

Welcome to the Sentinel Innovation and Methods Seminar Series

The webinar will begin momentarily

- Please visit www.sentinelinitiative.org for recordings of past sessions and details on upcoming webinars.
- Note: closed-captioning for today's webinar will be available on the recording posted at the link above.



Advanced Approaches for Evaluating Drug Safety in Pregnancy

Krista F. Huybrechts, MS PhD

Associate Professor of Medicine, Harvard Medical School
Associate Professor of Epidemiology, Harvard T.H. Chan School of Public Health
Co-Director Harvard Program in Perinatal and Pediatric Pharmacoepidemiology

Sentinel Innovation and Methods Seminar
April 26, 2021



“The last therapeutic orphan”*

- Pregnant women are de facto excluded from most clinical trials to protect the fetus from research-related risks.
- A drug’s structure and function does not predict its teratogenicity.
- Animal studies are seriously limited in their ability to predict human teratogenesis.
- When a new drug enters the market, there is little to no information about its safety during pregnancy.
- Urgent need to develop evidence in a timely manner so that serious problems can be quickly detected, or concerns alleviated.

* Wisner KL. *Am J Psychiatry* 2012;169(6):554-6.

Approaches to Drug Safety Surveillance in Pregnancy



Pregnancy Registry	Healthcare Utilization Database
Prospective data if enrolled before outcome	Prospective data recording
Ad hoc collection takes time and \$\$	Data exist (economy of cost and time)
Selected group of volunteers, limited follow up	Real world experience, dynamic population
Information on one, or few, drugs	Information on multiple drugs
Real use, outpatient and inpatient, Rx and OTC	Usually outpatient filling of prescription
Information on outcomes of interest	Information on multiple outcomes if reimbursed
Incomplete ascertainment of pregnancy losses	Incomplete ascertainment of pregnancy losses
Validation usually part of the design	May have access to validation
Key clinical factors collected in detail	Broad range of clinical factors, with less granularity
Can collect information on socio-demographics	Little information on socio-demographics
May have laboratory data if collected	May have laboratory data in subsample
Can collect key factors (e.g., gestational age, family history)	Key characteristics may be missing, e.g., LMP, no claims for it, use algorithm
Some use external reference, few allow CER	Internal control groups allow CER
Small populations	Huge source population
Can target new drugs (need to recruit users)	No information on new drugs

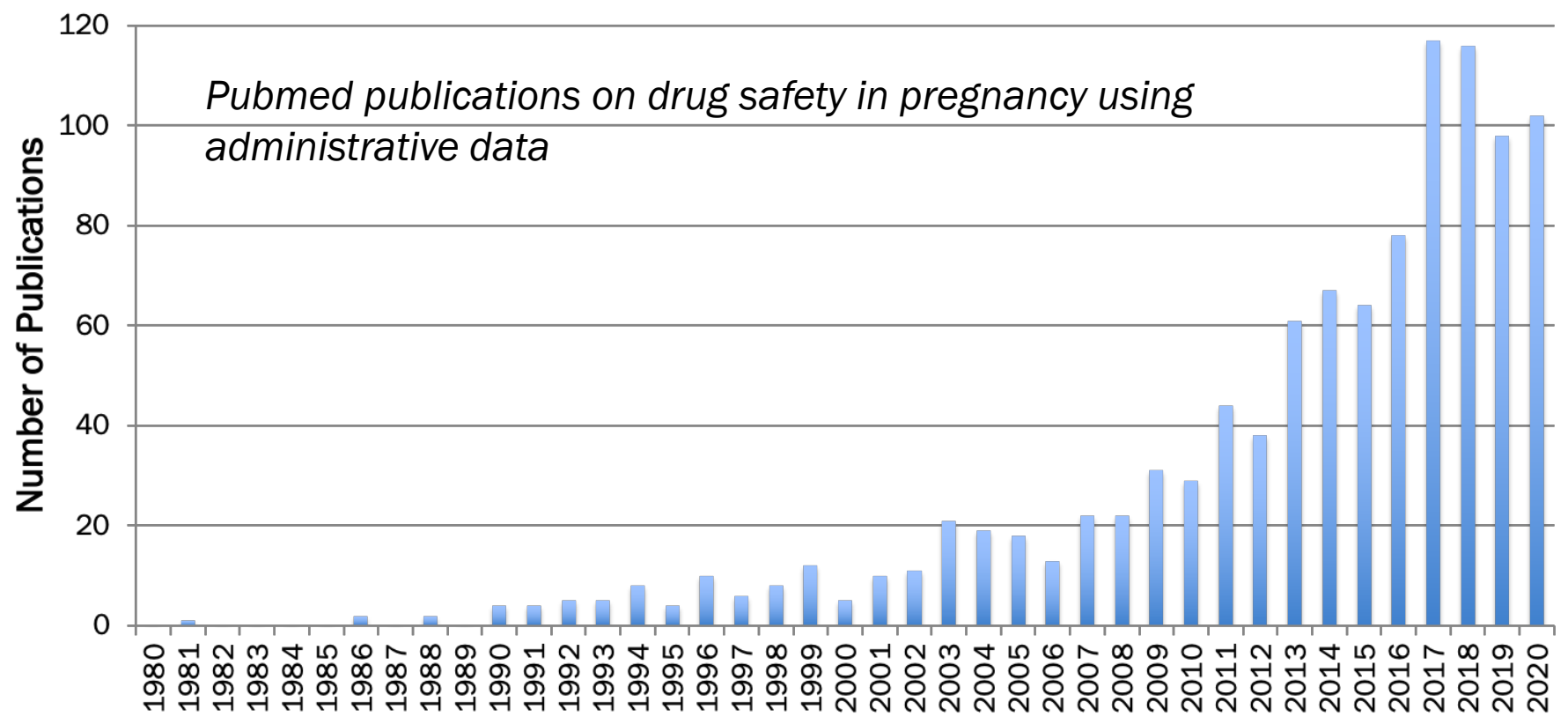


Approaches to Drug Safety Surveillance in Pregnancy



Pregnancy Registry | Healthcare Utilization Database

- Prospect
- Ad hoc
- Selecte
- Informa
- Real us
- Informa
- Incomp
- Validati
- Key clin
- Can col
- May ha
- Can col
- Some u
- Small p
- Can tar



- ity
- claims for

Huybrechts et al. *Pharmacoepidemiol Drug Saf.* 2019 ;28(7):906-922.



