PregMedNet: Multifaceted Maternal Medication Impacts on Neonatal Complications

Yeasul Kim[1,2,3], Ivana Marić[1,2,3], Chloe M. Kashiwagi[1,2,4], Lichy Han[1], Philip Chung[1], Jonathan D Reiss[2], Lindsay D. Butcher[2], Kaitlin J. Caoili[2], Eloïse Berson[1,3,5], Lei Xue[1,2,3], Camilo Espinosa[1,2,3,4], Tomin James[1,2,3], Sayane Shome[1,2,3], Feng Xie[1,2,3], Marc Ghanem[1,2,3], David Seong[1,4,6], Alan L Chang[1,2,3], S Momsen Reincke[1,2,3], Samson Mataraso[1,2,3], Chi-Hung Shu[1], Davide De Francesco[1,2,3], Martin Becker[7], Wasan M Kumar[8,9], Ron Wong[2], Brice Gaudilliere[1,2], Martin S Angst[1], Gary M Shaw[2], Brian T Bateman[1], David K Stevenson[2], Lance S Prince[2], Nima Aghaeepour[1,2,3,4]

[1]Department of Anesthesiology, Perioperative and Pain Medicine, Stanford School of Medicine
[2]Department of Pediatrics, Stanford School of Medicine
[3]Department of Biomedical Data Science, Stanford University
[4]Immunology Program, Stanford University School of Medicine, Stanford, CA, USA
[5]Department of Pathology, Stanford University School of Medicine, Stanford, CA ,USA
[6]Medical Scientist Training Program, Stanford University School of Medicine, Stanford, CA, USA
[7]Institute for Visual and Analytic Computing, University of Rostock, Rostock, Germany
[8]Medical Doctor Program, Stanford University School of Medicine, Stanford, CA, USA
[9]Graduate School of Business, Stanford University School of Medicine, Stanford, CA, USA

Abstract

While medication intake is common among pregnant women, medication safety remains underexplored, leading to unclear guidance for patients and healthcare professionals. PregMedNet addresses this gap by providing a multifaceted maternal medication safety framework based on systematic analysis of 1.19 million mother-baby dyads from U.S. claims databases. A novel confounding adjustment pipeline was applied to systematically control confounders for multiple medication-disease pairs, robustly identifying both known and novel maternal medication effects. Notably, one of the newly discovered associations was experimentally validated, demonstrating the reliability of claims data and machine learning for perinatal medication safety studies. Additionally, potential biological mechanisms of newly identified associations were generated using a graph learning method. These findings highlight PregMedNet's value in promoting safer medication use during pregnancy and maternal-neonatal outcomes.

Introduction

More than 80% of women use at least one medication during pregnancy, and the use of prescription medication has increased by over 60% over the last three decades.(1–3) However, medication safety is still underexplored in pregnant women, and pregnant women and their healthcare teams frequently encounter limited or contradictory guidance on medication safety.(4) More than one-third of newly approved drugs by the United States (US) Food and Drug Administration (FDA) during the last decade are out of compliance with the Pregnancy and Lactation Labeling Rule (PLLR), and less than 20% have adequate drug safety data on pregnancy and lactation based on human studies.(5)

One of the main reasons for the lack of medication safety evidence in pregnant women is the recognition of pregnant women as a vulnerable population and regulatory guidance has prioritized decreasing harm. Premarketing clinical trials for drug safety generally poses greater than minimal risk to pregnant women, which conflicts with FDA's and institutional review board's priority to provide special protections for pregnant women.(6–8) Because of this, pregnant women have been considered "therapeutic orphans" due to the lack of adequate drug safety and efficacy data.(9–12) The current drug safety data of pregnancy have been collected mainly through post-marketing surveillance,(13) especially through spontaneous reporting systems (SRSs).(14) Although SRSs are regarded as an effective tool for drug safety surveillance, they are potentially biased by under-reporting, may not be able to assess low-level risks, and are challenged to definitively quantify risks and side effects.(13–16)

Computational and data-driven approaches to assess drug safety have recently become viable fueled by rapid digitization of patient data resulting in large healthcare databases, and the development of machine learning.(10) Large administrative claims databases are increasingly being used to assess drug safety because they contain extensive samples of real-world patient information. Unlike electronic health records (EHR) which are typically confined to a single hospital system, claims databases include patients across institutions and geographic regions, thereby providing large enough sample sizes, rich information on potential confounders, and analytic cost-effectiveness. This makes them an invaluable tool for analysis of medication effects, particularly in underrepresented populations such as pregnant women. Previous studies have demonstrated that claims databases can be successfully utilized to evaluate drug safety in pregnant women, by developing algorithms to establish a pregnant cohort and applying statistical analysis.(17–20) However, these studies have typically concentrated on the effects of one or a few maternal medications on pregnancy outcomes.

We introduce PregMedNet, a platform dedicated to advancing maternal medication safety during pregnancy, with particular focus on neonatal complications. The impacts of multiple maternal medications used during pregnancy on neonatal outcomes were assessed based on the Merative™ Marketscan® Commercial Database,(21) a comprehensive collection of claims data across the U.S. covering more than 250 million patients. Utilizing retrospective, large-scale mother-baby dyads, the multifaceted maternal medication impacts during pregnancy on neonatal complications were calculated with high statistical power.

First, we analyzed the associations between 1152 maternal medications and 24 neonatal complications, encompassing all medications available in the database. To minimize bias, we then implemented a novel confounding adjustment pipeline, which automatically selects and adjusts for confounding factors without requiring manual selection for each medication-complication pair. Additionally, we analyzed drug-drug interactions (DDIs), uncovering additive effects of maternal drug combinations on neonatal complications. In total, we identified 1,446 odds ratios, 261 adjusted odds ratios, and 5 DDIs, capturing both known and novel maternal medication effects on neonatal complications.

Notably, we experimentally validated a newly identified association between maternal ondansetron use and neonatal bronchopulmonary dysplasia (BPD) *in vivo*, demonstrating the reliability of retrospective claims data and machine learning for maternal medication safety research.

Finally, we employed a knowledge graph and a graph learning method to uncover potential biological mechanisms underlying newly discovered associations. A case study of ondansetron and

BPD demonstrated how this approach generates testable hypotheses, aiming to expedite mechanism of action (MoA) discovery by narrowing down specific molecular targets for further investigation.

Our comprehensive findings on maternal medication impacts and underlying mechanisms can be explored through an interactive website: http://pregmednet.com.

Results

Mother-baby cohort: encompassing a wide-range of variables

A total of 1,190,257 mother-baby dyads were obtained between 2007 and 2021 from the Merative™ Marketscan® Commercial Database(21) (Supplementary Figures 1 & 2). The average maternal age was 31.18 ± 4.75 years, and the average gestational age was 38.58 ± 1.71 weeks with a preterm birth rate of 8.98% (Supplementary Table 1-1). The prevalence of each neonatal complication and maternal comorbidity is presented in Supplementary Table 2.

The discovery cohort included 817,402 dyads, used as the primary group for initial analysis and hypothesis generation. The validation cohort included 372,855 dyads, used to validate the hypotheses generated. Each cohort contained information about pregnant mothers in the 90 days prior to delivery, including 68 maternal comorbidities (Supplementary Figure 3), preterm birth, gestational age at birth, maternal age at birth, the region and geographic location of the healthcare facilities, and all maternal medications taken during the 90 days. Additionally, the cohort included information about the corresponding babies during the first 90 days after birth, covering both sex and incidence of 24 neonatal complications (Figure 1.c). All maternal medications taken were extracted and harmonized to the medications' compound-level identity along with route of administration information (Supplementary Figure 4). There were a total of 1,152 maternal medications in the discovery cohort and 980 in the validation cohort.

Multifaceted insights into perinatal medication safeties

PregMedNet offers comprehensive insights into the impacts of maternal medications on neonatal complications, examining (a) effects of single medications, (b) impacts of concomitant medications on neonatal outcomes, and (c) potential underlying mechanisms behind significant associations.

Individual maternal medication impacts on newborn outcomes

The associations between maternal medications and neonatal complications were investigated by estimating both unadjusted and adjusted odds ratios (ORs and aORs, respectively). Out of 27,648 medication-complication pairs, 1,446 statistically significant ORs (5.23%) and 261 statistically significant aORs (0.94%) were identified, capturing both already known and novel maternal medication effects on neonatal complications.

Network graphs visualize intricate patterns of associations

PregMedNet illustrates these associations using network graphs, with ORs presented in Figure 2 and aORs in Supplementary Figure 5. Medications and diseases are represented as nodes clustered by groups, with edges color-coded to indicate increased risk (gray, OR or aOR>1) or reduced risk (blue, OR or aOR<1) associations. The network graphs highlight intricate patterns between medications and diseases.

In the OR network graph (Figure 2), 1,153 edges (79.7%) are gray, while 293 edges (20.3%) are blue, out of a total of 1,446 edges. Focusing on diseases, neonatal respiratory diseases, particularly Respiratory Distress Syndrome (RDS) (140 gray edges) and BPD (135 gray edges), exhibit the most increased risk associations with maternal medications. In contrast, Postmaturity (87 blue edges) and Large for Gestational Age (49 blue edges) show the most reduced risk associations. Focusing on maternal medication groups, the Central Nervous System (CNS) group has the highest number of increased risk associations (300 gray edges), followed by the Hormone and Synthetic Substances group (212 gray edges). Conversely, the Vitamins & Combinations group shows the highest number of reduced risk associations (151 blue edges), followed by the Anti-Infective Agents group (40 blue

edges). Supplementary Figure 6 provides detailed information on the number of significant ORs per neonatal complication and maternal medication group.

In contrast, the aOR network graph (Supplementary Figure 5) reveals a shift toward decreased risk associations, with 117 edges (44.8%) gray and 144 edges (55.2%) blue out of 261 total edges. Focusing on diseases (Figure 3.a.), patterns with aORs align with ORs: neonatal respiratory complications show the highest increased risk associations, led by BPD (24 gray edges), followed by Other Respiratory (21 gray edges) and RDS (19 gray edges), while Postmaturity (30 blue edges) and large-for-gestational age (LGA) (23 blue edges) have the most decreased risk associations. Similarly, focusing on maternal medication groups (Figure 3.b.), CNS medications show the highest increased risk associations (50 gray edges), followed by Cardiovascular Agents and Hormone and Synthetic Substances (14 gray edges for each), while Vitamins & Combinations show the most decreased risk associations (102 blue edges), followed by Anti-Infective Agents (13 blue edges). Supplementary Figure 7 provides detailed counts of significant aORs per neonatal complication and maternal medication group.

Detailed information on associations: highlighted examples of ORs

PregMedNet provides detailed ORs and aORs for all medication-disease pairs through an interactive website, presenting both significant and non-significant results. Supplementary Figure 8 highlights key examples among 1,446 significant ORs.

Notably, within the CNS category, antidepressants and opioids are associated with elevated ORs for several neonatal complications, such as neonatal respiratory disorders and neonatal abstinence syndrome (NAS) (Supplementary Figures 8.a, 8.b). Similarly, antiemetics - commonly used during pregnancy(22) - are associated with elevated ORs for multiple neonatal complications (Supplementary Figure 8.c). In contrast, long-chain polyunsaturated fatty acids (LC-PUFAs) in the Vitamins & Combinations category are highlighted here, showing substantial reduced risk associations for multiple neonatal complications including neonatal respiratory disorders (Supplementary Figure 8.d). Medications in the Anti-Infective Agents category also exhibit reduced risk associations (Supplementary Figure 8.e), with oral valacyclovir linked to reduced risks across multiple neonatal complications, consistent with prior studies.(17)

aORs robustly obtained by systematic confounder adjustment

While ORs provide valuable insights, they are unadjusted for potential confounders, which could introduce bias or spurious associations. In PregMedNet, a method that systematically selects and adjusts for confounding variables across multiple disease-medication pairs was employed, robustly obtaining aORs. In particular, we focus on four key medication groups for detailed discussion in the following section: ondansetron, antidepressants, LC-PUFAs, and opioids. (Figure 3.c., 3.d.)

Ondansetron linked to increased neonatal respiratory risks

In PregMedNet, ondansetron is categorized into three types based on the route of administration: oral, oromucosal, and injection. Among these, orally administered ondansetron exhibits six statistically significant aORs, including NAS, three elevated risk associations for neonatal respiratory complications (RDS, BPD, and other respiratory diseases) and two reduced risk associations for postmaturity and LGA.

Ondansetron is a medication used to treat nausea and vomiting during pregnancy, with a prevalence exceeding 25% in the US.(23) The American College of Obstetricians and Gynecologists (ACOG) recommends cautious use of ondansetron during pregnancy due to insufficient safety data, as current studies are limited by small sample sizes and potential recall bias (24). The European Medicines Agency's Pharmacovigilance Risk Assessment Committee (EMA PRAC) advises against the use of ondansetron during the first trimester of pregnancy due to recent studies that have found some associations with congenital malformations.(25,26) Our study, based on a large-scale retrospective database, addresses some of these limitations and found that maternal use of ondansetron during the

last 90 days before delivery is associated with neonatal complications, particularly those related to the neonatal respiratory system.

Neonatal respiratory risks reconfirmed in perinatal SSRI use

Four SSRIs — citalopram, escitalopram, fluoxetine, and sertraline, all administered orally — show increased associations with neonatal respiratory complications (BPD, other respiratory disorders, transient tachypnea of newborn [TTN]) in PregMedNet, aligning with previous research results.(27,28)

Approximately 10% of pregnant women experience depression and 2-3% take antidepressants, yet their use during pregnancy is controversial due to potential adverse effects on neonates.(29) Serotonin reuptake inhibitors (SSRIs) are the most commonly used antidepressants during pregnancy, because they have a more favorable safety profile than other antidepressant drug classes. However, findings associate them with neonatal complications,(29–31) including an increased risk of respiratory complications in neonates.(27,28) PregMedNet reconfirms the increased risk of neonatal respiratory complications in perinatal SSRI-exposed babies based on aORs. Specifically, PregMedNet is able to identify associations at the level of individual medications rather than SSRI medications as a whole, thus providing more granular safety information for each medication.

LC-PUFAs linked to reduced risks across neonatal outcomes

LC-PUFAs, categorized under Vitamins & Combinations, demonstrate the most significant reduced risk associations with neonatal complications, such as BPD, other respiratory diseases, and TTN, with six out of seven LC-PUFAs showing reduced risk effects.

Fish oil and linoleic acid are particularly noteworthy for their reduced risk associations, each displaying nine reduced risk aORs related to various neonatal complications, including disorders of the respiratory system (BPD, other respiratory diseases, TTN), gestational age or size-related disorders (both LGA and SGA, postmaturity), hypoglycemia, jaundice, and neonatal death.

Previous studies have shown the associations between maternal LC-PUFAs levels during pregnancy and maternal health, including preeclampsia and postpartum depression,(32,33) as well as their association with children's health, including neurodevelopmental outcomes.(34) Furthermore, recent research suggests that LC-PUFAs may also decrease the incidence of asthma and wheezing in infants, although this evidence remains limited.(35) PregMedNet highlights the perinatal benefits of LC-PUFAs in addressing neonatal respiratory disorders, reinforcing previous findings while uncovering previously unknown associations, such as hypoglycemia and neonatal death.

Opioid use associated with multiple neonatal complications

As many as 6.6% of pregnant women use prescription opioids and there is concern these opioids can cause neonatal complications.(36) PregMedNet found associations between opioids and increased risks of serious neonatal complications such as BPD, RDS, NAS, and jaundice (Figure 3.c., 3.d.).

The medications highlighted in this section reveal both previously reported and novel associations as ORs and aORs, demonstrating PregMedNet's reliability and efficiency in identifying associations between maternal medications and neonatal complications. While this section highlights selected medication groups, comprehensive results, including all 1,446 ORs and 261 aORs, are available on our interactive website.

Systematic Analysis of Maternal Drug-Drug Interactions

PregMedNet also captures the additive effects of concurrent use of maternal medications on neonates through a systematic analysis of Drug-Drug Interactions (DDIs) for all possible maternal medications pairs out of 1,152 medications and their effects on 24 neonatal complications. OR_1, OR_2, and OR_{12} represent the ORs of medication 1, medication 2, and the concurrent use of two medications,

respectively, while β_3 represents the contribution of the DDI to the overall associations between the concurrent use of medications and the neonatal complications. Negative β_3 suggests that the interaction reduces the risk of neonatal complications, while the positive β_3 suggests it increases the risk. The magnitude of association is proportional to the absolute value of β_3. (Figure 4)

A total of 72 significant DDIs are identified in the discovery cohort and 15 in the validation cohort with 5 DDIs demonstrating significant β_3 and OR_{12} values in both the discovery and validation cohorts. All five medication pairs exhibit negative β_3 values, indicating synergy in reducing neonatal complications beyond the effects of each medication alone (Table 1).

The concurrent use of oral oxycodone and oral acetaminophen reduces NAS compared to using each medication alone. The DDI effect for this combination on NAS was β_3=-3.50, with OR_{12}=14.77, which was lower than the odds ratios for each medication alone (oxycodone: OR_1=99.97, acetaminophen: OR_2 =4.89). Prior research has shown that combining oxycodone and NSAIDs - another class of analgesic medications - improves pain management over oxycodone alone.(37) Our findings further highlight the added benefit of this combination regimen for pregnant women and their neonates, specifically in reducing the odds of NAS when both medications are used together.

Another medication pair to highlight is the neonatal benefits of concurrent use of maternal insulins. The DDI effect of combined use of maternal Insulin Human Isophane (NPH) (subcutaneous) and Insulin lispro (injection) for neonatal hypoglycemia is β_3 = -1.84, and the OR_{12}= 2.03, which is lower than single medication use (NPH: 2.19, Insulin lispro: 5.97). Similarly, the DDI effect of the combination of NPH (subcutaneous) and Insulin aspart (injection) on hypoglycemia is β_3=-1.52 with OR_{12} = 2.55, which is lower than the odds observed for Insulin Aspart alone (NPH: OR_1 =2.14, Insulin Aspart: OR_2 =5.42). These findings suggest that basal-bolus insulin regimens for the management of gestational diabetes (GDM),(38) result in superior maternal glucose control compared to using a single insulin regimen. Improved glucose regulation minimizes in-utero fetal hyperinsulinemia, a common complication in GDM, ultimately reducing the risk of neonatal hypoglycemia at birth.

Lastly, there are two additional medication pairs which show significant β_3 in both the discovery cohort and the validation cohort: Buprenorphine (sublingual) and ondansetron (oral) for NAS (β_3=-1.89), and NPH (subcutaneous) and Insulin aspart (injection) for RDS (β_3=-1.02).

The DDI analysis conducted through PregMedNet shows that certain well-known medication pairs provide additional benefits for neonatal complications when used concurrently in mothers. Although we have spotlighted five medication pairs with overlapping significance in both the discovery and validation cohorts, it is important to note that DDIs identified exclusively in one cohort still hold value. Further details on the additional DDIs are available on our interactive website.

Validation of Single and Concomitant Medication Effects

The ORs and aORs, along with DDIs results from the discovery cohorts, were validated in the validation cohort using the same analysis. Scatter plots and Pearson correlation coefficients (Figure 2.b and Supplementary Figure 9) show high correlations between the discovery cohorts and the validation cohorts: R^2 = 0.972 for ORs, R^2 = 0.747 for aORs, and R^2 = 0.920 for DDIs. These results confirm maternal medication associations in the discovery cohort were replicable in the validation cohort.

Maternal ondansetron impairs neonatal lung development *in vivo*

To further evaluate the robustness of the newly discovered associations, we experimentally investigate the novel association, the effect of maternal ondansetron on fetal and neonatal lung development *in vivo*. Ondansetron was of particular interest due to its high usage during pregnancy and uncertain perinatal safety.(24,25) Our analysis showed that maternal ondansetron use was associated with increased odds of neonatal BPD (aOR = 1.16, p-value = 1.18e-6), warranting further investigation.

Timed pregnant C57BL/6 female mice were dosed daily with 5 mg/kg bodyweight ondansetron by intraperitoneal injection from embryonic day 15 (E15) until delivery. Lung histology was assessed on postnatal day 11 (P11) to test if ondansetron exposure might impact saccular and alveolar stage lung development (Figure 5). Maternal ondansetron exposure did not affect neonatal survival or weight gain (Figures 5 b,c). However, the lungs of ondansetron exposed mice had increased alveolar mean linear intercept, indicative of reduced alveolar septation and interstitial septal thinning (39) (Figure 5d,e). These data support the connection between maternal ondansetron exposure during pregnancy and subsequent neonatal lung abnormalities.

Knowledge Graph Reveals Potential Underlying Mechanisms

PregMedNet identifies potential underlying mechanisms of the newly discovered associations between maternal medications and neonatal complications. This was achieved by combining a knowledge graph(40) and subsequently applying a graph learning method for the 261 significant aORs, which drew associations between 19 neonatal complications and 68 maternal medications. Of these, 15 neonatal complications and 63 maternal medications were mapped from PregMedNet to corresponding nodes in the knowledge graph, resulting in 153 disease-medication pairs for which mechanistic hypotheses are tested.

Three types of subgraphs were constructed: (a) one containing all biological interactions, (b) one containing all biological interactions without other diseases and medications involved, and (c) one with only proteins/genes between the medication and disease pair. Hypothesis testing involved extracting the shortest paths between the medication and the disease nodes and examining the involved biological nodes on these paths. This process was iterated across all 153 associations.

Notably, the mechanistic pathways underlying maternal ondansetron's effect on neonatal BPD are highlighted here as a case study. PregMedNet generates plausible hypotheses involving interleukin-1b (IL-1b), known as a key factor in BPD pathophysiology.(41,42) Hypotheses based on the protein/gene-only subgraph (Figure 6) are particularly insightful, as they only include direct molecular interactions, avoiding non-specific biological entities (e.g., generic "protein bindings"), while the full biological interaction subgraphs (Figure 10, 11) generates a broader range of hypotheses.

PregMedNet suggests that the shortest paths between ondansetron and BPD consist of 4-hop paths, meaning it takes four edges to reach the other side of the graph (Figure 6.a.). A total of 4 shortest paths were identified, involving 7 intermediary nodes. Each path highlights potential mechanisms, visually represented by the edges between the nodes in the graph.

One of these paths suggests that ondansetron and BPD are connected via an interaction between IL-1b and Cytochrome P450 3A5 (Figure 6.b). While there are prior research studies linking IL-1b to arrested lung development in both human and experimental BPD(41,42), ondansetron to multiple forms of cytochrome P450 (CYP450)(43,44), and CYP450 to IL-1b (45), to the best of our knowledge, no prior research has investigated the mechanism by which maternal ondansetron might impact neonatal BPD risk. As such, PregMedNet, in this instance, generates testable hypotheses by connecting proven yet isolated biological knowledge using graph learning methods. This approach reduces the hypotheses from countless possibilities to 4 specific ones to be tested.

While we highlighted one of the potential MoA here, all 153 associations and their potential mechanisms are available in our interactive website.

Discussion

PregMedNet provides a comprehensive maternal medication safety framework for previously underrepresented populations—pregnant women and neonates. Its contributions include: (a) uncovering maternal medication effects on neonatal complications, (b) systematically adjusting for confounding variables across multiple medication-disease pairs, (c) experimentally validating the reliability of generated hypotheses, and (d) identifying potential mechanisms of action (MoA) using large-scale, real-world mother-baby dyads and validated by *in vivo* experimentation.

PregMedNet's systemic analysis of multiple maternal medications and neonatal complications was achieved by developing mother-baby dyads using retrospective claims data. We advanced our previous algorithms(17) for linking mother and baby records, estimating gestational age and preterm birth, and incorporating multiple disease features based on ICD codes. While there are previous studies using claims databases,(46) most have focused on short-term outcomes like live birth and stillbirth(17,19,20,47,48) or examined a few maternal medication and neonatal complication pairs using traditional statistical testing.(26) In contrast, PregMedNet is able to analyze all noted maternal medications' impacts on more detailed neonatal outcomes, including 24 different complications. PregMedNet also studies maternal medication effects at finer resolution by distinguishing between drug routes of administration. This approach differentiates distinct use cases for the same medication, such as topical betamethasone for skin inflammation and intramuscular betamethasone for preventing complications in preterm labor.

