Sentinel

Triple challenges – Small sample size in both exposure and control groups to scan rare maternal outcomes in a signal identification approach: A simulation study

Thuy N Thai[1], Almut G Winterstein [2], Elizabeth A Suarez[3], Jiwei He[4], Yueqin Zhao[4], Di Zhang[4], Danijela Stojanovic[4], Jane Liedtka[4], Abby Anderson[4], José J Hernández-Muñoz[4], Monica Munoz[4], Wei Liu[4], Inna Dashevsky[1], Elizabeth Messenger-Jones[1], Elizabeth Siranosian[1], Judith C Maro[1]

[1] Harvard Pilgrim Health Care and Harvard Medical School; [2] University of Florida, [3] Rutgers University, [4] U.S. Food and Drug Administration

39th International Conference on Pharmacoepidemiology & Therapeutic Risk Management
Halifax, Nova Scotia, Canada

08/27/2023

Disclosures

- This project was supported by Task Order HHSF22301012T under Master Agreement HHSF223201400030I from the US Food and Drug Administration (FDA).

- The views expressed in this presentation represent those of the authors and do not necessarily represent the official views of the US FDA.

Overview of TreeScan™

https://www.treescan.org/

- A signal identification method
- Evaluates thousands of outcomes (1 exposure: N outcomes) simultaneously to identify potential adverse events
- Does not require selecting a specific exposure-outcome pairing for hypothesis testing
- Automatically adjusts for multiple scenarios

Kulldorff M, Fang Z, Walsh S. A tree-based scan statistic for database disease surveillance. Biometrics, 2003,59:323-331.
Kulldorff M, Dashevsky I, Avery TR, Chan KA, Davis RL, Graham D, Platt R, Andrade SE, Boudreau D, Gunter MJ, Herrinton LJ, Pawloski P, Raebel MA, Roblin D, Brown JS. Drug safety data mining with a tree-based scan statistic. Pharmacoepidemiology and Drug Safety, 2013, 22:517-523.
Maro JC, Nguyen MD, Dashevsky I, Baker MA, Kulldorff M. Statistical Power for Postlicensure Medical Product Safety Data Mining. *EGEMS (Wash DC)*. Jun 2017;5(1):6.

Overview of TreeScan™

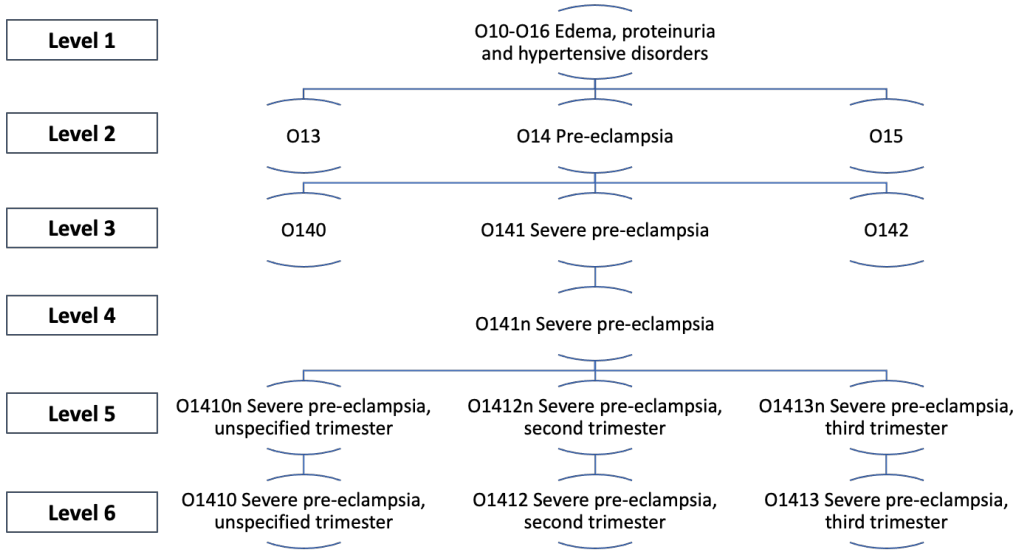


Figure 1. An example of a branch of the hierarchical maternal outcome tree

- Thousands of outcomes are structured under a hierarchical tree.
- Observed and expected event counts in the exposure group are calculated for each node.
- Scanning is conducted at multiple outcome levels.

Thuy Thai, Almut Winterstein, Elizabeth Suarez et al. *Use Of The Tree-Based Scan Statistic For Surveillance Of Maternal Outcomes Following Medication Use During Gestation*

Maro JC, Nguyen MD, Dashevsky I, Baker MA, Kulldorff M. Statistical Power for Postlicensure Medical Product Safety Data Mining. *EGEMS (Wash DC)*. Jun 2017;5(1):6.

Poisson TreeScan's assumption

- **Assumption**: Expected events among the exposure group, which are estimated from the control group's background incidence proportions, are assumed to be **known without error**.
- **The problem:** With a small control group, imprecise background incidence proportions may affect TreeScan's ability to detect potential signals or control Type I error.
- **Objective**: Given triple challenges in pregnancy-related studies (small exposure and control groups as well as rare outcomes), we evaluated impacts of the sample size of the control group on TreeScan analysis when screening for adverse maternal outcomes in a simulated data set.

