Does Makena work? A quasi-experimental approach to examine the drug treating preterm

Xuanhao He* Engy Ziedan †

Abstract

Preterm birth causes more than 26-billion-dollar in social costs in the US. Despite its heavy burden on society, the cause of preterm birth remains unknown and there is no cure for preterm birth. Makena was granted accelerated approval by the FDA in 2011 and is the only approved drug for preventing preterm birth. While the initial randomized control trial showed the drug efficacy in reducing recurrent preterm birth, the follow-up trial failed to confirm the efficacy. This article aims to throw light on the confusion in the drug efficacy using a quasi-experimental approach. It applies the difference-in-differences method to both Medicaid drug utilization data and Natality birth records. The results show that Medicaid coverage of Makena reduces the incidences of preterm birth and low birth weight while the prior authorization requirement has the opposite effects among high risk women. Further, women above 35 appear to benefit more from the coverage.

Keywords: Preterm birth; Makena (17P); Accelerated drug approval

JEL Codes: I18, I13, I12

*Anlaysis Group

†Department of Economics, Tulane University, New Orleans, LA 70118. Email: eziedan@tulane.edu

1 Introduction

Clinical evidence indicates that preterm birth (PTB) and Low birth weight (LBW) of less than 2,500 grams are markers for subsequent poor health and socio-economic adversity in both the short and long term (Aizer and Currie, 2014; Almond et al., 2018; Barker, 1990; Currie, 2009, 2011; Currie and Almond, 2011). Specifically, LBW's primary causes are PTB and intrauterine growth retardation (Kramer, 1987). Associated outcomes include increased risks of infant and child mortality, cognitive disadvantages, reduced educational outcomes, compromised labor market performance, and intergenerational LBW (Almond et al., 2005; Behrman and Rosenzweig, 2004; Bharadwaj et al., 2018; Black et al., 2007; Figlio et al., 2014; Oreopoulos et al., 2008; Royer, 2009). The financial implications are profound: PTB-associated costs in the US reach upwards of 26 billion dollars, averaging $51,600 per infant born preterm (Behrman et al., 2007). Similarly, the societal cost for very low birth weight (VLBW) infants is even steeper, averaging $60,421 each (Currie et al., 2019).

In light of these challenges, this paper delves into the evaluation of a drug endorsed by the FDA to mitigate the risks of PTB. Despite a plethora of attempts, the medical literature had long struggled to discover a substantive intervention to combat PTB. This scenario changed in 2003 when a randomized control trial (RCT) unveiled the efficacy of Makena's active ingredient, demonstrating a PTB reduction rate of 34% (Meis et al., 2003). This revelation resulted in the FDA's accelerated approval of Makena in 2011. Controversially, a subsequent trial in 2020 contested these findings, reporting a negligible impact of Makena on PTB rates (Blackwell et al., 2020). This inconsistency fueled debates concerning Makena's place in the pharmaceutical market and the therapeutic void its removal could create.

This disagreement among the RCTs is not unusual. A large share of FDA approved therapies find low to zero effectiveness in post market trials (Deyo, 2004). Researchers have recently pointed towards characteristics of the "compliers" selected by clinical RCTs as a main cause (Alsan et al., 2022; Rothwell, 2006). These studies suggest that clinical trials selecting the sickest patients may fail to produce technologies that increase the health of the

average patient. The opposite is also true in studies that advantageously select the healthiest patients, and may endorse therapies that in fact have very low marginal productivity on sicker individuals. The outcome of this discrepancy is that quasi-experimental assessments of health technologies could complement the RCT results.

This paper adopts a quasi-experimental approach to unravel the complexities surrounding Makena's efficacy. This choice is informed by multiple considerations. Notably, behavioral modifications influenced by FDA-approved drugs, like Makena, could obscure the genuine impact of such treatments. Moreover, the intrinsic rarity of PTBs across populations necessitates large enough samples to unearth statistically significant findings - a challenge for most RCTs on rare diseases. The era of Makena's introduction further complicates matters, as it coincided with the rollout of various potentially confounding policies in Medicaid and in particular the timing of the Medicaid expansion.

To address these empirical challenges, this study harnesses the variability in Medicaid's Makena coverage policies across states. Medicaid's involvement is particularly relevant for multiple reasons. For one, it caters to half of US births, thereby providing a sizable sample for enhanced statistical power. Additionally, Medicaid recipients, often at the intersection of lower income brackets and racial disparities, exhibit a heightened predisposition towards PTB (Lu and Halfon, 2003; Markus et al., 2017). Further enriching this analysis is the inherent variability across Medicaid agencies in terms of drug coverage initiation dates and the requirement for prior authorization.

Employing both event study and difference-in-differences (DID) methodologies, our research reveals that Medicaid's Makena coverage significantly amplifies Makena-related prescriptions and applications. However, pre-authorization requirements produce converse effects. Intriguingly, while Makena coverage markedly suppresses PTB and LBW rates, pre-authorization requirements exacerbate very PTB (VPTB) and VLBW frequencies among high-risk demographics.

This research's contributions are twofold. Firstly, to our knowledge it is the first quasi-

experimental assessment of Makena's efficacy, enhancing the existing knowledge derived from previous RCTs that yielded incongruous outcomes. While RCTs undoubtedly provide rigorous insights, their scope can sometimes be circumscribed by the specifics of trial conditions and demographic peculiarities. Our quasi-experimental approach not only vindicates the preliminary RCT insights on Makena but also spotlights the specific demographic segments that are most poised to benefit from Makena. Secondly, we test whether Makena materially influences maternal behavioral patterns related to conception decisions. An often cited hypothesis by state Medicaid agencies opposing listing Makine, is that PTB risk-mitigating solutions might inadvertently embolden potentially vulnerable women to pursue conception - otherwise known as expost Moral Hazard. We test this hypothesis directly.

2 Background

2.1 Compounded 17P

Compounding is a process of creating drugs to meet the unique needs of an individual patient when a manufactured drug is not available or appropriate for the patient, or must be altered in some way (e.g., delivery routes) (Galson, 2003). Before Makena's approval, its active ingredient, 17 alpha-hydroxyprogesterone caproate (17P) was available at compounding pharmacies. 17P is a synthetic form of progesterone that was initially approved by FDA under the trade name Delalutin in 1956. It was used to treat gynecologic and obstetrical conditions, such as habitual abortion. In 2003, the American College of Obstetricians and Gynecologists (ACOG) recognized the role of 17P in reducing PTB (American College of Obstetricians and Gynecologists, 2003). In 2008, both ACOG and the Society for Maternal Fetal Medicine (SMFM) recommended using 17P for reducing PTB (Society for Maternal Fetal Medicine Publications Committee, 2008). After that, some physicians started to prescribe compounded 17P for treating PTB.

Yet, compounded 17P was rarely used under Medicaid. To begin with, Medicaid generally does not cover compounded drugs.[1] Without unique codes, compounded drugs can not be reimbursed in the Medicaid drug system. Second, few women knew about 17P because compounded drugs cannot be advertised or marketed. The 1997 FDA Modernization Act (FDAMA) banned drug providers from promoting or advertising compounded products (Public Law 105-115, 1997). Third, compounded 17P can only be ordered from a small portion of special compounding pharmacies. In 2013, out of 56,000 community-based pharmacies in the US, about 7,500 (13%) are pharmacies specialized in compounding, and around 3,000 (5%) make sterile products such as 17P (Government Accountability Office, 2013).

