[NCKU, Taiwan] RWE Investigator Meeting (July 25th, 2025)

Safety Evaluation in RWE Context: Approaches for pharmacovigilance using multiple data sources

Ju Hwan (Joe) Kim, Pharm.D., Ph.D.
Research Professor
Department of Bio-health Regulatory Science
School of Pharmacy, Sungkyunkwan Univ. (SKKU)
South Korea

Disclosure Statement

- The content of this presentation is intended to provide an overview of a recent methodological approach for pharmacovigilance using real-world data (RWD).

- I have no conflicts of interest to disclose regarding the published articles or materials referenced in this presentation.

Contents

Approaches for pharmacovigilance using multiple data sources

1

- Conventional approach
- Machine learning approach
- Hierarchical data mining method
- Case examples

2

- Considerations for federated network approach and multi- database study

Approaches for Pharmacovigilance

- Methods for identifying safety signals in the longitudinal healthcare databases

  1) Pharmacoepidemiologic designs (e.g., case-only study, cohort study)

  2) Sequence symmetry analysis (SSA)

  3) Supervised machine learning (i.e., Gradient boosting machine, random forest)

  4) Tree-based scan statistic

Arnaud M, et al. Methods for safety signal detection in healthcare databases: a literature review. Expert Opin Drug Saf. 2017 Jun;16(6):721-732.

Approaches for Pharmacovigilance

- Pharmacoepidemiologic designs (e.g., case-only study, cohort study)
  - Exposure(s) and outcome(s) of interest need to be **pre-specified**
  - (1) Cohort study design for safety surveillance
    - New-user, active comparator design (comparator drug needs to be pre-specified)
    - # of cohorts depend on the # of outcomes under assessment (i.e., Incident outcome cohort)
    - Confounders in the exposure-outcome pathway need to be carefully selected

Approaches for Pharmacovigilance

- Pharmacoepidemiologic designs (e.g., case-only study, cohort study)
  - Exposure(s) and outcome(s) of interest need to be **pre-specified**
  - (2) Case-only study design for safety surveillance
    - Self-controlled case series, Case-crossover, Case-case-time-control designs
    - Ideal design in cases where…
      - No adequate comparator group
      - Transient exposure & outcome
      - Need to control for time-invariant confounders

Approaches for Pharmacovigilance

RESEARCH

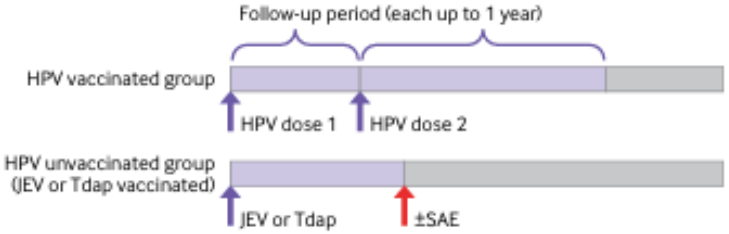
OPEN ACCESS

Check for updates

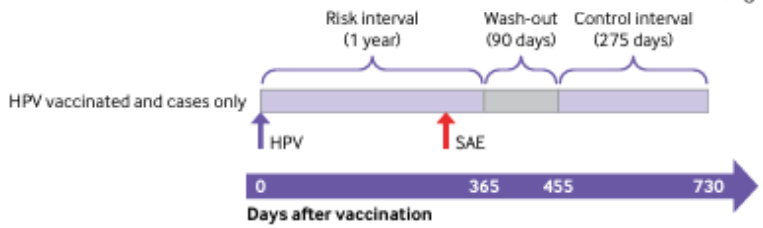
Association between human papillomavirus vaccination and serious adverse events in South Korean adolescent girls: nationwide cohort study

Dongwon Yoon,[1] Ji-Ho Lee,[1,2] Hyesung Lee,[1] Ju-Young Shin[1,3]