Approaches to Drug Safety Surveillance in Pregnancy



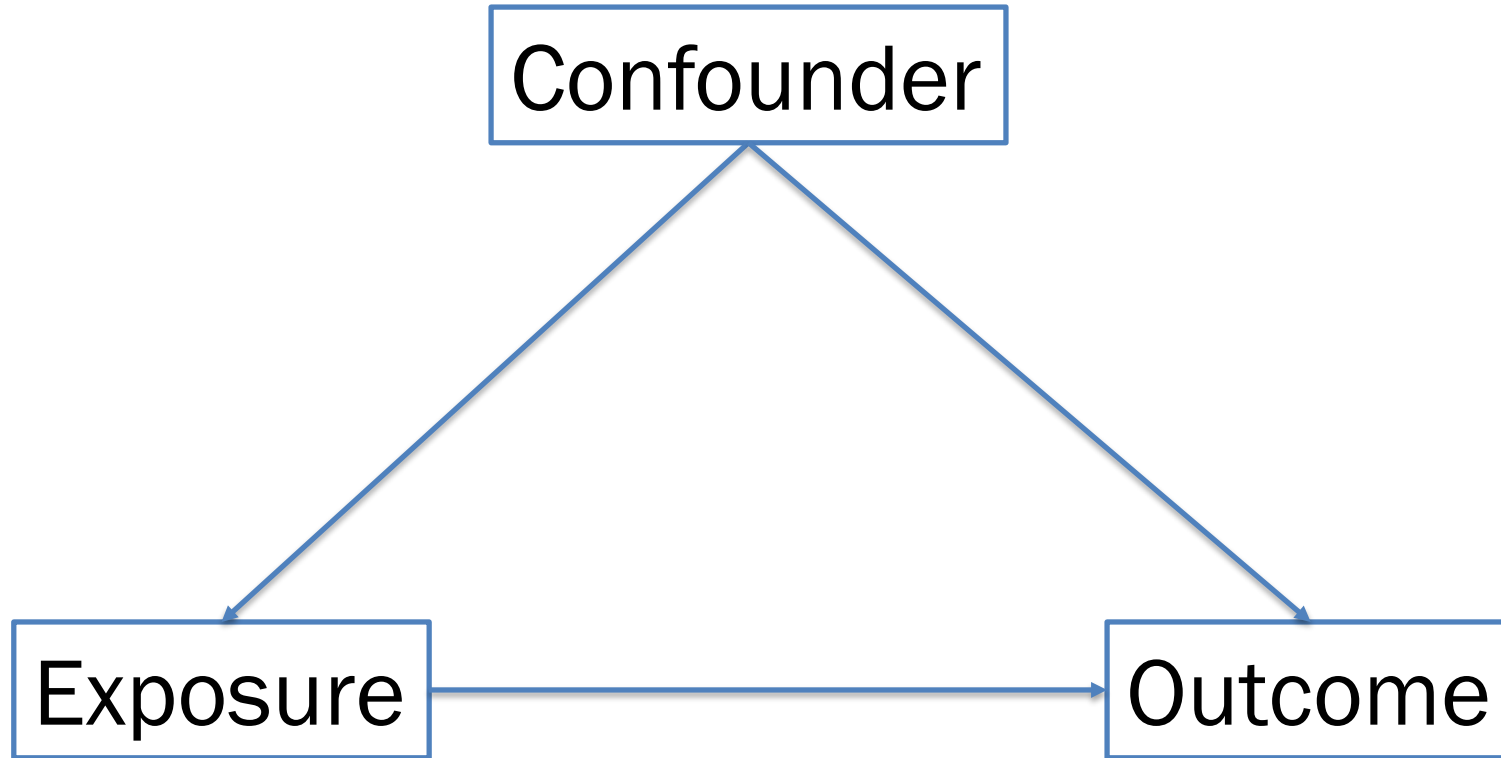
Pregnancy Registry	Healthcare Utilization Database
Prospective data if enrolled before outcome	Prospective data recording
Ad hoc collection takes time and \$\$	Data exist (economy of cost and time)
Selected group of volunteers, limited follow up	Real world experience, dynamic population
Information on one, or few, drugs	Information on multiple drugs
Real use, outpatient and inpatient, Rx and OTC	Usually outpatient filling of prescription
Information on outcomes of interest	Information on multiple outcomes if reimbursed
Incomplete ascertainment of pregnancy losses	Incomplete ascertainment of pregnancy losses
Validation usually part of the design	May have access to validation
Key clinical factors collected in detail	Broad range of clinical factors, with less granularity
Can collect information on socio-demographics	Little information on socio-demographics
May have laboratory data if collected	May have laboratory data in subsample
Can collect key factors (e.g., gestational age, family history)	Key characteristics may be missing, e.g., LMP, no claims for it, use algorithm
Some use external reference, few allow CER	Internal control groups allow CER
Small populations	Huge source population
Can target new drugs (need to recruit users)	No information on new drugs

Approaches to Drug Safety Surveillance in Pregnancy

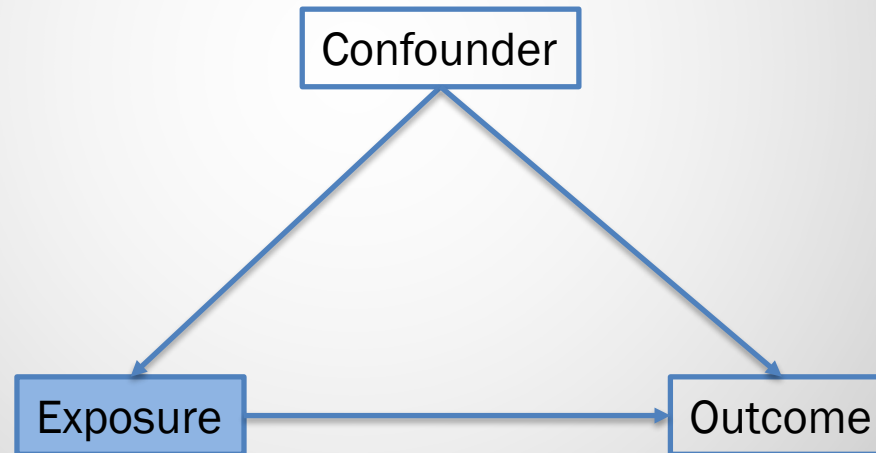


Pregnancy Exposure Registry	Healthcare Utilization Database
Prospective data if enrolled before outcome	Prospective data recording
Ad hoc collection takes time and \$\$	Data exist (economy of cost and time)
<p>Complementary Approaches</p> <p>~</p> <p>Today's Focus: Unique opportunities to advance the field of perinatal pharmacoepidemiology supported by healthcare utilization databases</p>	
Can collect key factors (e.g., gestational age, family history)	Key characteristics may be missing, e.g., LMP, no claims for it, use algorithm
Some use external reference, few allow CER	Internal control groups allow CER
Small populations	Huge source population
Can target new drugs (need to recruit users)	No information on new drugs

Outline



Exposure Classification



Etiologically relevant time window

- Many studies ignore the precise gestational timing of exposure
 - Use at any time
 - Use during a broad window
 - Reasons:
 - Uncertainty about the biological mechanism
 - Uncertainty about timing
 - Lack of power
 - Ascertaining exposure during the wrong window → exposure misclassification → bias towards a null finding
- ⇒ Pregnancy Etiologically Relevant Interval scoping (PERIscooping):
A method to detect risk associated with exposure *at specific time points* in pregnancy, without a priori specification of the etiologically relevant window

PERIscooping

- Compare **observed** number of outcomes for women exposed in a give risk window to **expected** counts under the null
 - Expected counts: Reassign observed outcomes to observed prescription histories through random permutation
 - Exposure risk windows: e.g., each separate day in pregnancy, consecutive windows or overlapping windows
- Inference based on **Monte Carlo hypothesis testing** that adjusts for the multiple testing
 - Generate *window-specific test statistic T* for observed data and 9,999 random replicates; rank according to T
 - p-value: rank of the observed data / 10,000
 - *Overall test statistic T* is the minimum p-value across all potential risk windows

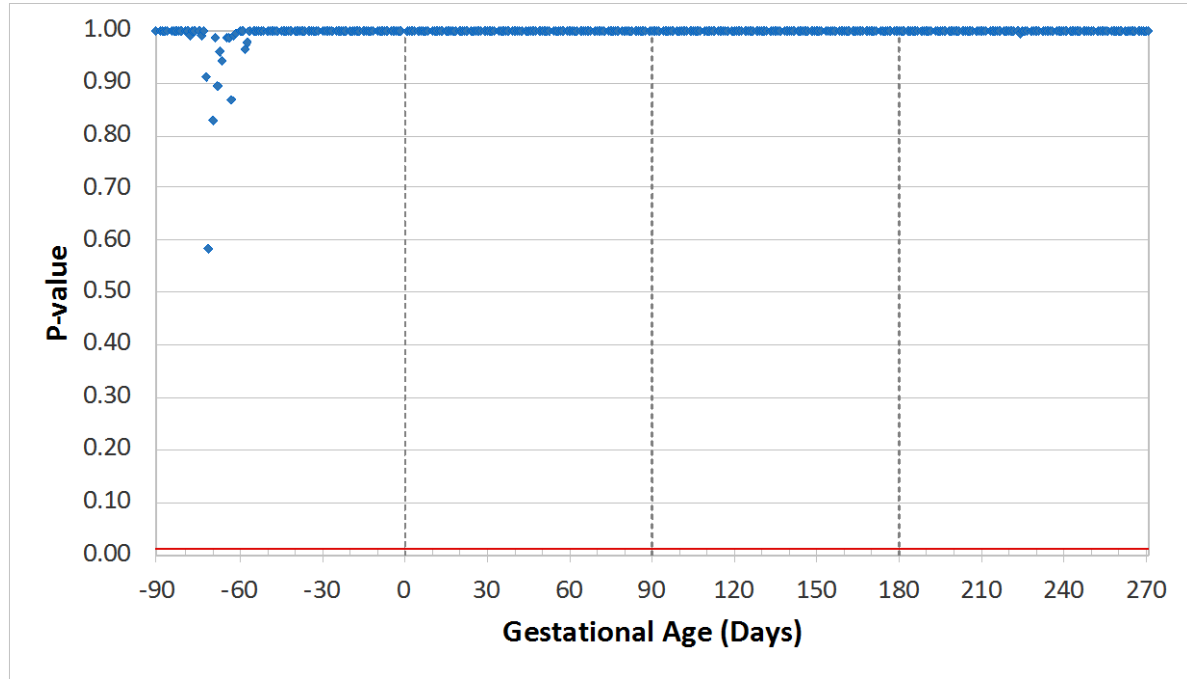
PERIscooping

STRENGTHS	LIMITATIONS
Approach 1: Exposed in a given window vs exposed in a different window Exposure duration stratified	
<ul style="list-style-type: none">• Confounding by indication removed through restriction to women exposed at some time in pregnancy• Confounding by disease severity addressed through stratification by duration of exposure	<ul style="list-style-type: none">• Less power
Approach 2: Exposed vs unexposed in a given window Risk window specific propensity score weighting	
<ul style="list-style-type: none">• Greater power	<ul style="list-style-type: none">• Greater potential for unmeasured confounding by disease indication and severity

