliance on this parameter as a surrogate of future preterm birth risk.

Covis proposes that the RCT would use an efficacy endpoint of increase in delivery time from randomization for 17-OHPC vs. placebo focused on <35 weeks. CDER has also stated that earlier preterm birth endpoints—in contrast to later preterm birth endpoints such as reduction in

[17] ACOG Committee Opinion, No. 700, *supra* note 409.

preterm birth <37 weeks—would not need confirmatory evidence of clinical benefit.[18] The strength of the <35 week endpoint is based on the extensive literature supporting the association between increase in gestational age at birth and reduction in neonatal morbidity/mortality. For example, Manuck et al.'s analysis of an obstetric cohort of 115,502 women and their neonates published in 2016 demonstrates that incidence rates of death, major neonatal morbidity, and minor neonatal morbidity decline significantly with each advancing week of gestation, from roughly 30 weeks to 36 weeks, as shown below. In particular, the relative reduction in the rates of the morbidities from week 32 to week 36 demonstrates the clinical benefit in extending gestational age of the neonate into week 35 or later.

Figure 15
***Incidence Rates by Gestational Age*[19]**



The original tables from Manuck et al. article showing the incidence of these morbidities and mortalities in aggregate as well as individually are reproduced here:

[18] *Id.* CDER has also described the delay in delivery observed in the Meis trial as an "intermediate clinical endpoint," indicating that the delay in delivery, in itself, provides a therapeutic effect. Expedited Programs Guidance at 19.

[19] *See* Tracy A. Manuck et al., *Preterm Neonatal Morbidity and Mortality by Gestational Age: A Contemporary Cohort*, *215* A. J. OBSTET. GYNECOL. 103.e1–103.e14 (Jul. 2016), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4921282/.

Table 19
Frequency of Death and Major, Intermediate, and Minor Morbidity

Outcome, n (%)	All N=8334	Delivery Gestational Age, Week														
		23 N=43	24 N=114	25 N=124	26 N=169	27 N=159	28 N=196	29 N=213	30 N=262	31 N=312	32 N=451	33 N=639	34 N=1058	35 N=1477	36 N=3117	p-value
Death	119 (1.4)	19 (44.2)	36 (31.6)	15 (12.1)	19 (11.2)	13 (8.2)	4 (2.0)	4 (1.9)	4 (1.5)	3 (1.0)	1 (0.2)	1 (0.2)	0	0	0	<.001
Major morbidity[a]	657 (7.9)	19 (44.2)	60 (52.6)	68 (54.8)	88 (52.1)	64 (40.3)	43 (21.9)	48 (22.5)	36 (13.7)	22 (7.1)	39 (8.7)	27 (4.2)	46 (4.4)	42 (2.8)	55 (1.8)	<.001
Minor morbidity[b]	3136 (37.6)	4 (9.3)	18 (15.8)	39 (31.5)	59 (34.9)	77 (48.4)	144 (73.5)	147 (69.0)	206 (78.6)	255 (81.7)	344 (76.3)	406 (63.5)	540 (51.0)	402 (27.2)	495 (15.9)	<.001
Survival without any of above morbidities	4422 (53.1)	1 (2.3)	0	2 (1.6)	3 (1.8)	5 (2.6)	5 (2.6)	14 (6.6)	16 (6.1)	32 (10.3)	67 (14.9)	205 (32.1)	472 (44.6)	1033 (69.9)	2567 (82.4)	-

a. Includes persistent pulmonary hypertension, intraventricular hemorrhage grade III/IV, seizures, hypoxic-ischemic encephalopathy, necrotizing enterocolitis stage II/IIII, bronchopulmonary dysplasia.
b. Includes intraventricular hemorrhage grade III/IV, necrotizing enterocolitis stage I, respiratory distress syndrome, hyperbilirubinemia requiring treatment, hypotension requiring treatment.

Table 20
Outcomes among Those with Major Morbidity

Outcome, n (%)	Delivery gestational age, week														
	23 N=19	**24 N=60**	**25 N=68**	**26 N=88**	**27 N=64**	**28 N=43**	**29 N=48**	**30 N=36**	**31 N=22**	**32 N=39**	**33 N=27**	**34 N=46**	**35 N=42**	**36 N=55**	**p-value**
Minor Morbidities															
Respiratory distress	15 (79.0)	42 (70.0)	41 (60.3)	49 (55.7)	38 (59.4)	24 (55.8)	26 (54.2)	15 (41.7)	10 (45.5)	17 (43.6)	10 (37.0)	14 (30.4)	17 (40.5)	16 (29.1)	<.001
Hyperbilirubinemia requiring treatment	16 (84.2)	51 (85.0)	57 (83.8)	73 (83.0)	61 (95.3)	35 (81.4)	42 (87.5)	32 (88.9)	20 (90.9)	33 (84.6)	23 (85.2)	34 (73.9)	25 (59.5)	29 (52.7)	<.001
Intraventricular hemorrhage grade I/II	3 (15.8)	14 (23.3)	17 (25.4)	18 (20.7)	14 (21.9)	8 (18.6)	10 (20.8)	5 (14.6)	2 (9.1)	1 (2.6)	4 (14.8)	3 (6.5)	1 (2.4)	3 (3.6)	<.001
Necrotizing enterocolitis stage I	1 (5.6)	3 (5.3)	1 (1.5)	5 (6.0)	4 (6.3)	2 (4.7)	0	0	0	0	0	0	1 (2.5)	0	.003
Hypotension requiring treatment	6 (31.6)	7 (11.7)	5 (7.4)	12 (13.6)	7 (10.9)	3 (7.0)	2 (4.2)	2 (5.6)	1 (4.6)	4 (10.3)	1 (3.7)	2 (4.4)	3 (7.1)	3 (5.5)	.007
Major Morbidities															
Persistent pulmonary hypertension	1 (5.3)	7 (11.7)	4 (5.9)	5 (5.7)	4 (6.3)	0	4 (8.3)	2 (5.6)	1 (4.6)	1 (2.6)	2 (7.4)	3 (6.5)	4 (9.5)	11 (20.0)	.08
Intraventricular hemorrhage grade III/IV	5 (26.3)	17 (28.3)	11 (16.4)	12 (13.8)	9 (14.1)	4 (9.3)	11 (22.9)	7 (20.0)	3 (13.6)	6 (15.8)	2 (7.4)	1 (2.2)	2 (4.8)	1 (1.8)	<.001
Necrotizing enterocolitis stage II/III	3 (16.7)	5 (8.8)	10 (14.9)	9 (10.8)	6 (9.4)	6 (14.0)	5 (10.4)	3 (8.6)	1 (4.6)	1 (2.6)	2 (7.4)	3 (6.5)	1 (2.5)	1 (1.8)	.001
Seizures	2 (10.5)	5 (8.3)	1 (1.5)	2 (2.3)	4 (6.3)	4 (9.3)	2 (4.2)	1 (2.8)	1 (4.6)	2 (5.1)	1 (3.7)	0	2 (4.8)	2 (3.6)	.26
Hypoxic-ischemic encephalopathy	8 (47.1)	15 (29.4)	14 (23.0)	17 (21.3)	19 (31.7)	11 (26.8)	22 (47.8)	17 (47.2)	15 (71.4)	31 (81.6)	21 (80.8)	37 (80.4)	36 (87.8)	48 (87.3)	<.001
Bronchopulmonary dysplasia	18 (94.7)	44 (73.3)	53 (77.9)	66 (75.0)	44 (68.8)	31 (72.1)	14 (29.2)	10 (27.8)	2 (9.1)	3 (7.7)	1 (3.7)	3 (6.5)	0	0	<.001

a. Includes persistent pulmonary hypertension, intraventricular hemorrhage grade III/IV, seizures, hypoxic-ischemic encephalopathy, necrotizing enterocolitis stage II/III, bronchopulmonary dysplasia.