Fourth, physicians are reluctant to prescribe compounded drugs due to its risk of safety and personal liability. Under the 1997 FDAMA, compounded drugs are exempted from three key regulations of manufactured drugs. One exemption is from the requirement to undergo the pre-market review for safety, effectiveness, and quality (Gudeman et al., 2013; Public Law 105-115, 1997). Another exemption is from the requirement of compounding sites being subject to good manufacturing practice guidelines, including regular inspections, quality control testing, and rejection of low-quality material (Code of Federal Regulations, 1985).[2] The other exemption is from the requirement to bear adequate direction for use. On the other hand, prescribing compounded drugs increases the potential risk of malpractice insurance invalidation and personal liability should unintended outcomes occur (Gudeman et al., 2013; Patel and Rumore, 2012; Sellers and Utian, 2012).

[1] There are a few exceptions in managed care organizations (MCO) that opt to cover compounded drugs.

[2] Following the passage of the Drug Quality and Security Act in 2013, the specialized compounding pharmacies are subject to current good manufacturing practice and increased federal oversight (Public Law 113-54, 2013). The oversight requires registration, twice yearly report on compounded drugs, adverse event reports, and certain criteria for both bulk drug substance and other ingredients.

2.2 Makena Introduction

Treating PTB is challenging. Previous interventions towards reducing PTB were not effective regardless of women's level of PTB risk. These include nutritional supplement during pregnancy (Buppasiri et al., 2015; Harper et al., 2010; Hauth et al., 2010), vaginal diseases treatment (Carey et al., 2000; Klebanoff et al., 2001), and periodontal disease therapy (Offenbacher et al., 2009). At least two reasons can explain this. One, PTB is a multi-factorial problem whose biological mechanisms are elusive (Buekens and Klebanoff, 2001; Goldenberg and Rouse, 1998; Iams, 2014; Slattery and Morrison, 2002). The other, psychosocial barriers restrict women's willingness and ability to participate in RCTs, increasing the difficulty to recruit and retain pregnant women (Frew et al., 2014).

In 2003, Meis et al. (2003) published the results of a RCT showing evidence of 17P on PTB (<37 weeks of gestation) reduction. Since the results are reasonably likely to predict clinical benefits to neonates, the FDA granted Makena accelerated approval on February, 2011 (BRUDAC Meeting, 2019a). Makena is indicated for reducing the recurrence of PTB for women with singleton gestation and prior spontaneous PTB. It is administered through weekly injections of 250 milligram beginning at 16−20 weeks of pregnancy until delivery (or 37 weeks of pregnancy whichever comes first). The FDA also designated Makena as an orphan drug for PTB prevention on January, 2007 (Food and Drug Administration, 2007). The orphan drug status offers seven years' market exclusivity for manufacturers developing drugs treating rare disease (affecting less than 200,000 individuals per year). Makena was available at a cost of $690 per dose or $14,000 per pregnancy (Patel and Rumore, 2012).[3]

[3]The initial wholesale acquisition cost was nearly $1,500 per dose ($30,000 per pregnancy). The price was cut shortly afterwards to accommodate the FDA's concern over limited access to the therapy.

2.3 Medicaid Coverage of Makena

The Medicaid drug rebate program requires a drug manufacturer to enter the National Drug Rebate Agreement (NDRA) as to be covered by Medicaid. Once the NDRA is entered, Medicaid programs can cover the drug as of the date of drug entering. Generally, Medicaid programs must cover the drug from the mandatory effective date. That is the first date of the calendar quarter that begins more than sixty days after the NDRA entering date (Public Law, 1993).

Post its FDA approval in 2011, Makena had a staggered introduction across states. Medicaid programs have their own drug formularies. The drug review process for placing a drug onto a formulary and starting to cover the drug varies across states.[4] A state may require a minimum period of time (e.g., six months) after the FDA approval before a drug can be reviewed by pharmacy and therapeutics (P&T) committee. Also, the committee meetings are held at a particular interval of time (e.g., quarterly) to determines whether or not to cover the new drug.

Established under Omnibus Budget Reconciliation Act (OBRA) of 1990 and 1993, PA programs require physicians or pharmacists to obtain approval from Medicaid agencies before dispensing a drug (Public Law, 1993). The laws established two requirements only for the program. One, PA requests must be responded via telephone or other telecommunication devices. Two, a 72-hour supply of a medication must be available in emergency situations.

This broad nature of federal requirement creates a large variation in Medicaid PA programs. First, states vary in terms of the entities that determine whether a medication is subject to PA, including but not limited to the Pharmacy Office, drug utilization review (DUR) board, and pharmacy and therapeutics (P&T) committee (Gifford et al., 2020; Tilly and Elam, 2003). Second, there are different reasons for PA requirement, such as drug cost

[4]Drug formulary is a list of generic and brand drugs covered under Medicaid. This is different from the preferred drug list (PDL) since not all drugs covered by Medicaid are reviewed on PDL.

exceeding a threshold, a drug being in the class of preferred drug list (PDL), or temporary PA requirement until full drug review (Gifford et al., 2020). Third, agencies managing a PA process vary across states, including Medicaid agencies, pharmaceutical benefit management companies, and university college of pharmacies (Tilly and Elam, 2003).

Based on the conversation with Medicaid agencies, the following describes the process of a state determining PA requirement in detail. "For medications that do not fall into a PDL drug class, such as Makena, decisions for requiring prior authorization are based on review of clinical information for the drug product, including published information on safety and efficacy (such as from product labeling), information from clinical and guideline compendia, and data from published studies. Drug cost information is also evaluated when making a determination regarding prior authorization. Prior authorization criteria for approval are drafted by the Pharmacy Office DUR team and reviewed at the Department's quarterly public DUR Board Meeting, where the DUR Board reviews and makes recommendations to the Department regarding criteria for approval. Stakeholders' input provided to the Department is also taken into consideration as part of the public DUR Meeting review and as part of final decisions made by the Department regarding prior authorization criteria."

3 Literature Review

3.1 RCTs on Makena(17P)

The first RCT on Makena (17P) was funded by the National Institute of Child Health and Human Development (NICHD) and conducted by Maternal-Fetal Medicine Units Network (MFMU) in 1999−2002 (Meis et al., 2003). This RCT enrolled 463 women aged 20 to 32 with a history of spontaneous PTB from nineteen academic centers in the US. The sample has a high-risk of having recurrent PTB: 60% are African Americans and 32% have prior spontaneous PTB. The primary efficacy end point is PTB. The trial found that weekly

intramuscular injection of 17P reduced the risk of PTB (<37 weeks of gestation at delivery) by 34%, the risk of MPTB (<35 weeks of gestation at delivery) by 33%, the risk of VPTB (<32 weeks of gestation at delivery) by 42%, and the risk of LBW by 34%. After adjusting for multiple analyses simultaneously, the estimate on PTB reduction is highly statistically significant while those on MPTB and VPTB are not.

As required by Makena's accelerated approval, the AMAG Pharmaceuticals conducted a confirmatory trial. Due to the limited number of women eligible for the trial in the US, the trail recruited 77% women aged 24 to 36 from 93 international centers (36% from Russia and 25% from Ukraine) during 2009−2018 (Blackwell et al., 2020). The participants have a much lower risk of having PTB again: 7% are Black/African Americans and 12% have prior spontaneous PTB. The co-primary efficacy end points are MPTB and neonatal morbidity/mortality index. In contrast to the previous trial, this trial found that weekly Makena (17P) injection did not affect the risk of PTB, MPTB, VPTB, birth weight, or neonatal morbidity and mortality among women. One potential explanation for this finding is the lack of statistical power due to low baseline risk of having PTB.