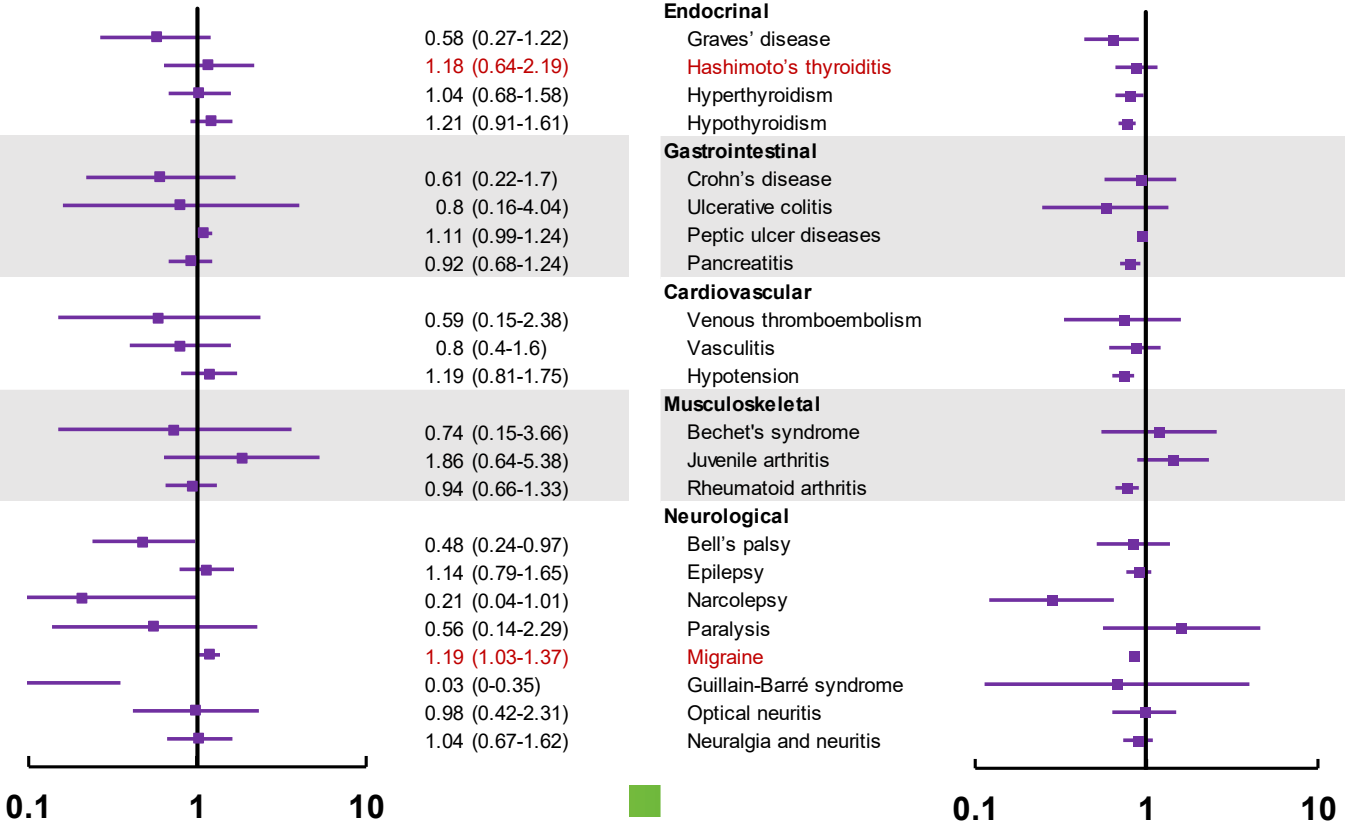
Primary analysis: cohort design



Secondary analysis: self-controlled risk interval design



	Cohort	SCRI
	Adjusted Rate Ratio (95% CI)	Adjusted Relative Risk (95% CI)
Endocrinal		
Graves' disease	0.58 (0.27-1.22)	0.63 (0.44-0.91)
Hashimoto's thyroiditis	1.18 (0.64-2.19)	0.87 (0.65-1.15)
Hyperthyroidism	1.04 (0.68-1.58)	0.8 (0.66-0.96)
Hypothyroidism	1.21 (0.91-1.61)	0.77 (0.69-0.87)
Gastrointestinal		
Crohn's disease	0.61 (0.22-1.7)	0.92 (0.57-1.48)
Ulcerative colitis	0.8 (0.16-4.04)	0.58 (0.25-1.36)
Peptic ulcer diseases	1.11 (0.99-1.24)	0.94 (0.89-0.99)
Pancreatitis	0.92 (0.68-1.24)	0.8 (0.69-0.91)
Cardiovascular		
Venous thromboembolism	0.59 (0.15-2.38)	0.73 (0.34-1.6)
Vasculitis	0.8 (0.4-1.6)	0.86 (0.61-1.21)
Hypotension	1.19 (0.81-1.75)	0.73 (0.63-0.85)
Musculoskeletal		
Bechet's syndrome	0.74 (0.15-3.66)	1.18 (0.54-2.57)
Juvenile arthritis	1.86 (0.64-5.38)	1.43 (0.89-2.3)
Rheumatoid arthritis	0.94 (0.66-1.33)	0.77 (0.66-0.9)
Neurological		
Bell's palsy	0.48 (0.24-0.97)	0.84 (0.51-1.36)
Epilepsy	1.14 (0.79-1.65)	0.9 (0.76-1.06)
Narcolepsy	0.21 (0.04-1.01)	0.28 (0.12-0.65)
Paralysis	0.56 (0.14-2.29)	1.6 (0.55-4.62)
Migraine	1.19 (1.03-1.37)	0.85 (0.8-0.9)
Guillain-Barré syndrome	0.03 (0-0.35)	0.67 (0.11-3.99)
Optical neuritis	0.98 (0.42-2.31)	0.98 (0.64-1.5)
Neuralgia and neuritis	1.04 (0.67-1.62)	0.89 (0.73-1.09)



Outcomes of interest need to be pre-specified

More suitable for signal evaluation / hypothesis testing

Yoon D, et al. BMJ 2021; 372:m4931

Approaches for Pharmacovigilance

- Sequence symmetry analysis (SSA) – prescription based
  - Case-only design; comparing temporal sequences of drug A → drug B
  - Some advantages over cohort study designs include:
    - Minimal dataset requirement, computationally efficient
    - Controls time-invariant confounders
    - Ideal design for safety signal detection / hypothesis generating
  - Major drawback:
    - Rely on an assumption of “event onset influences drug initiation”
    - Prescribing time trends need to be accounted for

Lai EC, et al. Eur J Epidemiol 2017; 32:567-582.