Manuck et al. Preterm neonatal morbidity and mortality. Am J Obstet Gynecol 2016.

Table 21
Outcomes among Those with Minor Morbidity

Outcome, n (%)	Delivery gestational age, week														
	23 N=19	**24 N=60**	**25 N=68**	**26 N=88**	**27 N=64**	**28 N=43**	**29 N=48**	**30 N=36**	**31 N=22**	**32 N=39**	**33 N=27**	**34 N=46**	**35 N=42**	**36 N=55**	**p-value**
Intraventricular hemorrhage grade I/II	1 (25.0)	4 (22.2)	14 (35.9)	12 (20.7)	14 (18.2)	20 (13.9)	23 (15.7)	41 (19.9)	35 (13.7)	31 (9.0)	16 (3.9)	13 (2.4)	4 (1.0)	3 (0.6)	<.001
Necrotizing enterocolitis stage I	0	0	4 (10.8)	4 (6.8)	5 (6.7)	8 (5.6)	4 (2.8)	5 (2.4)	3 (1.2)	6 (1.7)	5 (1.2)	6 (1.1)	1 (0.3)	3 (0.6)	<.001
Respiratory distress syndrome	4 (100.0)	10 (55.6)	21 (53.9)	32 (54.2)	38 (49.4)	72 (50.0)	62 (42.2)	85 (41.3)	76 (29.8)	77 (22.4)	96 (23.7)	119 (22.0)	78 (19.4)	77 (15.6)	<.001
Hyperbilirubinemia requiring treatment	3 (75.0)	14 (77.8)	34 (87.2)	53 (89.8)	72 (93.5)	133 (92.4)	141 (95.9)	194 (94.2)	244 (95.7)	322 (93.6)	377 (92.9)	505 (93.5)	368 (91.5)	436 (88.1)	.12
Hypotension requiring treatment	0	6 (33.3)	4 (10.3)	3 (5.1)	4 (5.2)	1 (0.7)	6 (4.1)	5 (2.4)	9 (3.5)	8 (2.3)	11 (2.7)	15 (2.8)	12 (3.0)	22 (4.4)	.08

a. Includes persistent pulmonary hypertension, intraventricular hemorrhage grade III/IV, seizures, hypoxic-ischemic encephalopathy, necrotizing enterocolitis stage II/III, bronchopulmonary dysplasia.

Manuck et al. Preterm neonatal morbidity and mortality. Am J Obstet Gynecol 2016.

Richter et al., which analyzed 754,763 singleton births in Washington State between 2004 and 2013, also demonstrate a diminishing rate of neonatal morbidities on a week-by-week basis between week 32 and 36:[1]

Table 22
Gestational-Age-Specific Rates of Adverse Neonatal Outcomes Among Singleton Preterm Infants, Washington State, USA (2004-2013)

Outcome and Gestational Age Category, Weeks	Rates Per 100 Live Births N (Rate) 2004 – 2006	Rates Per 100 Live Births N (Rate) 2011 – 2013	Rate Ratio (95% CI)	Adjusted Odds Ratio Per 1-year Change* (95% CI)
Neonatal death				
24 – 27	76 (15.5)	85 (14.2)	0.92 (0.67 to 1.25)	0.97 (0.92 to 1.03)
28 – 31	55 (4.9)	40 (3.0)	0.61 (0.41 to 0.92)	0.95 (0.89 to 1.01)
32 – 33	23 (1.6)	18 (1.0)	0.63 (0.34 to 1.16)	0.93 (0.84 to 1.02)
34 – 36	43 (0.4)	64 (0.5)	1.25 (0.85 to 1.84)	1.06 (1.00 to 1.13)
All (24 – 36)	197 (1.3)	207 (1.3)	1.00 (0.82 to 1.22)	0.99 (0.95 to 1.02)
Neonatal death/severe morbidity				
24 – 27	353 (72.2)	429 (71.7)	0.99 (0.86 to 1.14)	1.00 (0.96 to 1.04)
28 – 31	383 (33.7)	496 (36.6)	1.08 (0.95 to 1.24)	1.03 (1.00 to 1.06)
32 – 33	166 (11.3)	302 (16.3)	1.44 (1.19 to 1.74)	1.05 (1.02 to 1.08)
34 – 36	307 (2.5)	639 (5.4)	2.16 (1.89 to 2.47)	1.10 (1.08 to 1.12)
All (24 – 36)	1209 (7.9)	1866 (11.9)	1.51 (1.40 to 1.62)	1.06 (1.05 to 1.08)

Severe morbidity includes BPD, IVH grade ≥ 3, PVL, ROP, NEC, neonatal sepsis, convulsions of newborn and severe birth trauma.
Adjusted odds ratios express the average variable; adjusted for temporal changes in maternal age, BMI, race, education, smoking, marital status, parity, chronic hypertension, pre-pregnancy, diabetes, assisted conception, health insurance provider, gestational age, SGA infant, sex, and congenital abnormalities.
BMI body mass index; BPD, bronchopulmonary dysplasia; IVH, intraventricular hemorrhage; NEC, necrotizing enterocolitis; PVL, periventricular leukemia; ROP, retinopathy of prematurity; SGA, small-for-gestational age.

Table 23
Gestational-Age-Specific Rates of Neonatal Death by Subtype of Preterm Birth, Washington State, USA (2004-2013)

Gestational Age Category and Preterm Birth Subtype	Neonatal Death N (Per 100 Live Births) 2004 – 2006	Neonatal Death N (Per 100 Live Births) 2011 – 2013	Rate Ratio (95% CI)	Adjusted Odds Ratio Per 1-year Change* (95% CI)
24 – 27 weeks				
PPROM	27 (14.9)	18 (13.9)	0.92 (0.51 to 1.69)	1.05 (0.94 to 1.17)
Spontaneous labor	21 (15.6)	26 (13.8)	0.88 (0.50 to 1.57)	0.95 (0.86 to 1.06)
Clinician-initiated	28 (16.2)	41 (14.7)	0.91 (0.56 to 1.47)	0.94 (0.86 to 1.03)
28 – 31 weeks				
PPROM	14 (4.8)	5 (1.9)	0.40 (0.14 to 1.10)	0.92 (0.78 to 1.07)
Spontaneous labor	11 (3.1)	9 (2.2)	0.71 (0.30 to 1.71)	0.91 (0.77 to 1.06)
Clinician-initiated	30 (6.2)	26 (3.8)	0.61 (0.36 to 1.04)	0.96 (0.88 to 1.06)

[1] Lindsay A. Richter et al., *Temporal Trends in Neonatal Mortality and Morbidity Following Spontaneous and Clinician-Initiated Preterm Birth in Washington State, USA: A Population-Based Study*, 9 BMJ OPEN e023004 (2019).