3.2 RCTs and Quasi-experimental Designs

RCTs are experimental designs, in which the treatment and control groups are well-balanced and any observed post-treatment difference can be attributable to the treatment. RCTs are limited in several aspects. First, the advantage of internal validity of RCTs restricts the results' generalizability to other settings (Deaton and Cartwright, 2018). Second, with enough heterogeneity between those at marginal risk and at risk of PTB, the average treatment effects of the trial samples and general population can be quite different (Deaton and Cartwright, 2018; Longford and Nelder, 1999). Third, RCTs' results are limited by its sample size, selective participants, provision of care, and among others (Booth and Tannock, 2014; Nallamothu et al., 2008).

Quasi-experimental designs alleviate these limitations. Different from RCTs using randomly assigned treatment and control groups, quasi-experimental designs can be applied to the whole population (e.g., all pregnant women giving birth), removing the difficulty of extrapolating the results from ideal experimental settings to the real world. The heterogeneity within population allows for a population average estimate and estimates for different subpopulations that are often desired by policy makers. Further, quasi-experimental designs employ methodologies to minimize estimation bias.

3.3 Innovation, Health, and Health Behaviors

The introduction of Makena could cause changes in health behaviors. On the one hand, pregnant women may reduce investment in health and engage in risky health behaviors considering the medical technology lowers the cost of the behaviors. Previous studies found that HIV+ individuals who received the HIV drug — highly active antiretroviral therapy — had four additional sex partners (100% increase) in the US (Lakdawalla et al., 2006). Statin use was associated with increases in bmi, probability of being obese, alcohol use among men, and with mixed evidence on smoking and exercise in the US (Kaestner et al., 2014).

On the other hand, people could engage in health-promoting activities (e.g., exercise) since the marginal productivity of health investment is higher due to Makena availability. Previous studies found that Mothers' reception of tetanus, which reduces mortality in the first month of life, was associated with an increase in birth weight in Sub-Saharan Africa, suggesting an increase in nutrition input during pregnancy (Dow et al., 1999). Participation in male circumcision for HIV prevention was associated with a 10-20 percentage point decrease in the probability of having multiple sexual partners and an increase in condom use after one year in Sub-Saharan Africa (Wilson et al., 2014). Yet, these results are based on developing countries and may not be applicable to the US.

4 Data

4.1 Medicaid State Drug Utilization Data

I utilize three data sets. The first data set is the Medicaid state drug utilization data (SDUD) in 2002−2015 from the Centers for Medicare and Medicaid Services. States participating in the Medicaid Drug Rebate Program are required to report the SDUD to receive rebates from the federal government. The SDUD records the data of outpatient drugs covered by Medicaid fee-for-service (FFS) and managed care (MC) programs. This contains the National Drug Codes (NDC), the units of drugs utilized, the number of the prescriptions filled, and the amount of Medicaid reimbursement. I obtain a list of Makena's NDC codes from the National Drug Code Directory and convert the number of drug units into the number of drug doses (in 1,000s) based on Makena products' package descriptions. There were two Makena products on the market in the study period. One (NDC: 64011024702) contains four single-dose vials, with each dose containing 250 milligram hydroxyprogesterone caproate. The other (NDC: 64011024301) contains a five-dose vial with the same concentration. I also calculate the number of Makena prescriptions filled (in 1,000s). In 2015 alone, the seventeen states reported more than 19,000 Makena prescriptions filled and more than 350,000 doses of Makena reimbursed under Medicaid.

The SDUD has two limitations. One, it is under-reported. The Appendix Table 1 presents the years and quarters when states did not report drug utilization data. The non-reporting events mainly happened in 2011−2013 in states such as Mississippi and Wisconsin. The other, the SDUD is aggregated to the state-by-quarter level, making it impossible to identify individuals who received Makena. Two other data sets contain prescriptions and medical claims. One is the Medicaid Analytic eXtract (MAX) file. MAX contains enrollment information and medical claims of Medicaid recipients. While the data on Medicaid FFS recipients is more reliable in the MAX, the MC encounter data is vulnerable to errors (Li

et al., 2018; Office of Inspector General, 2009, 2015).[5] The alternative data set is produced by Truven Health Analytics and named the MarketScan Commercial Claims and Encounters Data (CCAE). Yet, CCAE records mostly women covered by privately insurance rather than those by Medicaid.

4.2 Birth Records

The primary data set is the birth records from the National Vital Statistics System of the National Center for Health Statistics. The main sample contains approximately 20 million birth whose mothers conceived between 2002 and 2015. The data set contains detailed measures of both mothers' and neonates' demographics and health, including mothers' age, race, marital status, educational attainment, gestational age, neonates' birth weight, and among others.

The main outcome variables are gestational age and birth weight. Consistent with the RCTs, I use three measures of gestational age, including the indicators of VPTB (<32 weeks' gestation), MPTB (<35 weeks' gestation), and PTB (<37 weeks' gestation). The gestational age was measured based on the last menses period (LMP) before 2014 and based on the obstetric estimate (OE) of gestation at delivery from 2014 onward. While both measures of gestational age are highly correlated, birth is less likely to be designated as PTB using the OE than the LMP measure (Martin et al., 2015). I also use three measures of birth weight, including birth weight in grams, the indicators of VLBW (<1,500 grams) and LBW (<2,500 grams). Birth weight is often used as a summary measure of neonates' health. LBW and VLBW are associated with a range of poor outcomes (e.g., (Almond et al., 2018)).

For ease of computational burden, I aggregate the birth records to the month-state-demographic group level. Data cells are constructed using mothers' age groups (18 – 24, 25 –

[5]This is due to two reasons. One, data reporting under MC is non-essential for MC plans and providers to receive capitation payment. The other, states have fewer incentives to report MC organizations' data compared to reporting FFS data.

34, 35 – 44), mothers' race and ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, and other or unknown), mothers' educational attainment (less than high school, high school, some college education, college degree or more, not reported, and not on certificate), marital status (married and unmarried), conception year and month, and state of residence.[6] There are 144 possible cells in each state and conception year-month and 411,264 potential cells in total. Due to the limited number of birth in the groups of educational attainment "not reported" and " not on certificate," the number of cells is 281,795. I use the number of women in each cell as the sample weight and clustered standard errors at the state level.

4.3 Medicaid Policies of Makena

The third data set is Makena coverage policies under Medicaid. I focus on three policy dates, the date when a Medicaid agency starts to cover Makena, the date when a PA requirement is implemented, and the date when the PA requirement is removed. I collect these dates using three methods. First, I search for online postings by Medicaid Pharmacy Offices, DUR board meeting reports, and among others. For instance, the Mississippi DUR Package of March 23, 2019 states that "Makena has been covered by Medicaid since it was FDA approved." A file named "Criteria for Prior Authorization" from Kansas KanCare and Medicaid Pharmacy Office shows the exact date when the PA requirement for Makena started (June 15, 2011).[7]

Second, I use internet archive, the Wayback Machine, to search for older versions of websites that retain earlier online postings from the agencies above. Third, I engage in direct phone conversations and emails with Medicaid agencies and Medicaid Pharmacy Office members, including managers and pharmacists, in Kansas, Pennsylvania, Wisconsin, and others. Using all three methods, I collect seventeen states with both coverage and PA dates available. Figure 1 plots the initial coverage dates. While most states started to cover

[6]One limitation with this level of aggregation is that 60% of the cells have fewer than 25 women. As a robustness check, the results from aggregating to conception year level instead of year and month are similar.