Another key contribution of PregMedNet is its framework for selecting and adjusting confounding variables for each maternal medication and neonatal complication pair, ensuring robust associations and eliminating the need for manual selection — an approach well-suited for large-scale claims data, as demonstrated by the 27,648 associations in our study. PregMedNet successfully identifies both established and novel associations of maternal medications on neonatal complications. Notably, it demonstrates its effectiveness by reconfirming previously known associations, such as the increased risk of neonatal respiratory complications in babies exposed to SSRIs perinatally(27) and the increased risk of NAS in babies exposed to opioids perinatally.(49) As demonstrated, PregMedNet exhibits the capability of a confounding adjustment framework in systematic analyses, which can be extended to a wide range of research studies, including those utilizing claims data or EHR.

Furthermore, PregMedNet is able to examine the additive effects of concurrent use of maternal medications on neonatal complications through retrospective DDIs analysis. Notably, it reveals that certain known medication pairs (acetaminophen/oxycodone, NPH/short-acting insulins) provide possibly synergistic reduced associations for neonatal complications. While there have been previous DDI research studies utilizing EHR, these studies have typically concentrated on already known side effects,(50,51) being based on data from a single institution,(50,52) or mainly focused on a limited number of medication pairs. (51,53,54) In contrast, PregMedNet agnostically tested the effects of drug combinations without relying on prior knowledge, using a large-scale claims database from multiple institutions. This approach allows for the discovery of novel interactions and ensures unbiased findings, which is particularly important as polypharmacy in pregnant women has risen during the last three decades.(1) To the best of our knowledge, this represents the first attempt to screen the effects of maternal drug combinations on neonates based on real-world, retrospective, large-scale claims data from multiple institutions.

Notably, PregMedNet demonstrated one of the hypotheses–maternal ondansetron effects on neonatal BPD – through laboratory experimentation. Our preliminary *in vivo* experiments in mice demonstrated the effects of maternal ondansetron administration on neonatal lung development. Increased mean linear intercept (MLI) could reflect defects in postnatal alveolar septation. Importantly, maternal ondansetron did not alter neonatal survival or weight gain, which can have a secondary effect on lung development.

Ondansetron is a potent and selective antagonist of the 5-HT3A serotonin receptor (55). In the developing human lung, the 5-HT3A receptor is expressed on epithelial, endothelial, and smooth

muscle cells throughout the canalicular, saccular, and alveolar stages of lung development (56). However, the role of serotonin signaling through the 5-HT3A receptor in guiding normal lung morphogenesis is not yet understood. The increased MLI observed in ondansetron-treated mice, as predicted by PregMedNet's hypothesis, suggests that the 5-HT3A receptor may play a role in normal lung morphogenesis.

As such, our *in vivo* experimental validation of a novel medication effect discovered by PregMedNet demonstrates its reliability. While previous research studies have investigated maternal medication safety using claims databases,(17,19,20,26,46–48) to the best of our knowledge, this is the first study to rigorously validate such hypotheses through *in vivo* experimentation. Notably, this experimentation also generated follow-up hypotheses regarding the previously unknown role of the 5-HT3A receptor in lung development, further highlighting PregMedNet's potential to provide a scientific rationale for subsequent hypothesis-driven research. This validation underscores the broader applicability of retrospective claims databases and machine learning approaches for generating reliable hypotheses, extending beyond maternal medication safety. Further investigations into other PregMedNet-generated hypotheses will elucidate mechanisms and identify potential alternatives or treatments.

Understanding a drug's MoA is critical for developing new treatments and preventing the side effects.(57,58) However, traditional approaches to uncovering MoAs are complex and resource-intensive tasks, spanning years, necessitates in-depth hypotheses for testing.(58) PregMedNet generates hypotheses of MoAs for newly discovered associations that have never been tested before by integrating real-world associations and a knowledge graph. It narrows down the number of molecular targets that need to be tested, significantly reducing the time required for further research and validation. A case study with ondansetron and BPD showcased the plausibility of the hypothesis. Although there are previous research studies using graph learning methods to identify drug treatment mechanisms(59) and polypharmacy drug side effects,(59) these efforts have relied solely on biological interactions from publicly available databases. PregMedNet distinguishes itself by connecting associations based on a large, real-world patient dataset to a knowledge graph to identify potential MoA. PregMedNet's capability of suggesting possible MoAs is expected to be useful to facilitate targeted scientific investigation to elucidate the underlying biological mechanism of action.

Our study was subject to several limitations inherent to the database, the Merative™ Marketscan® Commercial Database. These include the restricted inclusion of patients with private insurance, excluding those without coverage, and the focus on prescribed outpatient medications, which omits inpatient and over-the-counter medications. Critical demographic details such as race, smoking habits, alcohol use, weight, and height are also unavailable. Additionally, there are also limitations inherent to claims data, including challenges in excluding coding errors introduced during billing processes and difficulties in confirming patient compliance with prescribed medication regimens. The limited specificity of ICD and CPT codes presents further challenges, as rare diseases often lack dedicated codes and are grouped under broader categories. For example, in ICD-9, BPD is categorized under code 770.7, described as "Chronic Respiratory Disease arising in the perinatal period," while in ICD-10, it has a more specific code, P27.1, described as "Bronchopulmonary Dysplasia originating in the perinatal period". To address coding challenges and ensure accuracy, three clinicians reviewed the final list of codes after extracting the ICD and CPT codes. As such, we maximized the utility of the available data to provide robust insights into maternal medication safety despite these inherent limitations.

An additional limitation to consider is the complexity and variability of the confounding structure for each association, which may not be fully addressed in the systematic analysis provided by PregMedNet. This may lead to the inclusion of non-causal associations, such as confounding by the indication for the medication and its severity, both of which can vary between associations. To fully establish causal relationships, follow-up analyses aimed at causal inference that account for detailed confounding factors for each association will be necessary. Nonetheless, PregMedNet significantly narrowed the scope to 261 associations involving 68 maternal medications from an initial 27,648

associations, enabling prioritization for further causal investigation. Our robust experimental validation of a novel association further strengthens the reliability of our findings.

It is also important to consider that medications with statistically non-significant associations in our dataset are not necessarily safe; limited usage of medications already known to pose risks among expectant mothers may reduce the statistical power to assess these associations. We analyzed the frequency of medication use across Pregnancy Categories based on Australian categories for prescribing medicine in pregnancy(60) (see Supplementary Table 10), which shows that prescription frequency decreases as the risk category increases from A to X. Category X medications (high risk of causing permanent damage to the fetus) were prescribed to only 36 patients, whereas category A medications (drugs without any proven damage to the fetus) were prescribed to 213,019 patients. This distribution supports the notion that higher-risk medications are prescribed less frequently, resulting in less statistical power and explaining why they may not appear in the results. This trend is also reflected in the DDI results, where fewer statistically significant DDIs are observed due to small cohort sizes per drug combination. Thus, non-significant DDIs do not necessarily indicate a lack of effect. However, significant results from PregMedNet are highly reliable, providing risk scales for maternal medication effects on newborns that can guide decisions on maternal medication safety, a topic not heavily studied before.

Finally, while the knowledge graph used for MoA discovery aggregates comprehensive and verified biological associations, it lacks information on the relative strength and directionality of these associations and their specific cellular locations. Including this additional information in future directions may further improve the MoA hypotheses generated through graph learning.

PregMedNet offers a multifaceted and novel approach to investigating medication use during pregnancy and its effect on neonates, an area previously underexplored. Using retrospective mother-baby pairs from real-world data, the effects between maternal medications and newborn diseases were systematically investigated. A novel confounding adjustment pipeline was applied to systematically obtain the more robust associations. Notably, one of the newly discovered associations was experimentally validated, demonstrating the reliability of retrospective claims database and machine learning for the perinatal medication studies. Finally, hypotheses regarding biological mechanisms underlying newly discovered associations were generated. The extensive results of PregMedNet are available through an interactive website (http://pregmednet.com). We expect that PregMedNet will significantly contribute to advancing maternal and neonatal healthcare by providing insights that inform safer medication use during pregnancy and potentially reduce the risk of neonatal complications.

Methods

PregMedNet

PregMedNet is a platform designed to assess the safety of maternal medications during pregnancy and their impact on neonatal complications. PregMedNet provides a comprehensive analysis of the multifaceted associations between 1,152 maternal medications used during pregnancy and 24 neonatal complications, systematically examined retrospectively using large-scale claims data and a novel machine learning method. Notably, it experimentally validated one of the results, demonstrating the reliability of the PregMedNet. Additionally, possible underlying mechanisms of discovered correlations between diseases and medications are discovered using Knowledge Graph and graph learning methods. The entire results are open source and accessible through an interactive website (http://pregmednet.com).

1. Data Preparation

Mother and baby cohorts, along with their medical condition and medication histories, were extracted from the large-scale, de-identified claims data of the Merative™ Marketscan® Commercial Database.(21) This database houses de-identified medical, medication, and dental claim records for over 250 million patients in the US. We accessed the data through the Stanford Center for Population Health Sciences. A conceptual flow diagram of the data preparation process is shown in Supplemental Figure 1 with details in Supplemental Figure 2. The study was reviewed and approved by the international review board (IRB) at Stanford University (no. 39225).

1.a. Determining the Mother and Baby Cohort

The pregnancy cohort was created building upon methodology devised in our previous research involving mother-baby dyads.(17) We identified a cohort of women aged 12–55 years with singleton pregnancies using the International Classification of Diseases, Ninth Revision (ICD-9), ICD-10, and diagnosis-related group (DRG) codes from 2007-2021 (Supplementary Table 3) . We included only the first pregnancies to avoid dependencies among women with multiple pregnancies. Molar and ectopic pregnancies, as well as spontaneous and induced abortions, were excluded (Supplementary Table 4). Only mothers continuously enrolled in the insurance plan up to 90 days before the delivery date were included to ensure all maternal records are captured without any omissions.

The newborn babies' cohort was generated based on the newborn babies' codes from ICD-9 and ICD-10 from 2007-2021 (Supplementary Table 5). Only babies continuously enrolled in the insurance plan for 90 days after the delivery date are included to ensure newborn records are captured without any omissions.

1.b. Extracting Relevant Records: Medical Conditions & Medications

For each mother, all medical claims records from 90 days before the delivery date to the delivery date were extracted from both the inpatient and outpatient claims tables. From these records, 68 maternal conditions of interest (Supplementary Table 9) were extracted along with maternal age, gestational age at delivery, preterm birth, and the region and geographic location of the healthcare facility. All maternal medication records during the observation period were extracted from the prescription drug claims table as National Drug Codes (NDCs). These medications were further processed as described in the section 1.d. For each newborn, medical claims records from the delivery date up to 90 days post-delivery were extracted, and ICD codes were used to determine the presence of 24 neonatal complications of interest (Supplementary Table 8). The sex of each baby was also extracted. Maternal data and corresponding baby data were linked using a Family ID, indicating which baby was delivered by which mother (Figure 1.a).

1.c. Estimating Gestational Age and Preterm Birth

Gestational Age (GA) at delivery was determined using GA codes, which are ICD codes that represent the number of completed weeks of gestation (Supplementary Table 6). ICD codes indicating preterm birth (PTB) were designated as PTB codes (Supplementary Table 7). We searched maternal records for both GA and PTB codes spanning one month before to one month after the delivery date. This wide temporal window was necessary to account for the occasionally inaccurate timing of claims data.

GA estimation proceeded as follows:

1. If a single GA code was recorded in maternal records, that value was extracted as the GA. If multiple GA values are recorded, the highest recorded GA value was selected.
2. If no GA code was recorded in maternal records, but infants had a GA code in their records within one month of delivery, GA was inferred from the infant record.
3. If no GA code was recorded in both maternal and infant records, but PTB codes were recorded, a GA of 34 weeks was assigned. Otherwise, GA of 39 weeks was designated.

Preterm birth records were estimated similarly:

1. If PTB codes were recorded in maternal records, the pregnancy was associated with PTB.
2. If no PTB codes were recorded in maternal records, but GA code suggested less than 37 weeks of gestation, the pregnancy was considered to be delivered preterm. Otherwise, a full-term delivery was assumed.
3. If mothers had both PTB codes and GA codes, but GA was higher than 37 weeks, they were removed from the cohort.

1.d. Maternal Medications

Generic drug names, drug classes, and route of administration information from MarketScan Redbook(61) were additionally integrated to the maternal medication records using NDC. Medication records were further processed as follows (Supplementary Figure 4):

1. Categorized medication claims records by their route of administration, identifying 24 distinct routes.
2. For each route of administration, generic medication records were divided at the compound level using slashes and semi-colons.
3. Next, medication names were standardized using DrugBank IDs to account for compounds described differently (e.g., Oxycodone Hydrochloride vs. Oxycodone HCl). Medications without DrugBank IDs, totaling 139, were manually reconciled. Out of these, 31 medications were assigned manual IDs, while the rest were discarded.
4. Each final medication consisted of a harmonized medication name and route of administration such that the same medication taken by different routes of administration was considered separately. Medication administration was binarized so that each medication was designated as “taken” or “not taken”.

1.e. Final Mother-Baby Cohort

The final dataset included 1,152 maternal medications and 68 maternal covariates during the 90 days leading up to delivery, as well as data on 24 neonatal complications during the first 90 days post-delivery. It also included 6 additional attributes: maternal age, gestational age at delivery, preterm birth, baby’s sex, and the region and geographic location of the healthcare facility. Except for the demographic details, all other variables in the dataset were binary variables to indicate the presence or absence of a condition or medication.

The final dataset was split into a discovery cohort and validation cohort. Babies delivered between 2007 and 2015 were included in the discovery cohort, while those delivered between 2016 and 2021 were used for validation.

2. Odds Ratios: Maternal Medications and Neonatal Complications

The associations between maternal medications during pregnancy and neonatal complications were calculated using both unadjusted and adjusted odds ratios. Odds ratios were derived from logistic regression, while adjusted odds ratios were determined through Lasso Logistic Regression Without Penalizing Exposure Variable (LassoNoExp).(62,63) Potential confounders were selected using variance inflation factors (VIF) from among 68 maternal comorbidities and 6 additional attributes (maternal age, gestational age at delivery, preterm birth, and the region and geographic location, the sex of each baby) to mitigate the effects of multicollinearity. We next describe this approach in detail.

2.a. Odds Ratios

Odds ratios (ORs) were computed using univariate logistic regression for each medication-disease pair. In equation (1), $Y \in \{0,1\}^{817,402}$ represented the neonatal complication vector, and $M \in \{0,1\}^{817,402}$ denoted the medication vector where entry M_j equaled 1 if the mother indexed by j was exposed to the medication, and 0 otherwise.Therefore, the odds ratio for each pair was given by equation (2). This process was repeated for all possible pairs, yielding 27,648 ORs from 1,152 medications and 24 diseases. The Benjamini-Hochberg procedure(64) was subsequently applied to control the false discovery rate (FDR) at 5% to correct for multiple comparisons.

- Logistic Regressions between each medication and neonatal complication pair

$$log\left[\frac{P(Y=1|M)}{P(Y=0|M)}\right] = log(odds) = \beta_0 + \beta_1 M \quad (1)$$

- Odds Ratio

$$OR = e^{\beta_1} \quad (2)$$

2.b. Adjusted Odds Ratios

While ORs from 2.a. provided an estimate of the effects of maternal medication, they could be influenced by spurious associations due to unaccounted confounders. However, traditional methods of manually selecting confounders were impractical for our study of 27,648 associations, and uniformly adjusting for high-dimensional confounders also posed a challenge due to overadjustment bias(65) and multicollinearity(66). To obtain a more accurate representation of the maternal medication's impact on neonatal complications, we adjusted the odds ratios for 74 selected potential confounders. These adjusted odds ratios (aORs) were computed using Lasso Logistic Regression Without Penalizing Exposure Variable (LassoNoExp) after removing certain covariates to mitigate multicollinearity.

2.b.1. Managing Multicollinearity of Confounders

Multicollinearity refers to correlation of independent variables in multiple regression analysis. This poses challenges for both mathematical modeling and results interpretation. For a given multiple regression $y = X\beta + \epsilon$, the ordinary least squares (OLS) is given by $\hat{\beta} = (X^T X)^{-1} X^T y$. However, if columns of X are linearly dependent and therefore correlated, then $X^T X$ becomes non-invertible, making it impossible to complete the OLS formula. Highly correlated columns are also difficult to interpret because it becomes difficult to discern the individual contributions of each factor to predicting the dependent variable, and increases the variance of the regression coefficient causing it to become unstable.(67)

In our study, we initially included a total of k=74 potential confounders, encompassing maternal comorbidities and demographic variables, identified based on previous research related to maternal morbidity.(68) To reduce multicollinearity, we first selected independent variables (i.e. confounders) based on the Variance Inflation Factor (VIF).(69) VIF is a measure of multicollinearity that indicates the degree of intercorrelation among the independent variables. VIF of each confounder was calculated by fitting a linear regression model using one confounder as the dependent variable and all others as independent variables (3). Subsequently, the multiple correlation coefficient of ith confounder X_i, ${R_i}^2$,

was used to calculate VIF of ith confounder X_i, as given by (4). This process was repeated for all confounders, with a higher VIF indicating greater multicollinearity.

$$X_i = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k \quad (3)$$

$$VIF_i = \frac{1}{1-R_i^2} \quad (4)$$

In our model, we excluded six confounders that had either a VIF greater than 5(70) or zero vectors. This exclusion included specific substance use disorders (e.g., fentanyl, amphetamine, ecstasy, methamphetamine, methylphenidate) and region of the healthcare facility features. Nonetheless, we retained both gestational age and maternal age, despite their VIFs exceeding 5, due to their critical role in determining outcomes. After this process, a total of 68 confounders remained, which were used for adjustment.

2.b.2. Feature Engineering: Demographic Features

Maternal age, gestational age, and the geographic location of the healthcare facility features were further processed. Because maternal age and gestational age had a wider span of values compared to binary inputs which could have led to them being overemphasized in the model, they were scaled using a min-max scaler. The geographic location of the healthcare facility was a categorical variable with 54 categories, which can introduce additional multicollinearity. Thus, it was encoded using a binary encoding method, expressing 54 variables in 6 columns. The final dataset for adjusted odds ratios contained 73 columns representing confounders, along with 1,152 maternal medications and 24 neonatal complications.

2.b.3. Additional Adjustment for Postmaturity

When the outcome variable y was postmaturity, preterm birth and gestational age were excluded from the confounders due to their potential to introduce multicollinearity, as both were defined by gestational age. As a result, the total number of confounders for postmaturity was reduced to 66, compared to 68 for the other 23 neonatal complications. This adjustment led to a final dataset containing 71 columns representing confounders, along with 1,152 maternal medications, specific to postmaturity, distinguishing it from the datasets used for the other neonatal complications.

2.b.4. LASSONoExp: Adjusted Odds Ratios Calculation

We adjusted the confounders using LASSONoExp, an algorithm based on LASSO regression.(71,72) Lasso regression is effective for adjusting potential confounders and estimating treatment effects due to its ability to select variables among high-dimensional confounders and reduce overfitting.(73) LASSONoExp is the same as the LASSO, except that the regression coefficient of the exposure variable, in our case, maternal medication exposure, is not penalized in the model. This prevents the coefficient of maternal medication from yielding a zero value. Consequently, the exposure effect of the estimator is not biased, while the effects of the confounding variables are regularized.(62,63)

Let M denote the maternal medication of interest, as defined in Equation (1). As before, Y represents the neonatal complications of interest, and X_i (where $2 \leq i \leq P$) indicates the confounders that were selected from section 2.b.1 to 2.b.3 (in our case, $P = 71$ for postmaturity, and $P = 73$ for the rest of the neonatal complications) such that $X_{ij} = 1$ if the mother indexed by j had comorbidity i and 0 otherwise. The model is then given by

$$Y = \beta_0 \cdot 1 + \beta_1 M + \sum_{i=2}^{P} \beta_i X_i \quad (5)$$

Coefficients β_i are determined by

$$argmin_{\beta} \left(||y - Y||_2^2 + \lambda \sum_{j=2}^{P} |\beta_j| \right) \quad (6)$$

$$OR_{adjusted} = e^{\beta_1} \quad (7)$$

In LASSONoExp, the coefficients β_i, as denoted in (5), are obtained by solving the objective function (6) where hyperparameter λ represents a penalty weight for the confounders. In (6), the term y represents the ground truth (actual observed values), whereas Y signifies the predicted values obtained from the model. Here, β_1, the coefficient of maternal medication, is not penalized, while β_2 to β_P are penalized based on the penalty weight. Adjusted odds ratios are then given by (7).

We repeated this process for the 1,446 maternal medication and neonatal complication pairs that yielded statistically significant unadjusted odds ratios, as described in Section 2.a. The Python package *statsmodels* was used to fit the model. The penalty weight, λ, was selected for each maternal medication-neonatal complication pair based on 3-fold stratified cross-validation. This selection involved 10 coefficients ranging between 0.01 and 10, using the Python package *optum*. Finally, the Benjamini-Hochberg procedure was applied to control the false discovery rate (FDR) at 5%, correcting for multiple comparisons.

2.c. Network Graph

Unadjusted odds ratios and adjusted odds ratios, given by equations (2) and (7) respectively, were visualized as a network graph. This visualization was chosen to illustrate the pattern of interactions and to allow exploration of multiple correlations in a single plot, given that there are more than 1,000 statistically significant correlations. (See Figure 2, Supplementary Figure 5, and the interactive graph at http://pregmednet.com). In the figures, each node represents either neonatal complications or maternal medications, while the edges between them indicate statistically significant odds ratios. The size of the nodes is proportional to the number of edges connected to them, and the width of the edges is inversely proportional to the p-values of the odds ratios. The color of the nodes represents either the medication group or disease (with gray-colored nodes at the center), and the color of the edges indicates whether the odds ratios are positive (gray) or negative (blue). The locations of the nodes were estimated based on t-distributed Stochastic Neighbor Embeddings (t-SNE)(74) within each disease or medication group, demonstrating relative similarities in the patterns of medication use or disease outbreaks in mother-baby dyads. The Python package *scikit-learn* was used to calculate t-SNE, *networkx* to generate the network graph, and *bokeh* for creating an interactive graph.

2.d. Verification of Results in the Validation Cohort

To ensure the reproducibility of the odds ratios and adjusted odds ratios, we repeated the same process on the validation cohort and compared the results of unadjusted and adjusted odds ratios between the two cohorts by matching maternal medications and neonatal complications. After matching, scatter plots were created, and Pearson correlation coefficients were calculated between discovery and validation cohort odds ratios.

3. Drug-Drug Interactions (DDIs)

3.a. Derivation of DDIs

Drug-Drug Interactions (DDIs) occur when two or more medications react with each other affecting the outcome in a way that would otherwise not occur. DDIs are considered in pairs of maternal medications relative to each of the 24 neonatal complications. This is done to determine whether interactions between two medications could additionally affect neonatal complications. Gestational Age, one of the strongest confounders, is included and adjusted for in the analysis. DDI effects were captured using a logistic regression-based model,(75) given by:

$$Y = \beta_0 + \beta_1 M_1 + \beta_2 M_2 + \beta_3 M_1 M_2 + \beta_4 GA \quad (8)$$

where GA denotes gestational age, M_i ($i = 1,2$) denotes exposure to maternal medication i, and Y represents the neonatal complication of interest. M_1M_2 represents the interaction between medication 1 and medication 2, and its coefficient, β_3, indicates the direction and magnitude of this interaction. If β_3 is less than 0, it suggests that the interaction tends to reduce the risk of neonatal complications. Conversely, a β_3 greater than 0 indicates that the interaction increases the risk of neonatal complications. The higher absolute value of β_3 suggests stronger interactions.