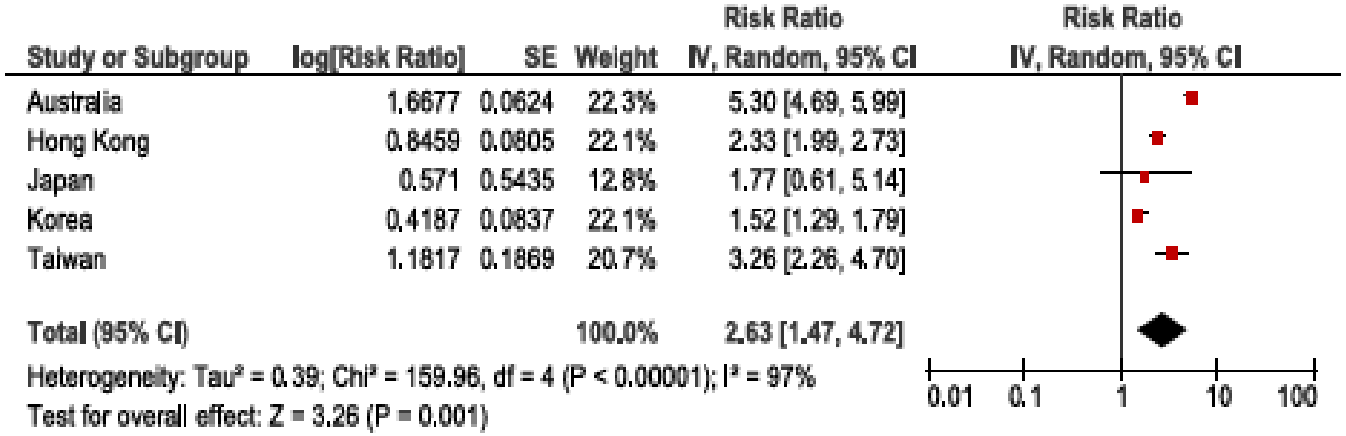
Approaches for Pharmacovigilance

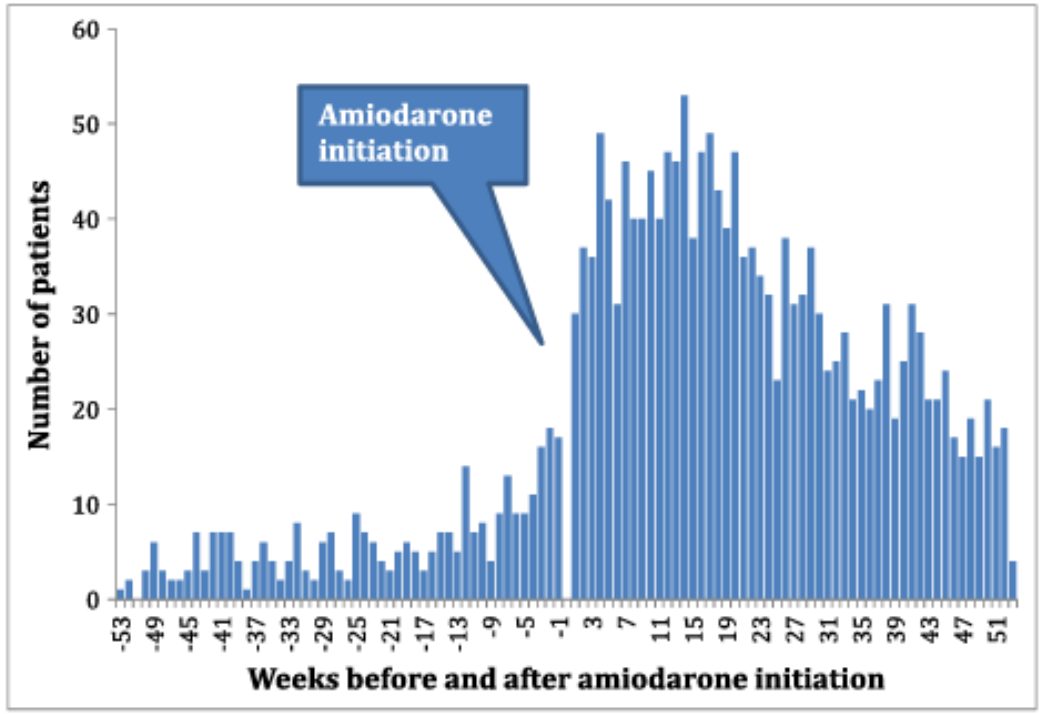
PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2015; **24**: 858–864
Published online 22 April 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.3780

ORIGINAL REPORT

Prescription sequence symmetry analysis: assessing risk, temporality, and consistency for adverse drug reactions across datasets in five countries

Nicole Pratt[1], Esther W. Chan[2], Nam-Kyong Choi[3], Michio Kimura[4], Tomomi Kimura[4], Kiyoshi Kubota[5], Edward Chia-Cheng Lai[6], Kenneth K. C. Man[2], Nobuhiro Ooba[5], Byung-Joo Park[7,8], Tsugumichi Sato[5], Ju-Young Shin[8], Ian C. K. Wong[2], Yea-Huei Kao Yang[6] and Elizabeth E. Roughead[1]*





Pratt N, et al. PDS 2015; 24: 858-864.