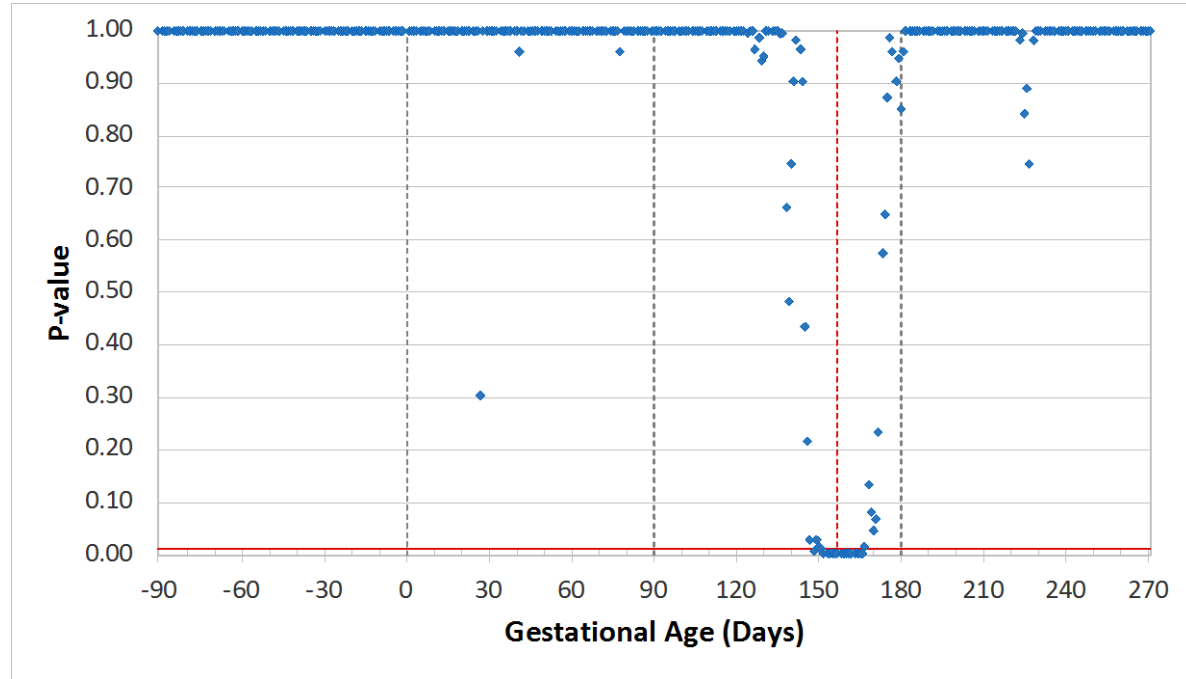
Negative and Positive Control Test Case



Simulated negative control dataset

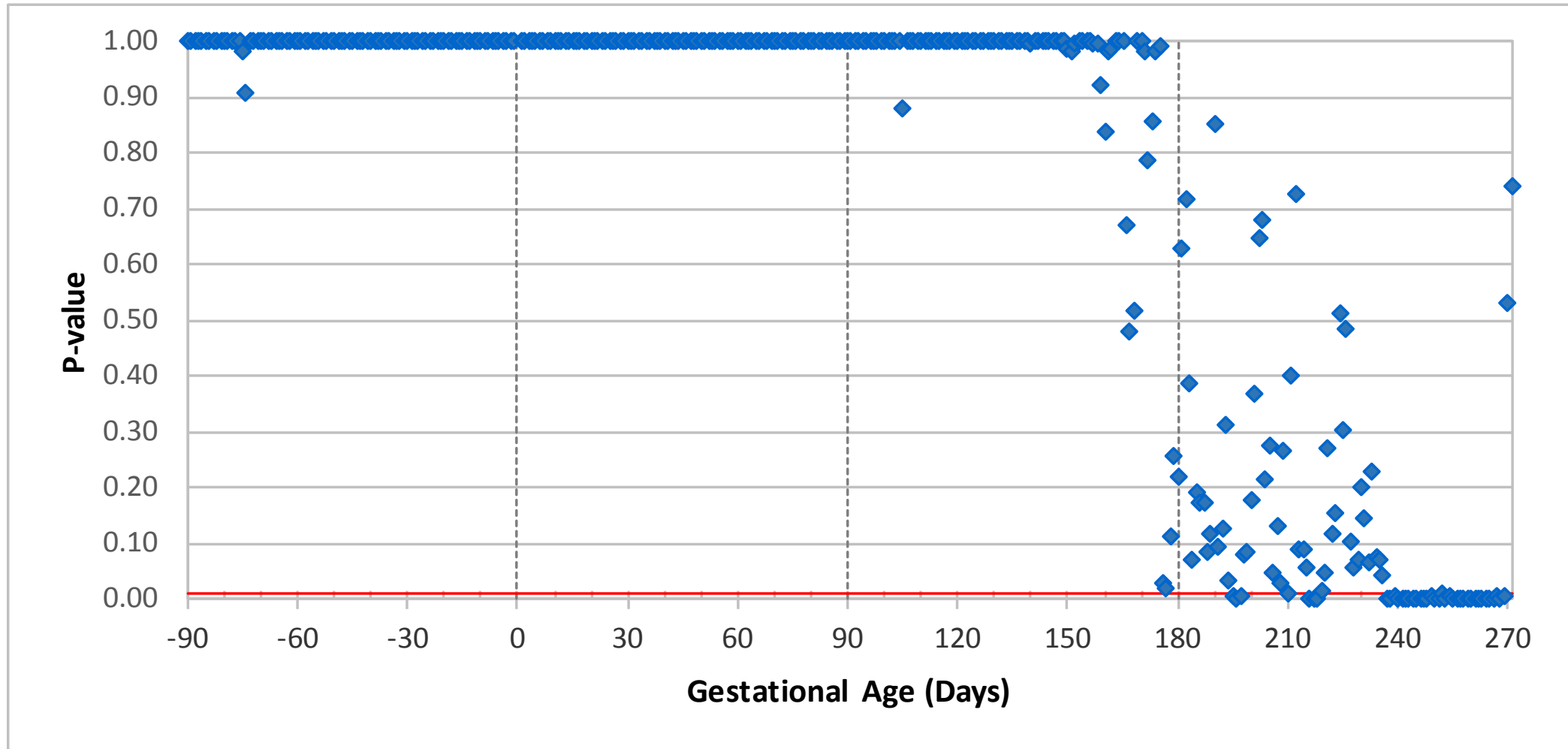


Simulated positive control dataset



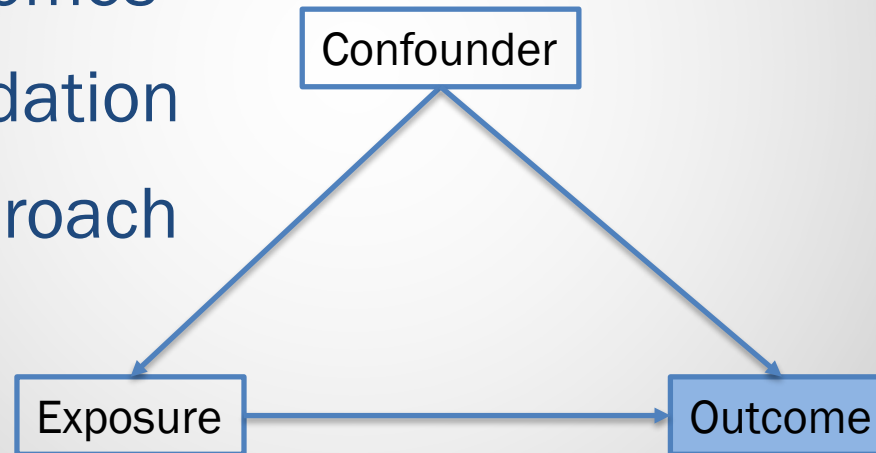
Data permuted to create a true increase in risk by assigning a higher proportion of outcomes on gestational day 157