We repeated this process for all maternal medication pairs among 1,152 medications for 24 neonatal complications. We only considered interactions where there were more than 10 mother-baby dyads where the mothers took both medications and the babies had neonatal complications. A total of 28,417 maternal medication pairs and neonatal complications out of the 15,911,424 all possible pairs (calculated as ${}_{1152}C_2 \times 24$) were analyzed. The Benjamini-Hochberg procedure was then applied to control the false discovery rate (FDR) at 5%, correcting for multiple comparisons. Ultimately, a total of 121 maternal medication pairs and neonatal complications showed statistically significant β_3.

3.b. Calculate the Odds of Drug Pairs

For the 121 maternal medication pairs and neonatal complications that exhibited statistically significant β_3 values computed from Section 3.a., odds ratios for concomitant medication use were calculated as:

$$OR_1 = exp(\beta_1) \quad (9)$$

$$OR_2 = exp(\beta_2) \quad (10)$$

$$OR_{12} = exp(\beta_3) * OR_1 * OR_2 \quad (11)$$

As with β_3, the Benjamini-Hochberg correction was subsequently applied to β_1 and β_2, respectively, to control the false discovery rate (FDR) at 5%, thereby correcting for multiple comparisons. Equation (11) was applied when all β_3, OR_1, OR_2 are statistically significant. Finally, we extracted the medication pairs that showed statistically significant OR_{12} among the 121 medication pairs, and then compared these results with the odds ratios of their corresponding single medication use cases.

3.c. Final DDIs Overlap: Discovery vs. Validation Cohorts

The same process was repeated within the validation cohort to identify significant values of β_3 and OR_{12}. DDIs for which both β_3 and OR_{12} were found to be significant in both the discovery and validation cohorts were extracted for further analysis.

4. *In vivo* ondansetron experiments

The association between maternal ondansetron and neonatal bronchopulmonary dysplasia, one of the associations identified in section 2, is tested experimentally.

4a. Mice

Female and male C57BL/6 mice were purchased from the Jackson Laboratory and maintained in-house. All animal experiments were approved by Stanford University's Institutional Animal Care and Use Committee under protocol 33897.

4b. Ondansetron administration

Ondansetron (Millipore Sigma 1478571) was dissolved in PBS without calcium and magnesium (Gibco 10010-049). C57BL/6 female and male mice were timed mated and pregnancy was confirmed by a 2 gram increase in weight by E10 (78). Starting on E15, pregnant mice were dosed daily with either 5 mg/kg body weight of ondansetron or PBS by intraperitoneal injection until delivery of pups.

4c. Lung collection for histology

On P11, pups were euthanized and the lungs were inflated with 4% PFA (Electron Microscopy Sciences 15713). Following inflation, the lungs were collected and fixed in 4% PFA for 1 hour at room temperature. The lungs were then transferred to 15% sucrose (Millipore Sigma S0389) and incubated overnight at 4°C. The following day, the lungs were transferred to 30% sucrose and incubated for 3 days at 4°C, followed by embedding in OCT mounting medium (Leica 14020108926) for cryo-sectioning. Lungs were sectioned at a thickness of 10μm.

4d. Hematoxylin and eosin staining

For hematoxylin and eosin staining, slides were brought to room temperature and fixed in 10% buffered formalin (StatLab FXBFOLT) for 20 minutes. Slides were immersed in Gill II hematoxylin (Newcomer Supply 1180D) for 15 seconds, rinsed with water, and immersed in Scott's bluing reagent (Ricca Chemical Company 66971) for 30 seconds. Slides were then rinsed with water, immersed in eosin Y (Millipore Sigma HT110116) for 2 minutes, and again rinsed with water. Slides were dehydrated with dips in ethanol (Fisher A962P-4) and cleared with dips in xylene (Fisher X5-4) before being mounted with non-aqueous mounting medium (Vector Laboratories H570060).

4e. Mean linear intercept quantification

For quantification, brightfield images were captured with the Leica Thunder microscope using the 20X objective. Images were quantified on FIJI using the Mean Linear Intercept plug-in (79). The image analysis parameter was set to 500 pixels between lines, and both horizontal and vertical chords were measured. All horizontal and vertical chords were averaged to get an image average value. For each pup, 3 representative lung sections were stained and imaged, with 5 images taken from each section for a total of 15 images quantified.

5. Knowledge Graph: Hypothesis Generation for New Medication Effects

The hypotheses of underlying mechanisms for newly discovered medication effects were explored using the Precision Medicine Knowledge Graph (PrimeKG).(76) This publicly available knowledge graph integrates 20 different biological databases, providing a holistic view of the interactions among 10 different types of biological entities. In PrimeKG, biological entities are represented as nodes, and their interactions as edges. Using this framework, we tested three different types of subgraph configurations.

First, we extracted a subgraph that included six node types (drugs, diseases, proteins/genes, biological processes, molecular functions, and cellular components) and nine different edge types (drug-protein, protein-protein, disease-protein, bioprocess-protein, molfunc-protein, cellcomp-protein, bioprocess-bioprocess, molfunc-molfunc, and cellcomp-cellcomp). From this subgraph, we identified and matched the two nodes of interest: in our case, a maternal medication and a neonatal complication pair that showed statistically significant adjusted odds ratios in section 2.b. Subsequently, all possible shortest paths between the two nodes were extracted using the Dijkstra algorithm.(77) The Python package *networkx* was used for this extraction. The list of shortest paths obtained was considered indicative of possible underlying mechanisms between the newly discovered interactions, suggesting how maternal medications might affect neonatal complications.

Next, we repeated the process, focusing on the drug and the disease nodes of interest without including other drugs and diseases. We first constructed a subgraph with four biological node types (proteins/genes, biological processes, molecular functions, and cellular components) and seven different types of edges, excluding drug-protein and disease-protein. We then added the two nodes of interest - a specific medication and a disease that we are examining - along with the edges connecting them to the proteins (drug-protein edges and disease-protein edges). This ensured that the subgraph

contained only the pair of drug and disease of interest, without involving other drugs and diseases. Finally, all possible shortest paths between the maternal medication and the neonatal complication were extracted.

Finally, we conducted the same analysis using a subgraph with only three node types–drugs, proteins/genes, and diseases and three different types of edges (drug-protein, protein-protein, disease-protein). As before, all possible shortest paths between the maternal medication and the neonatal complication were extracted as before.

This process was iterated for all newly discovered associations between maternal medications and neonatal complications identified in section 2.b. The resulting hypotheses were visualized as graphs using the Python packages *networkx* and *bokeh*.

Data Availability

The Merative™ Marketscan® Commercial Database is available to purchase by federal, nonprofit, academic, pharmaceutical, and other researchers. Use of the data is contingent on completing a data use agreement and purchasing the data needed to support the study. More information about licensing the Merative™ Marketscan® Commercial Database is available at: https://www.merative.com/documents/brief/marketscan-explainer-general.

Code Availability

The PregMedNet framework and custom computer code used in this study can be accessed on GitHub (https://github.com/yerakim824/PregMedNet).

Author Contribution

Y.K., I.M., G.M.S., D.K.S., and N.A. conceived the study. Y.K., I.M., G.M.S., B.T.B., D.K.S., L.S.P., and N.A. conceptualized and designed the study. Y.K., I.M., and G.M.S. collected and investigated claims data. Y.K. processed the claims data. Y.K. and L.H. developed and reviewed the medication cleanup method. Y.K., J.D.R., and W.M.K. extracted and reviewed the ICD codes. Y.K., I.M., and N.A. developed the computational methodology. Y.K. conducted computational experiments. Y.K. developed the software. L.D.B. and L.S.P. designed the in vivo experiments. C.M.K. performed ondansetron administration. C.M.K., L.D.B., and K.J.C. harvested lung tissue. C.M.K. sectioned, stained, and imaged lungs for mean linear intercept quantification. C.M.K. performed statistical analyses for the in vivo experiments. Y.K. wrote the original draft. C.M.K. wrote the in vivo experimentation section of the manuscript. I.M., P.C., J.D.R., L.S.P. and N.A. reviewed the original manuscript. All authors reviewed and edited the final manuscript.

Ethics Declaration

Competing Interests

N.A. is a cofounder of Takeoff AI, a member of the scientific advisory boards of January AI, Parallel Bio, Celine Therapeutics, and WellSim Biomedical Technologies, and a paid consultant for MaraBio Systems. B.G. is an advisory board member at SurgeCare. The remaining authors declare no competing interests.

Acknowledgement

Data for this project were accessed using the Stanford Center for Population Health Sciences Data Core. The PHS Data Core is supported by a National Institutes of Health National Center for Advancing Translational Science Clinical and Translational Science Award (UL1TR003142) and from Internal Stanford funding. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Certain data were supplied by Merative as part of one or more MarketScan Research Databases. Any analysis, interpretation, or conclusion based on these data is solely that of the authors and not Merative.

Figure Legends

Figure 1. Study Overview. (a) A total of 1,190,257 mother-newborn dyads from 2007 to 2021 are extracted from the Merative™ Marketscan® Commercial Database, a set of de-identified claims data accessed through the Stanford Center for Population Health Sciences. Maternal records from the last 90 days before delivery are examined to extract 68 maternal comorbidities, 6 additional attributes, and all prescribed medications. Similarly, newborn records from the first 90 days after birth are analyzed to identify 24 neonatal complications. The final dataset includes 1,152 maternal medications, 68 maternal comorbidities, 24 neonatal complications, and 6 additional attributes: maternal age, gestational age at delivery, preterm birth, baby's sex, and the region and geographic location of the healthcare facility. The dyads are divided into a discovery cohort (2007-2015) and a validation cohort (2016-2021) based on the year of birth. **(b)** Three experiments are conducted based on the database. First, the associations between maternal medications and neonatal complications are calculated using odds ratios based on logistic regression. Notably, confounders are selected and adjusted for each medication-disease pair using the LassoNoExp algorithm, which penalizes and selects confounding variables using Lasso regression while preserving maternal medication exposure information. Second, additive effects of concomitant maternal medications on neonatal complications are calculated through drug-drug interaction analysis. Lastly, potential underlying mechanisms of newly discovered drug-disease associations are discovered using a knowledge graph and a graph learning method. **(c)** A tetrachoric correlation plot of the 24 neonatal outcomes shows the relationships between the neonatal complications. The size of each node is correlated with disease prevalence, and the nodes are connected if the diseases are positively correlated. The edge thickness is proportional to the strength of the correlation.

Figure 2. A Network of the Associations Between Maternal Medications and Neonatal Complications, and Their Validation. (a) Odds Ratio (OR) results are displayed as a network graph, where nodes represent maternal medications or neonatal complications, and edges indicate statistically significant ORs. Disease nodes are centrally placed, while maternal medication nodes are peripheral. Medications are grouped and color-coded within each cluster. The graph shows significant ORs as edges; gray for ORs above 1 (positive correlations), and blue for ORs below 1 (negative correlations). There are 24 disease nodes, 1,152 medication nodes, and 1,446 edges—1,153 gray and 293 blue. Edge thickness correlates with the strength of the association (inverse of p-values), and node size with the number of connecting edges. **Abbreviations in disease nodes**: BPD: Bronchopulmonary Dysplasia, Other Resp.: Other Respiratory Diseases, RDS: Respiratory Distress Syndrome, SGA: Small for Gestational Age, PPHN: Persistent Pulmonary Hypertension, Hypogly.: Hypoglycemia, UTI: Urinary Tract Infection, Seiz.: Seizures, TTN: Transient Tachypnea, ROP: Retinopathy of Prematurity, Ker.: Kernicterus, GALD: Gestational Alloimmune Liver Disease, NAS: Neonatal Abstinence Syndrome, NEC: Necrotizing Enterocolitis, IVH: Intraventricular Hemorrhage, Death: Neonatal Death, LGA: Large for Gestational Age, AOP: Anemia of Prematurity. **Additional conditions in disease nodes not abbreviated**: Pneumonia, Postmaturity, Sepsis, Arrhythmia, Anemia, and Jaundice. **(b)** ORs from the discovery cohort were validated in the validation cohort. ORs for both cohorts were matched by maternal medications and neonatal complications and plotted on a scatterplot. This figure compares the log of significant ORs from the discovery (x-axis) to the

validation cohort (y-axis). Of the ORs, 772 matched between the cohorts—1,446 significant in the discovery cohort and 1,097 in the validation cohort. Pearson correlation yielded an R^2 of 0.972 with a p-value of 0.0.

Figure 3. Associations Between Four Maternal Medication Groups and Neonatal Complications: Highlights from the Adjusted Odds Ratio Network **(a)** The figure shows the number of aORs by neonatal complications. The gray line represents the total aORs, the navy line indicates reduced risk aORs, and the magenta indicates increased risk aORs. **(b)** The figure shows the number of aORs by maternal medication groups, with the line colors corresponding to those in (a). **(c)** This network graph highlights four maternal medication groups and their adjusted odds ratios (aOR) with neonatal complications. Nodes are consistent with those in Figure 2's unadjusted odds ratio graph, while edges represent statistically significant adjusted odds ratios. LassoNoExp was utilized to calculate these aORs, selecting and adjusting for maternal covariates out of a possible 73 for each drug-complication pair. Edges in gray indicate aORs greater than 1 (positive correlations), and those in blue indicate aORs less than 1 (negative correlations). The edge thickness corresponds to the association strength, inversely proportional to the p-values. Highlighted medication groups include Ondansetron, Antidepressants, Long-Chain Polyunsaturated Fatty Acids (LC-PUFAs), and Opioids. The full network can be examined in Supplementary Figure 6. **(d)** The subfigure details the values of statistically significant adjusted odds ratios and their confidence intervals for the four medication groups. Magenta dots and lines depict aORs higher than 1, signifying positive correlations, whereas navy blue dots and lines indicate aORs lower than 1, representing negative correlations.

Figure 4. Maternal Drug-Drug Interaction Effects on Neonatal Complications.
(a) Drug-Drug Interaction (DDI) effects from the concomitant use of maternal medications were calculated based on mother-baby dyads. Here, β_1, β_2, and β_3 represent the individual contributions of the medications and their interaction to the overall effect on neonatal complications. Specifically, β_1 represents the impact of medication 1, β_2 the impact of medication 2, and β_3 the interaction contribution between medications 1 and 2 to the collective effect on neonatal complications. **(b)** A network graph illustrates the DDIs as edges connecting the medication nodes. Notably, five DDIs significant in both discovery and validation cohorts are illustrated. All five edges are colored blue to indicate that the DDIs reduce neonatal complications, thereby offering additional reduced risk benefits through the combination of maternal medications.

Figure 5: Impact of maternal ondansetron exposure and neonatal lung development in mice. (a) Experimental timeline of maternal ondansetron administration and quantification of lung development. Timed pregnant C57BL/6 female mice were injected daily with ondansetron (5 mg/kg i.p.) or an equivalent volume of sterile PBS (control) from E15 until delivery. (b) Kaplan-Meier survival analysis and (c) neonatal weight of control and ondansetron exposed mice. N = 20 mice from 4 litters in the control group; n = 24 mice from 5 litters in the ondansetron group. (d) Representative H&E images of lungs from control and ondansetron exposed mice at postnatal day 11. (e) Mean linear intercept quantification of control and ondansetron exposed mice. For each mouse, 3 lung sections were stained, with 5 random images taken from each section for a total of 135-180 images from 9-12 mice. Darker, overlayed dots represent average mean linear intercept values for each individual mouse. **** indicates $P < 0.0001$ as determined using an unpaired t test.

Figure 6. Potential Underlying Mechanisms of Ondansetron Effects on Bronchopulmonary Dysplasia. (a) This figure illustrates the potential underlying mechanisms behind the newly discovered associations between maternal medications and neonatal complications, utilizing significant adjusted odds ratios from PregMedNet and the PrimeKG Knowledge Graph, which contains scientifically validated biological interactions. Specifically, the subgraph containing three different node types—drugs, proteins/genes, and diseases—is shown here to generate more reliable hypotheses than other types of subgraphs containing additional biological node types. Displayed is the subgraph of all nodes and edges that connect ondansetron to bronchopulmonary dysplasia via the shortest paths between the two nodes. This results in four paths with seven intermediary nodes. Each path represents a hypothesis for a possible underlying mechanism of action. **(b)** A particular path of interest suggests that ondansetron affects bronchopulmonary dysplasia through the protein interactions between CYP3A5 and IL-1b.

This hypothesis builds on previous research showing the relationship between ondansetron and CYP3A, and a separate relationship between bronchopulmonary dysplasia and IL-1b. While there is no prior published literature that relates CYP3A5 to IL-1b in the context of ondansetron and bronchopulmonary dysplasia, PregMedNet can generate hypotheses for novel relationships by connecting already proven, but isolated biological knowledge using graph learning methods.

Supplementary Figure 1. Flow Chart of the Data Construction Overview. This figure illustrates the high-level process of data cleaning and preparation for analysis derived from the Merative™ Marketscan® Commercial Database.

Supplementary Figure 2. Comprehensive Flowchart Detailing Data Construction. This figure details the extensive process of data cleaning and preparation undertaken using the Merative™ Marketscan® Commercial Database. It expands upon the methodologies outlined in Supplementary Figure 1, providing clarity on the data preparation steps for the study. The flowchart depicts the procedural actions executed to assemble and refine the dataset, leading to the creation of a table that encapsulates maternal medication use, maternal comorbidities, 6 other attributes (maternal age, gestational age at delivery, preterm birth, baby's gender, and the region and geographic location of the healthcare facility), and neonatal complications. This offers a thorough perspective at the individual patient level and serves to enhance the foundational information displayed in Figure 1(a).

Supplementary Figure 3. Correlation Plot of Potential Maternal Covariates. This tetrachoric correlation plot displays the interrelationships among 68 maternal covariates and preterm birth, illustrating potential connections between various diseases. The diameter of each node is proportionate to the prevalence of the disease it represents. Nodes are connected if their tetrachoric correlation exceeds 0.5, with the thickness of the edges reflecting the strength of the correlation. The abbreviations used in the figure are expanded as follows:

- ADHD: Attention-deficit hyperactivity disorders
- alcohol abuse: Alcohol Abuse
- alcohol withdrawal: Alcohol Withdrawal Syndrome
- Anemia: Anemia
- Anxiety: Anxiety
- APLS: Antiphospholipid Syndrome
- Asthama: Asthma
- Autoimmune: Autoimmune Disorders
- Bipolar: Bipolar Disorder
- catatonic: Catatonia
- Cesarean Section: Cesarean Section
- CHD: Congenital Heart Disease
- CHF: Congestive Heart Failure
- Chronic IHD: Chronic Ischemic Heart Disease
- Chronic Pain: Chronic Pain
- Chronic Renal dis.: Chronic Renal Disease
- Depression: Depression
- DM: Pregestational Diabetes
- drug abuse: Drug Abuse
- Eating dis.: Eating Disorders
- Eclampsia: Eclampsia
- Epilepsy: Epilepsy
- GDM: Gestational Diabetes
- Gestational HTN: Gestational Hypertension
- Headache: Headache
- HIV: HIV
- HTN: Pregestational Hypertension

- Hyperemesis Gravidarum: Hyperemesis Gravidarum
- Infertility: Female Infertility
- Migraine: Migraine
- Mild Preeclampsia: Mild or unspecified Preeclampsia
- Obesity: Overweight/Obesity
- Placenta Abruption: Placenta Abruption
- Placenta Previa: Placental Previa
- poisoning-analgesics: Other Unspecified Psychotropic Poisoning (analgesics)
- poisoning-anesthetics: Other Unspecified Psychotropic Poisoning (anesthetics)
- poisoning-antidepressants: Other Unspecified Psychotropic Poisoning (antidepressants)
- poisoning-antiepileptics: Other Unspecified Psychotropic Poisoning (antiepileptics)
- poisoning-antiparkinson: Other Unspecified Psychotropic Poisoning (anti-parkinson drug poisoning)
- poisoning-antipsychotics: Other Unspecified Psychotropic Poisoning (antipsychotics)
- Previous c-section: Previous Cesarean Delivery
- PTB: Preterm Birth
- Pul HTN: Pulmonary Hypertension
- Renal dis.: Renal disease
- SCD: Sickle Cell Disease
- Schizo.: Schizophrenia/Schizoaffective Disorder
- Severe Preeclampsia/Eclampsia: Severe Preeclampsia
- SLE: Systemic Lupus Erythematosus
- Sleep dis.: Sleep Disorder
- STD: Sexually Transmitted Diseases (STD)
- SUD-alcohol: Alcohol Use Disorder
- SUD-AMP: Psychostimulant Abuse/Dependence/Poisoning (amphetamine)
- SUD-barbiturate: Sedative Abuse/Dependence/Poisoning (barbiturate)
- SUD-BZD: Sedative Abuse/Dependence/Poisoning (benzodiazepine)
- SUD-cocaine: Cocaine Abuse/Dependcence/Poisoning
- SUD-ecstasy: Psychostimulant Abuse/Dependence/Poisoning (ecstasy)
- SUD-fentanyl: Opioid Use Disorder (Fentanyl)
- SUD-hallucinogen: Hallucinogen Abuse/Dependence/Poisoning (Hallucinogen, LSD, Unspecified Psychodysleptics)
- SUD-heroin: Opioid Use Disorder (Heroin)
- SUD-marijuana: Marijuana Use
- SUD-meth: Psychostimulant Abuse/Dependence/Poisoning (methamphetamine)
- SUD-methadone: Opioid Use Disorder (Methadone)
- SUD-MPH: Psychostimulant Abuse/Dependence/Poisoning (methylphenidate)
- SUD-opioids (all): Opioid Use Disorder (All)
- SUD-other psychotropic: Other Unspecified Psychotropic Poisoning (all)
- SUD-psychostimul. (all): Psychostimulant Abuse/Dependence/Poisoning (all)
- SUD-sedatives (all): Sedative Abuse/Dependence/Poisoning (all)
- SUD-smoking: Smoking
- VHD: Valvular Heart Disease

Supplementary Figure 4. Medication Name Harmonization This figure outlines the process employed to clean up and harmonized maternal medication data. 1) Maternal medication with generic names was divided at the compound level. 2) Identical medications with varying names were consolidated and harmonized by employing DrugBank IDs along with their common names. 3) Lastly, information regarding the route of administration was integrated into these standardized names.

Supplementary Figure 5. Network Graph of the Adjusted Odds Ratios Between Maternal Medications and Neonatal Complications. Network graph depiction of adjusted odds ratios (aORs) between maternal medications and neonatal complications. aORs are calculated using LassoNoExp, a method that selects confounders by penalizing the excess of potential confounders for each medication-disease pair. Each node in

the network graph represents either a maternal medication or a neonatal complication, consistent with the representation in Figure 2. Edges between nodes indicate statistically significant aORs, differentiated from unadjusted odds ratios presented Figure 2. PregMedNet identifies 261 statistically significant adjusted odds ratios (aORs) after adjustment, involving the associations of 68 maternal medications and 19 neonatal complications. Out of the total 261 edges, the 117 edges representing aORs > 1, indicating an increased risk association, are shown as gray edges, while the 144 edges representing aORs < 1, indicating a reduced risk association, are colored blue. The abbreviations for disease nodes are as defined in Figure 2.

Supplementary Figure 6. Distribution of Unadjusted Odd Ratios for each Neonatal Complications and Maternal Medication Group Bar graphs that show the count of statistically significant unadjusted odds ratios (OR). (a) The number of all statistically significant odds ratios associated with neonatal complications. BPD showed the most significant associations, followed by other respiratory disorders, RDS, TTN, and Postmaturity. (b) The number of statistically significant increased risk odds ratios associated with neonatal complications. RDS showed the most increased risk associations, followed by BPD, other respiratory disorders, TTN, and ROP. (c) The number of statistically significant reduced risk ORs associated with neonatal complications. Postmaturity has the most reduced risk edges followed by LGA and Jaundice. (d) The number of all significant ORs categorized by maternal medication groups. Medications in the Central Nervous System showed the most significant associations, followed by Hormones & Synthetic substances, Vitamins & Combinations, and Anti-Infective Agents. (e) The number of significant increased risk ORs categorized by maternal medication groups. Medications in the Central Nervous Systems show the most increased risk associations, followed by Hormones & Synthetic Substances, Anti-Infective Agents, and Cardiovascular Agents. (f) The number of significant reduced risk ORs categorized by maternal medication groups. The Vitamins & Combinations group shows the most reduced risk associations, followed by Anti-Infective Agents.