Gestational Age Category and Preterm Birth Subtype	Neonatal Death			Adjusted Odds Ratio Per 1-year Change* (95% CI)
	N (Per 100 Live Births)		Rate Ratio (95% CI)	
	2004 – 2006	2011 – 2013		
23 – 33 weeks				
PPROM	2 (0.5)	5 (1.2)	2.40 (0.47 to 12.37)	1.08 (0.80 to 1.45)
Spontaneous labor	7 (1.3)	5 (0.8)	0.62 (0.20 to 1.94)	0.97 (0.83 to 1.13)
Clinician-initiated	14 (2.5)	8 (1.0)	0.40 (0.17 to 0.95)	0.85 (0.74 to 0.97)
34 – 36 weeks				
PPROM	14 (0.7)	7 (0.4)	0.57 (0.32 to 1.42)	0.97 (0.84 to 1.12)
Spontaneous labor	7 (0.1)	22 (0.4)	4.00 (1.71 to 9.36)	1.08 (0.96 to 1.20)
Clinician-initiated	22 (0.5)	35 (0.8)	1.60 (0.94 to 2.73)	1.10 (1.01 to 1.20)
All (24 – 36 weeks)				
PPROM	57 (2.1)	35 (1.4)	0.67 (0.44 to 1.02)	1.00 (0.93 to 1.07)
Spontaneous labor	46 (0.7)	62 (0.9)	1.29 (0.88 to 1.88)	0.98 (0.92 to 1.04)
Clinician-initiated	94 (1.6)	110 (1.7)	1.06 (0.81 to 1.40)	0.98 (0.94 to 1.03)

Adjusted odds ratios express the average annual change in the odds of neonatal death.
* Calendar year was modelled as a continuous variable; adjusted for temporal changes in maternal age, BMI, race, education, smoking, marital status, parity, chronic hypertension, pre-pregnancy diabetes, assisted conception, health insurance provider, gestational age, SGA infant, sex, and congenital anomalies.
PPROM, preterm premature rupture of membranes.

As will be seen below, Covis proposes designing the gestational age endpoint to focus on the benefit of extending gestation through this critical period for which the morbidity/mortality benefits of extended time in utero have been established. Further support for the association between gestational age and morbidity/mortality in the 32-35 week period can also be found in other recent analyses including a large meta-analysis.[2]

For these reasons, Covis proposes that a further RCT will have either of two primary endpoints: (1) a measure of the improvement of gestational age in a time window that is strongly correlated with morbidity/mortality benefits for the neonate, with a preference for weeks gained from 17-OHPC treatment focused on <35 weeks or (2) a study primarily focusing on a neonatal morbidity/mortality endpoint. Notably, CDER has described reduction in preterm birth <32 or <35 weeks as a "well-established surrogate" or "established surrogate."[3] Accelerated approval based on an intermediate clinical endpoint means that FDA has recognized that the product does have a demonstrated therapeutic effect.

In case the recommendation is to conduct an outcome study on neonatal morbidity and mortality, Covis proposes to revisit the previously utilized definition of the morbidity/mortality endpoint with the Agency. For example, the Manuck article describes both major and minor morbidities, which are all associated with significant healthcare burden to the patient as well as the system. Covis proposes to have an in-depth discussion with the Agency on the selection of an

[2] *See* Margreet Teune et al., *A Systematic Review of Severe Morbidity in Infants Born Late Preterm*, 205 AM. J. OBSTET. GYNECOL. 374.e1 (2011), https://pubmed.ncbi.nlm.nih.gov/21864824/.

[3] *See* NDA 21945 Clinical Review, at 15, https://www.fda.gov/media/80892/download; S*ee* FDA Briefing Document, at 20, https://www.fda.gov/media/132003/download ("FDA determined that further study was needed to provide confirmatory evidence of the drug's efficacy in terms of direct clinical benefit on neonatal outcomes or through an established surrogate such as the rate of preterm birth prior to 35 and 32 weeks gestation").

appropriate subset of clinically relevant morbidities to include that allow for a clinical study that may be enrolled in a realistic timeframe (based on incidence of events, an assumed effect size and the implied sample size).

Table 16
Classification of Morbidities

Meis/PROLONG	Manuck et al. (2016)	
	Major Morbidities	**Minor Morbidities**
Neonatal death Grade 3 or 4 intraventricular hemorrhage Respiratory distress syndrome Bronchopulmonary dysplasia Necrotizing enterocolitis Sepsis	Persistent pulmonary hypertension Grade 3 or 4 intraventricular hemorrhage Seizures Hypoxic-ischemic encephalopathy Necrotizing enterocolitis stage II/III Bronchopulmonary dysplasia	Grade 1 or 2 intraventricular hemorrhage Necrotizing enterocolitis stage I Respiratory distress syndrome Hyperbilirubinemia requiring treatment Hypotension requiring treatment

Finally, in the case of the gestational age-based primary endpoints we would be looking for a trend that supports improvement in morbidity and mortality outcomes.

B. Proposed RCT Designs With Statistical Analysis Methodologies, Estimated Sample Size, And Time Frame For Completion

In the table below, we identify potential primary endpoints of a future RCT and proposed statistical analysis methodology, estimated sample size requirements and time frame.

Covis is confident that it is feasible to enroll patients at high risk of preterm birth, as defined in terms of the new inclusion criteria (i.e., <34 weeks prior spontaneous preterm birth with a validated recent prior birth gestational ages). Given the previously described differences in preterm birth by race, Covis believes it will be important to recruit and enrich for high-risk Black women in this study. Our preliminary assumption is that we may seek to enroll, depending on the endpoint used, about a third of the total subjects from the Black population in order to allow for a reasonable sample size and with stratification assuring balance for unconfounded exploration of the effects in Black subjects. The proposed study designs are based on the totality of understanding regarding the association of neonatal morbidity/mortality by gestational age and would provide confirmation of Makena's clinical benefit in women at high risk for a future preterm birth. Enrollment will turn on the number of sites and enrollment rate at each individual site, which is the subject of ongoing feasibility studies. A sample of the results from the survey are described below. This enrollment estimate is based on several independent workstreams:

1) A formal RCT feasibility assessment that has identified 19 potential non-academic sites in the US and an (at this point) additional 10 sites in European (e.g., France, Germany, Italy, Spain, UK) countries where a similar high-risk subpopulation may be available to recruit. The ex-US feasibility assessment is still ongoing and may identify additional sites in the near term. The estimate is that these sites may be able to contribute 0.5-1 patient with prior spontaneous preterm birth <34 weeks per month per site, therefore accounting for

approximately 100 patients per year assuming some reduction in the number of actual sites that participate after they review the full protocol.