Opioid Use and the Risk of Neonatal Abstinence Syndrome

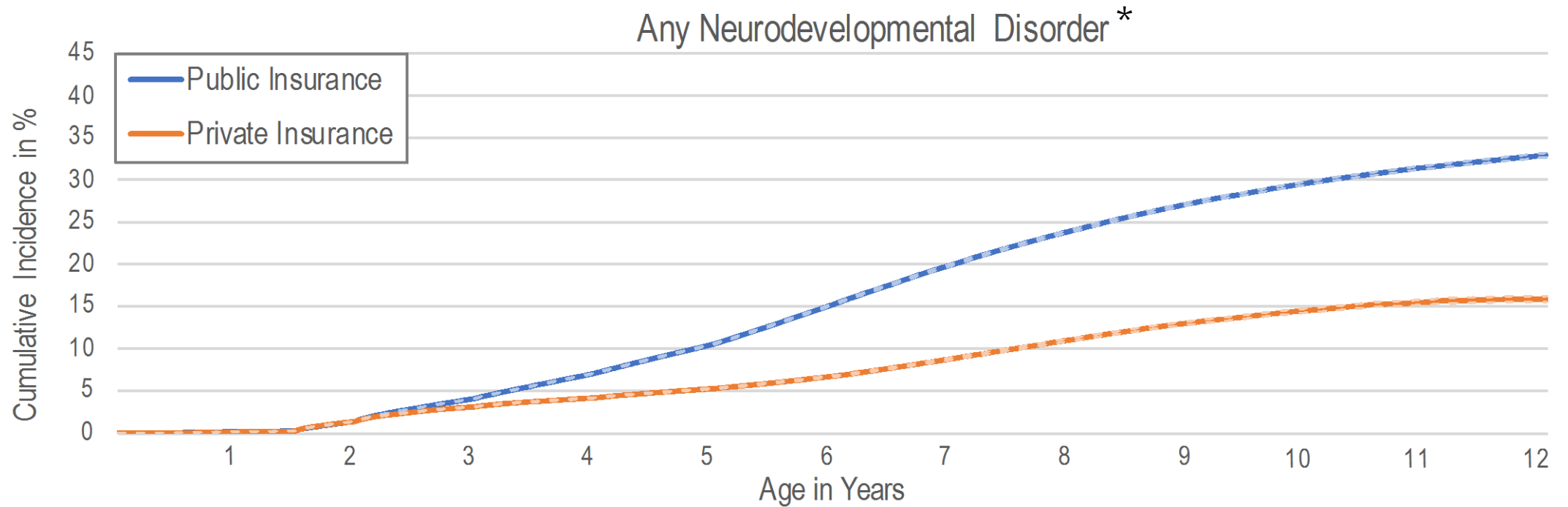


Outcome Ascertainment

- Types of outcomes
- Outcome validation
- Scanning approach



Long-term outcomes

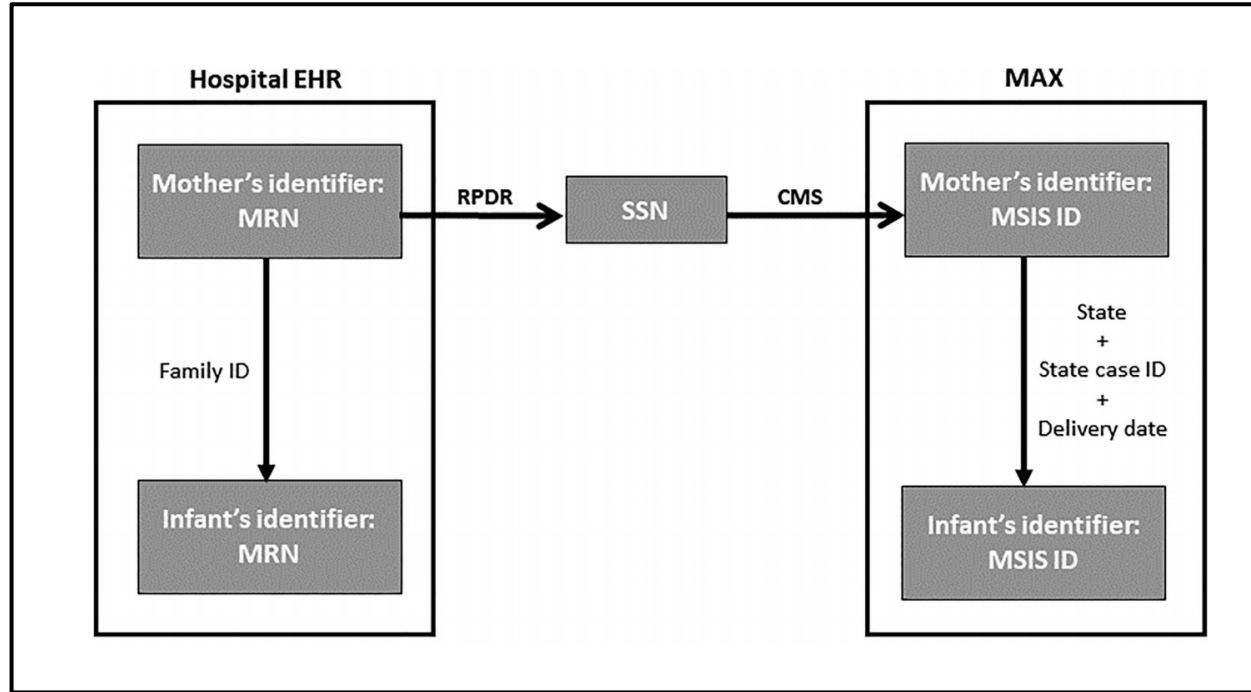


Number at Risk

Public Insurance	1,665,752	1,106,543	784,612	559,445	402,350	280,201	189,471	127,480	81,157	46,682	25,710	9,925
Private Insurance	913,523	589,023	390,323	260,931	176,368	117,453	78,158	50,897	30,025	14,705	5,233	109

* Includes autism spectrum disorder, ADHD, learning disability, developmental speech/language disorder, developmental coordination disorder, intellectual disability, behavioral disorder

Outcome Validation Studies

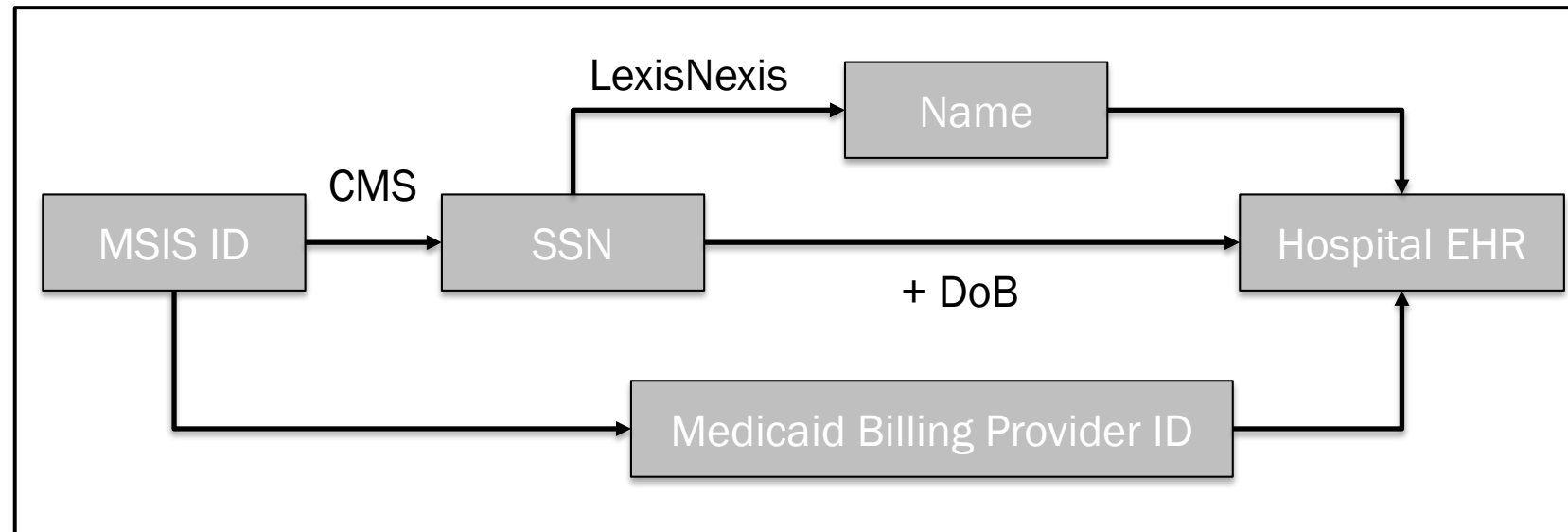


- Validation within [local healthcare system](#)
 - Maternal & infant outcomes
 - Non-live birth outcomes

He M et al. Pharmacoepidemiol Drug Saf 2020;29: 419-426

- Validation at [national level](#)
 - Maternal & infant outcomes
 - Gestational Age

Palmsten et al. Pharmacoepidemiol Drug Saf 2014;23: 646-655



Validity of Claims-Based Algorithms



Outcome	N Records Reviewed	N True Positives	PPV (95% CI)
Autism	50	47	0.94 (0.83 - 0.99)
ADHD	50	44	0.88 (0.76 - 0.95)
Learning Difficulty	50	49	0.98 (0.89 - 1.00)
Developmental Speech or Language Disorder	50	49	0.98 (0.89 - 1.00)
Intellectual disability	50	41	0.82 (0.69 - 0.91)
Developmental Coordination Disorder (DCD)	50	19 45	0.38 (0.25 - 0.53) 0.90 (0.82 - 0.98)
Behavioral Disorder	50	46	0.92 (0.81 - 0.98)

Outcomes: Comprehensive Safety Surveillance

- Most research focuses on a single or selected outcomes
 - By design
 - As a result of selective publication of associations in the context of multiple comparisons

Benefits of
treatment



All possible adverse
effects for mother and
newborn infant

Need for a safety surveillance approach that allows for the simultaneous evaluation of a comprehensive range of adverse maternal, fetal and neonatal outcomes.