Supplementary Figure 7. Distribution of Adjusted Odd Ratios for each Neonatal Complications and Maternal Medication Group Bar graphs that show the count of statistically significant adjusted odds ratios (aOR). (a) The number of all statistically significant aORs associated with neonatal complications. BPD showed the most significant associations, followed by other respiratory disorders, Jaundice, Postmaturity, and TTN. (b) The number of statistically significant increased risk aORs associated with neonatal complications. BPD showed the most increased risk associations, followed by other respiratory disorders, RDS, Jaundice, and TTN. (c) The number of statistically significant reduced risk aORs associated with neonatal complications. Postmaturity has the most reduced risk edges followed by LGA and Jaundice. (d) The number of all significant aORs categorized by maternal medication groups. Medications in the Vitamins & Combinations showed the most significant associations, followed by Central Nervous System, and Anti-Infective Agents. (e) The number of significant increased risk aORs categorized by maternal medication groups. Medications in the Central Nervous Systems show the most increased risk associations, followed by Cardiovascular Agents, Hormones & Synthetic Substances, and Anti-Infective Agents. (f) The number of significant reduced risk aORs categorized by maternal medication groups. The Vitamins & Combinations group shows the most reduced risk associations, followed by Anti-Infective Agents.

Supplementary Figure 8. Unadjusted odds ratios for six different medication groups. (a) Antidepressants. (b) Opioids. (c) Antiemetics. (d) Long-Chain PolyUnsaturated Fatty Acids in Vitamins & Combinations Group. (e) reduced risk odds ratios in the Anti-Infective Groups.

Supplementary Figure 9. Results from the discovery cohort are correlated with the results from the validation cohort. The results of the association between the two cohorts are matched based on maternal medications and neonatal complications. Scatterplots and Pearson correlation coefficients are generated for these matched results. **(a)** A scatterplot of significant adjusted odds ratios in the discovery cohort (x-axis) and the validation cohort (y-axis). In total, 81 adjusted odds ratios are matched between the two cohorts, out of 261 from the discovery cohort, and 167 from the validation cohort. The Pearson correlation analysis yields an R^2 value of 0.747 with p-value of 1.246e-15. **(b)** A scatterplot of significant drug-drug interactions (beta3) in the discovery cohort (x-axis) and the validation cohort (y-axis). In total 5 adjusted DDIs are matched between the two cohorts

out of 72 from the discovery cohort and 15 from the validation cohort. The Pearson correlation analysis yields an R^2 value of 0.920 with p-value 2.682e-2.

Supplementary Figure 10. Potential Underlying Mechanisms of Ondansetron Effects on Bronchopulmonary Dysplasia with Inclusion of All Biological Interactions. (a) Hypothesized associations between maternal medications and neonatal complications are generated by combining PregMedNet results with a Knowledge Graph. This is an expanded version of Figure 5. While Figure 5 only includes proteins/genes nodes, this subgraph depiction includes all biological node types, including other disease and medication nodes. The figure depicts the shortest paths between bronchopulmonary dysplasia and ondansetron. The subgraph includes 191 four-hop paths with 81 nodes involved in the interactions between the medication and the disease. **(b)** The path involving IL-1B and CYP3A5 is also found in this subgraph, but IL-1B and CYP3A5 are connected by an unspecific node, 'protein binding.'

Supplementary Figure 11. Potential Underlying Mechanisms of Ondansetron Effects on Bronchopulmonary Dysplasia with Inclusion of All Biological Nodes, Focusing on the Drug and Disease of Interest. (a) Hypothesized associations between maternal medications and neonatal complications are generated by combining PregMedNet results with a Knowledge Graph. This figure is an expanded version of Figure 5 and Supplementary Figure10. While Supplementary Figure 10 includes all biological node types, including other diseases and medications, this subgraph focuses only on the disease and medication of interest, excluding others. The figure depicts the shortest paths between bronchopulmonary dysplasia and ondansetron. The subgraph includes 23 four-hop paths with 21 nodes involved in the interactions between the medication and the disease. **(b)** The same path involving IL-1B and CYP3A5 is also found in this subgraph.

Tables

Disease	Medication 1	Medication 2	β_3	OR_1	OR_2	OR_{12}
NAS	**Acetaminophen, Oral**	**Oxycodone, Oral**	-3.50	4.89	99.97	14.77
Hypoglycemia	**NPH, subcutaneous**	**Insulin Lispro, Injection**	-1.84	2.19	5.92	2.07
Hypoglycemia	**NPH, subcutaneous**	**Insulin Aspart, Injection**	-1.52	2.14	5.42	2.55
NAS	**Buprenorphine, sublingual**	**Ondansetron, Oral**	-1.89	607.34	5.17	474.43
RDS	**NPH, subcutaneous**	**Insulin Aspart, Injection**	-1.02	1.44	2.25	1.17

Table 1. Maternal Drug-Drug Interaction Effects on Neonatal Complications. This table details the five

significant Drug-Drug Interactions (DDI). The term β_3 represents the contribution of the DDI to the overall effect of medication pairs on a neonatal complication. OR_1, OR_2, and OR_{12} represent the effect of medication 1, medication 2, and the concurrent use of two medications on a neonatal complication, respectively. A negative β_3 indicates the reduced risk effects of the DDI on neonatal complications.

References

1. Mitchell AA, Gilboa SM, Werler MM, Kelley KE, Louik C, Hernández-Díaz S, et al. Medication use during pregnancy, with particular focus on prescription drugs: 1976-2008. Am J Obstet Gynecol. 2011 Jul;205(1):51.e1-8.
2. Lupattelli A, Spigset O, Twigg MJ, Zagorodnikova K, Mårdby AC, Moretti ME, et al. Medication use in pregnancy: a cross-sectional, multinational web-based study. BMJ Open. 2014 Feb 17;4(2):e004365.
3. Palmsten K, Hernández-Díaz S, Chambers CD, Mogun H, Lai S, Gilmer TP, et al. The most commonly dispensed prescription medications among pregnant women enrolled in the U.S. medicaid program. Obstet Gynecol. 2015 Sep;126(3):465–73.
4. Nörby U, Noël-Cuppers B, Hristoskova S, Desai M, Härmark L, Steel M, et al. Online information discrepancies regarding safety of medicine use during pregnancy and lactation: an IMI ConcePTION study. Expert Opin Drug Saf. 2021 Sep;20(9):1117–24.
5. Byrne JJ, Saucedo AM, Spong CY. Evaluation of drug labels following the 2015 pregnancy and lactation labeling rule. JAMA Netw Open. 2020 Aug 3;3(8):e2015094.
6. Wood ME, Andrade SE, Toh S. Safe expectations: current state and future directions for medication safety in pregnancy research. Clin Ther. 2019 Dec;41(12):2467–76.
7. Grimsrud KN, Sherwin CMT, Constance JE, Tak C, Zuppa AF, Spigarelli MG, et al. Special population considerations and regulatory affairs for clinical research. Clin Res Regul Aff. 2015;32(2):47–56.
8. Ren Z, Bremer AA, Pawlyk AC. Drug development research in pregnant and lactating women. Am J Obstet Gynecol. 2021 Jul;225(1):33–42.
9. Tannenbaum C, Day D, Matera Alliance. Age and sex in drug development and testing for adults. Pharmacol Res. 2017 Jul;121:83–93.
10. Basile AO, Yahi A, Tatonetti NP. Artificial intelligence for drug toxicity and safety. Trends Pharmacol Sci. 2019 Sep;40(9):624–35.
11. Scaffidi J, Mol BW, Keelan JA. The pregnant women as a drug orphan: a global survey of registered clinical trials of pharmacological interventions in pregnancy. BJOG. 2017 Jan;124(1):132–40.

12. McKiever M, Frey H, Costantine MM. Challenges in conducting clinical research studies in pregnant women. J Pharmacokinet Pharmacodyn. 2020 Aug;47(4):287–93.

13. Richardson JL, Moore A, Bromley RL, Stellfeld M, Geissbühler Y, Bluett-Duncan M, et al. Core Data Elements for Pregnancy Pharmacovigilance Studies Using Primary Source Data Collection Methods: Recommendations from the IMI ConcePTION Project. Drug Saf. 2023 May;46(5):479–91.

14. Alomar M, Tawfiq AM, Hassan N, Palaian S. Post marketing surveillance of suspected adverse drug reactions through spontaneous reporting: current status, challenges and the future. Ther Adv Drug Saf. 2020 Aug 10;11:2042098620938595.

15. Varallo FR, Planeta CS, Mastroianni P de C. Effectiveness of pharmacovigilance: multifaceted educational intervention related to the knowledge, skills and attitudes of multidisciplinary hospital staff. Clinics (Sao Paulo). 2017 Jan 1;72(1):51–7.

16. Goldman SA. Limitations and strengths of spontaneous reports data. Clin Ther. 1998;20 Suppl C:C40-4.

17. Marić I, Winn VD, Borisenko E, Weber KA, Wong RJ, Aziz N, et al. Data-driven queries between medications and spontaneous preterm birth among 2.5 million pregnancies. Birth Defects Res. 2019 Oct 1;111(16):1145–53.

18. Ailes EC, Simeone RM, Dawson AL, Petersen EE, Gilboa SM. Using insurance claims data to identify and estimate critical periods in pregnancy: An application to antidepressants. Birth Defects Res Part A Clin Mol Teratol. 2016 Nov;106(11):927–34.

19. MacDonald SC, Cohen JM, Panchaud A, McElrath TF, Huybrechts KF, Hernández-Díaz S. Identifying pregnancies in insurance claims data: Methods and application to retinoid teratogenic surveillance. Pharmacoepidemiol Drug Saf. 2019 Sep;28(9):1211–21.

20. Huybrechts KF, Bateman BT, Hernández-Díaz S. Use of real-world evidence from healthcare utilization data to evaluate drug safety during pregnancy. Pharmacoepidemiol Drug Saf. 2019 Jul;28(7):906–22.

21. Stanford Center for Population Health Sciences. MarketScan Databases. Redivis. 2024;

22. Fisher A, Paterson JM, Winquist B, Wu F, Reynier P, Suissa S, et al. Patterns of antiemetic medication use during pregnancy: A multi-country retrospective cohort study. PLoS ONE. 2022 Dec 1;17(12):e0277623.

23. Slattery J, Quinten C, Candore G, Pinheiro L, Flynn R, Kurz X, et al. Ondansetron use in nausea and vomiting during pregnancy: A descriptive analysis of prescription patterns and patient characteristics in UK general practice. Br J Clin Pharmacol. 2022 Oct;88(10):4526–39.

24. Committee on Practice Bulletins-Obstetrics. ACOG practice bulletin no. 189: nausea and vomiting of pregnancy. Obstet Gynecol. 2018 Jan;131(1):e15–30.

25. van der Mijle AE, Woestenberg PJ. The use of ondansetron in the first trimester: Any differences after the EMA PRAC recommendation to not use it? Neurotoxicol Teratol. 2023 Jul;98:107217.

26. Huybrechts KF, Hernández-Díaz S, Straub L, Gray KJ, Zhu Y, Patorno E, et al. Association of Maternal First-Trimester Ondansetron Use With Cardiac Malformations and Oral Clefts in Offspring. JAMA. 2018 Dec 18;320(23):2429–37.

27. Gover A, Endrawes K, Molad M, Lavie-Nevo K, Riskin A. The effect of SSRI exposure in pregnancy on early respiratory and metabolic adaptation in infants born preterm. Children (Basel). 2023 Mar 4;10(3).

28. Bandoli G, Chambers CD, Wells A, Palmsten K. Prenatal antidepressant use and risk of adverse neonatal outcomes. Pediatrics. 2020 Jul;146(1).

29. Dubovicky M, Belovicova K, Csatlosova K, Bogi E. Risks of using SSRI / SNRI antidepressants during pregnancy and lactation. Interdiscip Toxicol. 2017 Sep;10(1):30–4.

30. Tran H, Robb AS. SSRI use during pregnancy. Semin Perinatol. 2015 Nov;39(7):545–7.

31. Alwan S, Friedman JM, Chambers C. Safety of selective serotonin reuptake inhibitors in pregnancy: A review of current evidence. CNS Drugs. 2016 Jun;30(6):499–515.

32. Irwinda R, Hiksas R, Siregar AA, Saroyo YB, Wibowo N. Long-chain polyunsaturated fatty acid (LC-PUFA) status in severe preeclampsia and preterm birth: a cross sectional study. Sci Rep. 2021 Jul 19;11(1):14701.

33. Hsu M-C, Tung C-Y, Chen H-E. Omega-3 polyunsaturated fatty acid supplementation in prevention and treatment of maternal depression: Putative mechanism and recommendation. J Affect Disord. 2018 Oct 1;238:47–61.

34. Ryan AS, Astwood JD, Gautier S, Kuratko CN, Nelson EB, Salem N. Effects of long-chain polyunsaturated fatty acid supplementation on neurodevelopment in childhood: a review of human studies. Prostaglandins Leukot Essent Fatty Acids. 2010 Jun;82(4–6):305–14.

35. Bärebring L, Nwaru BI, Lamberg-Allardt C, Thorisdottir B, Ramel A, Söderlund F, et al. Supplementation with long chain n-3 fatty acids during pregnancy, lactation, or infancy in relation to risk of asthma and atopic disease during childhood: a systematic review and meta-analysis of randomized controlled clinical trials. Food Nutr Res. 2022 Oct 11;66.

36. Ko JY, D'Angelo DV, Haight SC, Morrow B, Cox S, Salvesen von Essen B, et al. Vital Signs: Prescription Opioid Pain Reliever Use During Pregnancy - 34 U.S. Jurisdictions, 2019. MMWR Morb Mortal Wkly Rep. 2020 Jul 17;69(28):897–903.

37. Derry S, Derry CJ, Moore RA. Single dose oral ibuprofen plus oxycodone for acute postoperative pain in adults. Cochrane Database Syst Rev. 2013 Jun

26;2013(6):CD010289.

38. Celeste Durnwald, MD. Gestational diabetes mellitus: Glucose management and maternal prognosis. Waltham, MA.: UpToDate; 2024.

39. Cyr-Depauw C, Hurskainen M, Vadivel A, Mižíková I, Lesage F, Thébaud B. Characterization of the innate immune response in a novel murine model mimicking bronchopulmonary dysplasia. Pediatr Res. 2021 Mar;89(4):803–13.

40. Chandak P, Huang K, Zitnik M. Building a knowledge graph to enable precision medicine. BioRxiv. 2022 May 1;

41. Green EA, Garrick SP, Peterson B, Berger PJ, Galinsky R, Hunt RW, et al. The Role of the Interleukin-1 Family in Complications of Prematurity. Int J Mol Sci. 2023 Feb 1;24(3).

42. Bry K, Lappalainen U. Pathogenesis of bronchopulmonary dysplasia: the role of interleukin 1beta in the regulation of inflammation-mediated pulmonary retinoic acid pathways in transgenic mice. Semin Perinatol. 2006 Jun;30(3):121–8.

43. Dixon CM, Colthup PV, Serabjit-Singh CJ, Kerr BM, Boehlert CC, Park GR, et al. Multiple forms of cytochrome P450 are involved in the metabolism of ondansetron in humans. Drug Metab Dispos. 1995 Nov;23(11):1225–30.

44. Lewis DFV, Ito Y, Eddershaw PJ, Dickins M, Goldfarb PS. An evaluation of ondansetron binding interactions with human cytochrome P450 enzymes CYP3A4 and CYP2D6. Drug Metab Lett. 2010 Jan;4(1):25–30.

45. Shedlofsky SI, Swim AT, Robinson JM, Gallicchio VS, Cohen DA, McClain CJ. Interleukin-1 (IL-1) depresses cytochrome P450 levels and activities in mice. Life Sci. 1987 Jun 15;40(24):2331–6.

46. Weaver J, Hardin JH, Blacketer C, Krumme AA, Jacobson MH, Ryan PB. Development and evaluation of an algorithm to link mothers and infants in two US commercial healthcare claims databases for pharmacoepidemiology research. BMC Med Res Methodol. 2023 Oct 21;23(1):246.

47. Moll K, Wong HL, Fingar K, Hobbi S, Sheng M, Burrell TA, et al. Validating Claims-Based Algorithms Determining Pregnancy Outcomes and Gestational Age Using a Linked Claims-Electronic Medical Record Database. Drug Saf. 2021 Nov;44(11):1151–64.

48. Bertoia ML, Phiri K, Clifford CR, Doherty M, Zhou L, Wang LT, et al. Identification of pregnancies and infants within a US commercial healthcare administrative claims database. Pharmacoepidemiol Drug Saf. 2022 Aug;31(8):863–74.

49. Stover MW, Davis JM. Opioids in pregnancy and neonatal abstinence syndrome. Semin Perinatol. 2015 Nov;39(7):561–5.

50. Datta A, Flynn NR, Barnette DA, Woeltje KF, Miller GP, Swamidass SJ. Machine learning liver-injuring drug interactions with non-steroidal anti-inflammatory drugs (NSAIDs) from a

retrospective electronic health record (EHR) cohort. PLoS Comput Biol. 2021 Jul 6;17(7):e1009053.

51. Ellis JJ, Sadosky AB, Ten Eyck LL, Mudumby P, Cappelleri JC, Ndehi L, et al. A retrospective, matched cohort study of potential drug-drug interaction prevalence and opioid utilization in a diabetic peripheral neuropathy population initiated on pregabalin or duloxetine. BMC Health Serv Res. 2015 Apr 15;15:159.
52. Choi YH, Lee IH, Yang M, Cho YS, Jo YH, Bae HJ, et al. Clinical significance of potential drug-drug interactions in a pediatric intensive care unit: A single-center retrospective study. PLoS ONE. 2021 Feb 8;16(2):e0246754.
53. Hirai T, Kondo Y, Sakazaki Y, Seki A, Ishitsuka Y, Iwamoto T. Drug-drug interaction signals between loop diuretics and teicoplanin during acute kidney injury evaluated using Japanese spontaneous adverse drug event reports. Sci Rep. 2023 Aug 26;13(1):13989.
54. Wu P, Nelson SD, Zhao J, Stone CA, Feng Q, Chen Q, et al. DDIWAS: High-throughput electronic health record-based screening of drug-drug interactions. J Am Med Inform Assoc. 2021 Jul 14;28(7):1421–30.
55. Wilde MI, Markham A. Ondansetron. A review of its pharmacology and preliminary clinical findings in novel applications. Drugs. 1996 Nov;52(5):773–94.
56. Nikolić J, Vukojević K, Šoljić V, Mišković J, Orlović Vlaho M, Saraga-Babić M, et al. Expression Patterns of Serotonin Receptors 5-HT1A, 5-HT2A, and 5-HT3A during Human Fetal Lung Development. Int J Mol Sci. 2023 Feb 3;24(3).
57. Mechanism matters. Nat Med. 2010 Apr;16(4):347–347.
58. Davis RL. Mechanism of action and target identification: A matter of timing in drug discovery. iScience. 2020 Aug 21;23(9):101487.
59. Ruiz C, Zitnik M, Leskovec J. Identification of disease treatment mechanisms through the multiscale interactome. Nat Commun. 2021 Mar 19;12(1):1796.
60. Australian categorisation system for prescribing medicines in pregnancy.
61. Stanford Center for Population Health Sciences. MarketScan Redbook. Redivis. 2023;
62. Chen Q, Nian H, Zhu Y, Talbot HK, Griffin MR, Harrell FE. Too many covariates and too few cases? - a comparative study. Stat Med. 2016 Nov 10;35(25):4546–58.
63. Huang H, Gao Y, Zhang H, Li B. Weighted Lasso Estimates for Sparse Logistic Regression: Non-asymptotic Properties with Measurement Error. arXiv. 2020;
64. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995 Jan;57(1):289–300.
65. Lu H, Cole SR, Platt RW, Schisterman EF. Revisiting Overadjustment Bias.

Epidemiology. 2021 Sep 1;32(5):e22–3.

66. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in regression analyses conducted in epidemiologic studies. Epidemiology (Sunnyvale). 2016 Apr;6(2).

67. Miles J. Tolerance and Variance Inflation Factor. In: Kenett RS, Longford NT, Piegorsch WW, Ruggeri F, editors. Wiley StatsRef: Statistics Reference Online [Internet]. 1st ed. Wiley; 2014 [cited 2023 Dec 19]. Available from: https://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat06593

68. Bateman BT, Mhyre JM, Hernandez-Diaz S, Huybrechts KF, Fischer MA, Creanga AA, et al. Development of a comorbidity index for use in obstetric patients. Obstet Gynecol. 2013 Nov;122(5):957–65.

69. Zuur AF, Ieno EN, Elphick CS. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution. 2010 Mar;1(1):3–14.

70. Craney TA, Surles JG. Model-Dependent Variance Inflation Factor Cutoff Values. Qual Eng. 2002 Mar 25;14(3):391–403.

71. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996 Jan;58(1):267–88.

72. Hastie T, Tibshirani R, Wainwright M. Statistical Inference. Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman and Hall/CRC; 2015. p. 155–82.

73. Bloniarz A, Liu H, Zhang C-H, Sekhon JS, Yu B. Lasso adjustments of treatment effect estimates in randomized experiments. Proc Natl Acad Sci USA. 2016 Jul 5;113(27):7383–90.

74. Maaten L van der, Hinton G. Visualizing Data using t-SNE. Journal of Machine Learning Research. 2008;

75. Noguchi Y, Tachi T, Teramachi H. Review of Statistical Methodologies for Detecting Drug-Drug Interactions Using Spontaneous Reporting Systems. Front Pharmacol. 2019 Nov 8;10:1319.

76. Chandak P, Huang K, Zitnik M. Building a knowledge graph to enable precision medicine. Sci Data. 2023 Feb 2;10(1):67.

77. Dijkstra EW. A note on two problems in connexion with graphs. Numer Math. 1959 Dec;1(1):269–71.

78. Heyne GW, Plisch EH, Melberg CG, Sandgren EP, Peter JA, Lipinski RJ. A simple and reliable method for early pregnancy detection in inbred mice. J Am Assoc Lab Anim Sci. 2015 Jul;54(4):368–71.

79. Crowley G, Kwon S, Caraher EJ, Haider SH, Lam R, Batra P, et al. Quantitative lung morphology: semi-automated measurement of mean linear intercept. BMC Pulm Med. 2019 Nov 9;19(1):206.