Approaches for Pharmacovigilance

- Supervised machine learning (i.e., Gradient boosting machine, random forest)
  - Data-driven approach for predicting new safety signals
  - Input dataset for machine learning training
    - All drug-event pairs in the database requires labeling (positive control, negative control, unknown)
    - Training based on the features of positive and negative controls (i.e., label data)
    - Validation with the subsets of label data to determine optimal threshold for signal detection
  - Generates predictive probability of being safety signal amongst the “Unknown” drug-AE pairs

Approaches for Pharmacovigilance

- Input dataset of every possible drug-event pairs
- Each pair need to be labeled as:
  - ① **Positive:**

    drug-event with established casual relationship
  - ② **Negative:**

    drug-event with no casual relationship
  - ③ **Unknown:**

    Potential event of interest

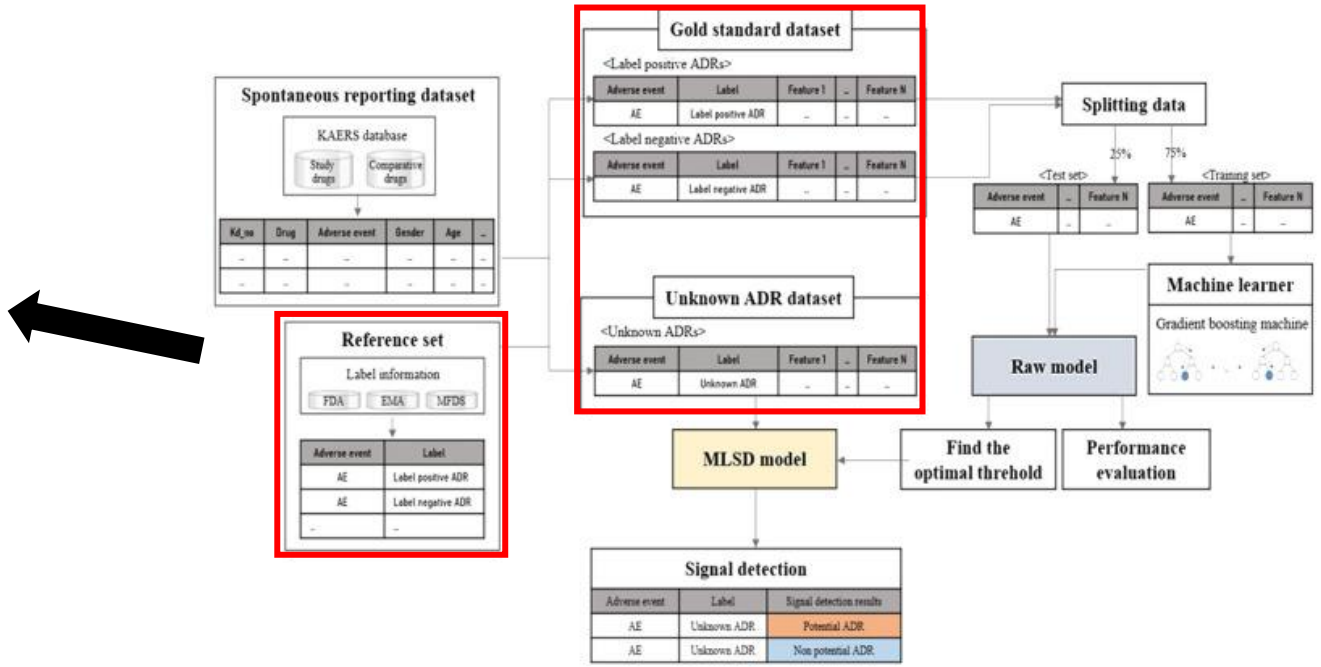


TABLE 3 | Distribution of labels for two input data.

Study drugs	Suspected ADRs	Gold standard		Unknown ADRs‡
		Label-positive ADRs*	Label-negative ADRs†	
Nivolumab	136 (100%)	70 (51%)	15 (11%)	51 (38%)
Docetaxel	486 (100%)	267 (55%)	71 (15%)	148 (30%)

ADR, adverse drug reaction.

**Label-positive ADRs are defined as the adverse events (AEs) of respective drug that were listed in the labels of the Ministry of Food and Drug Safety, U. S. Food and Drug Administration, and European Medicines Agency.*