TreeScan™ Approach in a Nutshell

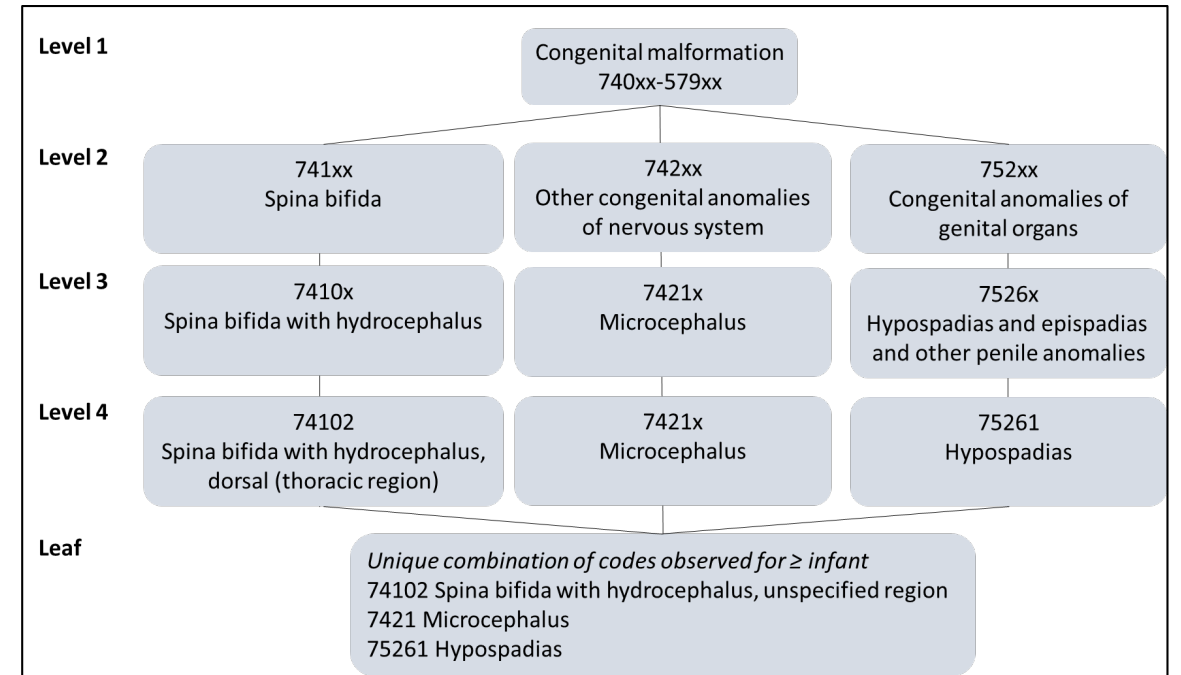


- **Scan** a hierarchical tree of (groups of clinically related) outcomes for associations with the exposure of interest
- Account for the **multiple testing** of correlated hypotheses
- Highlight potential problems that warrant further, thorough investigation.
 - Adverse event “signal” \neq causal relationship

The “tree” in TreeScan



- Classification system that hierarchically groups coded clinical concepts into clinically related categories
 - ICD, Multi-Level Clinical Classification (MLCC) for ICD codes, Medical Dictionary for Regulatory Activities (MedDRA) classification system
- Each grouping represents an outcome “node” in the hierarchical tree
- Maximizes power to detect clinically related outcomes



The “scan” in TreeScan



- Test statistic
- Different probability models for different data types
 - E.g., conditional and unconditional versions for Bernoulli/binomial and Poisson generated data
- Test hypothesis:
 - H_0 : no difference in risk of adverse events in any outcome node in the tree
 - H_1 : there is **at least one node** in the tree where the risk of adverse events is higher in the exposed group than in the comparator group (one-sided)
- Multiplicity-adjusted p-values that accurately reflect the type I error rate

Statistical Alert \neq Safety Signal



- Statistical alerts help prioritize associations unlikely to have occurred by chance
- Residual confounding can produce spurious alerts
- Potential signals of concern should be followed by a tailored cohort study:
 - Step 1: [Using the original data source](#) to assess whether the observed association remains with tailored design and confounding adjustment
 - Step 2: For associations that persist, further evaluate robustness of the finding by implementing the study in [independent data](#)

Test Case: Prescription opioids and Neonatal Abstinence Syndrome

Opioids

Exposure	Late pregnancy exposure, relatively common
Expected outcome	Neonatal abstinence syndrome, relatively common
Confounding adjustment method	Propensity score matching
Scan statistic	Unconditional Bernoulli
Hierarchical outcome tree	Pruned Multi-level Clinical Classification Software, no birth defects
Washout to identify incident outcomes	0 days
Outcome counts	Any unique position with

Pruned the tree; removed:

- Congenital malformations
- Codes unlikely to be an adverse reaction caused by drugs (e.g., well care visit, live birth)
- Codes that did not represent incident events (e.g., family history of alcoholism)
- Conditions with long latency/induction periods (e.g., cancer).

Huybrechts et al. Am J Epidemiol. 2021 Jan 11:kwaa288.

Test Case: Prescription opioids and Neonatal Abstinence Syndrome

Opioids

Exposure	Late pregnancy exposure, relatively common
Expected outcome	Neonatal abstinence syndrome, relatively common
Confounding adjustment method	Propensity score matching
Scan statistic	Unconditional Bernoulli
Hierarchical outcome tree	Pruned Multi-level Clinical Classification Software, no birth defects
Washout to identify incident outcomes	0 days
Outcome counts	Any unique occurrence of a code in any care setting or diagnosis position within 90 days on or following delivery

Tested 9,044 hierarchical outcome nodes at every level of the tree above the leaf level

Results: Opioids



- Source cohort: N = 53,771 exposed; N = 1,360,039 unexposed
- After 1:5 PS matching: N = 24,080 exposed; N = 120,400 unexposed
- The only tree branch on which there were statistical alerts at $p < 0.05$ were related to the expected safety concerns of drug withdrawal in the newborn
- No false positive alerts at the statistical alerting threshold of 0.05.

Node Identifier	Node Description	P-value	Risk (R)		RR	RD	Outcomes				O/E	O - E
			R _{Exp}	R _{Ref}			Observed (O), Expected (E)		O	O _{Exp}		
05	Mental Illness	0.001	72.0	32.6	2.2	39.4	2,857	875	1,982	571	1.5	303.6
05.12	Substance-Related Disorders	0.001	52.8	12.3	4.3	40.4	1,390	641	749	278	2.3	363
05.12.00	Substance-Related Disorders	0.001	52.8	12.3	4.3	40.4	1,390	641	749	278	2.3	363
05.12.00.00	Substance-Related Disorders	0.001	52.8	12.3	4.3	40.4	1,390	641	749	278	2.3	363
7795	Drug Withdrawal Syndrome In Newborn	0.001	35.2	5.8	6.1	29.5	778	428	350	156	2.8	272.4
76072	Narcotics Affecting Fetus Or Newborn Via Placenta Or Breast Milk	0.001	19.6	3.6	5.4	16	458	238	220	92	2.6	146.4
2920	Drug Withdrawal	0.001	4.8	1.1	4.5	3.7	123	58	65	25	2.4	33.4
30400	Opioid Type Dependence Unspecified Use	0.001	1.0	0.0	20.0	0.9	15	12	<11	3	4.0	9.0
06	<i>Diseases of the nervous system and sense organs</i>											
06.08	<i>Ear conditions</i>											
06.08.03	<i>Other ear and sense organ disorders</i>											
06.08.03.00	<i>Other ear and sense organ disorders</i>											
3899	<i>Unspecified Hearing Loss</i>	0.07	20.8	12.3	1.7	8.5	1,002	253	749	200	1.3	52.6
15	<i>Conditions originating in the perinatal period</i>											
15.07	<i>Other perinatal conditions</i>											
15.07.04	<i>Other and unspecified perinatal conditions</i>											
15.07.04.00	<i>Other and unspecified perinatal conditions</i>											
76079	<i>Other Noxious Influences Affecting Fetus Or Newborn Via Placenta Or Breast</i>	0.15	15.5	8.9	1.7	6.6	726	188	538	145	1.3	42.8