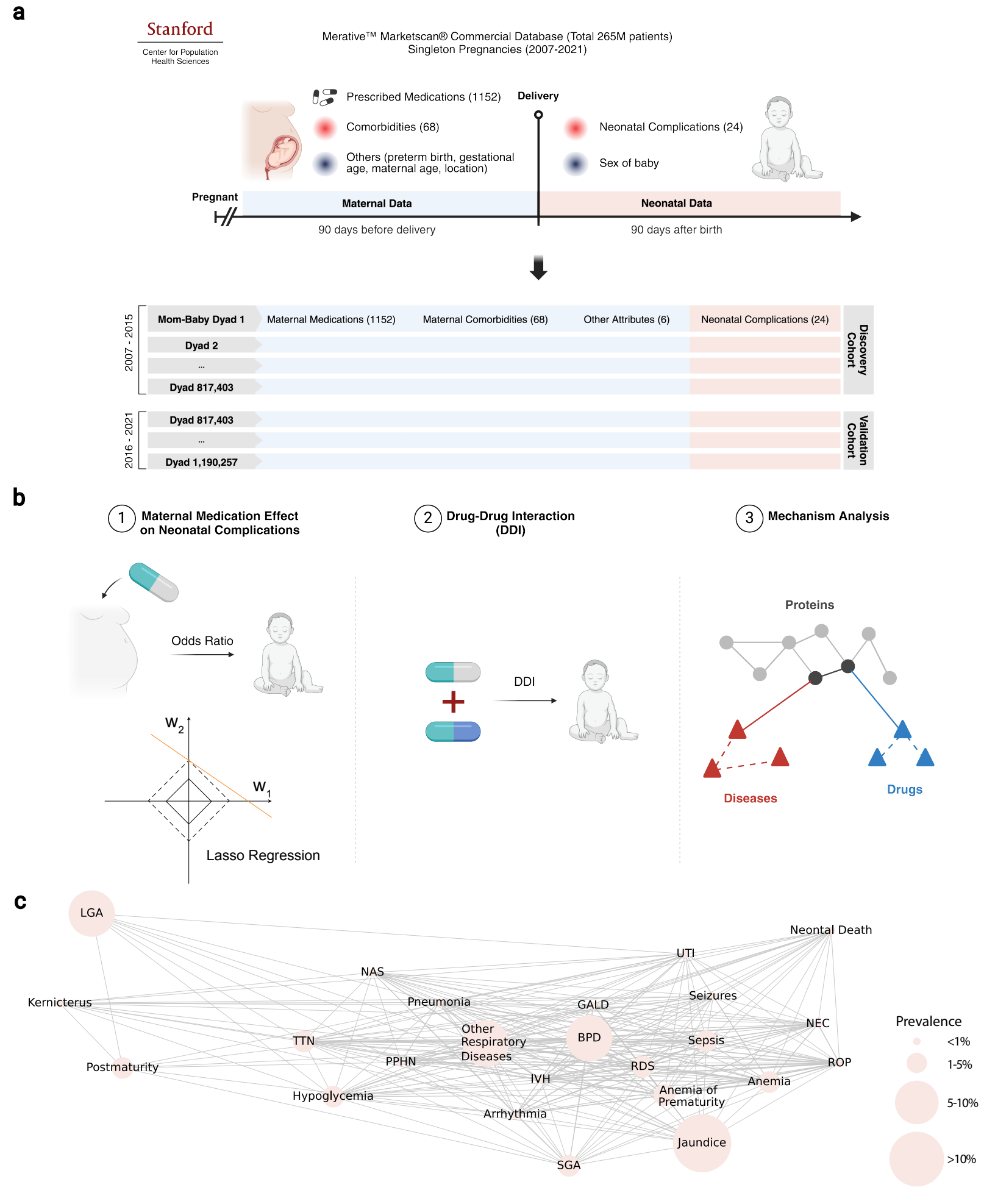
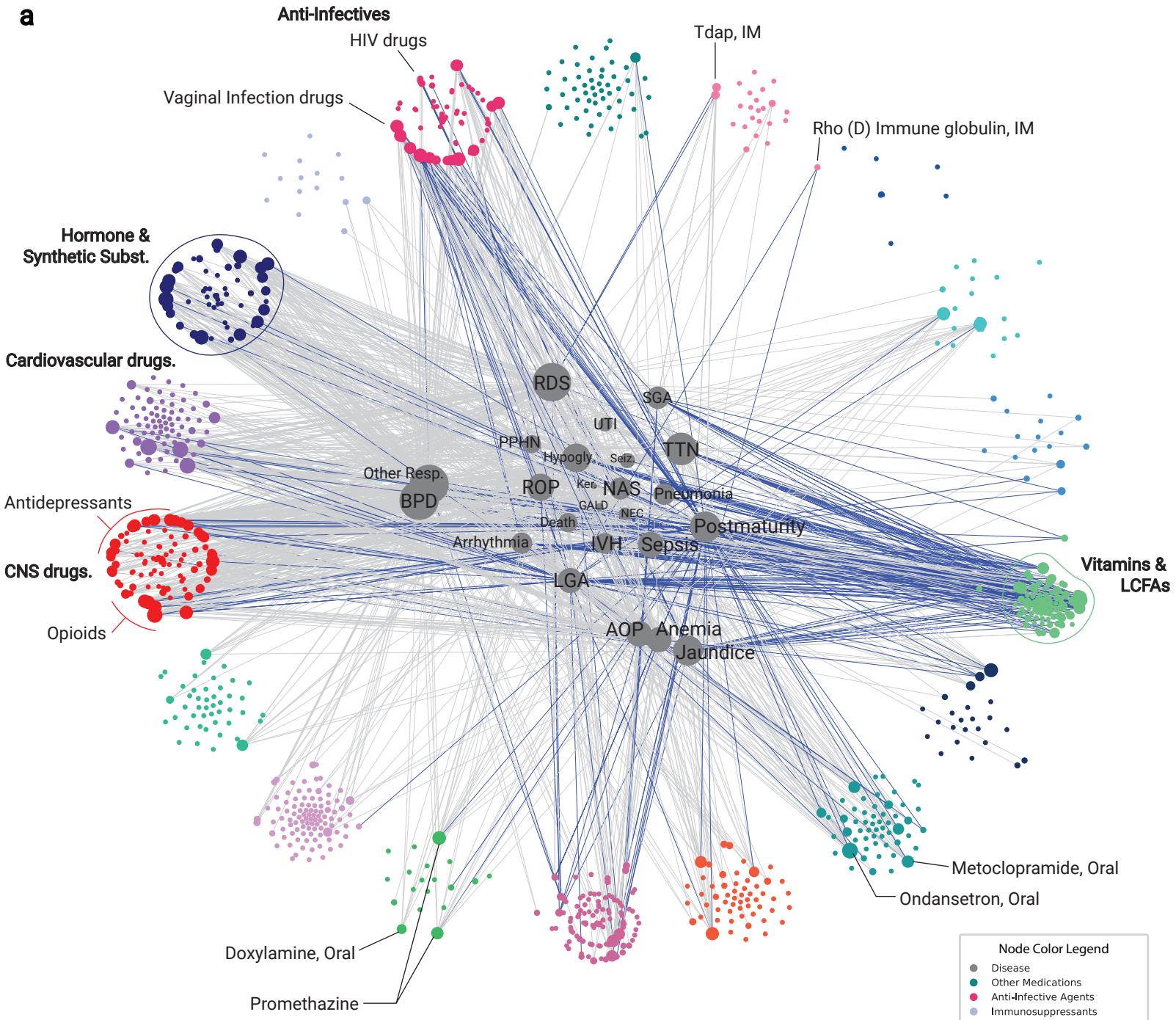
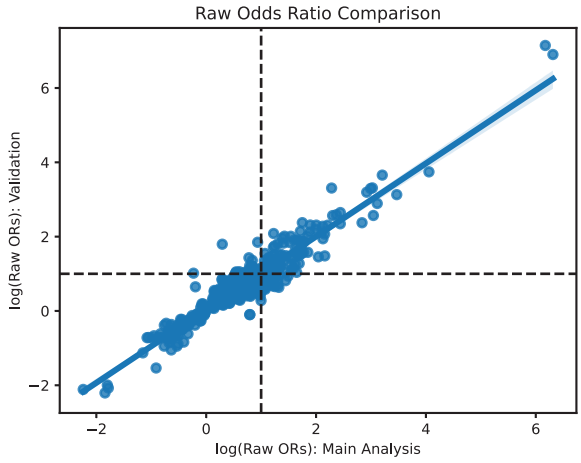

a
Stanford
Center for Population Health Sciences
Merative™ Marketscan® Commercial Database (Total 265M patients)
Singleton Pregnancies (2007-2021)
Prescribed Medications (1152)
Comorbidities (68)
Others (preterm birth, gestational age, maternal age, location)
Delivery
Neonatal Complications (24)
Sex of baby
Pregnant
Maternal Data
Neonatal Data
90 days before delivery
90 days after birth
2007 - 2015
Mom-Baby Dyad 1
Maternal Medications (1152)
Maternal Comorbidities (68)
Other Attributes (6)
Neonatal Complications (24)
Dyad 2
...
Dyad 817,403
Discovery Cohort
2016 - 2021
Dyad 817,403
...
Dyad 1,190,257
Validation Cohort
b
1 Maternal Medication Effect on Neonatal Complications
2 Drug-Drug Interaction (DDI)
3 Mechanism Analysis
Odds Ratio
W2
W1
Lasso Regression
DDI
Proteins
Diseases
Drugs
c
LGA
Neontal Death
UTI
NAS
Kernicterus
Pneumonia
GALD
Seizures
NEC
TTN
Other Respiratory Diseases
BPD
Sepsis
Postmaturity
PPHN
RDS
ROP
IVH
Anemia
Hypoglycemia
Anemia of Prematurity
Arrhythmia
Jaundice
SGA
Prevalence
<1%
1-5%
5-10%
>10%

Figure 1. Study Overview. (a) A total of 1,190,257 mother-newborn dyads from 2007 to 2021 are extracted from the Merative ™ Marketscan® Commercial Database, a set of de-identified claims data accessed through the Stanford Center for Population Health Sciences. Maternal records from the last 90 days before pregnancy are examined to extract 68 maternal comorbidities, 6 additional attributes, and all prescribed medications. Similarly, newborn records from the first 90 days after birth are analyzed to identify 24 neonatal complications. The final dataset includes 1,152 maternal medications, 68 maternal comorbidities, 24 neonatal complications, and 6 additional attributes: maternal age, gestational age at delivery, preterm birth, baby's sex, and the region and geographic location of the healthcare facility. The dyads are divided into a discovery cohort (2007-2015) and a validation cohort (2016-2021) based on the year of birth. (b) Three experiments are conducted based on the database. First, the associations between maternal medications and neonatal complications are calculated using odds ratios based on logistic regression. Notably, confounders are selected and adjusted for each medication-disease pair using the LassoNoExp algorithm, which penalizes and selects confounding variables using Lasso regression while preserving maternal medication exposure information. Second, additive effects of concomitant maternal medications on neonatal complications are calculated through drug-drug interaction analysis. Lastly, potential underlying mechanisms of newly discovered drug-disease associations are discovered using a knowledge graph and a graph learning method. (c) A tetrachoric correlation plot of the 24 neonatal outcomes shows the relationships between the neonatal complications. The size of each node is correlated with disease prevalence, and the nodes are connected if the diseases are positively correlated. The edge thickness is proportional to the strength of the correlation.thickness is proportional to the strength of the correlation.


a
Anti-Infectives
HIV drugs
Vaginal Infection drugs
Tdap, IM
Rho (D) Immune globulin, IM
Hormone & Synthetic Subst.
Cardiovascular drugs.
Antidepressants
CNS drugs.
Opioids
Vitamins & LCFAs
RDS
SGA
UTI
PPHN
Hypogly.
Seiz.
TTN
Other Resp.
BPD
ROP
Ker.
NAS
Pneumonia
GALD
NEC
Death
Postmaturity
Arrhythmia
IVH
Sepsis
LGA
AOP
Anemia
Jaundice
Metoclopramide, Oral
Ondansetron, Oral
Doxylamine, Oral
Promethazine
Node Color Legend
Disease
Other Medications
Anti-Infective Agents
Immunosuppressants
Hormones & Synthetic Subst
Cardiovascular Agents
Central Nervous System
Electrolytic, Caloric, Water
Eye, Ear, Nose Throat
Antihistamines & Comb.
Skin & Mucous Membrane
Autonomic Drugs
Gastrointestinal Drugs
Blood Form/Coagul Agents
Vitamins & Comb
Respiratory Tract Agents
Pharmaceutical Aids/Adjuvants
Antineoplastic Agents
Serums, Toxoids, Vaccines


b
Raw Odds Ratio Comparison
log(Raw ORs): Validation
log(Raw ORs): Main Analysis
−2
0
2
4
6

r^2=0.972

Figure 2. A Network of the Associations Between Maternal Medications and Neonatal Complications, and Their Validation. (a) Odds Ratio (OR) results are displayed as a network graph, where nodes represent maternal medications or neonatal complications, and edges indicate statistically significant ORs. Disease nodes are centrally placed, while maternal medication nodes are peripheral. Medications are grouped and color-coded within each cluster. The graph shows significant ORs as edges; gray for ORs above 1 (positive correlations), and blue for ORs below 1 (negative correlations). There are 24 disease nodes, 1,152 medication nodes, and 1,446 edges—1,153 gray and 293 blue. Edge thickness correlates with the strength of the association (inverse of p-values), and node size with the number of connecting edges. Abbreviations in disease nodes: BPD: Bronchopulmonary Dysplasia, Other Resp.: Other Respiratory Diseases, RDS: Respiratory Distress Syndrome, SGA: Small for Gestational Age, PPHN: Persistent Pulmonary Hypertension, Hypogly.: Hypoglycemia, UTI: Urinary Tract Infection, Seiz.: Seizures, TTN: Transient Tachypnea, ROP: Retinopathy of Prematurity, Ker.: Kernicterus, GALD: Gestational Alloimmune Liver Disease, NAS: Neonatal Abstinence Syndrome, NEC: Necrotizing Enterocolitis, IVH: Intraventricular Hemorrhage, Death: Neonatal Death, LGA: Large for Gestational Age, AOP: Anemia of Prematurity. Additional conditions in disease nodes not abbreviated: Pneumonia, Postmaturity, Sepsis, Arrhythmia, Anemia, and Jaundice. (b) ORs from the discovery cohort were validated in the validation cohort. ORs for both cohorts were matched by maternal medications and neonatal complications and plotted on a scatterplot. This figure compares the log of significant ORs from the discovery (x-axis) to the validation cohort (y-axis). Of the ORs, 772 matched between the cohorts—1,446 significant in the discovery cohort and 1,097 in the validation cohort. Pearson correlation yielded an R^2 of 0.972 with a p-value of 0.0.

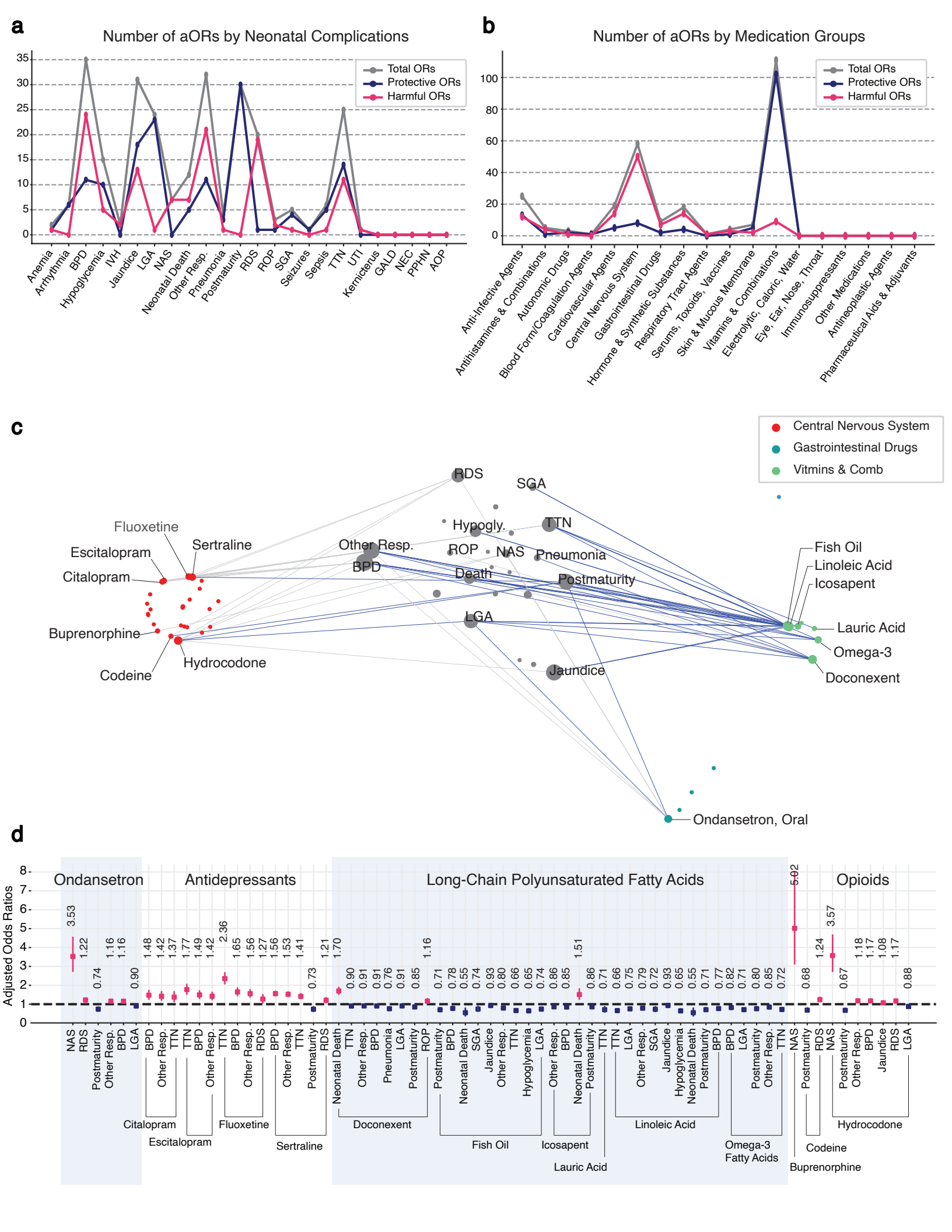

a
Number of aORs by Neonatal Complications
Total ORs
Protective ORs
Harmful ORs
Anemia
Arrhythmia
BPD
Hypoglycemia
IVH
Jaundice
LGA
NAS
Neonatal Death
Other Resp.
Pneumonia
Postmaturity
RDS
ROP
SGA
Seizures
Sepsis
TTN
UTI
Kerniciterus
GALD
NEC
PPHN
AOP
b
Number of aORs by Medication Groups
Total ORs
Protective ORs
Harmful ORs
Anti-Infective Agents
Antihistamines & Combinations
Autonomic Drugs
Blood Form/Coagulation Agents
Cardiovascular Agents
Central Nervous System
Gastrointestinal Drugs
Hormone & Synthetic Substances
Respiratory Tract Agents
Serums, Toxoids, Vaccines
Skin & Mucous Membrane
Vitamins & Combinations
Electrolytic, Caloric, Water
Eye, Ear, Nose, Throat
Immunosuppressants
Other Medications
Antineoplastic Agents
Pharmaceutical Aids & Adjuvants
c
Central Nervous System
Gastrointestinal Drugs
Vitmins & Comb
RDS
SGA
Hypogly.
TTN
Fluoxetine
Sertraline
Escitalopram
Citalopram
Other Resp.
BPD
ROP
NAS
Pneumonia
Death
Postmaturity
LGA
Buprenorphine
Codeine
Hydrocodone
Jaundice
Fish Oil
Linoleic Acid
Icosapent
Lauric Acid
Omega-3
Doconexent
Ondansetron, Oral
d
Adjusted Odds Ratios
Ondansetron
Antidepressants
Long-Chain Polyunsaturated Fatty Acids
Opioids
Citalopram
Escitalopram
Fluoxetine
Sertraline
Doconexent
Fish Oil
Icosapent
Lauric Acid
Linoleic Acid
Omega-3 Fatty Acids
Codeine
Buprenorphine
Hydrocodone

Figure 3. Associations Between Four Maternal Medication Groups and Neonatal Complications: Highlights from the Adjusted Odds Ratio Network (a) The figure shows the number of aORs by neonatal complications. The gray line represents the total aORs, the navy line indicates reduced risk aORs, and the magenta indicates increased risk aORs. (b) The figure shows the number of aORs by maternal medication groups, with the line colors corresponding to those in (a). (c) This network graph highlights four maternal medication groups and their adjusted odds ratios (aOR) with neonatal complications. Nodes are consistent with those in Figure 2's unadjusted odds ratio graph, while edges represent statistically significant adjusted odds ratios. LassoNoExp was utilized to calculate these aORs, selecting and adjusting for maternal covariates out of a possible 73 for each drug-complication pair. Edges in gray indicate aORs greater than 1 (positive correlations), and those in blue indicate aORs less than 1 (negative correlations). The edge thickness corresponds to the association strength, inversely proportional to the p-values. Highlighted medication groups include Ondansetron, Antidepressants, Long-Chain Polyunsaturated Fatty Acids (LC-PUFAs), and Opioids. The full network can be examined in Supplementary Figure 6. (d) The subfigure details the values of statistically significant adjusted odds ratios and their confidence intervals for the four medication groups. Magenta dots and lines depict aORs higher than 1, signifying positive correlations, whereas navy blue dots and lines indicate aORs lower than 1, representing negative correlations.

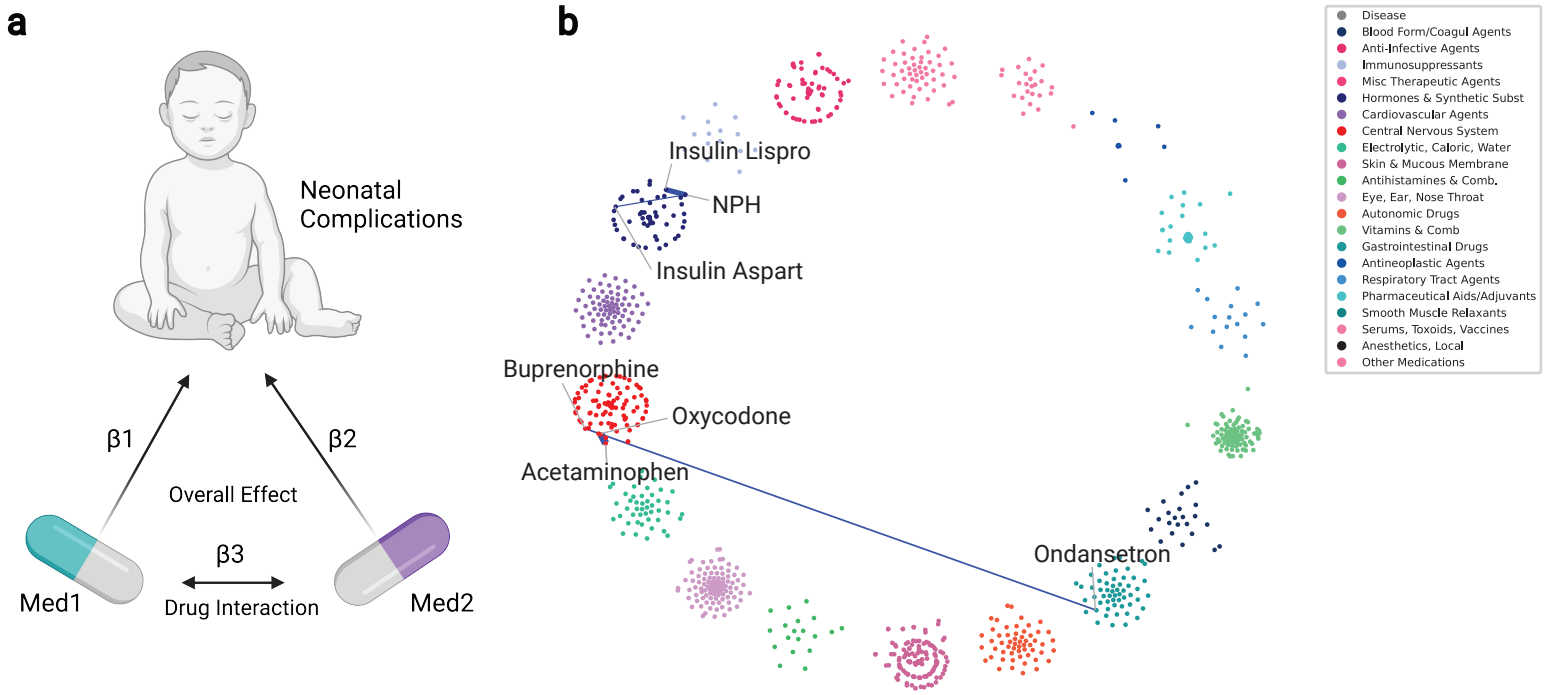


Figure 4. Maternal Drug-Drug Interaction Effects on Neonatal Complications.
(a) Drug-Drug Interaction (DDI) effects from the concomitant use of maternal medications were calculated based on mother-baby dyads. Here, β_1, β_2, and β_3 represent the individual contributions of the medications and their interaction to the overall effect on neonatal complications. Specifically, β_1 represents the impact of medication 1, β_2 the impact of medication 2, and β_3 the interaction contribution between medications 1 and 2 to the collective effect on neonatal complications. (b) A network graph illustrates the DDIs as edges connecting the medication nodes. Notably, five DDIs significant in both discovery and validation cohorts are illustrated. All five edges are colored blue, indicating that the DDIs reduce neonatal complications, thereby offering additional protective benefits through the combination of maternal medications.