[7]The file can be found through the following link: `https://www.kdheks.gov/hcf/pharmacy/PA_Criteria/Makena_PA_Criteria.pdf`.

Makena before 2012, Massachusetts and Nebraska did not cover Makena until mid-2013. Figure 2 shows that PA requirements started at the same time as coverage and stayed in most states except three states. Wisconsin removed its PA requirement in April 2014. Alabama and North Carolina Medicaid do not have PA requirements for Makena.

5 Method

I use the variation in the timing of Medicaid Makena coverage policies across states to study three questions. First, whether Makena coverage leads to an increase in Makena utilization and whether PA requirement for Makena causes a reduction in Makena utilization? Second, what are the effects of the coverage policies on birth outcomes? Third, are there any effects on the types of women giving birth? I execute event study and the DID designs to answer these questions.

5.1 Event Study

The event study specifications are based on a balanced panel of 17 states observed for 14 years. I use $s = 1, ..., 17$ to index the states and use $t = 1, ..., 14$ to index years. In total, the sample contains $17 \times 14 = 238$ state-year observations. Let PA_s be the PA policy year in state s that is the first year when the Medicaid PA requirement for Makena is effective for more than half of the year.[8] And let $PAT_{st} = t - PA_s$ measure the number of years between year t and the PA policy year. Next, I set $PAT_{st} = 0$ for states whose Medicaid never have Makena PA requirement and set $PAT_{st} = 0$ for Wisconsin in 2014 and 2015 since the PA

[8]The results are essentially the same when defining PA_s as the first year when the PA requirement is effective for the whole year. And I do not conduct event study analysis for Makena coverage under Medicaid due to the limited variation in policy timing across states.

requirement ended in April 2014. I estimate the following model using the SDUD.

$$Y_{st} = \alpha_0 + \alpha_1 Coverage_{st} + \sum_{k=-4, k\neq -1}^{4} \beta_k 1(PAT_{st} = k) + \mu_{st} + \eta_s + \tau_t + \epsilon_{st} \quad (1)$$

In model 1, Y_{st} stands for the number of Makena injections or prescriptions. $Coverage_{st}$ measures the share of months in which Medicaid Makena coverage policy is effective in state s and year t. β_k are event study coefficients that trace out deviations from the common state trend in the years leading up to and following the PA policy year. The reference period is the year before the PA policy year, when $PAT_{st} = -1$. μ_{st} is a set of state-level controls, including the Medicaid income eligibility thresholds of pregnant women, the share of population in poverty, and the average unemployment rate. η_s is a set of state fixed effects, capturing time invariant differences in the outcomes' levels across states. τ_t is a set of calendar year fixed effects, which control for the trends in the outcomes common across all states. ϵ_{st} is a residual error term. Robust standard errors are clustered at the state level.

Next, I estimate the following event study model using the birth records.

$$Y_{ist} = \gamma_0 + \gamma_1 Coverage_{ist}^{expo} + \sum_{k=-4, k\neq -1}^{4} \delta_k 1(PAT_{st} = k) + \phi_{ist} + \lambda_{st} + \eta_s + \upsilon_t + \epsilon_{ist} \quad (2)$$

Model 2 is different from model 1 in five aspects. First, Y_{ist} is a set of outcomes related to gestational age and birth weight. Second, $Coverage_{ist}^{expo}$ measures the extent to which a woman i is exposed to Medicaid Makena coverage policy during her pregnancy. It is defined as the share of the expected gestational period (ten months starting from the last menses month) in which the coverage is effective. For instance, a woman, who conceives in January 2013 in a state that started to cover Makena in May 2013, receives an exposure of 0.6 since six months of her gestation are subject to Makena coverage.

Third, ϕ_{ist} is a set of individual-level controls, including indicators for mothers' age groups, marital status, race and ethnicity categories, educational attainment levels, male

infants, and multiple birth. Fourth, λ_{st} is a set of state-level controls, including the Medicaid income eligibility thresholds of pregnant women, the share of population in poverty, and the average unemployment rate in the ten months after the conception month. Fifth, υ_t is a set of conception year and month fixed effects. δ_k are event study coefficients.

5.2 Difference-in-differences

The event study models are flexible to estimate the effects of Medicaid policies on Makena usage and birth outcomes separately for each period. However, the estimates may be statistically imprecise due to limited observations in each period. Next, I apply the DID method to estimate the average effects of Medicaid policies on the outcomes. I begin with the regressions on the SDUD data set in the following form. The treatment group of Medicaid coverage contains the states in the years that the coverage is effective while the control group contains the rest. Similarly, the control group of the PA requirement contains the states in the years when the requirement is not effective, which include Alabama and North Carolina in all years, and Wisconsin from the later half of 2014 to 2015.

$$Y_{st} = \alpha_0 + \alpha_1 Coverage_{st} + \beta_d PA_{st} + \mu_{st} + \eta_s + \tau_t + \epsilon_{st} \tag{3}$$

Model 3 is similar to Model 1 with one exception. PA_{st} measures the share of months when Medicaid PA requirement is effective in state s and year t. α_1 and β_d are the DID coefficients. Specifically, α_1 measures the effect of exposure to Medicaid coverage for a whole year on Makena utilization and β_d shows the effect of exposure to Medicaid PA requirement for a whole year on Makena usage. In the results, I present the estimates of Model 3 and the estimates after excluding μ_{st} from Model 3. These controls are meant to alleviate potential concern over the endogeneity in Medicaid Makena policies' adoption. Results that robust to removing them would strengthen the confidence in the Medicaid policies' exogeneity. Further, I fit models that allow for state-specific linear yearly trends. I next estimate the

reduced form regressions using the birth records. The treatment group contains the birth exposed to either Medicaid coverage or PA requirement policy during pregnancy while the control group contains the rest. The model follows the following form.

$$Y_{ist} = \gamma_0 + \gamma_1 Coverage_{ist}^{expo} + \delta_d PA_{ist}^{expo} + \phi_{ist} + \lambda_{st} + \eta_s + \upsilon_t + \epsilon_{ist} \tag{4}$$

Model 4 is similar to Model 2 except that PA_{ist}^{expo} is a measure of a woman i's exposure to Medicaid PA requirement during pregnancy. Similar to $Coverage_{ist}^{expo}$, it is defined as the share of the expected gestational period in which the PA requirement is effective. γ_1 and δ_d are the DID coefficients, measuring the effects of full exposure to Medicaid coverage and PA requirement during pregnancy on the outcomes. I present results of Model 4, of Model 4 with a set of interactions between baseline demographics and conception years. These demographics include mothers' age groups, marital status, race and ethnicity categories, and educational attainment levels. The interactions capture the common trend of each demographic category in the outcomes across all states, alleviating concern over alternative policies affecting same demographic groups simultaneously. Further, I present the results of Model 4 with state-specific linear trends.

5.3 Preterm Birth Risk Prediction

Full sample estimates may suffer from attenuation bias since not all pregnant women are qualified for Makena that targets women with high PTB risk. So, I present the estimates of the sub-samples with high PTB risk, which are defined as the probability of having PTB being higher than 12% and 20%. I predict the probability of having a PTB using the following model and birth whose mothers conceived by between 2002 and 2009.

$$Preterm_{ist} = \theta_{ist} + \eta_s + \zeta_t + \epsilon_{ist} \tag{5}$$

In Model 5, θ_{ist} is a set of baseline characteristics, including the indicators of mothers' ages, marital status, race and ethnicity categories, educational attainment levels, and male infants. ζ_t is a set of conception month fixed effects.[9] I predict $\widehat{Preterm}_{ist}$ from Model 5 and match it to all birth records based on θ_{ist}, η_s, and ζ_t. This way I essentially create a measure of expected PTB risk for each birth.