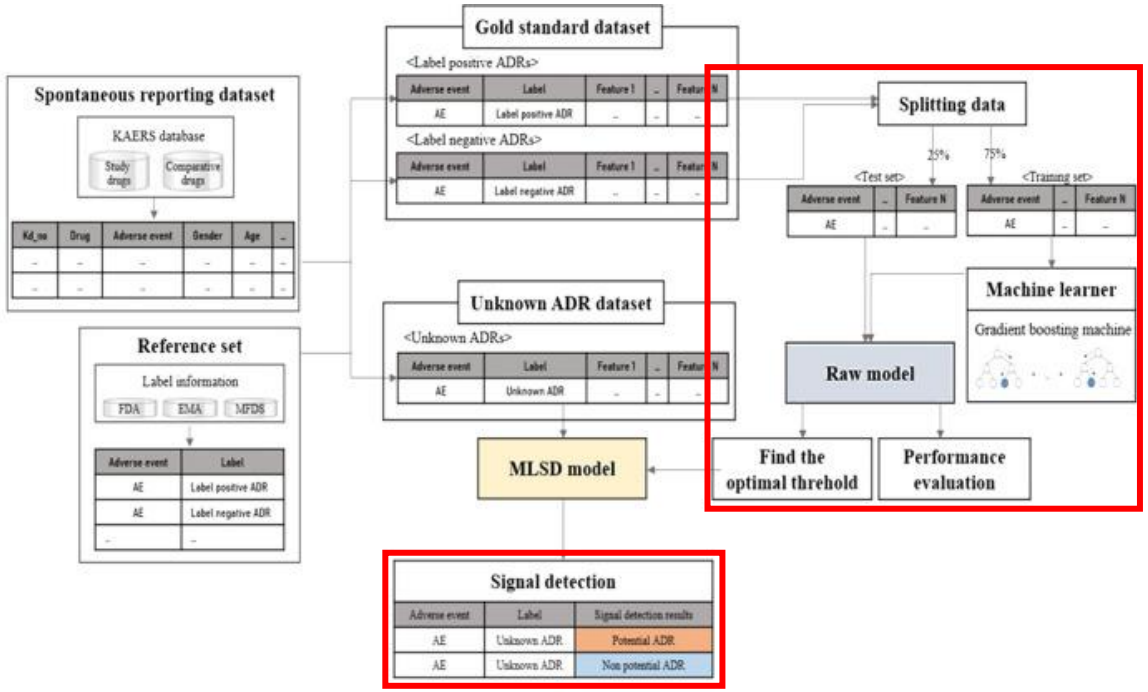
†Label-negative ADRs are defined as the AEs not listed in the labels of the target drug and any other drugs in the same therapeutic class and considered unlikely to be ADR signals by experts.

‡Unknown ADRs are drug-AE pairs neither label-positive ADRs nor label-negative ADRs.

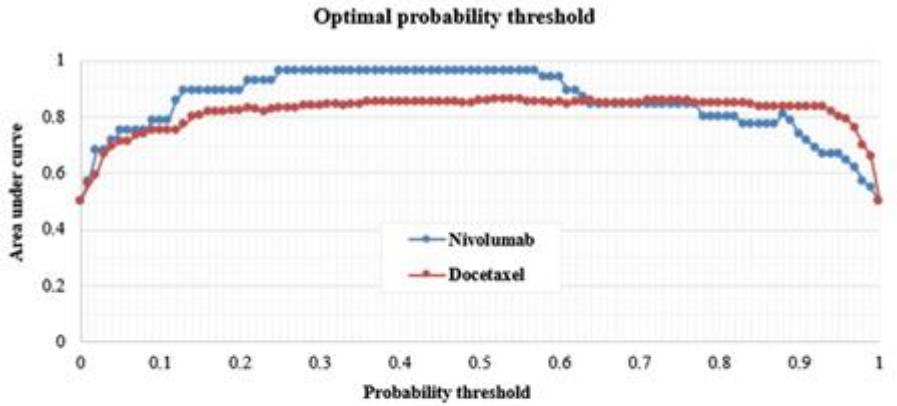
dverse event reporting system database; FDA, U.S Food and
ug safety; AE, adverse event; ADR, adverse drug reaction;

Bae J, et al. Front Pharmcol 2021; 11:602365.

Approaches for Pharmacovigilance



Abbreviations: KIDS-KD, Korea Institute of Drug Safety and Risk Manage[illegible]-Korea adverse event reporting system database; FDA, U.S Food and Drug Administration; EMA, European Medicines agency; MFDS, ministry o[illegible]d and drug safety; AE, adverse event; ADR, adverse drug reaction; MLSD, machine learning signal detection



	Optimal threshold	Area under curve	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Nivolumab	0.57	0.9643	97%	100%	93%	95%	100%
Docetaxel	0.55	0.8626	87%	83%	89%	86%	87%

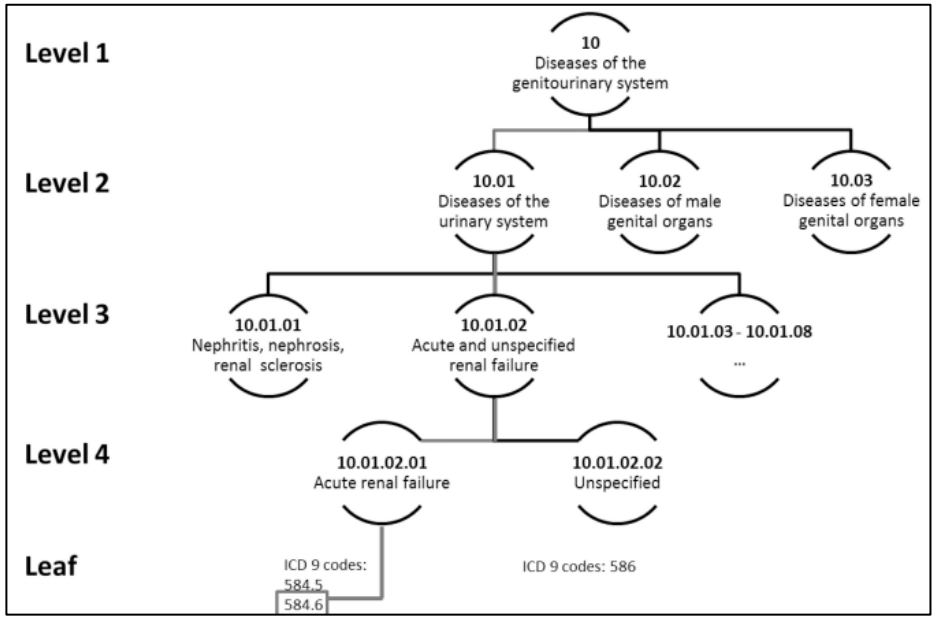
*The values of area under curve, accuracy, sensitivity, specificity, positive predictive value, and negative predictive value on the optimal threshold are calculated for each study drug.