2) A survey conducted within the Dorsata practice network indicating a willingness to participate in a placebo-controlled RCT with 17-OHPC. In the consenting practices, there are about 1200 patients per year who have pregnancies and who have had a prior spontaneous preterm birth <34 weeks. We may estimate between 5-15% of these patients may consent to enroll and screen successfully into the study, suggesting the addition of another 60-180 patients per year from this network.

3) Outreach has been made to prospective investigators at 12 academic MFMU network sites that participated in Meis. Prospective investigators at these sites express support for site participation in a new RCT of 17-OHPC. Of 12 investigators, 11 indicated interest in participating in a new trial, and one suggested that a new trial is not warranted because Meis affirmatively settled the question of 17-OHPC's efficacy.[4] This survey also provided some additional insights on the views of these prospective investigators on key features of the protocol. In particular, they overwhelmingly felt that any protocol today would need to include some aspect of rescue for patients presenting with shortening cervix post-randomization, e.g., the option to perform a cerclage on eligible patients. When queried about their views on the endpoints, half the respondents indicated interest in the type of "weeks gained" analysis that we have presented in this supplement. The details of the survey are in Attachment D.

4) In addition to the above, Covis conducted a systematic survey in 400 providers and a similar number of pregnant patients with prior preterm birth history to understand attitudes for enrolling in a placebo-controlled RCT. Covis subsequently conducted a follow-up survey, which included responses from 150 new providers. These surveys, described below in Attachment A and Attachment C to the Appendix, confirmed the willingness of physicians to contribute patients to the RCT and for patients to consent to such a study at significant rates. Importantly we also probed the recruiting effect having an approved product in the RCT has relative to an unapproved product. The willingness of both patients and providers to participate in a study with an approved product was higher than for an unapproved product. Importantly, providers willingness to participate in a study dramatically reduces when a product's approval is withdrawn (which situation we will be in if Makena's Accelerated Approval were withdrawn).

[4] Sean C. Blackwell et al., *17-OHPC to Prevent Recurrent Preterm Birth in Singleton Gestations* (PROLONG Study): A Multicenter, International, Randomized Double-Blind Trial, 37 AM. J. PERINATOL 127-36 (Oct. 2019), https://www.thieme-connect.com/products/ejournals/pdf/10.1055/s-0039-3400227.pdf

Please see Attachment B to the Appendix that shows the details underlying the assumptions including the assumed effect sizes for each option (treatment effect or hazard ratio).

Table 17
Proposed RCT Studies

Potential study	Rationale	Estimated sample size
Impact of treatment with Makena vs. placebo on time from randomization to delivery (weeks)	Established marked neonatal benefit on measures of morbidity/mortality with increase in weeks of gestation	▪ 1:1 randomization – 382 subjects ▪ 2:1 randomization – 429 subjects
Impact of treatment with Makena vs. placebo in achieving full term birth (comparison of percentage of patients achieving 37 weeks)	Moving a patient from a baseline of <34 weeks to full term has clear benefit to the neonate based on the reduction in morbidity event rates	▪ 2:1 randomization – 1152 subjects ▪ 3:1 randomization – 1364 subjects
Impact of treatment with Makena vs. placebo on preterm birth rate < 35 weeks (comparison of percentage of patients)	Established marked increase in morbidity/mortality in infants born prior to week 35	▪ 2:1 randomization – 1458 subjects ▪ 3:1 randomization – 1728 subjects
Impact of treatment with Makena vs. placebo on preterm birth rate < 35 weeks (measured as time-to-event censored at 35 weeks)	Established marked increase in morbidity/mortality in infants born prior to week 35	▪ 2:1 randomization – 1860 subjects ▪ 3:1 randomization – 2280 subjects
Impact of treatment with Makena vs. placebo on preterm birth rate <35 weeks (measured as weeks of gestational age less than 35 weeks)[5]	Established marked increase in morbidity/mortality in infants born prior to week 35 Method weighs delay of very early preterm birth more than mid/late preterm birth	▪ 2:1 randomization – 1413 subjects ▪ 3:1 randomization – 1660 subjects
Impact of treatment with Makena vs. placebo on a composite neonatal morbidity/mortality index	The definition of composite neonatal index subject to discussion with CDER, including potential expansion to include "minor" morbidities	Based on 15%-20% incidence rate of neonatal events ▪ 2:1 randomization – 2676 subjects ▪ 3:1 randomization – 3172 subjects

[5] A patient who reaches at least 34 weeks of gestational age would receive a score of 0, while a patient who gives birth at an earlier gestational age than 34 weeks would be scored based on the difference in weeks (as a continuous variable). For example, a patient who gives birth at 33 weeks would receive a score of 1 (week) or 7 (days).

Please note that the first several studies all have identical inclusion criterion, and all feature a gestational age/Preterm Birth endpoint. The sole difference between the options – which leads to the difference in sample size estimates is the analysis methodology. These, in concept range from the idea of taking patients with prior early SPTB history and assessing weeks of gestational age gained <35 weeks, to gestational age of prolonging a pregnancy beyond 35 weeks, the difference between the arms of getting patients to full term, the rate of events < 35 weeks, and a time-to-event-based methodology. Based on the analyses described in Sections I.D of this Appendix, Covis proposes conducting a new study that has a single primary endpoint of increase in delivery time from randomization for 17-OHPC vs. placebo focused on <35 weeks. This endpoint has several features that make it the preferred endpoint:

- It is more sensitive than a categorical endpoint such as PTB rate <35, allowing for the ability to test the hypothesis with a smaller sample size – this is critical given the historical challenges with recruiting subjects into a placebo-controlled study for this indication.

- It provides a clearer understanding of where the benefit to the neonate is accruing, i.e., the endpoint measures the difference in gestational age up to <35 weeks.

- The capping mechanism to 35 weeks ensures that any increase in GA is associated with clinical benefit to the neonate in a manner that can be quantified based on known incidence of morbidity/mortality by gestational age.

- Adding the MFMU network sites to the already identified sites (non-academic and Dorsata) we believe that the recruitment of ~400 patients is highly feasible and should be possible to complete within a 3-6 year timeframe based on previous calculations. Given CDER's concerns about keeping the product on the market for an additional 8-10 years while a confirmatory study is conducted, this endpoint has a high likelihood of delivering a clear view on the efficacy of 17-OHPC in a relatively short duration.