6 Results

6.1 The Results on Makena Usage

Figure 3 shows the estimated event study coefficients and 95% confidence intervals on Makena prescriptions (left panel) and injections (right panel). Since no Makena was available before 2011 when Medicaid starts to cover and require PA for the drug, the pre-PA effects are zero. In the post-PA periods, the effects are negative and statistically different from zero at the one percent level. The magnitude of the coefficients suggests that the PA requirement reduces Makena prescriptions by approximately an additional 300 per year. It also implies that the requirement reduces Makena injections by about 10,000 in the first year and an additional 5,000 per year after that.

Table 1 presents the DID estimates on Makena usage. I estimate three models by gradually adding the state-level controls and state-specific linear trends. The estimates on Coverage and PA are statistically significant in models without state-specific linear trends.[10] The magnitude of the estimates are largely similar across specifications even after including linear trends. It implies that Medicaid coverage is associated with more than 900 Makena prescriptions and around 15,000 Makena injections while the PA requirement is associated with more than 1,000 fewer prescriptions and more than 16,000 fewer injections. All of these

[9] I do not use conception year fixed effects in prediction since yearly trend of PTB before 2010 may be different from the trend of PTB after 2010.

[10] The estimates lose significance after including trends. This is expected since the state-specific linear trends are highly correlated with the gradual increase in Makena usage after coverage starts.

are substantial changes from the national average level of 60 prescriptions and 960 doses in 2012.[11] The joint significance tests confirm that both Medicaid coverage and PA requirement are strong predictors of Makena prescriptions and dosage.

6.2 The Event Study Results on Gestational Age

The previous section shows that the adoption of these Medicaid policies creates plausibly exogenous variation in Makena usage that can be used to estimate Makena's effect on birth outcomes and maternal behaviors. In the Figure 4, the first row shows the event study coefficients of the PA on gestational age using all birth. The left panel shows that neither pre-PA and post-PA effects on VPTB is significantly different from zero, while the standard errors are large. So, I can not draw conclusions of the PA effect on VPTB. Similarly, the middle panel shows no clear effect of the PA on MPTB. In the right panel shows the pre-policy effects are almost zero, supporting the assumption of no differential pre-trends or anticipation effects in PTB correlated with the timing of the PA. In contrast, the post-PA effects are positive and statistically different from zero at one percent level in the two years after the policy. The estimates imply that the PA requirement increases the probability of PTB by 0.4 percentage points (pp) in the first two years of the PA requirement. This effect gradually reduces to zero after that.[12]

The first row of the Figure 5 shows the event study coefficients using the birth whose predicted PTB risk is above 12%. The results show that the post-PA effects on VPTB and MPTB are larger and more similar to those on PTB. Due to fewer observations, the post-PA effects on PTB are statistically different from zero at 5 percent level in two years after

[11] The variation used to estimate the PA effect comes from Alabama, North Carolina, and Wisconsin (after ending PA) only. The limited number of observations in these states causes larger estimates on PA than those on Coverage. The estimates on Medicaid coverage are relatively smaller but similar once the PA requirement is excluded from all the Models.

[12] The small observations used to estimate the PA coefficients may explain why the effect goes away after two years.

the PA.[13] The coefficients' magnitude suggests that the PA requirement increases the PTB probability by about 0.7 pp in the three post-policy years. The first row of Figure 6 shows the event study coefficients of PA among birth with preterm risk above 20%. The estimates are not statistically significant due to the limited number of observations.

The first row of the Figure 7 shows the event study estimates using the birth whose mothers are below 35. The results are similar to that of Figure 4 using all birth in both trends and magnitude. The fist row of the Figure 8 shows the estimates using the birth whose mothers are above 35. In compatible with a causal impact, the pre-PA effects are not statistically different from zero across all panels. The post-PA effects on VPTB are stable and statistically significant in the first and third post-policy year. The magnitude suggests the PA requirement increases the probability of VPTB by about 0.2 pp. Also, the effects on MPTB are significant at 1 percent level in the first and third post-PA years, suggesting that the requirement increases MPTB rate by around 0.45 pp in the three post-PA years. Moreover, the post-PA effects on PTB are significant at 1 percent level in the three years following policy year, raising PTB rate by 0.5 pp. The effect goes to zero after that.

6.3 The Reduced Form Results on Gestational Age

The previous section shows no differential trends before the PA requirement across states that establishes PA and those do not. Given that the core DID assumptions are plausible and the effects are relatively stable over the post-PA period, I next fit the DID model. I use the DID Model to examine Medicaid coverage and PA requirement simultaneously. I present reduced form estimates using three model specifications. Model 1 is the baseline model, Equation (4). Model 2 adds the interactions between mothers' demographics and conception year indicators to control for potential shifts in the groups of mothers giving birth. Model 3 adds state-specific linear yearly trends.

[13]The p-value for the third post-PA policy effect is 0.54.

The Table 2a reports reduced form estimates of Medicaid Coverage and PA requirement. From left to right, each panel shows the results using the full sample, the sample where mothers' predicted PTB risk is less than or equal to 12%, the sample where mothers' predicted PTB risk is more than 12%, and the sample where mothers' predicted PTB risk is more than 20%. From top to bottom, each part presents the estimates from Model 1 to Model 3. The left panel shows that Medicaid coverage of the drug is negatively associated with the probabilities of having PTB. The estimates are consistent across specifications, suggesting that Medicaid coverage reduces the probability of PTB by 0.6-0.7 pp (4.8-5.6% reduction relative to the baseline level in 2009). The estimates on the probability of having MPTB is statistically significant in Model 3 only, which shows a 0.3 pp (5.8%) reduction.

Next, I separate the sample into three groups based on the predicted risk of PTB. Compared to the left panel, the next panel shows that the estimates on Coverage are smaller and less statistically significant using the sample in which mothers' predicted PTB risk is less than or equal to 12%. The next panel presents the estimates of the sample with predicted PTB risk more than 12%. While being less significant, the estimates are substantially larger than those using lower risk sample. The magnitude suggests that Medicaid coverage reduces the probability of PTB by 1.1-1.3 pp (7.4-8.8% reduction). Still, the estimates on the probability of MPTB are significant in Model 3 only, showing a 0.5 pp (7.8%) reduction. The right panel shows the estimates are even larger using the sample with predicted PTB risk more than 20%. The magnitude implies that Medicaid coverage reduces the PTB probability by 2.9-3.5 pp (13.3-16.1%).

Further, the Table 2a shows that the PA requirement are statistically significant in the sample with the highest predicted PTB risk, above 20%. The estimates on the probability of VPTB are consistent across model specifications and suggest a 0.3-0.5 pp (6.3-10.4%) increase. The Chow Test results show that the estimates are all statistically larger than those using the sample with predicted PTB risk above 12%.

The Table 2b presents the estimates using women below 35 and above 35. The estimates using the sample of women below 35 are similar to those using the full sample. Using the sample of women above 35, the estimates show that Medicaid coverage reduces the probability of PTB by 0.7-0.9 pp (4.8-6.2% reduction). Also, the estimates of Coverage on MPTB are significant across all specifications, showing that the coverage reduces the probability of MPTB by 0.4-0.5 pp (6.5-8.1% reduction).