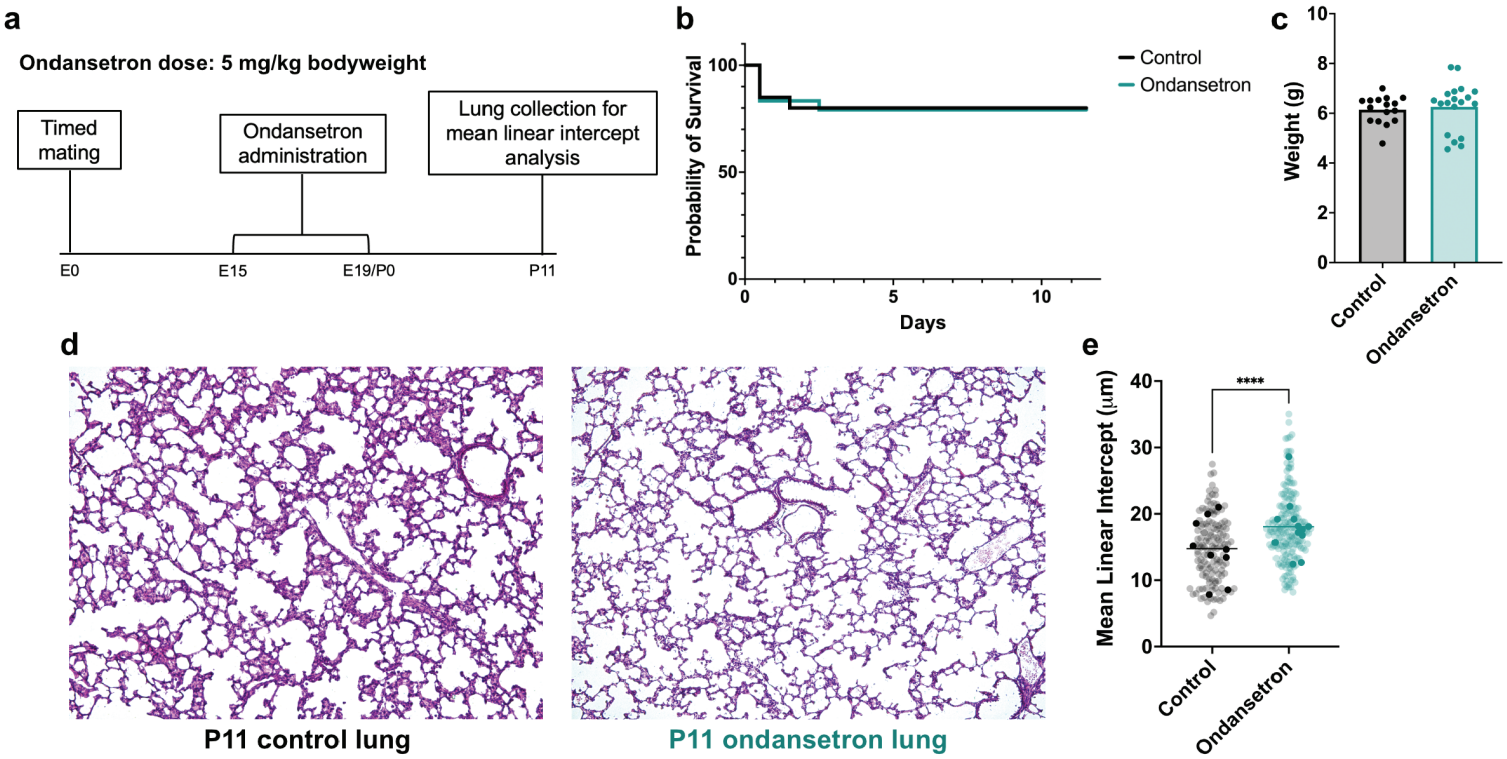


Figure 5: Impact of maternal ondansetron exposure and neonatal lung development in mice. (a) Experimental timeline of maternal ondansetron administration and quantification of lung development. Timed pregnant C57BL/6 female mice were injected daily with ondansetron (5 mg/kg i.p.) or an equivalent volume of sterile PBS (control) from E15 until delivery. (b) Kaplan-Meier survival analysis and (c) neonatal weight of control and ondansetron exposed mice. N = 20 mice from 4 litters in the control group; n = 24 mice from 5 litters in the ondansetron group. (d) Representative H&E images of lungs from control and ondansetron exposed mice at postnatal day 11. (e) Mean linear intercept quantification of control and ondansetron exposed mice. For each mouse, 3 lung sections were stained, with 5 random images taken from each section for a total of 135-180 images from 9-12 mice. Darker, overlayed dots represent average mean linear intercept values for each individual mouse. **** indicates $P < 0.0001$ as determined using an unpaired t test.

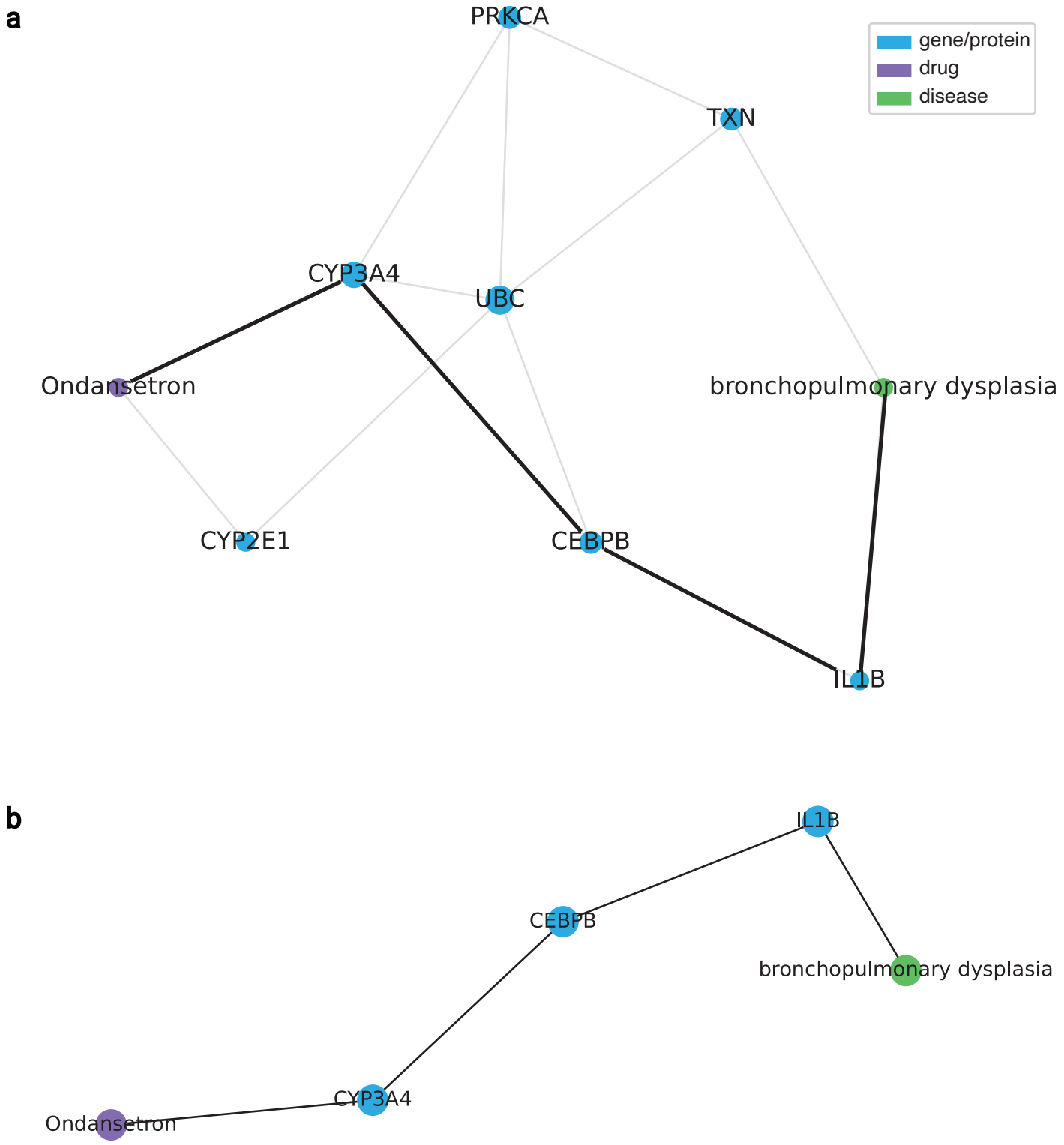
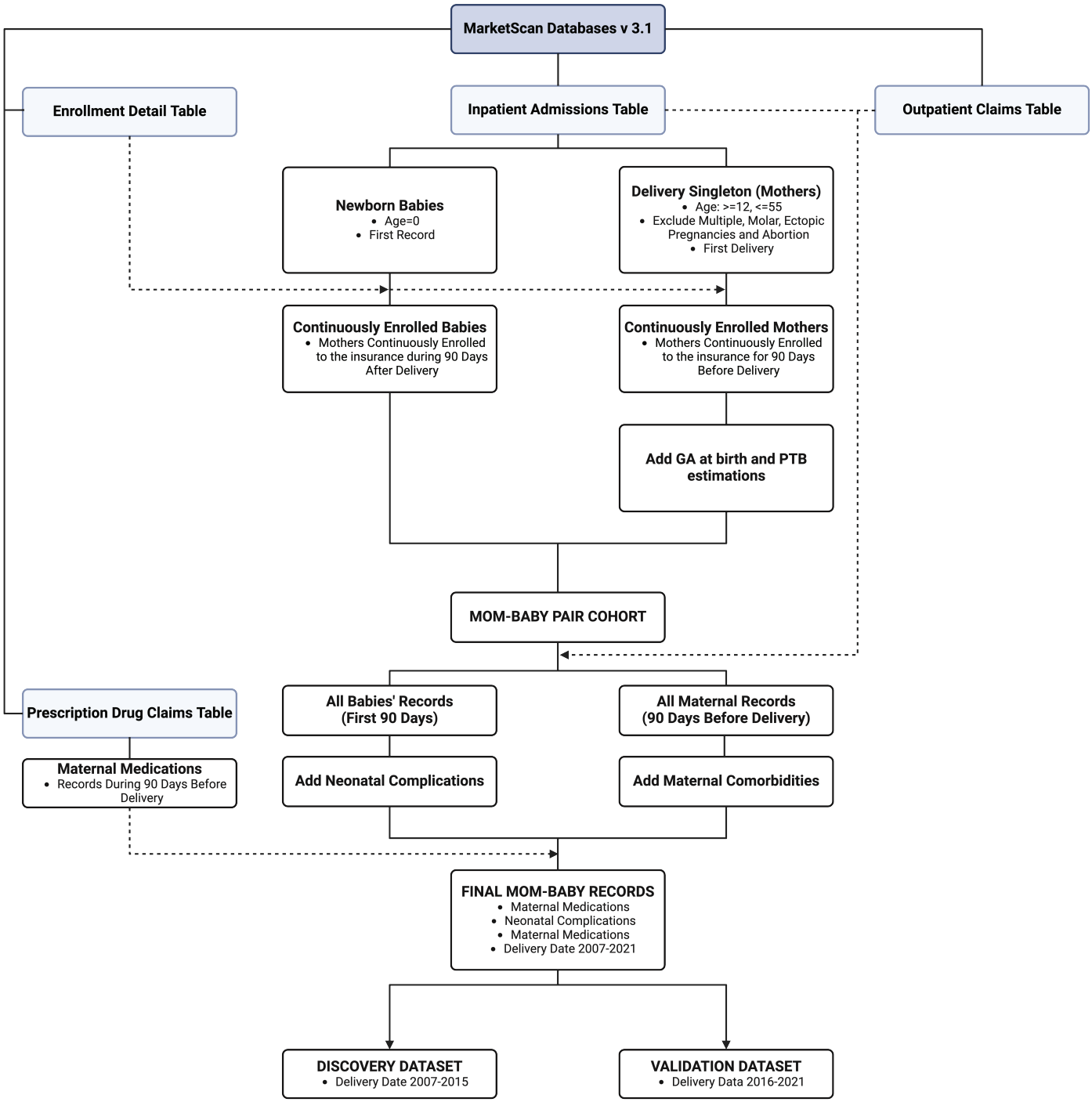


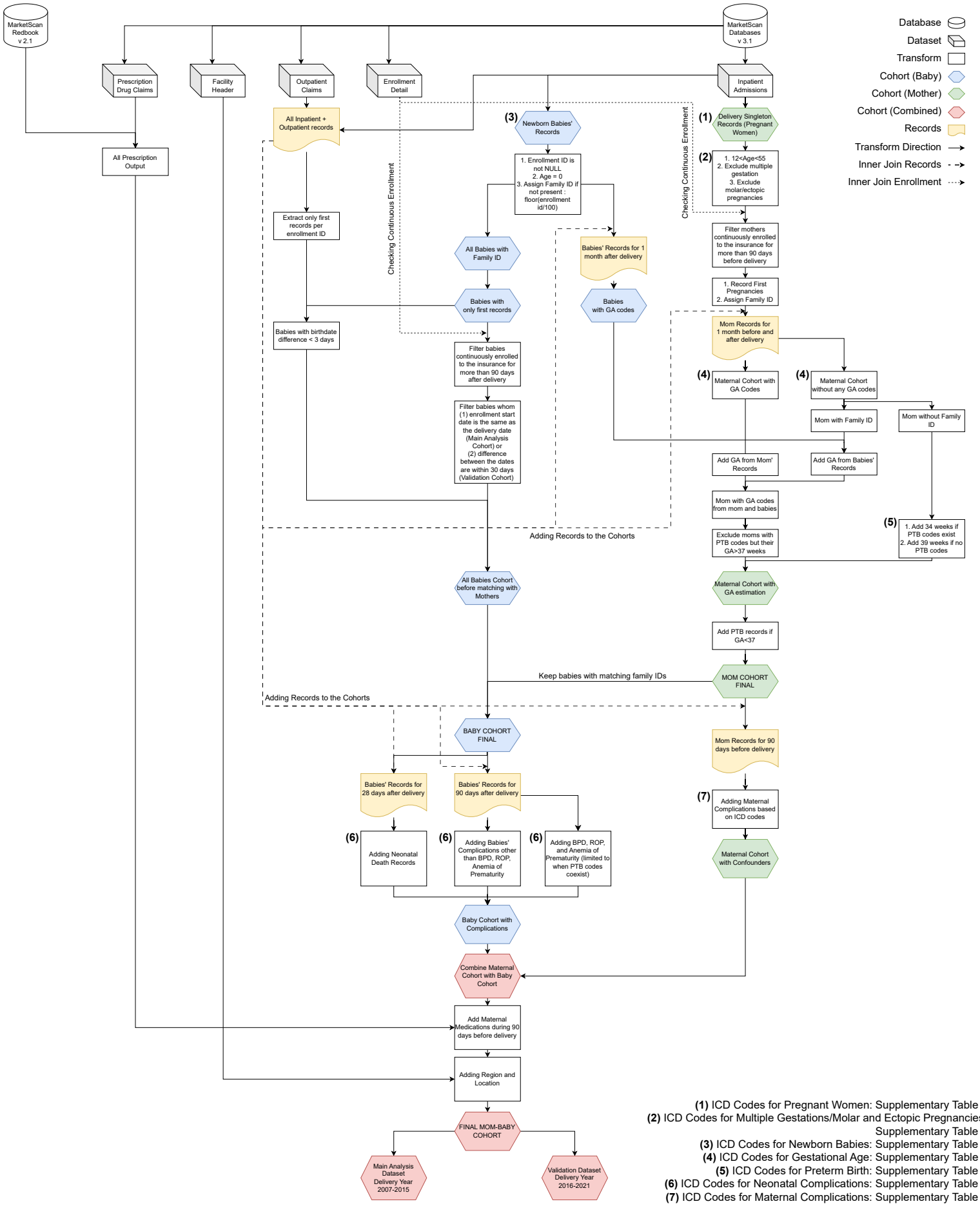
Figure 6. Potential Underlying Mechanisms of Ondansetron Effects on Bronchopulmonary Dysplasia. (a) This figure illustrates the potential underlying mechanisms behind the newly discovered associations between maternal medications and neonatal complications, utilizing significant adjusted odds ratios from PregMedNet and the PrimeKG Knowledge Graph, which contains scientifically validated biological interactions. Specifically, the subgraph containing three different node types—drugs, proteins/genes, and diseases—is shown here to generate more reliable hypotheses than other types of subgraphs containing additional biological node types. Displayed is the subgraph of all nodes and edges that connect ondansetron to bronchopulmonary dysplasia via the shortest paths between the two nodes. This results in four paths with seven intermediary nodes. Each path represents a hypothesis for a possible underlying mechanism of action. (b) A particular path of interest suggests that ondansetron affects bronchopulmonary dysplasia through the protein interactions between CYP3A5 and IL-1b. This hypothesis builds on previous research showing the relationship between ondansetron and CYP3A, and a separate relationship between bronchopulmonary dysplasia and IL-1b. While there is no prior published literature that relates CYP3A5 to IL-1b in the context of ondansetron and bronchopulmonary dysplasia, PregMedNet can generate hypotheses for novel relationships by connecting already proven, but isolated biological knowledge using graph learning methods.

Disease	Medication 1	Medication 2	β_3	OR_1	OR_2	OR_{12}
NAS	**Acetaminophen, Oral**	**Oxycodone, Oral**	-3.50	4.89	99.97	14.77
Hypoglycemia	**NPH, subcutaneous**	**Insulin Lispro, Injection**	-1.84	2.19	5.92	2.07
Hypoglycemia	**NPH, subcutaneous**	**Insulin Aspart, Injection**	-1.52	2.14	5.42	2.55
NAS	**Buprenorphine, sublingual**	**Ondansetron, Oral**	-1.89	607.34	5.17	474.43
RDS	**NPH, subcutaneous**	**Insulin Aspart, Injection**	-1.02	1.44	2.25	1.17

Table 1. Maternal Drug-Drug Interaction Effects on Neonatal Complications. This table details the five significant Drug-Drug Interactions (DDI). The term β_3 represents the contribution of the DDI to the overall effect of medication pairs on a neonatal complication. OR_1, OR_2, and OR_{12} represent the effect of medication 1, medication 2, and the concurrent use of two medications on a neonatal complication, respectively. A negative β_3 indicates the reduced risk effects of the DDI on neonatal complications.

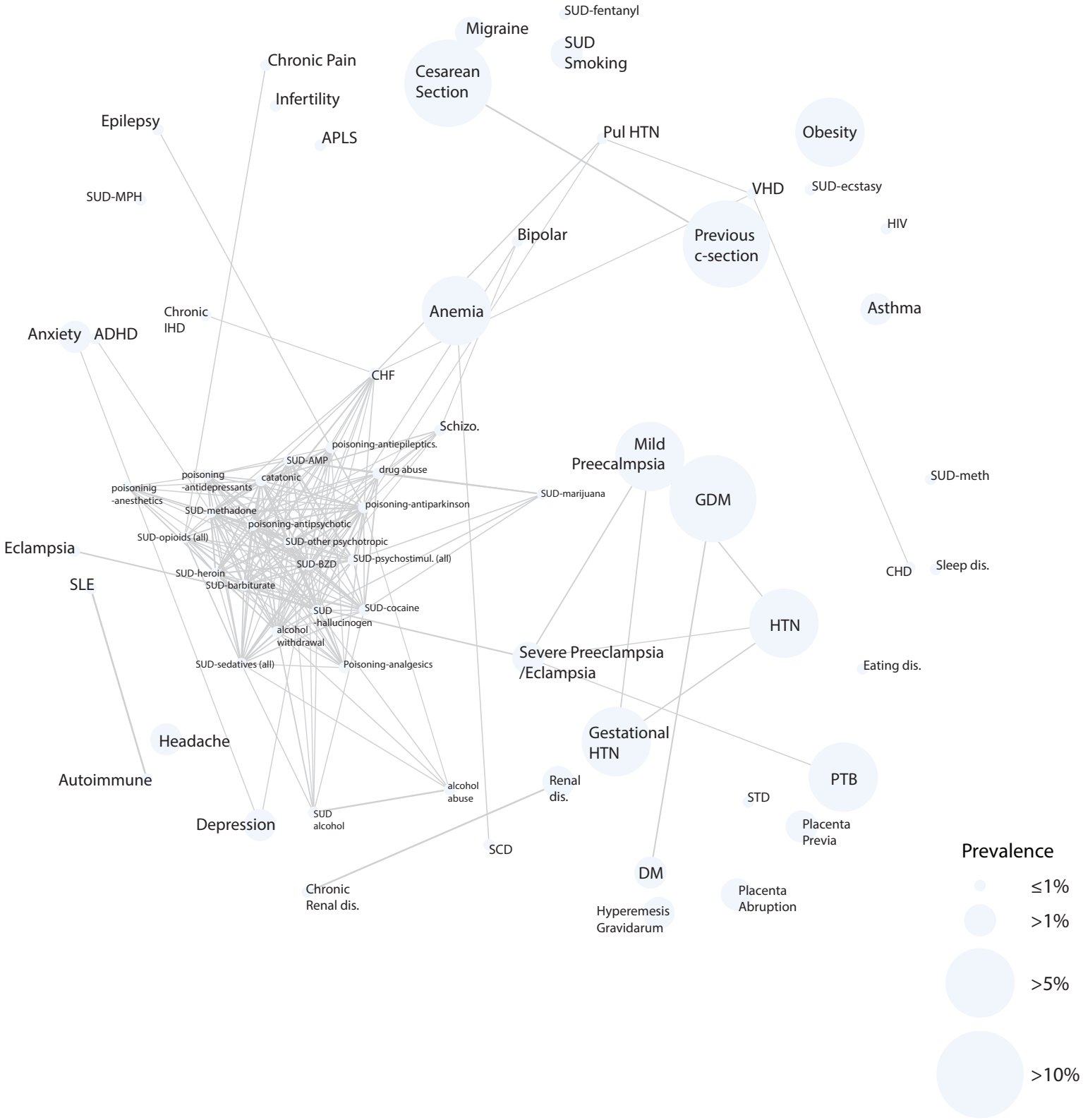


Supplementary Figure 1. Flow Chart of the Data Construction Overview. This figure illustrates the high-level process of data cleaning and preparation for analysis derived from the Merative™ Marketscan® Commercial Database.