This option is designed such that the analysis methodology does not give either arm any credit for births after week 35, counting them all the same. Therefore, any differences in the treatment arms would require an improvement of preterm birth outcomes at gestational ages prior to week 35. The numerical difference between arms can then be interpreted directly in terms of the likely benefit in morbidity/mortality of the neonate. Covis welcomes an opportunity to discuss these analytical approaches with the Agency as they present potentially distinct insights on the therapeutic benefit of the intervention.

III. PROPOSED OBSERVATIONAL STUDY

Since becoming Makena's sponsor in 2021, Covis has been working with experts to evaluate several databases and registries to investigate the feasibility of conducting well-designed observational studies that could provide meaningful data on the clinical benefit of Makena. An observational study would not be a substitute for an RCT but would, instead, provide additional data on the drug-induced benefit of Makena in high-risk patients, possible interactions with known risk factors for preterm birth, and further insight into the relationship

between gestational age and neonatal mortality and morbidity. Through its diligence, Covis has identified a partner, Dorsata, whose platform Covis believes can achieve all of these goals.[6] The principal elements of Dorsata's offering have been summarized above.

Covis' scientific advisory panel agreed that an observational study utilizing data from the Dorsata database could effectively supplement data from an RCT while simultaneously offering the advantage of a data readout in a shorter time frame than could be expected with an RCT. Key considerations were the need to define eligibility criteria that reflected the highest risk patient populations suggested by risk modeling data. Some of the analysis methodologies involve looking at prior preterm birth history to determine a baseline value from which to measure improvement. In crafting such a definition some care will need to be taken to understand how to analyze the impact of treatment relative to previous baseline if the patient were treated previously with 17-OHPC. In keeping with the eligibility criteria for an RCT, and to strengthen the likelihood that gestational age was appropriately captured for qualifying prior preterm births, only patients with at least one prior preterm birth within the five years immediately prior to the pregnancy under investigation would be included.

Potentially useful study designs include either a retrospective or prospective cohort study of patients stratified by history of prior preterm birth. The effect of Makena treatment compared to no treatment would be compared in two patient groups—patients with a prior spontaneous preterm birth at <34 weeks gestation and patients with a prior spontaneous preterm birth at >34 weeks gestation. The risk of confounding posed by observational study designs would also be addressed (e.g., by propensity score matching across the treatment groups). This would be done using variables that correlate to well-defined risks factors for spontaneous preterm birth, including number of prior spontaneous preterm births, specific comorbidities, socioeconomic status (defined for example by whether a patient is on Medicaid), and race. An evaluation identifying favorable trends associated with Makena treatment could potentially be conducted within 12-24 months. This approach would provide further data on the potential clinical benefit of Makena enhanced by an evaluation of differential benefit based on pre-existing risk. Further, the Dorsata dataset may also provide complementary data on the well-documented association between gestational age and the risk of neonatal morbidity and mortality. A prospective study would be enabled by the ability of the Dorsata system to capture detailed medical history information for all 17-OHPC-eligible patients and monitor these patients over the course of the preterm observational period (including tracking their medications). Covis would also plan to collect data on neonatal outcomes for all patients in any prospective observational study. Consequently, the same level of detailed information as might be provided by an RCT would be available for a prospectively conducted study that utilizes the Dorsata system for both treated and untreated patients.

* * * * *

In sum, since becoming Makena's sponsor in 2021, Covis has engaged with numerous scientific and medical experts to evaluate RCT and observational study options that could further

[6] *See generally* Dorsata, https://dorsata.com (last visited Sep. 14, 2022).

confirm the clinical benefit of Makena. Covis welcomes the opportunity to further engage with the Agency on a data-driven path forward for this important study.

ATTACHMENT A TO THE APPENDIX

PROVIDER AND PATIENT PERCEPTIONS OF CLINICAL TRIAL FEASIBILITY SURVEY METHODOLOGY AND DATA INTERPRETATION

RESEARCH OBJECTIVES

Covis Pharma commissioned a survey to better understand patient and provider perceptions of placebo-controlled clinical studies among pregnant women, and to support decision-making related to a data-driven path forward. The survey was designed to better understand:

- General awareness and perceptions of clinical trials;
- Perceptions of clinical trials specific to pregnancy and preterm birth;
- Attitudinal barriers and motivators, including factors influencing decision-making and hesitancy;
- Information needs and trusted sources; and
- Willingness to recommend or participate in clinical trials.
- Survey responders were not asked specifically about Makena or 17-OHPC.

Two surveys were conducted among two separate and discrete survey populations: 1) patients defined as women between the age of 18-45 with a history of singleton spontaneous preterm birth (SPTB); and 2) health care providers (OB/GYNs, family and general practitioners, and maternal-fetal medicine specialists) who report routinely treating patients at-risk for SPTB. The methodology for each survey is described in the table below.

Table 24
Research Methodology

	PATIENTS	**HEALTHCARE PROVIDERS**
Sample Universe/ Frame	Given the very low incidence of the survey population, a non-probability sample methodology was required. A large, global consumer research panel, Dynata, supplied a targeted sample frame of online research panelists profiled as women 18-45 who are currently or have been pregnant with a history of singleton spontaneous preterm birth ("SPTB").	Given the low incidence of the survey population, a non-probability sample methodology was required. The largest market research panel of health care providers, Survey Healthcare Global (SHG), supplied a targeted sample frame of OB/GYNs, general/family practice physicians, and maternal-fetal medicine specialists.
Sample Design/ Stratific-ation	Sample was pre-stratified to target only panelists known to be women 18-45 who are currently or were previously pregnant. **The total sample size (completed interviews with eligible participants meeting definition of survey population) is n=325; n=31 currently pregnant and at-risk for SPTB; n=294 previous SPTB.**	Sample was pre-stratified to target specific HCPs (specific groups sampled proportional to population on AMA data). **The total sample size (completed interviews with eligible HCPs screened for routinely treating patients at-risk for SPTB) is n=400; n=357 OB/GYNs; n=44 Maternal-Fetal Medicine Specialists.**

Data Collection	Dynata's router protocol applies randomization algorithms and targeting of existing and newly recruited panelists. These panelists were routed to the online survey. The online survey took an average of 18 minutes to complete. Participants receive credits toward incentives they receive for participation across multiple surveys. No personally identifiable data were asked in the survey instrument nor were routed to Covis through Dynata. **The survey was conducted between June 15th and July 6th, 2022**.	SHG targeted panelists were selected at random (PTP allocation) and sent to an online survey. SHG validates panelists using AMA and other provider directories to maintain sample integrity. The online survey took an average of 15.5 minutes to complete. Respondents received an incentive of $80 for completing the online survey. **The survey was conducted between June 15th and July 6th, 2022**.
Analysis	Data collected online were imported into SPSS for data coding and analysis. Given the niche demographic composition of the survey population, the data were not post-stratified or weighted. Although any errors introduced due to sampling cannot be accounted, all statistical testing (2-sided t-tests for subgroup analysis) assumes random distribution of the population.	Data collected online were imported into SPSS for data coding and analysis. The data were not stratified or weighted post-data collection since the population parameters represented those who routinely treat SPTB patients. Although any errors introduced due to sampling cannot be accounted for, all statistical testing (2-sided t-tests for subgroup analysis) assumes a random distribution of the population.