6.4 The Event Study Results on Birth Weight

Next, I present the event study coefficients on VLBW and LBW. In the second row of the Figure 4, the left panel shows no differential pre-trends or anticipation effects before the PA requirement but positive effects on VLBW after the PA requirement. The magnitude is stable in the five years after the requirement and implies a 0.07 pp increase in VLBW rate. The right panel shows positive effects on LBW in the first three years after the policy only. The magnitude suggests a 0.2 pp increase in LBW in the first two years and 0.1 pp increase in the third year. Similarly, the results using the birth with PTB risk above 12% show that the requirement increases VLBW by 0.15 pp in the five post-policy years and LBW by 0.4 pp in the first three post-policy years (the second row of the Figure 5). Besides, the results using the birth with above 20% predicted PTB risk are statistically insignificant due to a limited number of observations (the second row of the Figure 6).

Then, I show the coefficients separated by mothers' age. The second row of the Figure 7 shows that less statistically significant effects on VLBW using the birth whose mothers are below 35. The second row of the Figure 8 shows the coefficients on the birth whose mothers are above 35. Consistent with a causal story, the left panel shows zero pre-policy effects but positive and statistically significant post-policy effects on VLBW at 1 percent level. The positive effects gradually increase for four years before stabilizing in the fifth post-policy year. The magnitude implies a 0.25 pp increase per year on average. To the contrary, the

right panel shows insignificant effects of the PA requirement on LBW.

6.5 Reduced Form Results on Birth Weight

The Table 3a present the DID estimates of Medicaid policies on birth weight. I investigate birth weight in grams and the indicators of LBW and VLBW. The left panel shows consistent effects of Medicaid coverage on LBW, reducing the probability of LBW by 0.3 pp (3.6%). The effects of Coverage on VLBW is statistically significant in Model 2 only. The second panel implies that Medicaid coverage has no effect on LBW with below 12% predicted PTB risk. The third panel presents statistically larger but insignificant estimates of Medicaid coverage on LBW using the sample of birth with above 12% predicted PTB risk. Similarly, the right panel shows even larger but insignificant estimates of the coverage on LBW using the sample of birth with predicted preterm risk above 20%. More interestingly, consistent with the effects on VPTB, the PA requirement increases the probability of VLBW by 0.2-0.4 pp (5.6-11.1%). This effect stays robust across three model specifications. The Table 3b shows the estimates separated by mothers' age. The results show similar estimates of Medicaid coverage using either age group of mothers to those using all mothers. The estimates of the PA requirement generally show positive effects on LBW and VLBW, while the effect is inconsistent across model specifications.

6.6 The Effects among Low-educated Mothers

I next restricted the sample to be women with 12 years of education or less, who are more likely to be covered by Medicaid.[14] In the Table 4a, the left panel confirms the results using all women that the effect of Medicaid coverage is more robust on PTB than MPTB. The estimates suggest that Medicaid coverage decreases the probability of PTB by 1.1-1.2

[14]On average, 65% of these women are covered by Medicaid. I do not use the sample of women covered by Medicaid since delivery payment source data is not available until 2011.

pp (8.0-8.8%).[15] The next three panels show that the results using low-educated mothers grouped by predicted risk of PTB are similar to the results using all women. In the Table 4b, the results confirm those using all women that Medicaid coverage decreases the probability of LBW. The estimates show a 0.5 pp (5.4%) reduction in the LBW rate.

6.7 Evidence on Ex-ante Moral Hazard?

In this section, I test if Medicaid policies affect women's decision to conceive. While Makena is indicated for treating PTB among women with singleton pregnancy and histories of spontaneous PTB, it is advertised as potentially being effective in treating pregnant women with PTB risk, such as African Americans.[16] The introduction of Makena may encourage women with high risk of PTB to conceive again, shifting the characteristics of conceiving mothers.

Table 5 shows the estimates of Medicaid policies on the characteristics of women conceiving. The characteristics tested include the probabilities of having prior PTB, of being above 35, unmarried, white, African American, and Hispanic, of having no more than 12 years of education, of having the first live birth, the second live birth, the third live birth, and the fourth or later live birth. Again, I use three model specifications. Contrary to my expectation, the results show that Medicaid coverage of the drug decreases the proportion of African Americans giving birth by 2.0-2.3 pp (12.4-14.3% reduction). But the effect disappears after including state-specific trends. Overall, Medicaid policies have no significant effects on mothers' characteristics, suggesting little effect of *ex ante* moral hazard.

6.8 Placebo Tests

As a robustness check, I test the main results using placebo Medicaid policies. Assuming that nothing other than Medicaid policies of Makena in 2011-2015 is driving the estimates, I

[15] The effect on MPTB is statistically significant in Model 3 only, showing a 0.5 pp (8.5%) reduction.
[16] See the following link for details: `https://makena.com/`.

would find no effects of placebo Medicaid policies on Makena utilization and birth outcomes. I generate random Makena coverage and PA policies in the following steps. First, I assign random dates between 2011 and 2015 to states as placebo initial Medicaid coverage dates. Assuming Medicaid requires PA for prescribing Makena once the drug coverage starts, I use the placebo initial coverage dates as placebo PA requirement start dates. Second, I assign random PA end dates between 2011 and 2020 to states, which classify states into three PA status. A state is classified as having no placebo PA requirement if the PA end date is before PA start date. A state has a PA requirement through 2015 if the PA ends after 2015. The rest states have PA requirements ending within study period. Third, I create placebo coverage and PA exposure based on the random dates.

I start with placebo tests on Makena utilization using both placebo coverage and placebo PA. The results show none of the placebo coverage and placebo PA estimates on Makena prescriptions and doses are statistically significant across all model specifications (the Table 6). The magnitude is substantially smaller than that of real Medicaid policies. Also, the sign of the estimates on Makena doses is in opposite to that of actual Medicaid policies. I then run placebo tests on birth outcomes using all mothers (the left panel) and low-educated mothers (the right panel of the Table 7). Using all mothers, the estimates of placebo coverage on PTB are substantially smaller and statistically insignificant compared to the estimates of real coverage. Similarly, the estimates of placebo coverage on LBW are insignificant and have opposite signs to those of actual coverage. Using low-educated mothers, the estimates of placebo coverage on PTB are almost zero and insignificant. The estimates of placebo coverage on LBW are positive and significant using the Model 1. This suggests that, if anything, the unobservable factors during the study period would make it more difficult to detect significant effects on birth outcomes. In sum, these results supports the causal effect of Medicaid policies on Makena usage and birth outcomes.

7 Conclusion

Using quasi-experimental method, this paper examines the effects of Medicaid policies of Makena on both the drug utilization and birth outcomes. The results show clear evidence that Medicaid coverage of Makena significantly increases the number of drug prescriptions and doses while the PA requirement creates a large barrier for access to the drug. Consistent with the drug's indication for reducing PTB risk, Medicaid coverage of Makena decreases the probabilities of PTB and LBW, while the PA requirement increases the probabilities of VPTB and VLBW among pregnant women with PTB risk above 20%. Also, the effect of Medicaid coverage on PTB (<37 weeks of gestation) is more robust than that on MPTB (<35 weeks of gestation). Only among women above 35, Medicaid coverage has strong effects on both PTB and MPTB rates. Further, this paper finds little evidence that Makena's FDA approval changes women's conceiving behavior.