Methods

- The simulation was based on background incidence proportions that were estimated using a pregnancy cohort in the Merative® MarketScan® Database 2015-2019
  - Single livebirth pregnancies
  - Among pregnant persons aged 10-54 years old
  - Exposed to oral macrolides or oral penicillins
- Maternal outcomes:
  - The tree: codes from the ICD-10-CM Chapter Pregnancy, Childbirth, and the Puerperium (O00-O9A)
  - Evaluated one day after index date (medication date) to 30 days after delivery
- The "true" background incidence proportion for each outcome =

$$\frac{total\ incident\ outcomes\ in\ macrolide\ and\ penicillin\ users}{Total\ pregnant\ persons\ exposed\ to\ macrolides\ or\ penicillins}$$

Methods – Parameters for the simulation study

	Control group sample size	Exposure group sample size	Outcome	Relative risk (RR)
Null scenarios	1000, 2500, 5000, 10,000 and 50,000	1000, 2500, 5000, and 10,000	All outcomes were set with RR=1.0	RR=1.0
Elevated risk scenarios	1000, 2500, 5000, 10,000 and 50,000	1000, 2500, 5000, and 10,000	The following outcomes were set to RR=1.5 or RR=2.0 ○ O47.9 incidence proportion=0.02 ○ O14.13 incidence proportion=0.01 ○ O14.95 incidence proportion=0.005 All other leaves with RR=1.0	RR=1.5 and RR=2.0

We evaluated a total of 140 scenarios.

Steps to set up each scenario (1000 iterations per scenario)

Step	
1	Calculate the expected count of exposed outcomes
2	Calculate the observed count of exposed outcomes
3	Implement Poisson TreeScan analysis
4	Aggregate the TreeScan outputs over 1000 iterations

- Step 1:
  - When using true background incidence (truth)
    $$Expected\ count = true\ background\ incidence \times N\ exposed$$
  - When using the simulated imprecise background incidence
    $$Expected\ count = incidence\ based\ on\ Poisson\ draw \times N\ exposed$$
- Step 2: A random draw of: $true\ background\ incidence \times N\ exposed \times RR$
- Step 3: Evaluate the observed/expected counts using an alert threshold of $p<0.05$
- Step 4:
  - When evaluating the null scenarios:
    - Overall power = proportion of iterations with at least 1 alert among 1000 iterations
    - Mean number false alerts among 1000 iterations
  - When evaluating the elevated-risk scenarios: node-specific power = the proportion of iterations having the injected-node as an alert

Type I error and mean number of false alerts in null scenarios

Scenarios	Exposure size	Control size	Ratio of exposure and control sizes	Type I error	Mean number of false alerts
Base case (truth)				0.055	0.063
Imprecise incidence proportion	1000	1000	1:1	0.882	2.507
		2500	1:2.5	0.533	0.836
		5000	1:5	0.265	0.372
		10,000	1:10	0.189	0.232
		50,000	1:50	0.073	0.082
Base case (truth)				0.054	0.06
Imprecise incidence proportion	2500	1000	2.5:1	1.000	10.64
		2500	1:1	0.948	3.45
		5000	1:2	0.614	1.11
		10,000	1:4	0.380	0.57
		50,000	1:20	0.113	0.13

In the base case, the Type I error is maintained ~ 0.05 and the number of false alert was <1.

Type I error and mean number of false alerts in null scenarios

Scenarios	Exposure size	Control size	Ratio of exposure and control sizes	Type I error	Mean number of false alerts
Base case (truth)	1000			0.055	0.063
Imprecise incidence proportion		1000	1:1	0.882	2.507
		2500	1:2.5	0.533	0.836
		5000	1:5	0.265	0.372
		10,000	1:10	0.189	0.232
		50,000	1:50	0.073	0.082
Base case (truth)	2500			0.054	0.06
Imprecise incidence proportion		1000	2.5:1	1.000	10.64
		2500	1:1	0.948	3.45
		5000	1:2	0.614	1.11
		10,000	1:4	0.380	0.57
		50,000	1:20	0.113	0.13

- In the imprecise incidence proportions, there was an inflation in Type I error.
- When the exposure size was fixed, Type I error decreases as the control size increases.