(1) ICD Codes for Pregnant Women: Supplementary Table 3
(2) ICD Codes for Multiple Gestations/Molar and Ectopic Pregnancies: Supplementary Table 4
(3) ICD Codes for Newborn Babies: Supplementary Table 5
(4) ICD Codes for Gestational Age: Supplementary Table 6
(5) ICD Codes for Preterm Birth: Supplementary Table 7
(6) ICD Codes for Neonatal Complications: Supplementary Table 8
(7) ICD Codes for Maternal Complications: Supplementary Table 9

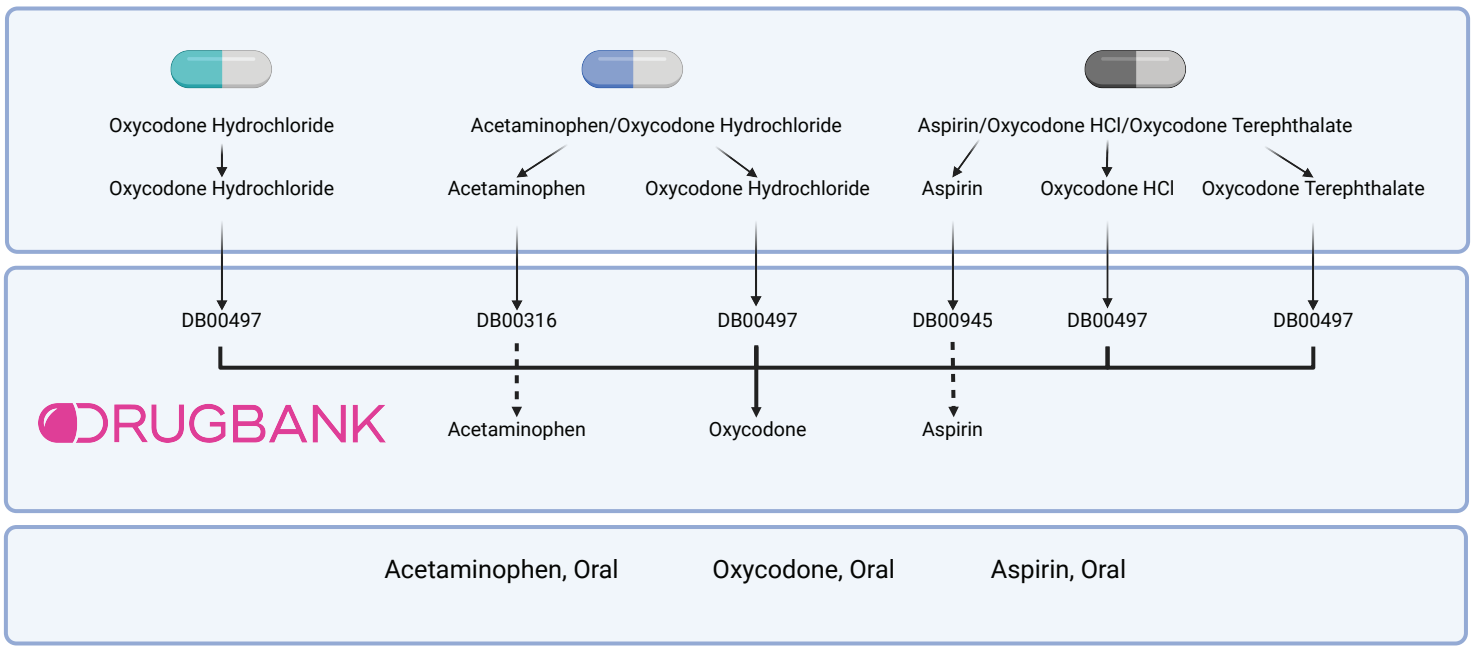
Supplementary Figure 2. Comprehensive Flowchart Detailing Data Construction. This figure details the extensive process of data cleaning and preparation undertaken using the Merative™ Marketscan® Commercial Database. It expands upon the methodologies outlined in Supplementary Figure 1, providing clarity on the data preparation steps for the study. The flowchart depicts the procedural actions executed to assemble and refine the dataset, leading to the creation of a table that encapsulates maternal medication use, maternal comorbidities, 6 other attributes (maternal age, gestational age at delivery, preterm birth, baby's gender, and the region and geographic location of the healthcare facility), and neonatal complications. This offers a thorough perspective at the individual patient level and serves to enhance the foundational information displayed in Figure 1(a).



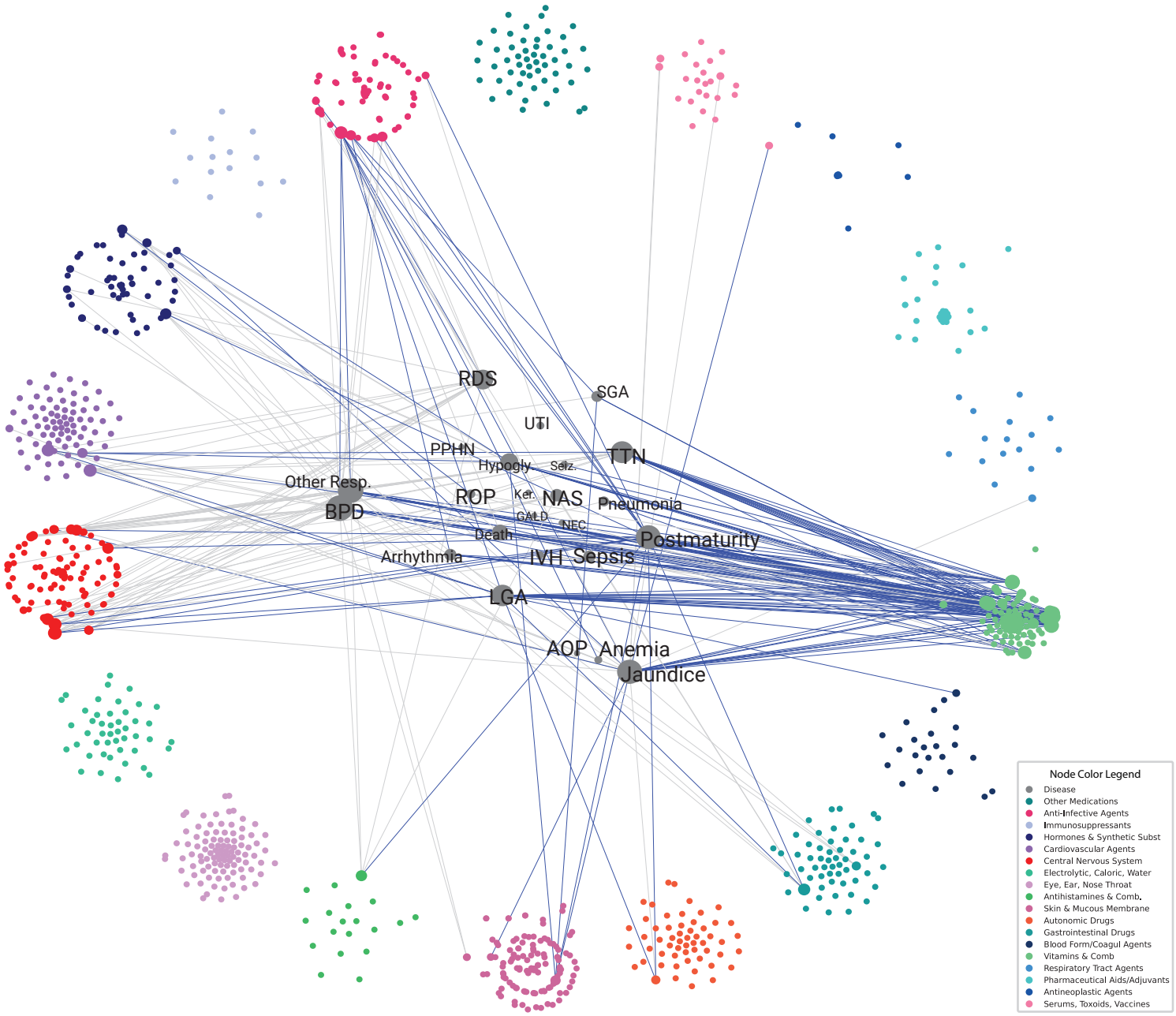
<u>Supplementary Figure 3. Correlation Plot of Potential Maternal Covariates.</u> This tetrachoric correlation plot displays the interrelationships among 68 maternal covariates and preterm birth, illustrating potential connections between various diseases. The diameter of each node is proportionate to the prevalence of the disease it represents. Nodes are connected if their tetrachoric correlation exceeds 0.5, with the thickness of the edges reflecting the strength of the correlation. The abbreviations used in the figure are expanded as follows:

- ADHD: Attention-deficit hyperactivity disorders
- alcohol abuse: Alcohol Abuse
- alcohol withdrawal: Alcohol Withdrawal Syndrome
- Anemia: Anemia
- Anxiety: Anxiety
- APLS: Antiphospholipid Syndrome
- Asthama: Asthma

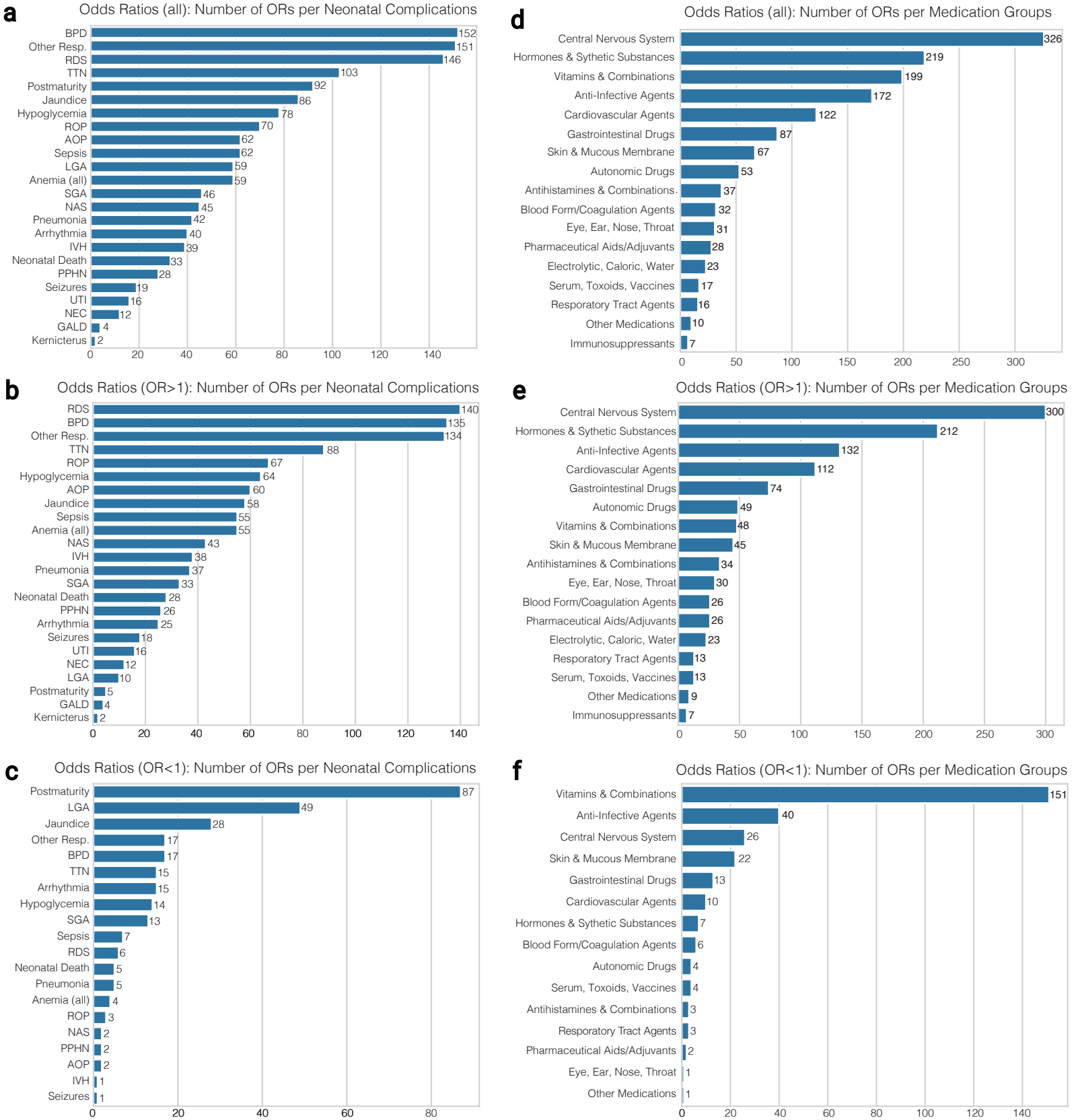
- Autoimmune: Autoimmune Disorders
- Bipolar: Bipolar Disorder
- catatonic: Catatonia
- Cesarean Section: Cesarean Section
- CHD: Congenital Heart Disease
- CHF: Congestive Heart Failure
- Chronic IHD: Chronic Ischemic Heart Disease
- Chronic Pain: Chronic Pain
- Chronic Renal dis.: Chronic Renal Disease
- Depression: Depression
- DM: Pregestational Diabetes
- drug abuse: Drug Abuse
- Eating dis.: Eating Disorders
- Eclampsia: Eclampsia
- Epilepsy: Epilepsy
- GDM: Gestational Diabetes
- Gestational HTN: Gestational Hypertension
- Headache: Headache
- HIV: HIV
- HTN: Pregestational Hypertension
- Hyperemesis Gravidarum: Hyperemesis Gravidarum
- Infertility: Female Infertility
- Migraine: Migraine
- Mild Preeclampsia: Mild or unspecified Preeclampsia
- Obesity: Overweight/Obesity
- Placenta Abruption: Placenta Abruption
- Placenta Previa: Placental Previa
- poisoning-analgesics: Other Unspecified Psychotropic Poisoning (analgesics)
- poisoning-anesthetics: Other Unspecified Psychotropic Poisoning (anesthetics)
- poisoning-antidepressants: Other Unspecified Psychotropic Poisoning (antidepressants)
- poisoning-antiepileptics: Other Unspecified Psychotropic Poisoning (antiepileptics)
- poisoning-antiparkinson: Other Unspecified Psychotropic Poisoning (anti-parkinson drug poisoning)
- poisoning-antipsychotics: Other Unspecified Psychotropic Poisoning (antipsychotics)
- Previous c-section: Previous Cesarean Delivery
- PTB: Preterm Birth
- Pul HTN: Pulmonary Hypertension
- Renal dis.: Renal disease
- SCD: Sickle Cell Disease
- Schizo.: Schizophrenia/Schizoaffective Disorder
- Severe Preeclampsia/Eclampsia: Severe Preeclampsia
- SLE: Systemic Lupus Erythematosus
- Sleep dis.: Sleep Disorder
- STD: Sexually Transmitted Diseases (STD)
- SUD-alcohol: Alcohol Use Disorder
- SUD-AMP: Psychostimulant Abuse/Dependence/Poisoning (amphetamine)
- SUD-barbiturate: Sedative Abuse/Dependence/Poisoning (barbiturate)
- SUD-BZD: Sedative Abuse/Dependence/Poisoning (benzodiazepine)
- SUD-cocaine: Cocaine Abuse/Dependcence/Poisoning
- SUD-ecstasy: Psychostimulant Abuse/Dependence/Poisoning (ecstasy)
- SUD-fentanyl: Opioid Use Disorder (Fentanyl)
- SUD-hallucinogen: Hallucinogen Abuse/Dependence/Poisoning (Hallucinogen, LSD, Unspecified Psychodysleptics)
- SUD-heroin: Opioid Use Disorder (Heroin)
- SUD-marijuana: Marijuana Use
- SUD-meth: Psychostimulant Abuse/Dependence/Poisoning (methamphetamine)
- SUD-methadone: Opioid Use Disorder (Methadone)
- SUD-MPH: Psychostimulant Abuse/Dependence/Poisoning (methylphenidate)
- SUD-opioids (all): Opioid Use Disorder (All)
- SUD-other psychotropic: Other Unspecified Psychotropic Poisoning (all)
- SUD-psychostimul. (all): Psychostimulant Abuse/Dependence/Poisoning (all)
- SUD-sedatives (all): Sedative Abuse/Dependence/Poisoning (all)
- SUD-smoking: Smoking
- VHD: Valvular Heart Disease



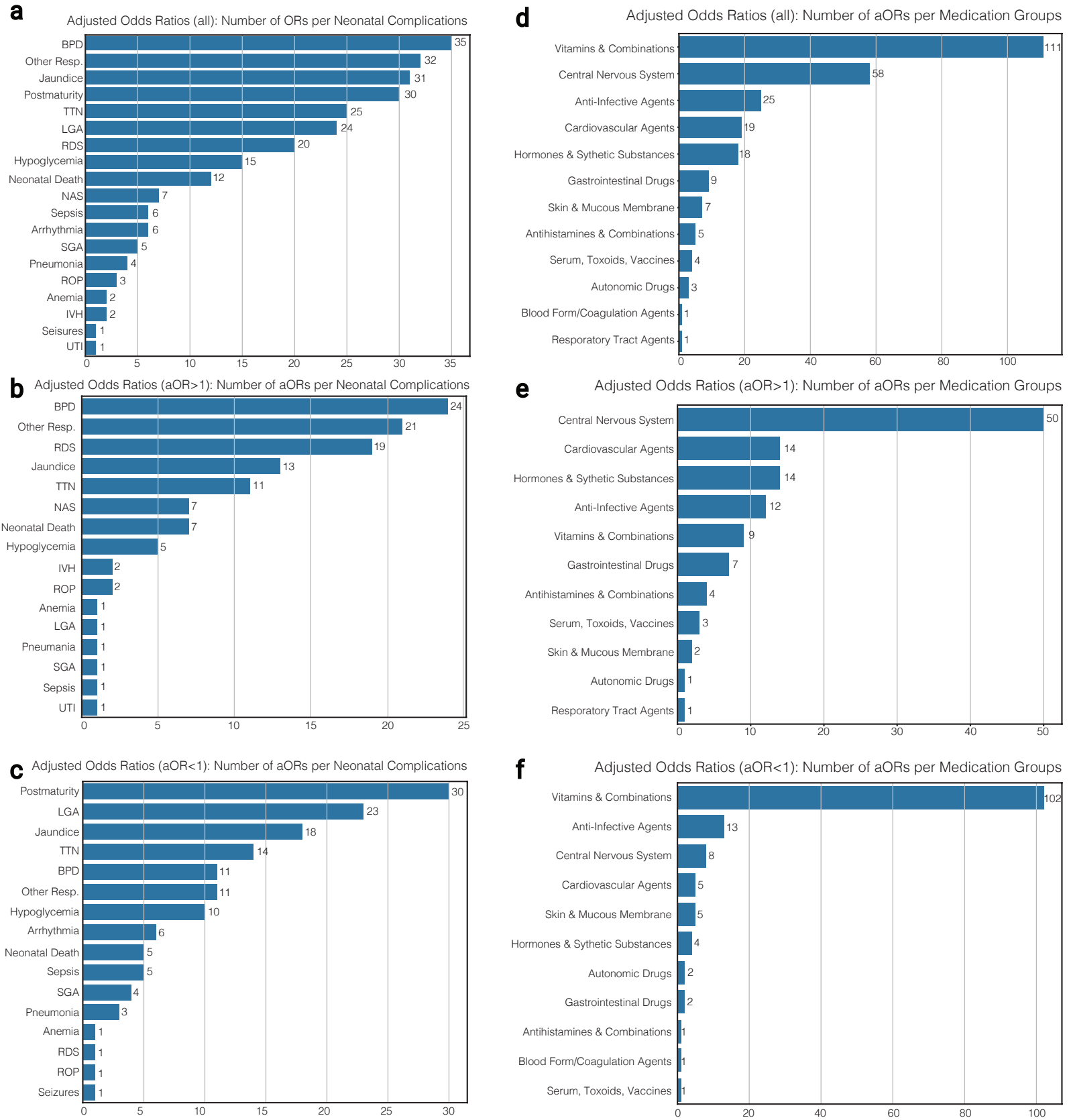
Supplementary Figure 4. Medication Name Harmonization. This figure outlines the process employed to clean up and harmonized maternal medication data. 1) Maternal medication with generic names was divided at the compound level. 2) Identical medications with varying names were consolidated and harmonized by employing DrugBank IDs along with their common names. 3) Lastly, information regarding the route of administration was integrated into these standardized names.



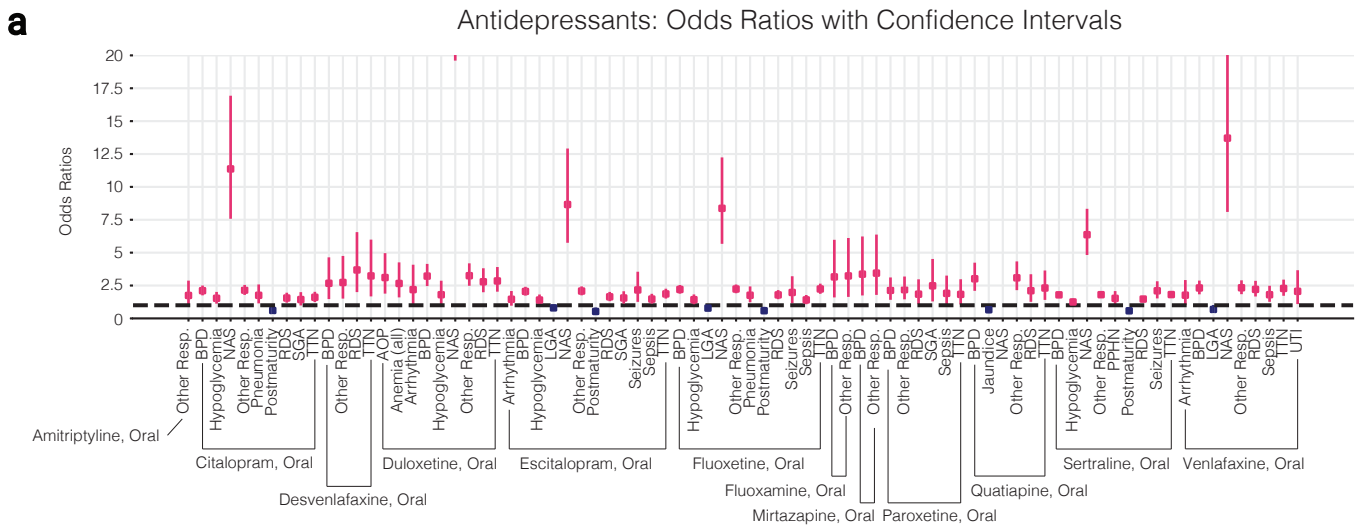
<u>Supplementary Figure 5. Network Graph of the Adjusted Odds Ratios Between Maternal Medications and Neonatal Complications.</u> Network graph depiction of adjusted odds ratios (aORs) between maternal medications and neonatal complications. aORs are calculated using LassoNoExp, a method that selects confounders by penalizing the excess of potential confounders for each medication-disease pair. Each node in the network graph represents either a maternal medication or a neonatal complication, consistent with the representation in Figure 2. Edges between nodes indicate statistically significant aORs, differentiated from unadjusted odds ratios presented Figure 2. PregMedNet identifies 261 statistically significant adjusted odds ratios (aORs) after adjustment, involving the associations of 68 maternal medications and 19 neonatal complications. Out of the total 261 edges, the 117 edges representing aORs > 1, indicating an increased risk association, are shown as gray edges, while the 144 edges representing aORs < 1, indicating a reduced risk association, are colored blue. The abbreviations for disease nodes are as defined in Figure 2.