What we really care about is the effect of Medicaid policies on the pregnant women who actually received Makena under Medicaid. The main results using all birth suggest that Medicaid coverage reduces PTB rate by 4.8% (0.006 / 0.125). Ideally, I would scale the estimate with the proportion of women who were Medicaid beneficiaries and who received Makena in the 17 states, which, however, is unknown. Therefore, I made the following assumptions. I assume that the estimate represents the average effect in the US and there were 200,000 women receiving Makena each year based on the orphan drug status designation. Assuming half of the Makena recipients were on Medicaid, there were 500,000 female Medicaid beneficiaries receiving Makena between 2011–2015. Given the total number of women giving birth during this period was around 7,000,000, the proportion of female Medicaid beneficiaries receiving Makena was 7.14%. Scaled by this proportion, the estimate means that the Medicaid coverage policy reduces PTB by around 67% among those who received Makena.

How do we understand the estimate? Does Makena actually reduce PTBs? The sign of the estimate suggests it does. But the estimate may not be the unbiased estimate for Makena's efficacy for several reasons. To begin with, the estimates from the 17 states may be different from those in the US. Indeed, additional analyses using sub-groups of the states show heterogeneous effects across states. Also, Makena is administered by doctors during office visits or by healthcare professionals during home visits. Having Makena injections is associated with an increase in prenatal care visits, which has been shown to positively affect birth outcomes (Issel et al., 2011; Joyce, 1999; Roman et al., 2014). Further, Makena availability may be associated with a change in healthy behaviors during pregnancy due to the price distortion. Additionally, Medicaid drug policies in MCOs may be different from those under FFS and Medicaid prescription drug coverage may be carved into MCOs. However, I expect this to be less of a concern since MCOs must provide benefits required by Medicaid agencies and it is unlikely for MCOs to cover such an expensive drug before the agencies start requiring that.

From a policy perspective, a quasi-experimental study is in urgent need since RCTs in the US may not be feasible. The FDA has been grappling with the question of whether to pull Makena off the market since the publication of the confirmatory trial. In an FDA Advisory Committee meeting to review both RCTs on October 2019, nine members voted to recommend FDA to pursue withdrawal of Makena approval though seven members voted to leave the product on the market while conducting a new confirmatory trial (BRUDAC Meeting, 2019b). Conducting another trial in the US, however, may not be feasible considering that 17P has been the standard of care for seventeen years, which makes it difficult to identify a population of high-risk women with no access to 17P and makes it hard for physicians or patients to accept a placebo. Enrolling low-risk women would require a much larger sample size than the second RCT in order to obtain enough statistical power and might not be comparable to the results of the first RCT. To this end, a quasi-experimental study will not only fill in the shortage of RCTs but also contribute to the understanding

of 17P efficacy on a heterogeneous population of women, including those at marginal risk of PTB. More importantly, this article suggests that Makena would be more effective in reducing short gestational birth if the drug targets pregnant women aged above 35.

References

Anna Aizer and Janet Currie. The intergenerational transmission of inequality: maternal disadvantage and health at birth. *Science*, 344(6186):856–861, 2014.

Douglas Almond, Kenneth Y Chay, and David S Lee. The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083, 2005.

Douglas Almond, Janet Currie, and Valentina Duque. Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446, 2018.

Marcella Alsan, Maya Durvasula, Harsh Gupta, Joshua Schwartzstein, and Heidi L Williams. Representation and extrapolation: evidence from clinical trials. Technical report, National Bureau of Economic Research, 2022.

American College of Obstetricians and Gynecologists. Acog committee opinion. use of progesterone to reduce preterm birth. *Obstetrics and gynecology*, 102(5 Pt 1):1115, 2003.

David J Barker. The fetal and infant origins of adult disease. *BMJ: British Medical Journal*, 301(6761):1111, 1990.

Jere R Behrman and Mark R Rosenzweig. Returns to birthweight. *Review of Economics and statistics*, 86(2):586–601, 2004.

Richard E Behrman, Adrienne Stith Butler, et al. *Preterm birth: causes, consequences, and prevention*, volume 772. National academies press Washington, DC, 2007.

Prashant Bharadwaj, Juan Pedro Eberhard, and Christopher A Neilson. Health at birth, parental investments, and academic outcomes. *Journal of Labor Economics*, 36(2):349–394, 2018.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122 (1):409–439, 2007.

Sean C Blackwell, Suneet P Chauhan, Cynthia Gyamfi-Bannerman, Joseph R Biggio, Brenna L Hughes, Judette M Louis, Tracy A Manuck, Hugh S Miller, Anita F Das, George R Saade, et al. 17-ohpc to prevent recurrent preterm birth in singleton gestations (prolong study): a multicenter, international, randomized double-blind trial. *American Journal of Perinatology*, 37(2):127–136, 2020.

CM Booth and IF Tannock. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *British journal of cancer*, 110(3): 551–555, 2014.

BRUDAC Meeting. Fda briefing document nda 021945: Hydroxyprogesterone caproate injection (trade name makena), 2019a.

BRUDAC Meeting. Transcript for the october 29, 2019 meeting of the bone, reproductive and urologic drugs advisory committee (brudac)., 2019b.

Pierre Buekens and Mark Klebanoff. Preterm birth research: from disillusion to the search for new mechanisms. *Paediatric and Perinatal Epidemiology*, 15:159–161, 2001.

Pranom Buppasiri, Pisake Lumbiganon, Jadsada Thinkhamrop, Chetta Ngamjarus, Malinee Laopaiboon, and Nancy Medley. Calcium supplementation (other than for preventing or treating hypertension) for improving pregnancy and infant outcomes. *Cochrane Database of Systematic Reviews*, 2015.

J Christopher Carey, Mark A Klebanoff, John C Hauth, Sharon L Hillier, Elizabeth A Thom, JM Ernest, R Phillip Heine, Robert P Nugent, Molly L Fischer, Kenneth J Leveno, et al. Metronidazole to prevent preterm delivery in pregnant women with asymptomatic bacterial vaginosis. *New England Journal of Medicine*, 342(8):534–540, 2000.

Code of Federal Regulations. Title 21: Part 211 current good manufacturing practice for finished pharmaceuticals, 1985.

Janet Currie. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of economic literature*, 47(1):87–122, 2009.

Janet Currie. Inequality at birth: Some causes and consequences. *American Economic Review*, 101(3):1–22, 2011.

Janet Currie and Douglas Almond. Human capital development before age five. In *Handbook of labor economics*, volume 4, pages 1315–1486. Elsevier, 2011.

Janet Currie, Michael Mueller-Smith, and Maya Rossin-Slater. Violence while in utero: The impact of assaults during pregnancy on birth outcomes. *National Bureau of Economic Research Working Papers*, 2019.

Angus Deaton and Nancy Cartwright. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210:2–21, 2018.

Richard A Deyo. Gaps, tensions, and conflicts in the fda approval process: implications for clinical practice. *The Journal of the American Board of Family Practice*, 17(2):142–149, 2004.

William H Dow, Tomas J Philipson, and Xavier Sala-i Martin. Longevity complementarities under competing risks. *American Economic Review*, 89(5):1358–1371, 1999.

David Figlio, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. The effects of poor neonatal health on children's cognitive development. *American Economic Review*, 104 (12):3921–55, 2014.

Food and Drug Administration. Orphan drug designations and approvals, 2007.

Paula M Frew, Diane S Saint-Victor, Margaret Brewinski Isaacs, Sonnie Kim, Geeta K Swamy, Jeanne S Sheffield, Kathryn M Edwards, Tonya Villafana, Ouda Kamagate, and Kevin Ault. Recruitment and retention of pregnant women into clinical research trials:

an overview of challenges, facilitators, and best practices. *Clinical Infectious Diseases*, 59 (suppl_7):S400–S407, 2014.