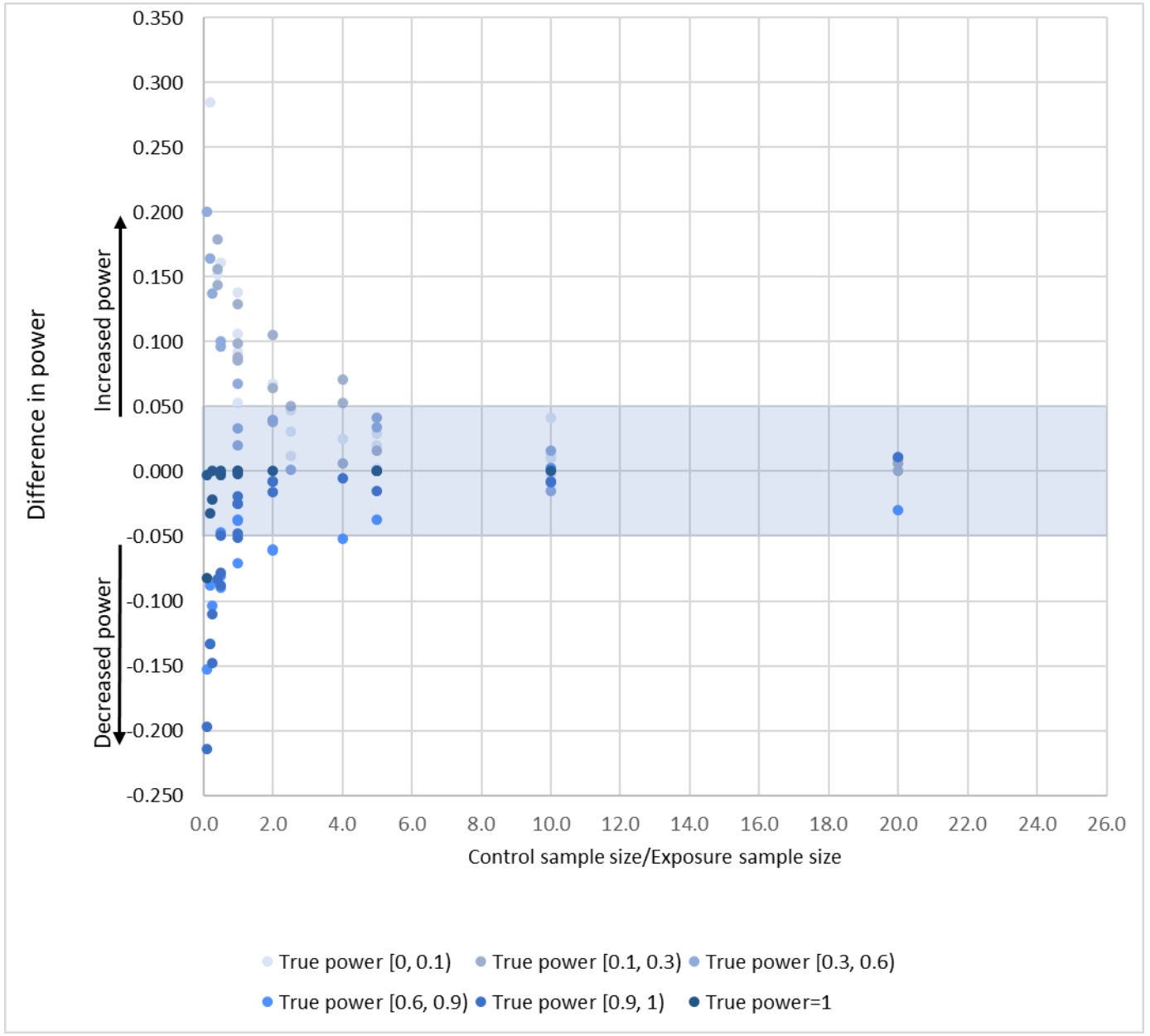
Type I error and mean number of false alerts in null scenarios

Scenarios	Exposure size	Control size	Ratio of exposure and control sizes	Type I error	Mean number of false alerts
Base case (truth)	1000			0.055	0.063
Imprecise incidence proportion		1000	1:1	0.882	2.507
		2500	1:2.5	0.533	0.836
		5000	1:5	0.265	0.372
		10,000	1:10	0.189	0.232
		50,000	1:50	0.073	0.082
Base case (truth)	2500			0.054	0.06
Imprecise incidence proportion		1000	2.5:1	1.000	10.64
		2500	1:1	0.948	3.45
		5000	1:2	0.614	1.11
		10,000	1:4	0.380	0.57
		50,000	1:20	0.113	0.13

The mean number of false alerts was larger when there was a higher ratio of the exposure relative to the control group size.

Results under the elevated relative risk scenarios

- Light blue shading indicates a tolerable range in difference of power between the imprecise vs. true incidence proportion scenarios at 0.05.
- When the control size was small, the difference in power was outside the tolerable range.
- When the control size was substantially larger than the exposure size: the difference in power was within the tolerable range.



Conclusions

- We found that imprecise incidence proportions generated by a small control group resulted in:
  - Under the null scenario: inflation of Type I error and increase number of false alerts.
  - Under the elevated risk scenario: changes in observed outcome-specific power in both directions (higher/lower than in the base case).
- Ideally, the control size should be several times larger than the exposure size to limit the number of false positive alerts and retain statistical power for true alerts.
- Our results suggest that users of Poisson tree-based scan statistics should proceed cautiously with small control groups and take measures to mitigate small control group sizes where possible.

Thank You

Thuy Thai

thuy_thai@harvardpilgrim.org