Considerations



False Positives

- Cannot dismiss potential adverse effects identified simply because a known biological explanation has not been established:
 - Pathophysiology of many adverse pregnancy outcomes is not fully understood
 - Biologic mechanisms for many accepted human teratogens remain unknown
- Approach controls the overall error rate:
 - Current practice of no adjustment for multiple testing, results in a much higher type I error rate than the experiment-wide alpha level
- P-values are used as a means to rank and prioritize alerts for further investigation, not to decide whether there is a causal association

False Negatives

- Multiplicity adjustment less conservative than for other methods (e.g., Bonferroni)
- Optimize tree:
 - Targeted towards pregnancy outcomes
 - Importance of “pruning” tree
- Do not strictly focus on statistical significance threshold
 - Outcomes that do not alert may still have low likelihood under the null
 - Evaluate pattern of outcomes unlikely to be observed if there was no relationship with exposure

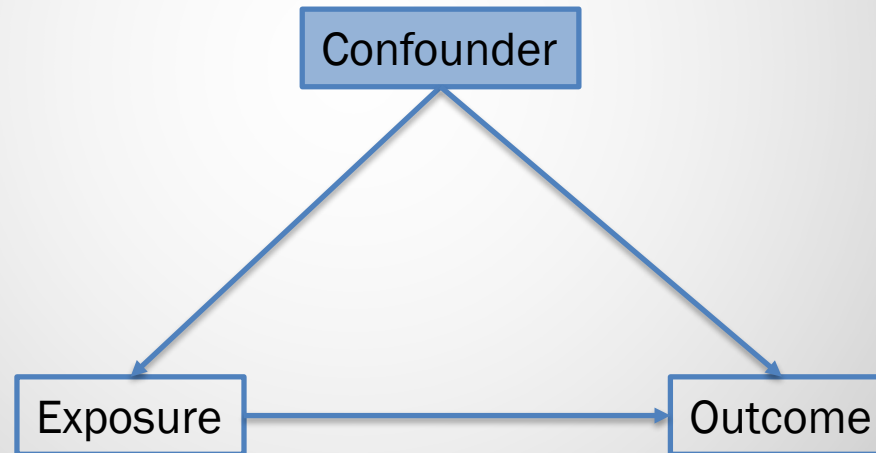


Conclusion



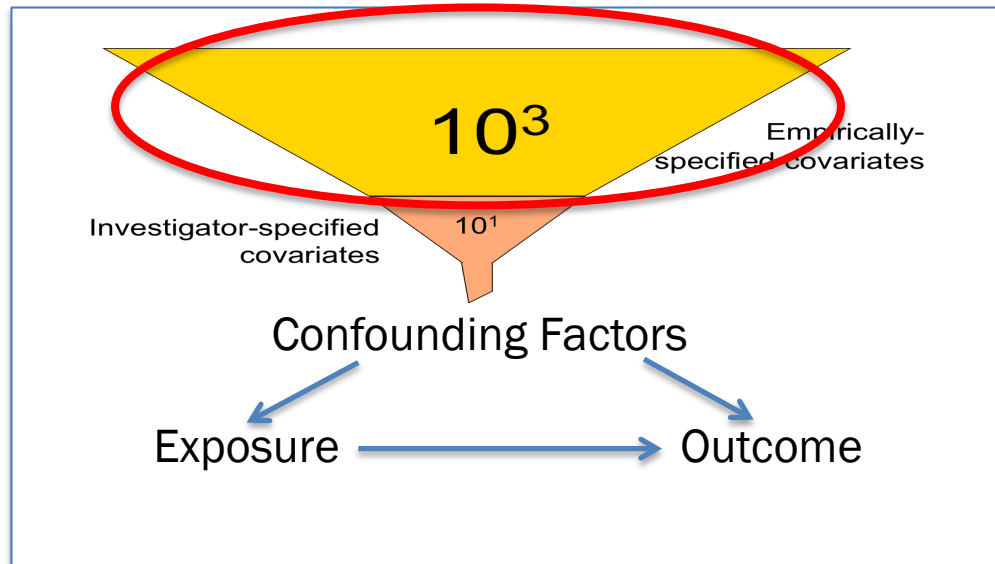
- Based on this [initial evaluation](#), TreeScan based approaches in pregnancy appear promising
- Consider further refinement of the methods:
 - Outcome trees with hierarchical groupings informed by embryology or shared disease processes
 - Improved confounding control
 - Methods to deal with different pregnancy durations

Confounding Adjustment



Confounding Adjustment

- Use the richness of the data to identify large number of potential **risk factors** for the outcome or **proxies** for them → summary confounding score
- Lack robust information on certain variables (e.g., BMI, OTC medications, smoking, illicit drug use, SES, lifestyle factors)
- Attempt to mitigate through the generous inclusion of potential proxies

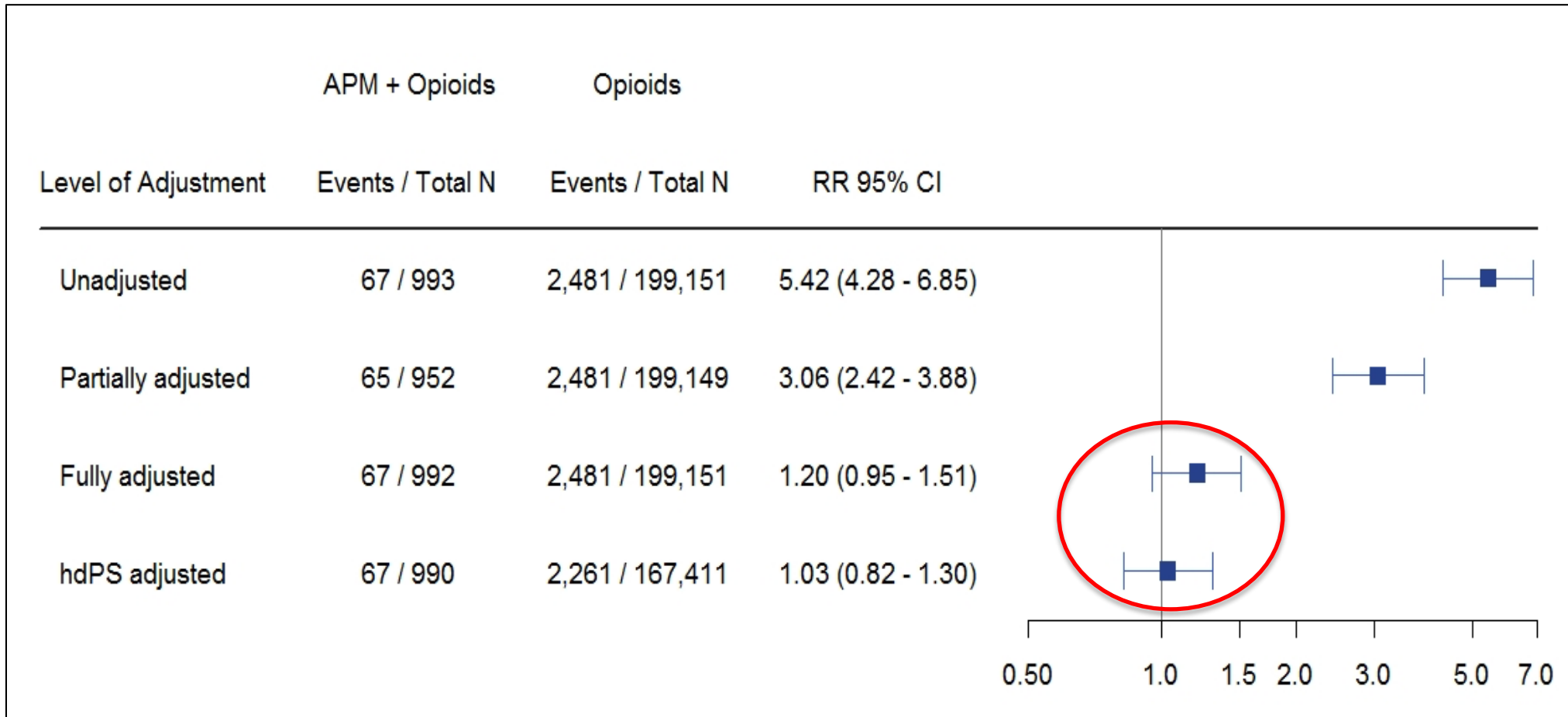


High-dimensional PS

- Empirically identify candidate covariates from thousands of codes, prioritize covariates based on confounding potential, and integrate them into a PS ($N \approx 200$)
- Demonstrated to improve confounding control in some circumstances