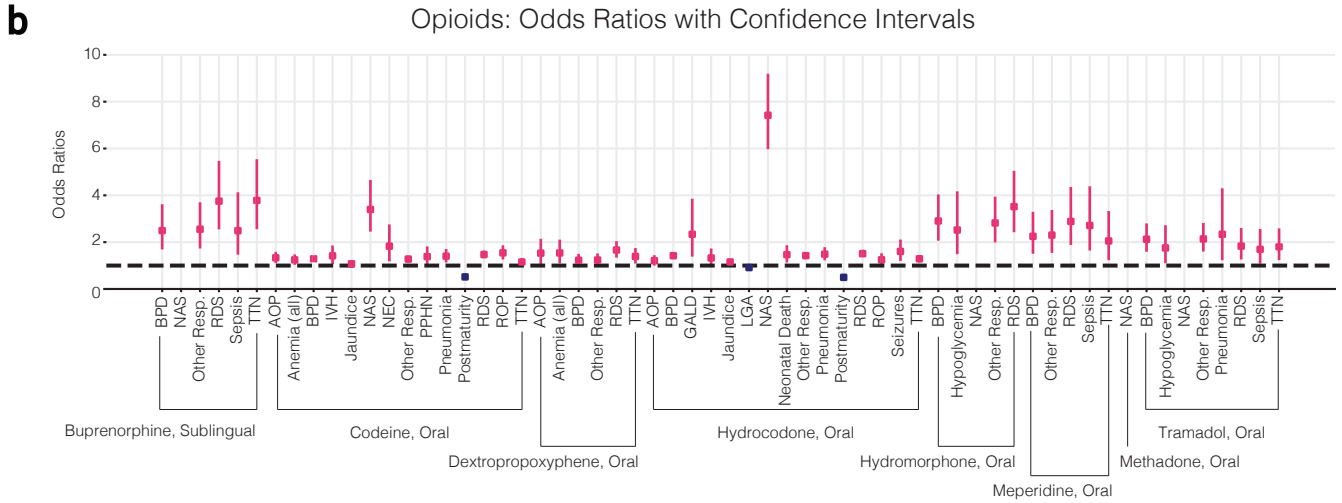


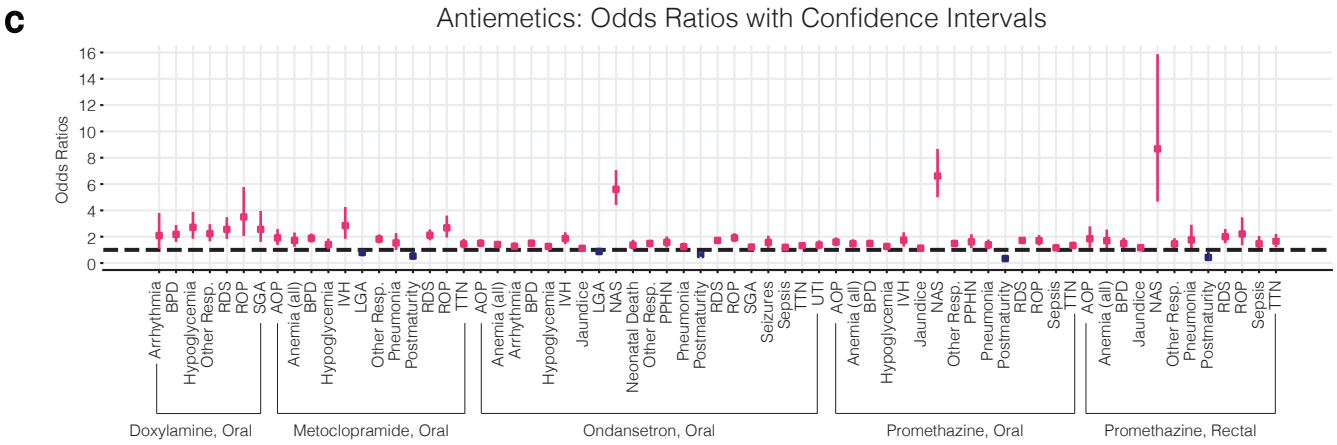
Supplementary Figure 6. Distribution of Unadjusted Odd Ratios for each Neonatal Complications and Maternal Medicaton Group. Bar graphs that show the count of statistically significant unadjusted odds ratios (OR). (a) The number of all statistically significant odds ratios associated with neonatal complications. BPD showed the most significant associations, followed by other respiratory disorders, RDS, TTN, and Postmaturity. (b) The number of statistically significant increased risk odds ratios associated with neonatal complications. RDS showed the most increased risk associations, followed by BPD, other respiratory disorders, TTN, and ROP. (c) The number of statistically significant reduced risk ORs associated with neonatal complications. Postmaturity has the most reduced risk edges followed by LGA and Jaundice. (d) The number of all significant ORs categorized by maternal medication groups. Medications in the Central Nervous System showed the most significant associations, followed by Hormones & Synthetic substances, Vitamins & Combinations, and Anti-Infective Agents. (e) The number of significant increased risk ORs categorized by maternal medication groups. Medications in the Central Nervous Systems show the most increased risk associations, followed by Hormones & Synthetic Substances, Anti-Infective Agents, and Cardiovascular Agents. (f) The number of significant reduced risk ORs categorized by maternal medication groups. The Vitamins & Combinations group shows the most reduced risk associations, followed by Anti-Infective Agents.

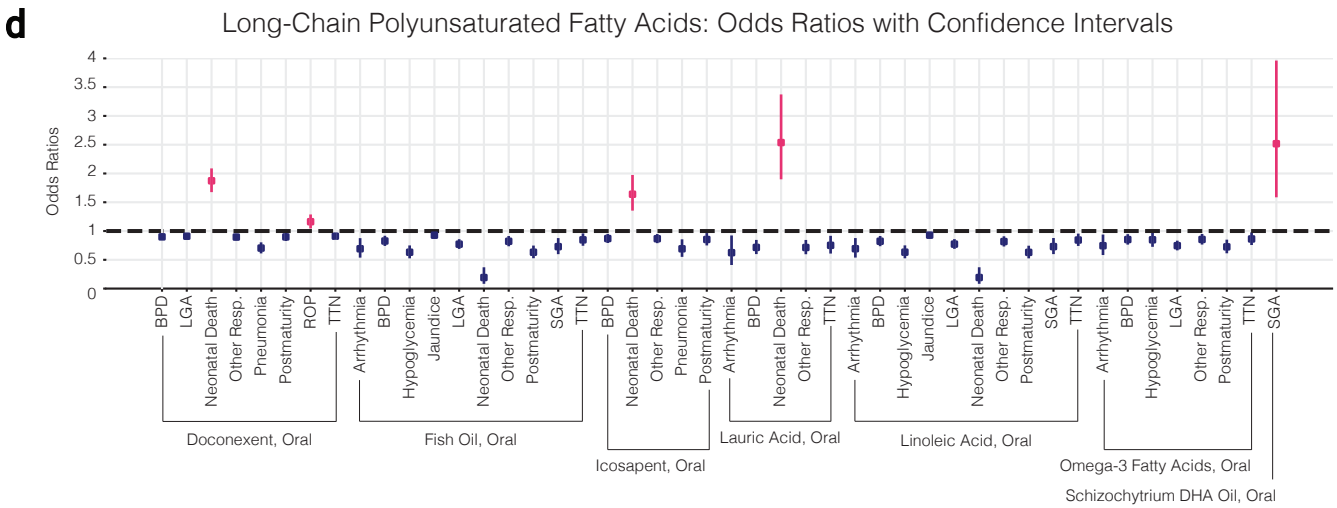


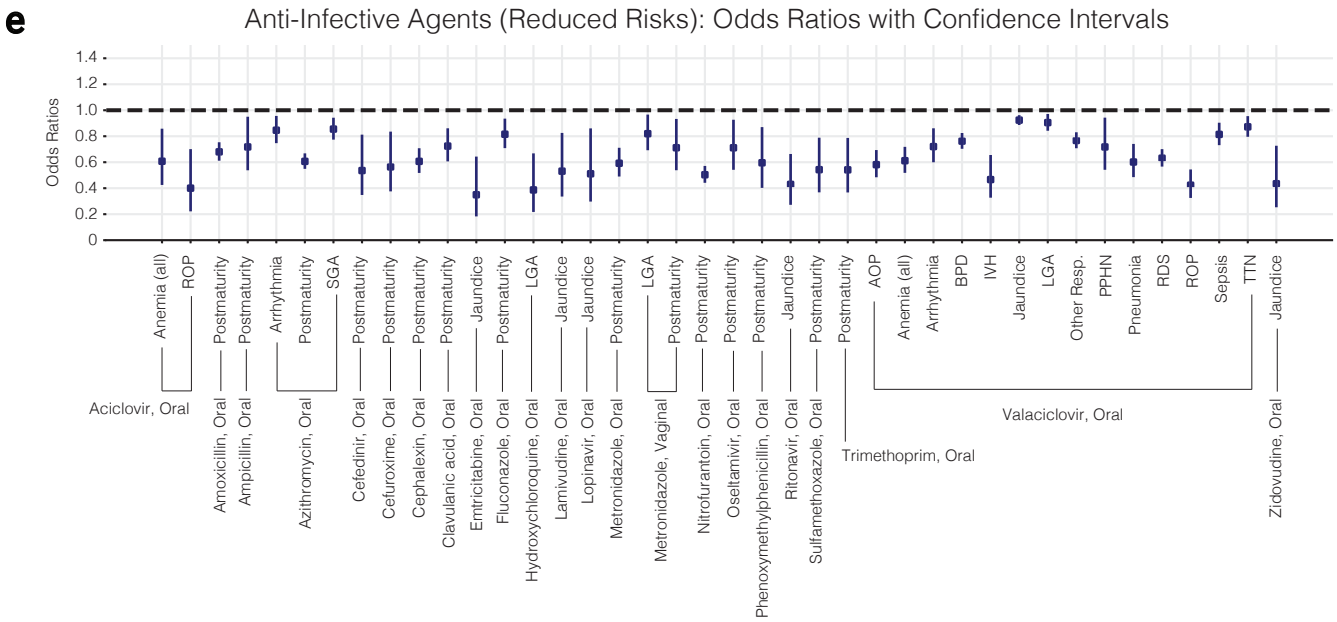
Supplementary Figure 7. Distribution of Adjusted Odd Ratios for each Neonatal Complications and Maternal Medicaton Group Bar graphs that show the count of statistically significant adjusted odds ratios (aOR). (a) The number of all statistically significant aORs associated with neonatal complications. BPD showed the most significant associations, followed by other respiratory disorders, Jaundice, Postmaturity, and TTN. (b) The number of statistically significant increased risk aORs associated with neonatal complications. BPD showed the most increased risk associations, followed by other respiratory disorders, RDS, Jaundice, and TTN. (c) The number of statistically significant reduced risk aORs associated with neonatal complications. Postmaturity has the most reduced risk edges followed by LGA and Jaundice. (d) The number of all significant aORs categorized by maternal medication groups. Medications in the Vitamins & Combinations showed the most significant associations, followed by Central Nervous System, and Anti-Infective Agents. (e) The number of significant increased risk aORs categorized by maternal medication groups. Medications in the Central Nervous Systems show the most increased risk associations, followed by Cardiovascular Agents, Hormones & Synthetic Substances, and Anti-Infective Agents. (f) The number of significant reduced risk aORs categorized by maternal medication groups. The Vitamins & Combinations group shows the most reduced risk associations, followed by Anti-Infective Agents.


a
Antidepressants: Odds Ratios with Confidence Intervals
Odds Ratios
Amitriptyline, Oral
Citalopram, Oral
Desvenlafaxine, Oral
Duloxetine, Oral
Escitalopram, Oral
Fluoxetine, Oral
Fluoxamine, Oral
Mirtazapine, Oral
Paroxetine, Oral
Quatiapine, Oral
Sertraline, Oral
Venlafaxine, Oral

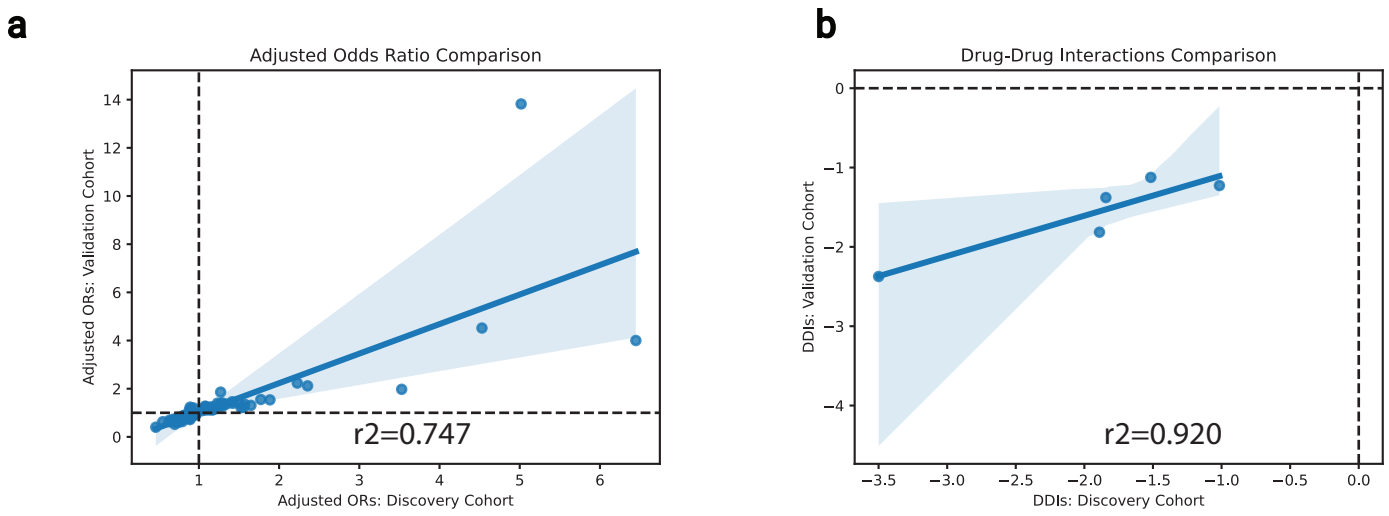

b
Opioids: Odds Ratios with Confidence Intervals
Odds Ratios
Buprenorphine, Sublingual
Codeine, Oral
Dextropropoxyphene, Oral
Hydrocodone, Oral
Hydromorphone, Oral
Meperidine, Oral
Methadone, Oral
Tramadol, Oral


c
Antiemetics: Odds Ratios with Confidence Intervals
Odds Ratios
Doxylamine, Oral
Metoclopramide, Oral
Ondansetron, Oral
Promethazine, Oral
Promethazine, Rectal

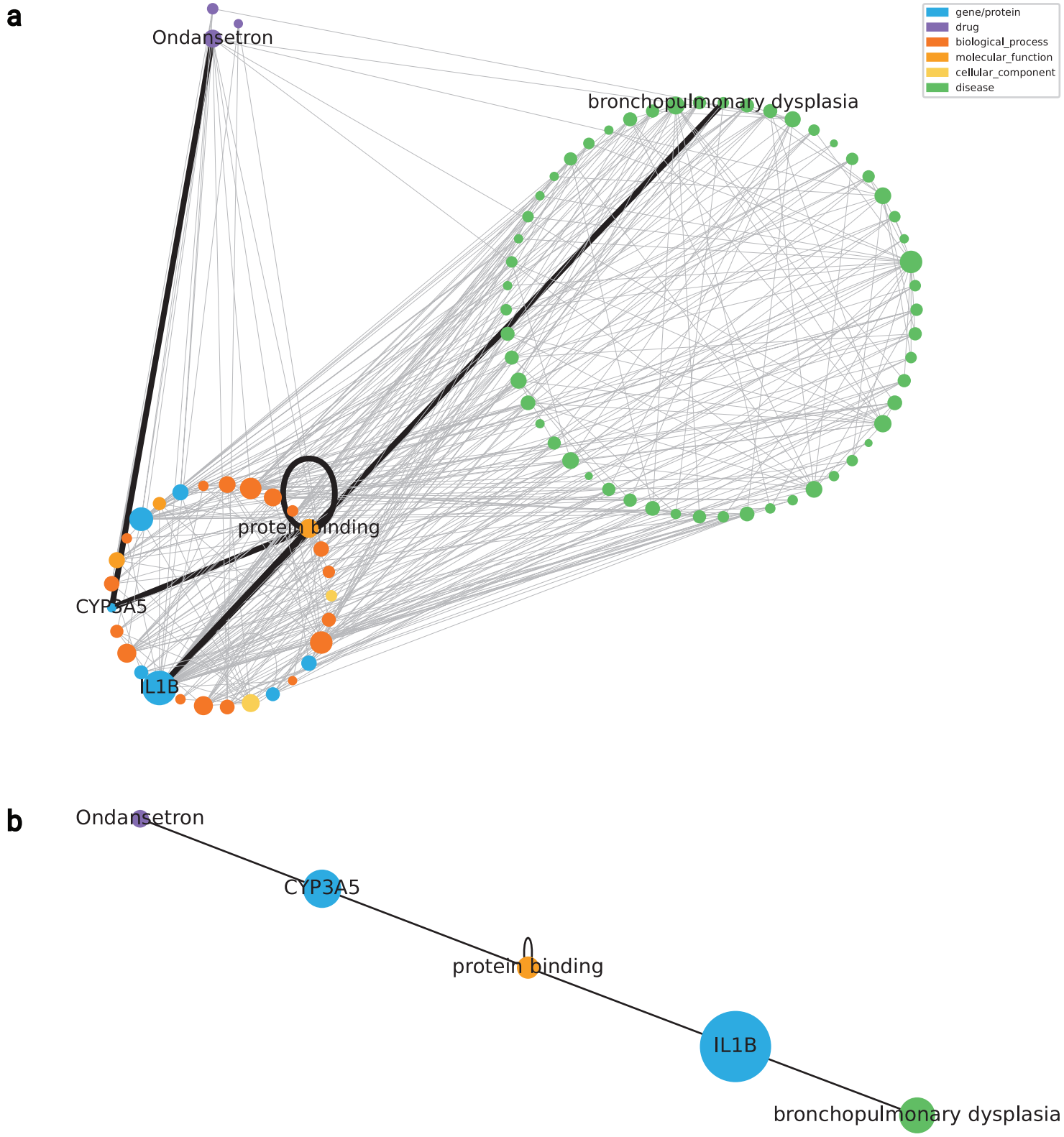




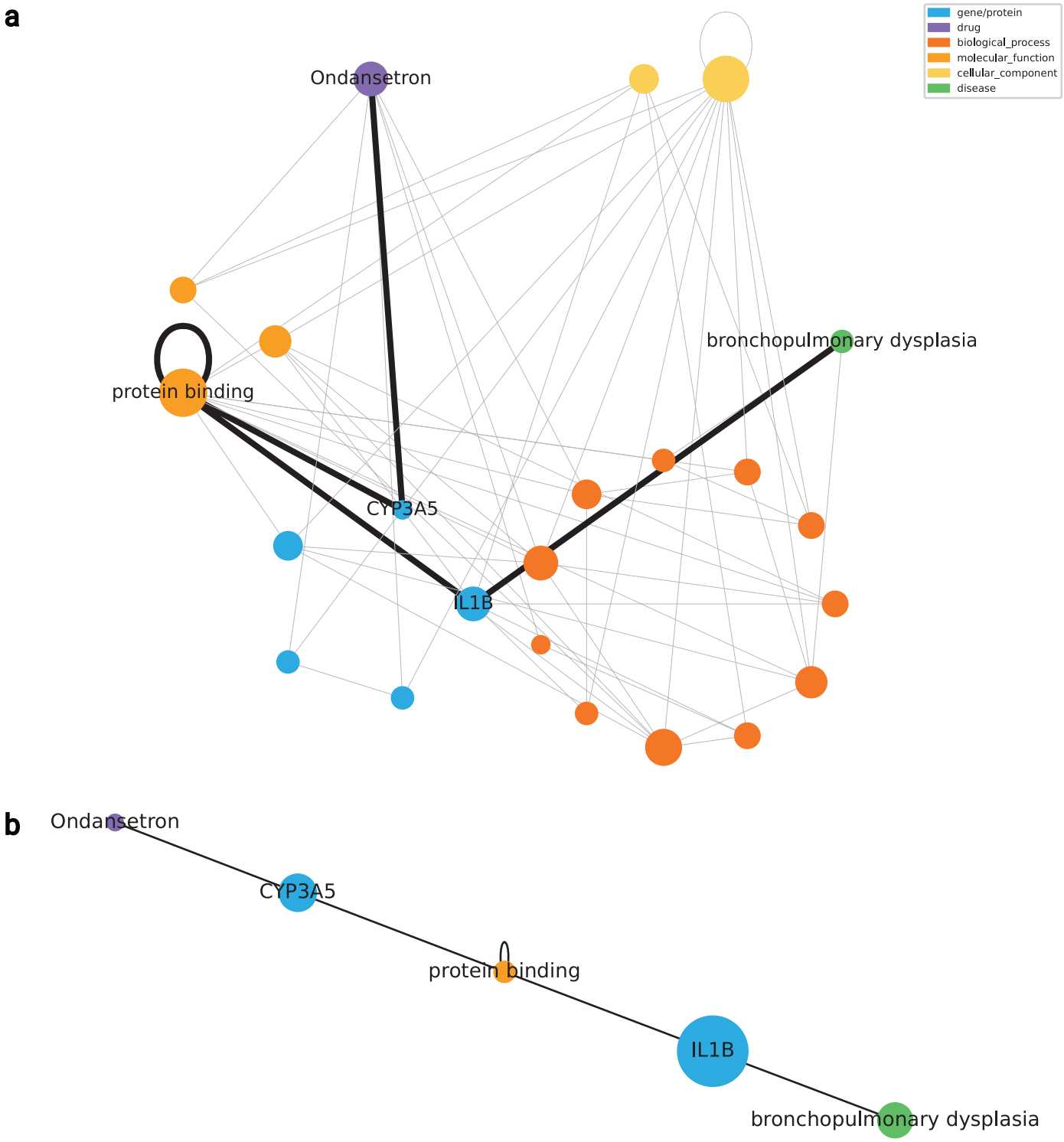
<u>Supplementary Figure 8. Unadjusted odds ratios for six different medication groups.</u> (a) Antidepressants. (b) Opioids. (c) Antiemetics. (d) Long-Chain PolyUnsaturated Fatty Acids in Vitamins & Combinations Group. (e) reduced risk odds ratios in the Anti-Infective Groups.



<u>Supplementary Figure 9. Results from the discovery cohort are correlated with the results from the validation cohort.</u> The results of the association between the two cohorts are matched based on maternal medications and neonatal complications. Scatterplots and Pearson correlation coefficients are generated for these matched results. (a) A scatterplot of significant adjusted odds ratios in the discovery cohort (x-axis) and the validation cohort (y-axis). In total, 81 adjusted odds ratios are matched between the two cohorts, out of 261 from the discovery cohort, and 167 from the validation cohort. The Pearson correlation analysis yields an R2 value of 0.747 with p-value of 1.246e-15. (b) A scatterplot of significant drug-drug interactions (beta3) in the discovery cohort (x-axis) and the validation cohort (y-axis). In total 5 adjusted DDIs are matched between the two cohorts out of 72 from the discovery cohort and 15 from the validation cohort. The Pearson correlation analysis yields an R2 value of 0.920 with p-value 2.682e-2.



Supplementary Figure 10. Potential Underlying Mechanisms of Ondansetron Effects on Bronchopulmonary Dysplasia with Inclusion of All Biological Interactions. (a) Hypothesized associations between maternal medications and neonatal complications are generated by combining PregMedNet results with a Knowledge Graph. This is an expanded version of Figure 5. While Figure 5 only includes proteins/genes nodes, this subgraph depiction includes all biological node types, including other disease and medication nodes. The figure depicts the shortest paths between bronchopulmonary dysplasia and ondansetron. The subgraph includes 191 four-hop paths with 81 nodes involved in the interactions between the medication and the disease. (b) The path involving IL-1B and CYP3A5 is also found in this subgraph, but IL-1B and CYP3A5 are connected by an unspecific node, 'protein binding.'



Supplementary Figure 11. Potential Underlying Mechanisms of Ondansetron Effects on Bronchopulmonary Dysplasia with Inclusion of All Biological Nodes, Focusing on the Drug and Disease of Interest. (a) Hypothesized associations between maternal medications and neonatal complications are generated by combining PregMedNet results with a Knowledge Graph. This figure is an expanded version of Figure 5 and Supplementary Figure 9. While Figure 9 includes all biological node types, including other diseases and medications, this subgraph focuses only on the disease and medication of interest, excluding others. The figure depicts the shortest paths between bronchopulmonary dysplasia and ondansetron. The subgraph includes 23 four-hop paths with 21 nodes involved in the interactions between the medication and the disease. (b) The same path involving IL-1B and CYP3A5 is also found in this subgraph.