DATA INTERPRETATION

Physicians who treat patients at risk for spontaneous preterm birth are willing to recommend their patients participate in clinical trials, but also place a great deal of importance on therapies already approved by the U.S. Food & Drug Administration (FDA).

A large majority of physicians surveyed (78%) say they are likely to recommend a pregnant patient enroll in a placebo-controlled study comparing the efficacy of a product vs. placebo *only when FDA has approved the product*. Furthermore, an even larger majority of physicians (88%) say that it is important for treatment options to be approved by FDA before recommending them to their pregnant patients.

Among patients at-risk for SPTB, almost all (95%) say it is important that treatment options to reduce the risk of another preterm birth be approved by FDA. In addition, these patients are more likely to take a prescription drug during pregnancy that is intended to treat preterm birth and is being studied by researchers when it is already approved by FDA, compared to a drug that FDA has not approved (68% vs. 37%).

Given their history of spontaneous preterm birth, most patients (68%) say they are likely to participate in a clinical trial while pregnant if it was designed to study treatment options to reduce the risk of preterm birth. Physicians play an essential role in recommending clinical trials to patients. More than half (60%) of the patients surveyed agree that they are open to participating in a clinical trial only if their provider recommends it.

ATTACHMENT B TO THE APPENDIX

COVIS SAMPLE SIZE RESULTS

General specifications:

1. Superiority study comparing 17-OHPC to placebo
2. Allocation: 2:1 or 3:1 (17-OHPC : placebo)
3. One-sided alpha = 0.025
4. Power = 90%

Part 1: Endpoint: Gestational Age (GA) ≥ 37 weeks
Sample size calculations were performed based on the general specifications plus the following specifications:

1. True Placebo percentage = 70%
2. True 17-OHPC percentage = 79% (30% reduction in percentage with GA < 37 weeks)
3. Use of Fisher's Exact Test

The results of the sample calculations are presented in Table 25.

Table 25
Required Sample Sizes for GA ≥ 37 weeks

Allocation (17-OHPC:Control)	Sample Size		
	17-OHPC	Placebo	Total
2:1	768	384	1152
3:1	1023	341	1364

Thus, for example, for a 2:1 allocation, the required total sample size for this endpoint is 1152 subjects (768 17-OHPC subjects and 384 placebo subjects).

Part 2: Endpoint: GA < 35 weeks
Sample size calculations were performed based on the general specifications plus the following specifications:

1. True Placebo percentage = 25%
2. Percentage reduction with 17-OHPC = 30%
3. Use of Fisher's Exact Test

The results of the sample calculations are presented in Table 26.

Table 26
Required Sample Sizes for GA < 35 weeks

Allocation (17-OHPC:Control)	Sample Size		
	17-OHPC	Placebo	Total
2:1	972	486	1458
3:1	1296	432	1728

Thus, for example, for a 2:1 allocation, the required total sample size for this endpoint is 1458 subjects (972 17-OHPC subjects and 486 placebo subjects).

Part 3: Endpoint: Time to birth up to 35 weeks
Sample size calculations were performed based on the general specifications plus the following specifications:

1. Cap births after 35 weeks at 35 weeks
2. 30% reduction in hazard rate of Placebo with use of 17-OHPC
3. Calculate required number of events (a birth prior to 35 weeks)
4. Use assumed rates of 25% and 17.5% for Placebo and 17-OHPC, respectively, to convert number of events to number of patients

The results of the sample calculations are presented in Table 27.

Table 27
Required Sample Sizes for Time to Birth up to 35 Weeks

Allocation (17-OHPC:Control)	Sample Size		
	17-OHPC	Placebo	Total
2:1	1240	620	1860
3:1	1710	570	2280

Thus, for example, for a 2:1 allocation, the required total sample size for this endpoint is 1860 subjects (1240 17-OHPC subjects and 620 placebo subjects).

Part 4: Endpoint: Number of weeks of GA less than 35 weeks
Sample size calculations were performed based on the general specifications plus the following specifications:

1. If GA ≥ 35 weeks, value is 0. If GA < 35 weeks, value is 35 – GA.
2. Assume true Placebo percentage < 35 weeks = 25%
3. Assume distribution < 35 weeks is proportional to distribution in Meis for population with qualifying visit <34 weeks
4. Assume 30% reduction in each category <35 weeks with 17-OHPC
5. Calculate percentages in each category of the endpoint for placebo and for 17-OHPC
6. Use of Wilcoxon rank sum test

The results of the sample calculations are presented in Table 28.

Table 28
Required Sample Sizes for Number of Weeks of GA Less than 35 Weeks

Allocation (17-OHPC:Control)	Sample Size		
	17-OHPC	Placebo	Total
2:1	942	471	1413
3:1	1248	416	1664

Thus, for example, for a 2:1 allocation, the required total sample size for this endpoint is 1413 subjects (942 17-OHPC subjects and 471 placebo subjects).

Part 5: Endpoint: Expanded neonatal morbidity/mortality composite
Sample size calculations were performed based on the general specifications plus the following specifications:

1. Assume true Placebo percentage = 15%
2. Assume true 17-OHPC reduction = 30%
3. Use of Fisher's exact test

The results of the sample calculations are presented in Table 29.

Table 29
Required Sample Sizes for Expanded Neonatal Morbidity/Mortality Composite

Allocation (17-OHPC:Control)	Sample Size		
	17-OHPC	Placebo	Total
2:1	1784	892	2676
3:1	2379	793	3172

Thus, for example, for a 2:1 allocation, the required total sample size for this endpoint is 2676 subjects (1784 17-OHPC subjects and 892 placebo subjects).

Part 6: Endpoint: Time from randomization to delivery (weeks), <35 weeks
Sample size calculations were performed based on the general specifications plus the following specifications:

1. Two-sample t-test
2. Two-sided Alpha = 0.5
3. Difference in Means (17-OHPC - Placebo): 1.0 or 2.0
4. Common SD: 3.0
5. Power = 90%

The results of the sample calculations are presented in Table 30.

Table 30
Required Sample Sizes for Time from randomization to delivery (weeks), <35 weeks

Difference in Means	Allocation (17-OHPC:Control)	Sample Size		
		17-OHPC	Placebo	Total
1.0	1:1	191	191	382
	2:1	286	143	429
2.0	1:1	49	49	98
	2:1	74	37	111

Thus, for example, for a difference in means of 2.0 and a 2:1 allocation, the required total sample size is 111 subjects (74 17-OHPC subjects and 37 Placebo subjects).