Confounding: hdPS Adjustment

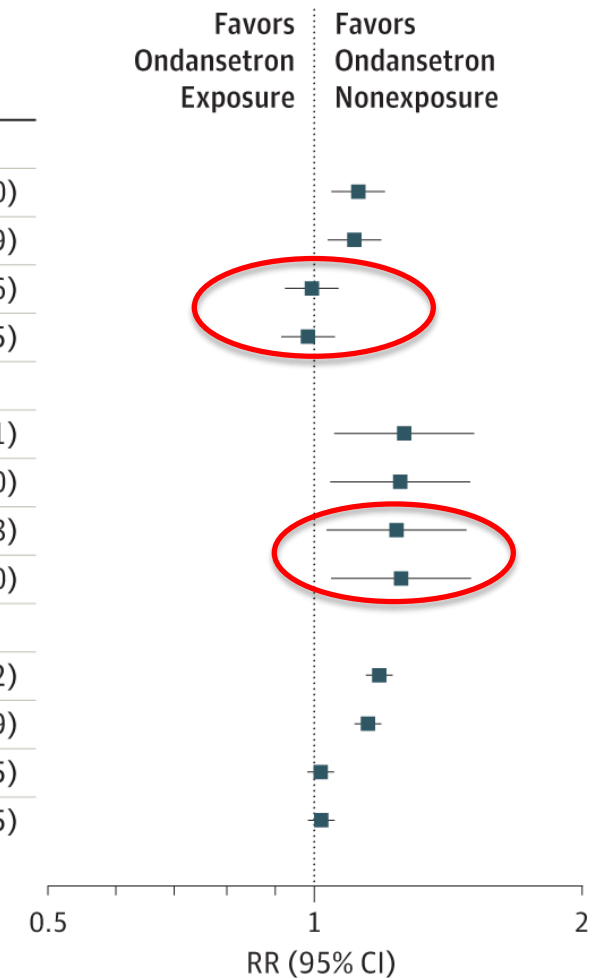
Risk of neonatal drug withdrawal:



Confounding: hdPS Adjustment

Ondansetron and the risk of congenital malformations:

Level of Adjustment	Exposed to Ondansetron		Unexposed to Ondansetron		RR (95% CI)
	No. of Events	Total No. of Infants	No. of Events	Total No. of Infants	
Cardiac malformations (primary outcome)					
Unadjusted	835	88 467	14 577	1 727 947	1.12 (1.04-1.20)
Propensity score stratified (level 1)	835	88 467	14 577	1 727 947	1.11 (1.03-1.19)
Propensity score stratified (level 2)	835	88 446	14 573	1 727 546	0.99 (0.93-1.06)
High-dimensional propensity score stratified	835	88 467	14 577	1 727 925	0.98 (0.92-1.05)
Oral clefts (primary outcome)					
Unadjusted	124	88 467	1 921	1 727 947	1.26 (1.05-1.51)
Propensity score stratified (level 1)	124	88 467	1 921	1 727 947	1.25 (1.04-1.50)
Propensity score stratified (level 2)	124	88 446	1 920	1 727 546	1.24 (1.03-1.48)
High-dimensional propensity score stratified	124	88 467	1 921	1 727 925	1.25 (1.04-1.50)
Any congenital malformation (secondary outcome)					
Unadjusted	3 277	88 467	54 174	1 727 947	1.18 (1.14-1.22)
Propensity score stratified (level 1)	3 277	88 467	54 174	1 727 947	1.15 (1.11-1.19)
Propensity score stratified (level 2)	3 275	88 446	54 163	1 727 546	1.01 (0.98-1.05)
High-dimensional propensity score stratified	3 277	88 467	54 174	1 727 925	1.02 (0.98-1.05)



Multi-Site Collaborations

Multi-Site Collaborations



- When exposure to the specific drug of interest involves a small fraction of the pregnant population, even these large cohorts are constrained in their information.
- Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP)
- International Pregnancy Safety Study (InPreSS) consortium
 - Denmark, Finland, Iceland, Norway, Sweden, US
 - Follow-up on a positive association identified in a single study
 - Common protocol; but allow deviations to take advantage of the best available information in each country's data

InPreSS: Follow-up on positive association

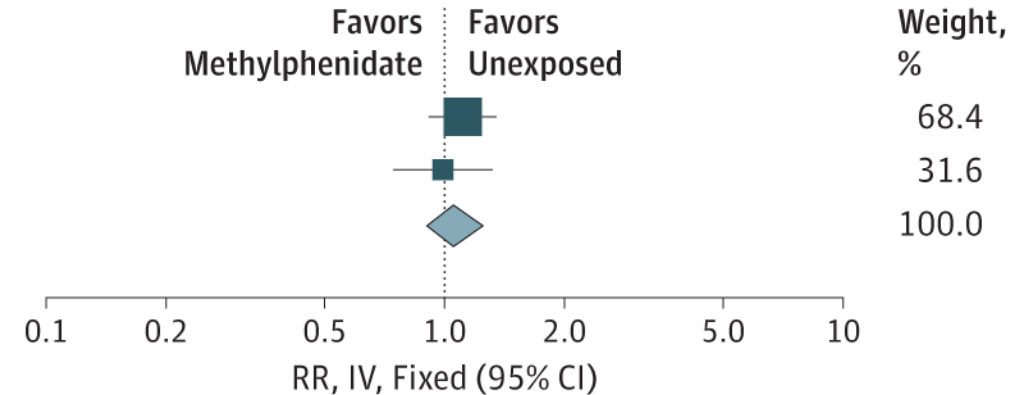
Methylphenidate and Amphetamine Use in Pregnancy and Risk for Congenital Malformations

A Any congenital malformation

Data Source	log(RR)	SE	RR (95% CI)
United States	0.10436002	0.09987	1.11 (0.91-1.35)
Nordic	-0.01005034	0.14677	0.99 (0.74-1.32)
Total (95% CI)			1.07 (0.91-1.26)

Heterogeneity: $\chi^2 = 0.42_1, P = .52; I^2 = 0\%$

Test for overall effect: $z = 0.83, P = .41$

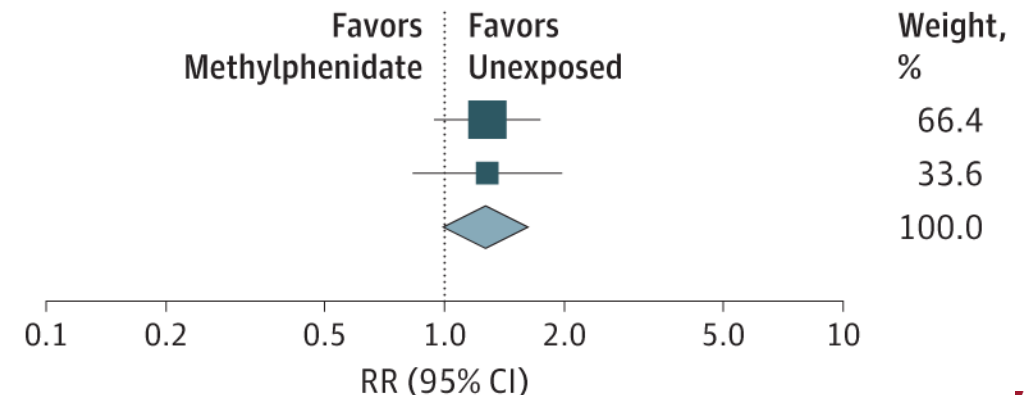


B Cardiovascular malformations

Data Source	log(RR)	SE	RR (95% CI)
United States	0.2468601	0.15665	1.28 (0.94-1.74)
Nordic	0.2468601	0.21999	1.28 (0.83-1.97)
Total (95% CI)			1.28 (1.00-1.64)