Serious cardiac AEs[a]	No. of reports	Probability[b] (threshold > 0.5)	PRR05[c]	ROR05[c]	IC05[c]	EBGM05[c]
Bradycardia	17	0.99	1.60	1.60	0.48	1.68
Pericardial effusion	15	0.98	5.44	5.43	1.97	5.23
Cardio-respiratory arrest	8	0.69	1.71	1.71	0.37	1.79
Pulseless electrical activity	8	0.89	15.95	15.94	2.29	16.07
Cardiorenal syndrome	8	0.19	161.11	160.97	2.80	145.14
Cardiotoxicity	7	0.83	10.59	10.59	1.90	9.63
Cardiomyopathy	4	0.07	1.29	1.29	−0.33	1.37

Bae J, et al. Front Pharmcol 2021; 11:602365.

Approaches for Pharmacovigilance

- **Tree-based scan statistic (TreeScan)**
  - Data mining method that scans for disproportionate reporting (observed > expected #s) of outcomes
  - Uses hierarchical classification trees* to scan for outcomes at different granularity



Example of a **multi-level clinical classification tree** of ICD-9 codes
(Ref: Wang SV, et al. Epidemiology 2018; 29(6): 895-903.)

Example of **ICD-10** code hierarchical structure
(Ref: Tan G, et al. American Journal of Epidemiology 2025; 194(7): 1999-2011.)

Approaches for Pharmacovigilance

- **Key features of TreeScan**
  - No need to pre-specify drug-outcome pairs for hypothesis generating
  - Hierarchical screening → ideal for determining the granularity for outcome definition

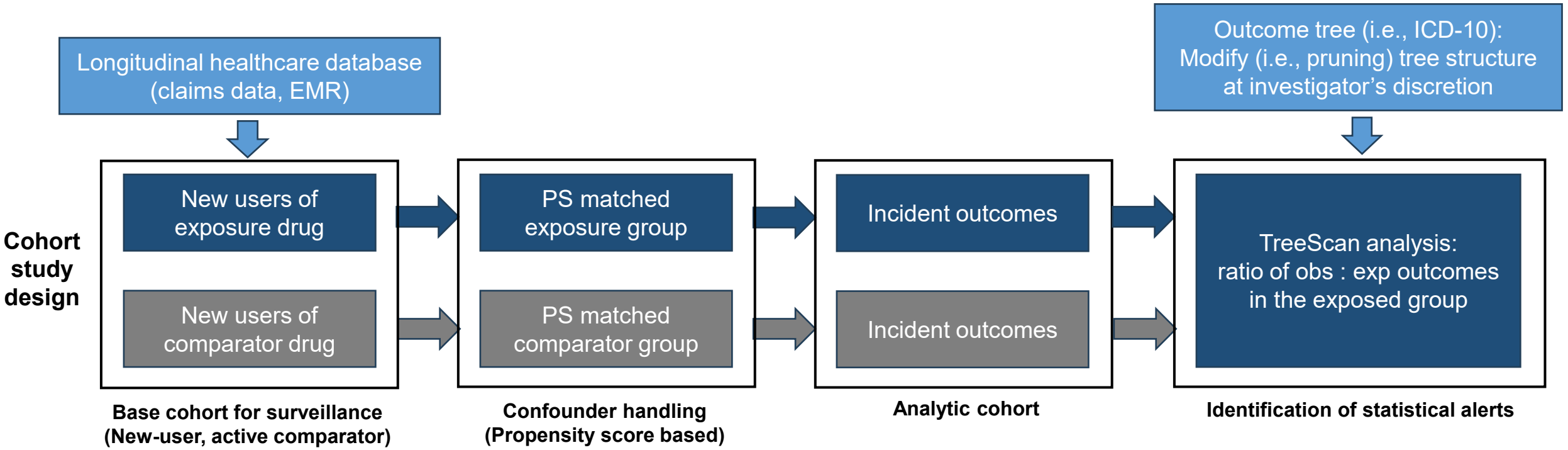
    (e.g., cardiovascular disorder > ischemic heart diseases > myocardial infarction)
  - Adjusts for multiple testing using Monte Carlo hypothesis testing (multiplicity-adjusted p-values)
  - Incorporate temporal risk windows

    → allows for specifying observation period for cohort or self-controlled case series analysis

Kulldorff M, et al. Data safety mining with a tree-based scan statistic. PDS 2013; 22: 517-523.