From this table, if we conservatively anticipate a 1 week difference between 17-OHPC and placebo, and we maintain the 2:1 randomization of the Meis and PROLONG studies, a sample size of 429 subjects appears sufficient to test the hypothesis with a two-sided alpha = 0.05 and 90% power.

ATTACHMENT C TO THE APPENDIX

PROVIDER PERCEPTIONS OF THERAPEUTIC PRACTICES TO TREAT PATIENTS AT RISK FOR SPONTANEOUS PRETERM BIRTH SURVEY METHODOLOGY AND DATA INTERPRETATION

RESEARCH OBJECTIVES

Covis Pharma commissioned a follow-up survey to better understand provider perceptions of therapeutic practices related to the treatment of pregnant patients at risk for spontaneous preterm birth (SPTB), and to support decision-making related to a data-driven path forward. The survey was designed to better understand:

- The diagnosis and care of patients at risk for singleton spontaneous preterm birth
- Primary risk factors and treatment options offered to patients specific to spontaneous preterm birth; and
- Willingness to recommend participation in clinical trials.

The follow-up survey was conducted among health care providers (OB/GYNs, maternal-fetal medicine specialists) who report routinely treating patients at-risk for SPTB. The methodology for the survey is described in the table below.

Table 31
Research Methodology

	HEALTHCARE PROVIDERS
Sample Universe/ Frame	Given the low incidence of the survey population, a non-probability sample methodology was required. The largest market research panel of health care providers, Survey Healthcare Global (SHG), supplied a targeted sample frame of OB/GYNs, and maternal-fetal medicine specialists.
Sample Design/ Stratification	Sample was pre-stratified to target specific HCPs (specific groups sampled proportional to population on AMA data). **The total sample size (completed interviews with eligible HCPs screened for routinely treating patients at-risk for SPTB) is n=322; n=286 OB/GYNs; n=19 Maternal-Fetal Medicine Specialists; n = 17 OB/GYNs + Maternal-Fetal Medicine Specialists. Of the n = 322; n= 172 of the providers participated in the first survey on perceptions of clinical trial feasibility (See Appendix A) and consented to be recontacted to take the follow-up survey. An additional n=150 were new respondents who participated in the survey.**
Data Collection	SHG targeted panelists were selected at random (PTP allocation) and sent to an online survey. SHG validates panelists using AMA and other provider directories to maintain sample integrity. The online survey took an average of 6.5 minutes to complete. Respondents received an incentive of $80 for completing the online survey. **The survey was conducted between August 27th and September 10th, 2022**.

Analysis	Data collected online were imported into SPSS for data coding and analysis. The data were not stratified or weighted post-data collection since the population parameters represented those who routinely treat SPTB patients. Although any errors introduced due to sampling cannot be accounted for, all statistical testing (2-sided t-tests for subgroup analysis) assumes a random distribution of the population.

DATA INTERPRETATION

When physicians who treat patients at risk for spontaneous preterm birth (SPTB) were asked about how they treated their most recent patient at risk for singleton SPTB, the vast majority of these physicians report recommending progesterone medication (84%), significantly more often than other treatment methods (28% bed rest: 34% surgical procedure like cerclage). However, less than half report recommending progesterone by injection (40%), while a similar proportion are recommending vaginal progesterone (39%). Progesterone by injection is the most commonly recommended treatment when prior SPTB is the primary risk factor (46%).

When asked about treatment options across all their patients at risk for SPTB over the past year, physicians report prescribing – on average – progesterone medication administered by injection with 35% of their patients with a history of prior SPTB. Physicians prescribe vaginal progesterone (on average) 32% of the time, cerclage 15% of the time, and bed rest 20% of the time (among patients at risk due to previous SPTB).

Of the patients that physicians define as those at highest risk for SPTB, progesterone administered by injection is the most commonly prescribed treatment. Still, injectable progesterone is prescribed only 38% of the time for those patients that doctors deem at highest risk. In comparison, vaginal progesterone is prescribed 33% of the time.

More than a third (36%) of doctors say they are recommending progesterone by injection less often than they did three years ago, despite being the most prescribed therapy for highest risk patients. Nearly half (47%) say they are recommending vaginal progesterone more often than they did three years ago. Most doctors do not report a change over the past three years to other therapies (i.e., bed rest, cerclage, oral progesterone).

A large majority of physicians (80%) say they are likely to recommend a pregnant patient enroll in a placebo-controlled study when the product is FDA approved. The reported likelihood drops by more than half (39%) if the product has not been approved by the FDA and even lower (15%) if the product has had its marketing approval withdrawn.

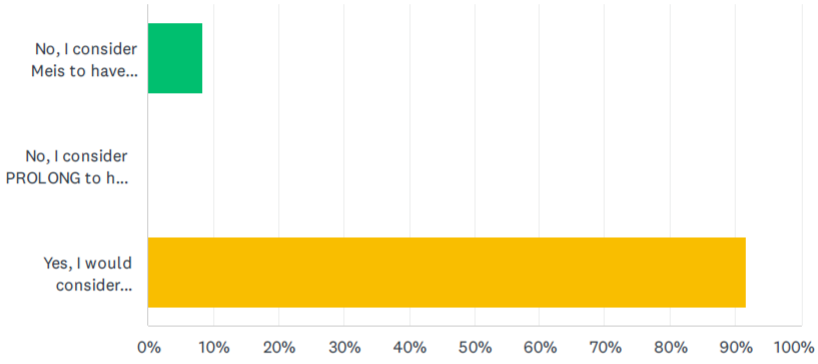
ATTACHMENT D TO THE APPENDIX

SURVEY OF MFMU NETWORK SITES

Dr. Sean Blackwell, an investigator of the PROLONG study and author of Blackwell et al., "*17-OHPC to Prevent Recurrent Preterm Birth in Singleton Gestations (PROLONG Study): A Multicenter, International, Randomized Double-Blind Trial*," 37 AM. J. PERINATOL. 127-36 (Oct. 2019), conducted a survey with his colleagues from 12 MFMU network trial sites that had participated in the Meis trial. These individuals represent prospective trial investigators from these institutions, and their positions reflected in this survey are expected to be representative of their institutions.

Q1 What is your level of interest in participating in another 17-OHPC trial? It would be only in the US, placebo-controlled, and involve women with singleton pregnancy and prior SPTB.