Heterogeneity: $\chi^2 = 0.00_1, P > .99; I^2 = 0\%$

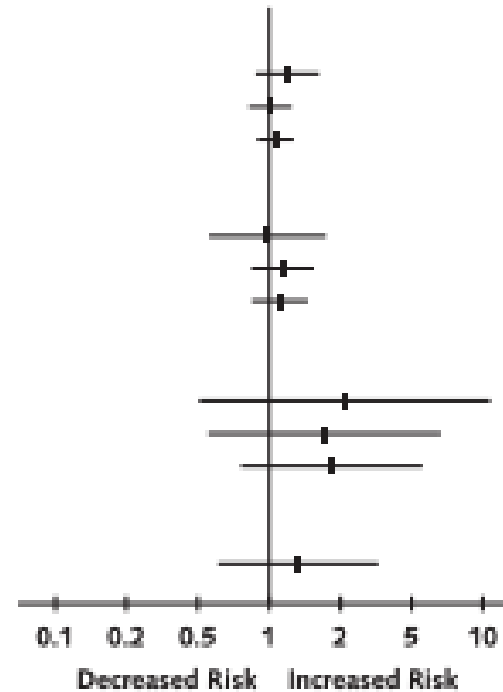
Test for overall effect: $z = 1.93, P = .05$



InPreSS: Uncommon exposure

β-blocker Use in Pregnancy and the Risk for Congenital Malformations

Outcome	Events/Total, n/N (%)		Adjusted RR (95% CI)
	Exposed	Unexposed	
Any congenital malformation			
Nordic	48/682 (7.0)	152/2895 (5.3)	1.22 (0.88 to 1.71)
US MAX	78/1668 (4.7)	534/13 232 (4.0)	1.01 (0.80 to 1.27)
Pooled	126/2350 (5.4)	686/16 127 (4.3)	1.07 (0.89 to 1.30)
Cardiac malformations			
Nordic	15/682 (2.2)	55/2895 (1.9)	0.98 (0.52 to 1.84)
US MAX	37/1668 (2.2)	224/13 232 (1.7)	1.16 (0.82 to 1.63)
Pooled	52/2350 (2.2)	279/16 127 (1.7)	1.12 (0.83 to 1.51)
Cleft lip/palate			
Nordic	3/682 (0.4)	4/2895 (0.1)	2.26 (0.47 to 10.8)
US MAX	<11/1668 (<0.7)	13/13 232 (0.1)	1.81 (0.52 to 6.33)
Pooled	<14/2350 (<0.6)	17/16 127 (0.1)	1.97 (0.74 to 5.25)
Central nervous system malformations			
US MAX	<11/1668 (<0.7)	36/13 232 (0.3)	1.37 (0.58 to 3.25)



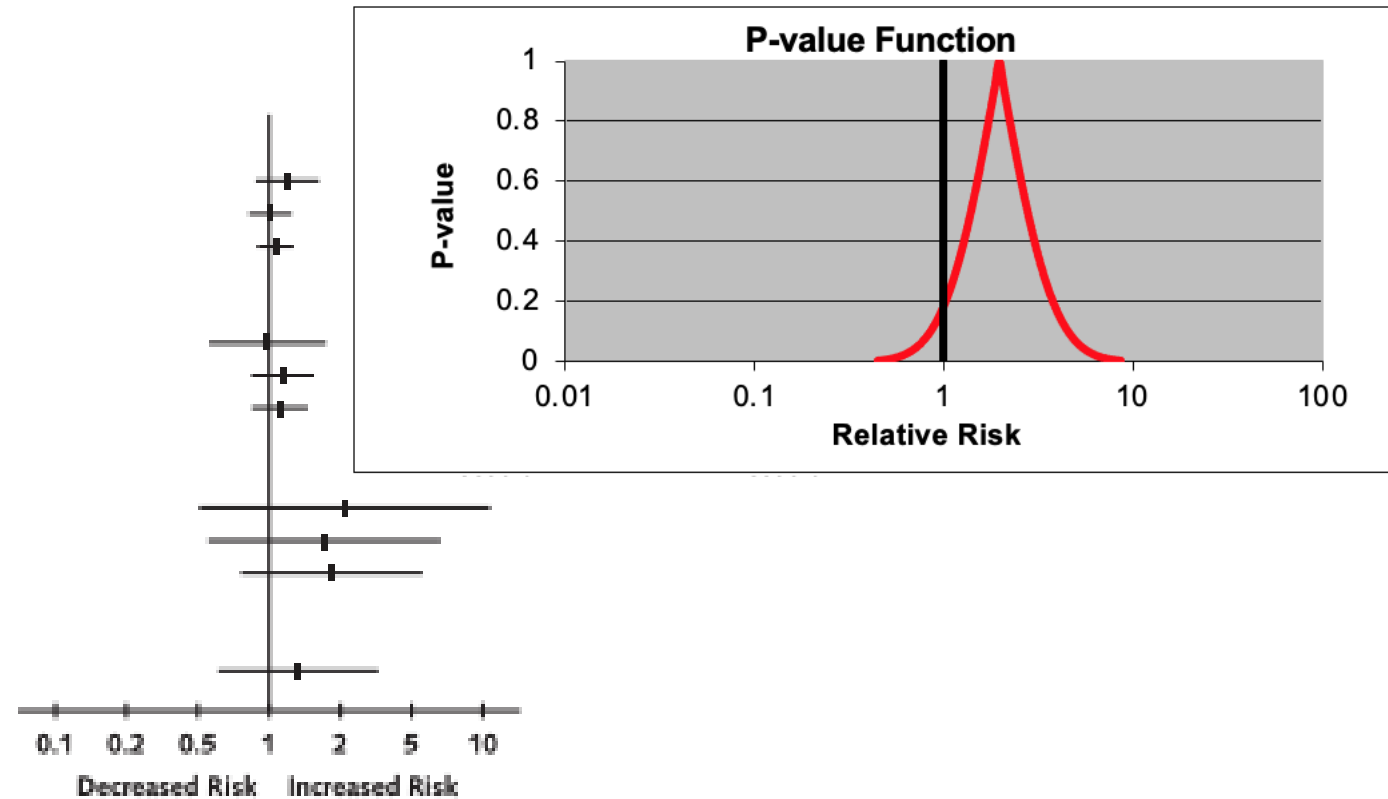
- First trimester exposure to β -blockers
- Background: Meta-analysis reported significantly increased risks for cardiac defects, cleft lip or palate, and neural tube defects.
- Cohort: Pregnant women with a diagnosis of hypertension

InPreSS: Uncommon exposure

β-blocker Use in Pregnancy and the Risk for Congenital Malformations



Outcome	Events/Total, n/N (%)		Adjusted RR (95% CI)
	Exposed	Unexposed	
Any congenital malformation			
Nordic	48/682 (7.0)	152/2895 (5.3)	1.22 (0.88 to 1.71)
US MAX	78/1668 (4.7)	534/13 232 (4.0)	1.01 (0.80 to 1.27)
Pooled	126/2350 (5.4)	686/16 127 (4.3)	1.07 (0.89 to 1.30)
Cardiac malformations			
Nordic	15/682 (2.2)	55/2895 (1.9)	0.98 (0.52 to 1.84)
US MAX	37/1668 (2.2)	224/13 232 (1.7)	1.16 (0.82 to 1.63)
Pooled	52/2350 (2.2)	279/16 127 (1.7)	1.12 (0.83 to 1.51)
Cleft lip/palate			
Nordic	3/682 (0.4)	4/2895 (0.1)	2.26 (0.47 to 10.8)
US MAX	<11/1668 (<0.7)	13/13 232 (0.1)	1.81 (0.52 to 6.33)
Pooled	<14/2350 (<0.6)	17/16 127 (0.1)	1.97 (0.74 to 5.25)
Central nervous system malformations			
US MAX	<11/1668 (<0.7)	36/13 232 (0.3)	1.37 (0.58 to 3.25)



Conclusions

Conclusions

- Goals:
 1. Quickly detect problems when they exist
 2. Show the absence of strong harmful effects when there are none
- Strength in the use of complementary approaches: pregnancy exposure registries, case-control surveillance, healthcare utilization databases
- Unique opportunities to further advance the field of perinatal pharmacoepidemiology: methods development, multi-site collaborations
- Value of linkages to external databases with additional clinical information: birth/death certificates, laboratory tests, electronic medical records



Acknowledgements

Harvard Program in Perinatal and Pediatric Pharmacoepidemiology Core Team:

- Sonia Hernández-Díaz, MD DrPH
- Brian T. Bateman, MD MS
- Yanmin Zhu, MS PhD
- Kathryn Gray, MD, PhD
- Elisabeth Suarez, PhD
- Loreen Straub, MD MSc
- Shirley Wang, PhD ScM
- Martin Kulldorff, PhD