Steven K Galson. Federal and state role in pharmacy compounding and reconstitution: exploring the right mix to protect patients. In *hearing before the Senate Committee on Health, Education, Labor, and Pensions, 108th Congress (October 23, 2003)*, 2003.

Kathleen Gifford, Anne Winter, Linda Wiant, Rachel Dolan, Marina Tian, and Rachel Garfield. How state medicaid programs are managing prescription drug costs: Results from a state medicaid pharmacy survey for state fiscal years 2019 and 2020. *Kaiser Family Foundation. Retrieved September 2020*, 2020.

Robert L Goldenberg and Dwight J Rouse. Prevention of premature birth. *New England Journal of Medicine*, 339(5):313–320, 1998.

Government Accountability Office. Drug compounding: Clear authority and more reliable data needed to strengthen fda oversight: July 2013. Technical report, United States Government Accountability Office, 2013.

Jennifer Gudeman, Michael Jozwiakowski, John Chollet, and Michael Randell. Potential risks of pharmacy compounding. *Drugs in R&d*, 13(1):1–8, 2013.

Margaret Harper, Elizabeth Thom, Mark A Klebanoff, John Thorp Jr, Yoram Sorokin, Michael W Varner, Ronald J Wapner, Steve N Caritis, Jay D Iams, Marshall W Carpenter, et al. Omega-3 fatty acid supplementation to prevent recurrent preterm birth: a randomized controlled trial. *Obstetrics and gynecology*, 115(2 0 1):234, 2010.

John C Hauth, Rebecca G Clifton, James M Roberts, Catherine Y Spong, Leslie Myatt, Kenneth J Leveno, Gail D Pearson, Michael W Varner, John M Thorp Jr, Brian M Mercer, et al. Vitamin c and e supplementation to prevent spontaneous preterm birth. *Obstetrics and gynecology*, 116(3):653, 2010.

Jay D Iams. Prevention of preterm parturition. *New England Journal of Medicine*, 370(3): 254–261, 2014.

L Michele Issel, Sarah G Forrestal, Jaime Slaughter, Anna Wiencrot, and Arden Handler. A review of prenatal home-visiting effectiveness for improving birth outcomes. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 40(2):157–165, 2011.

Theodore Joyce. Impact of augmented prenatal care on birth outcomes of medicaid recipients in new york city. *Journal of Health Economics*, 18(1):31–67, 1999.

Robert Kaestner, Michael Darden, and Darius Lakdawalla. Are investments in disease prevention complements? the case of statins and health behaviors. *Journal of health economics*, 36:151–163, 2014.

Mark A Klebanoff, J Christopher Carey, John C Hauth, Sharon L Hillier, Robert P Nugent, Elizabeth A Thom, JM Ernest, R Phillip Heine, Ronald J Wapner, Wayne Trout, et al. Failure of metronidazole to prevent preterm delivery among pregnant women with asymptomatic trichomonas vaginalis infection. *New England Journal of Medicine*, 345(7): 487–493, 2001.

Michael S Kramer. Intrauterine growth and gestational duration determinants. *Pediatrics*, 80(4):502–511, 1987.

Darius Lakdawalla, Neeraj Sood, and Dana Goldman. Hiv breakthroughs and risky sexual behavior. *The Quarterly Journal of Economics*, 121(3):1063–1102, 2006.

Yan Li, Yanmin Zhu, Chao Chen, Xi Wang, Yoonyoung Choi, Carl Henriksen, and Almut G Winterstein. Internal validation of medicaid analytic extract (max) data capture for comprehensive managed care plan enrollees from 2007 to 2010. *Pharmacoepidemiology and Drug Safety*, 27(10):1067–1076, 2018.

Nicholas T Longford and John A Nelder. Statistics versus statistical science in the regulatory process. *Statistics in medicine*, 18(17-18):2311–2320, 1999.

Michael C Lu and Neal Halfon. Racial and ethnic disparities in birth outcomes: a life-course perspective. *Maternal and child health journal*, 7(1):13–30, 2003.

Anne Rossier Markus, Shannon Krohe, Nicole Garro, Maya Gerstein, and Cynthia Pellegrini. Examining the association between medicaid coverage and preterm births using 2010–2013 national vital statistics birth data. *Journal of Children and Poverty*, 23(1):79–94, 2017.

Joyce A Martin, MJ Osterman, Sharon E Kirmeyer, and EC Gregory. Measuring gestational age in vital statistics data: transitioning to the obstetric estimate. *National Vital Statistics Reports: From the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, 64(5):1–20, 2015.

Paul J Meis, Mark Klebanoff, Elizabeth Thom, Mitchell P Dombrowski, Baha Sibai, Atef H Moawad, Catherine Y Spong, John C Hauth, Menachem Miodovnik, Michael W Varner, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. *New England Journal of Medicine*, 348(24):2379–2385, 2003.

Brahmajee K Nallamothu, Rodney A Hayward, and Eric R Bates. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. *Circulation*, 118(12):1294–1303, 2008.

Steven Offenbacher, James D Beck, Heather L Jared, Sally M Mauriello, Luisto C Mendoza, David J Couper, Dawn D Stewart, Amy P Murtha, David L Cochran, Donald J Dudley, et al. Effects of periodontal therapy on rate of preterm delivery a randomized controlled trial. *Obstetrics and gynecology*, 114(3):551, 2009.

Office of Inspector General. Medicaid managed care encounter data: collection and use. oei-07-06-00540., 2009.

Office of Inspector General. Not all states reported medicaid managed care encounter data as required. oei-07-13-00120, 2015.

Philip Oreopoulos, Mark Stabile, Randy Walld, and Leslie L Roos. Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. *Journal of Human Resources*, 43(1):88–138, 2008.

Yesha Patel and Martha M Rumore. Hydroxyprogesterone caproate injection (makena) one year later: to compound or not to compound that is the question. *Pharmacy and Therapeutics*, 37(7):405, 2012.

Public Law. 42 u.s.c. chapt. 7, subchapter 19, section 1396r–8. payment for covered outpatient drugs, 1993.

Public Law 105-115. Food and drug administration modernization act of 1997, 1997.

Public Law 113-54. Drug quality and security act, 2013.

LeeAnne Roman, Jennifer E Raffo, Qi Zhu, and Cristian I Meghea. A statewide medicaid enhanced prenatal care program: impact on birth outcomes. *JAMA pediatrics*, 168(3): 220–227, 2014.

Peter M Rothwell. Factors that can affect the external validity of randomised controlled trials. *PLoS clinical trials*, 1(1):e9, 2006.

Heather Royer. Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85, 2009.

Sarah Sellers and Wulf H Utian. Pharmacy compounding primer for physicians. *Drugs*, 72 (16):2043–2050, 2012.

Michael M Slattery and John J Morrison. Preterm delivery. *The Lancet*, 360(9344):1489–1497, 2002.

Society for Maternal Fetal Medicine Publications Committee. Acog committee opinion number 419 october 2008 (replaces no. 291, november 2003). use of progesterone to reduce preterm birth. *Obstetrics and gynecology*, 112(4):963, 2008.

Jane Tilly and Linda Elam. *Prior authorization for Medicaid prescription drugs in five states: lessons for policy makers*. Kaiser Commission on Medicaid and the Uninsured, 2003.

Nicholas L Wilson, Wentao Xiong, and Christine L Mattson. Is sex like driving? hiv prevention and risk compensation. *Journal of Development Economics*, 106:78–91, 2014.