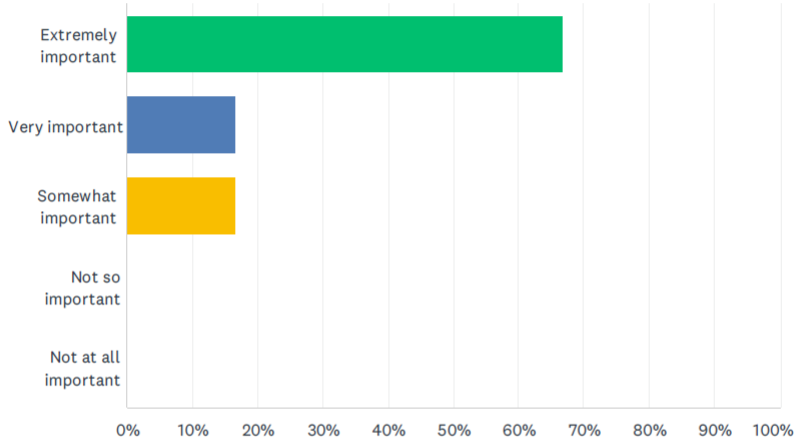
Answered: 12 Skipped: 0



ANSWER CHOICES	RESPONSES	
No, I consider Meis to have settled the matter of efficacy (in the affirmative) and I don't want to participate in another clinical trial.	8.33%	1
No, I consider PROLONG to have settled the matter of efficacy (in the negative) and I don't want to participate in another clinical trial.	0.00%	0
Yes, I would consider participating in another RCT with 17-OHPC in order to more clearly establish its role in women with prior SPTB	91.67%	11
TOTAL		12

Q2 If another RCT was conducted in women with a prior SPTB (17-OHPC vs placebo), in your opinion, how important is the following study design issue?After randomization, a short cervix developed (transvaginal ultrasound <= 25 mm) and the protocol allows for cerclage placement.

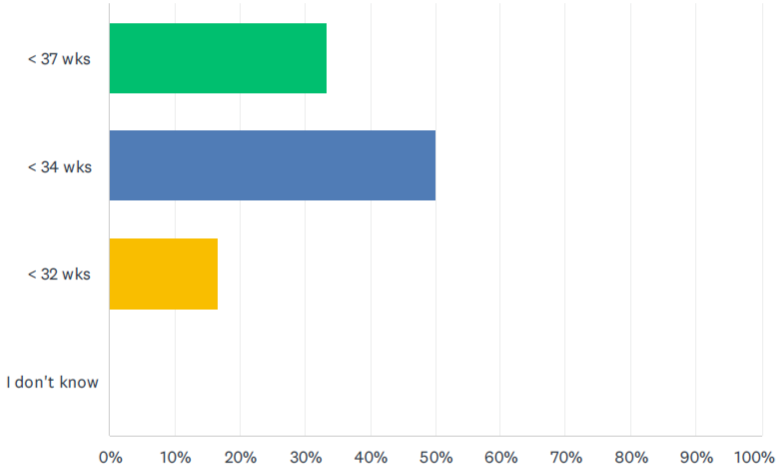
Answered: 12 Skipped: 0



ANSWER CHOICES	RESPONSES	
Extremely important	66.67%	8
Very important	16.67%	2
Somewhat important	16.67%	2
Not so important	0.00%	0
Not at all important	0.00%	0
TOTAL		12

Q3 In order to increase the "risk profile" of women eligible for the RCT, having a lower gestational age threshold for a qualifying SPTB has been discussed. This may identify women more likely to respond to 17OHPC. In both Meis and PROLONG, women qualified after a prior SPTB < 37 wks. What is your opinion on the best GA (wks) entry threshold?

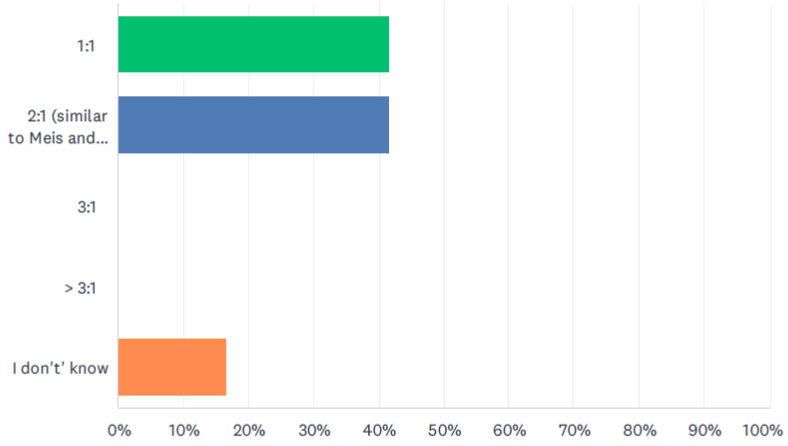
Answered: 12 Skipped: 0



ANSWER CHOICES	RESPONSES	
< 37 wks	33.33%	4
< 34 wks	50.00%	6
< 32 wks	16.67%	2
I don't know	0.00%	0
TOTAL		12

Q4 Other study design changes have been suggested to improve the RCT and increase the chance for more definitive conclusion of whether 17 OHPC is efficacious or not.In both Meis and PROLONG, the randomization ratio were 2:1 (2 with 17-OHPC to 1 placebo).What is your opinion on this issue? The best ratio is:

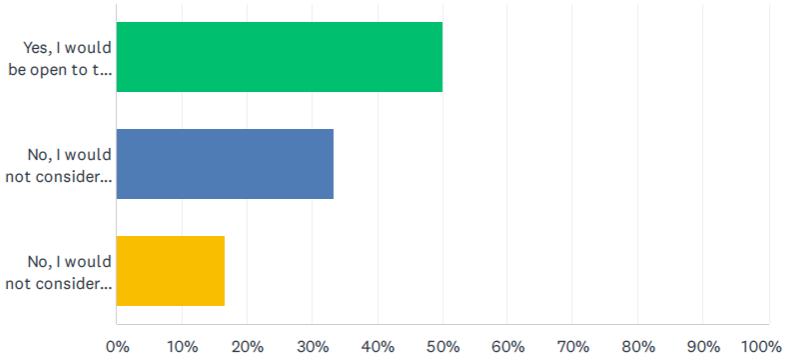
Answered: 12 Skipped: 0



ANSWER CHOICES	RESPONSES	
1:1	41.67%	5
2:1 (similar to Meis and PROLONG)	41.67%	5
3:1	0.00%	0
> 3:1	0.00%	0
I don't' know	16.67%	2
TOTAL		12

Q5 Another study design change has been suggested with the purpose of facilitating feasibility and completing a trial in the US. Given the challenges of cost and recruitment, there has been a proposal to evaluate a primary outcome that is a" delay in delivery" vs. a specific GA threshold. The primary outcome for Meis was PTB < 35 wks and PROLONG had co-primary outcomes < 35 wks and a composite of neonatal morbidity.What is your opinion regarding evaluating the primary outcome in a different manner? Would you consider a "delay in delivery" that had clinical meaning (e.g. 7 days difference between placebo vs. 17 OHPC)?This delay in delivery could be a continuous outcome (days) or viewed as a "time to event" metric.

Answered: 12 Skipped: 0



ANSWER CHOICES	RESPONSES	
Yes, I would be open to this approach, assuming there was adequate data to support the measure	50.00%	6
No, I would not consider this approach and I believe a GA threshold is required (e.g. < 34 wks)	33.33%	4
No, I would not consider this approach and I believe only an outcome of neonatal morbidity and mortality is required	16.67%	2
TOTAL		12