Approaches for Pharmacovigilance

- TreeScan as a proposed approach for PV in multi-database study



Approaches for Pharmacovigilance

(1) TreeScan for PV using RWD

Received: 20 October 2020 | Revised: 17 January 2021 | Accepted: 23 January 2021
DOI: 10.1002/edm2.237

Endocrinology, Diabetes & Metabolism — Open Access — WILEY

ORIGINAL RESEARCH ARTICLE

A novel data mining application to detect safety signals for newly approved medications in routine care of patients with diabetes

Michael Fralick[1,2] | Martin Kulldorff[1] | Donald Redelmeier[3,4] | Shirley V. Wang[1] | Seanna Vine[1] | Sebastian Schneeweiss[1] | Elisabetta Patorno[1]

Methods: In a U. S. commercial claims dataset (29 March 2013–30 Sept 2015), two pairwise cohorts of patients over 18 years of age with type 2 diabetes (T2D) who were newly dispensed canagliflozin or an active comparator, that is a dipeptidyl peptidase 4 inhibitor (DPP4) or a glucagon-like peptide 1 receptor agonist (GLP1), were identified and propensity score-matched. We used variable ratio matching with up to four people receiving a DPP4 or GLP1 for each person receiving canagliflozin. We identified potential safety signals using a hierarchical tree-based scan statistic data mining method with the hierarchical outcome tree constructed based on international classification of disease coding. We screened for incident adverse events where there were more outcomes observed among canagliflozin vs. comparator initiators than expected by chance, after adjusting for multiple testing.

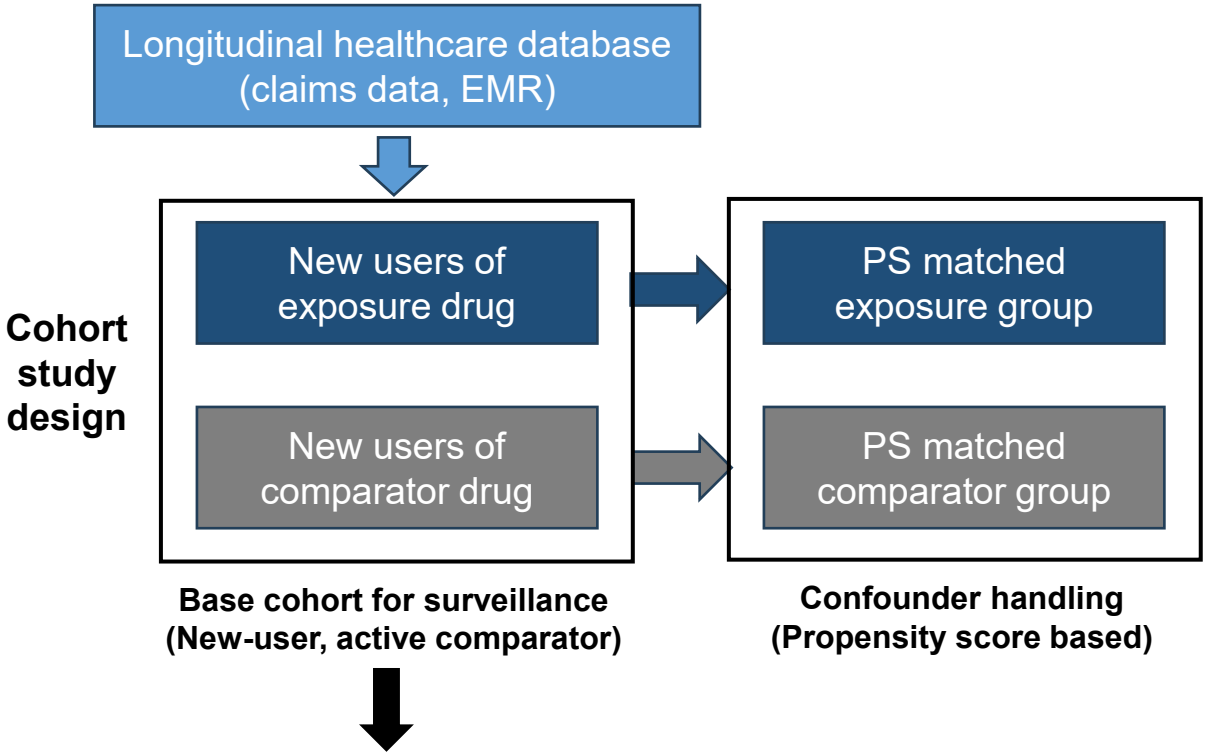
Results: We identified two pairwise propensity score variable ratio matched cohorts of 44,733 canagliflozin vs. 99,458 DPP4 initiators, and 55,974 canagliflozin vs. 74,727 GLP1 initiators. When we screened inpatient and emergency room diagnoses, diabetic ketoacidosis was the only severe adverse event associated with canagliflozin initiation with $p < .05$ in both cohorts. When outpatient diagnoses were also considered, signals for female and male genital infections emerged in both cohorts ($p < .05$).

Conclusions and relevance: In a large population-based study, we identified known but no other adverse events associated with canagliflozin, providing reassurance on its safety among adult patients with T2D and suggesting the tree-based scan statistic method is a useful post-marketing safety monitoring tool for newly approved medications.

Fralick M, et al. Endocrinol Diab Metab 2021; 4e00237.

Approaches for Pharmacovigilance

(1) TreeScan for PV using RWD



- Patients with type 2 diabetes (T2DM)
- Newly prescribed canagliflozin, DPP4i or GLP-1 RA
- 180-day washout period for defining “new users”

- Baseline covariates for propensity score matching (PSM)
  - Chronic medical conditions
  - Markers of diabetes severity
  - Healthcare utilization
  - Drug utilizations (diabetes and non-diabetes)

- Pairwise PSM (Cana vs DPP4i; Cana vs GLP-1RA)
  - Nearest neighbor matching (caliper 0.05)
  - Variable ratio matching (up to 4 comparators)
  - Standardized difference of 0.1 as balance cut-off

Fralick M, et al. Endocrinol Diab Metab 2021; 4e00237.

Approaches for Pharmacovigilance

(1) TreeScan for PV using RWD

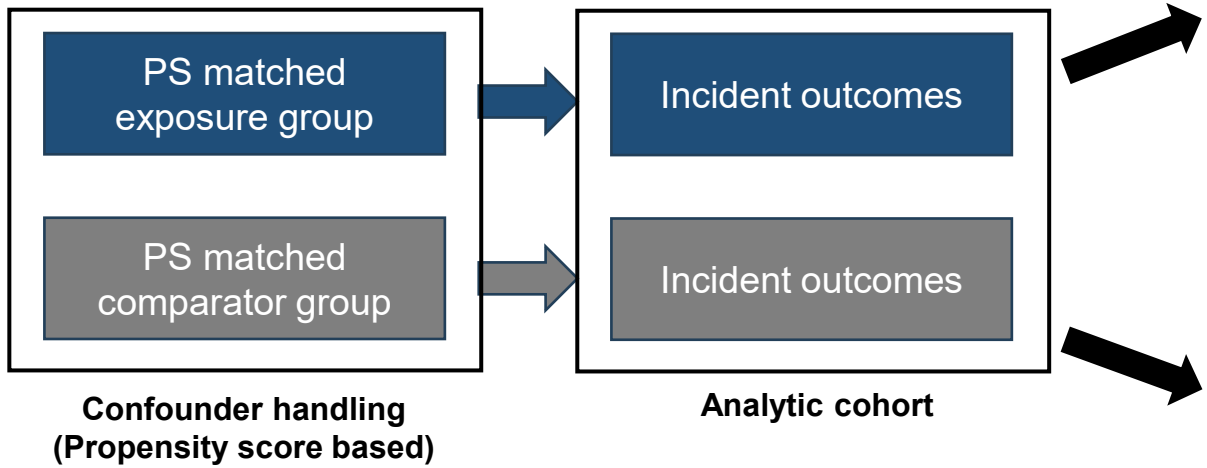
Outcome tree (i.e., ICD-10):
Modify (i.e., pruning) tree structure
at investigator's discretion

- **Five-level Hierarchical tree of potential outcomes (ICD-9 codes)**
  - **Level 1 (Broadest):** Entire disease category (e.g., ICD-9: 001-139 → ICD-10: A00-B99)
    - **Level 2:** subgroups of disease/injury (e.g., ICD-9: 130-136)
      - **Level 3:** ICD-9 codes without a decimal value (e.g., ICD-9 010 [Primary tuberculosis] → ICD-10: A15)
        - **Level 4:** ICD-9 codes with one decimal value (e.g., ICD-9 010.0 → ICD-10: A15.9)
          - **Level 5:** ICD-9 codes with 2+ decimal value

Fralick M, et al. Endocrinol Diab Metab 2021; 4e00237.

Approaches for Pharmacovigilance

(1) TreeScan for PV using RWD



- Incident cases ascertainment using the 180-day period
- Prevalent cases determined at **"level 2"**
- In case of more than one incident outcomes on the same day…

  → select one with lower frequencies

- Incident outcomes assessed at level 2 – 5
- Level 1 not considered due to broad nature of the categories

Fralick M, et al. Endocrinol Diab Metab 2021; 4e00237.

Approaches for Pharmacovigilance

TABLE 2 Signals for potential adverse events based on inpatient or emergency department diagnoses among canagliflozin initiators vs. initiators of other diabetes drugs in two pairwise propensity score-matched cohorts

		Canagliflozin vs. DPP4i						Canagliflozin vs. GLP-1RA					
Potential adverse event (ICD-9 code)[a]	Treelevel	N events Canagliflozin	N events DPP4i	RR	RD[b]	LLR	*p*	N events Canagliflozin	N events GLP-1RA	RR	RD[b]	LLR	*p*
Diabetes with ketoacidosis (250.1)	4	66	60	2.2	2.4	8.0	.043	92	53	2.3	3.1	13.0	.0006
Diabetes, type 2 with ketoacidosis (250.12)	5	41	32	2.6	1.7	6.6	.230	56	30	2.4	2.0	8.4	.032

Abbreviations: DPP-4i, dipeptidyl peptidase 4 inhibitors; GLP-1RA, glucagon-like peptide-1 receptor agonists; LLR, log-likelihood ratio; *p*, *p*-value; RD, rate difference; RR, rate ratio.

[a]Based on inpatient or emergency department diagnoses (any position).

[b]Per 1,000 person years.

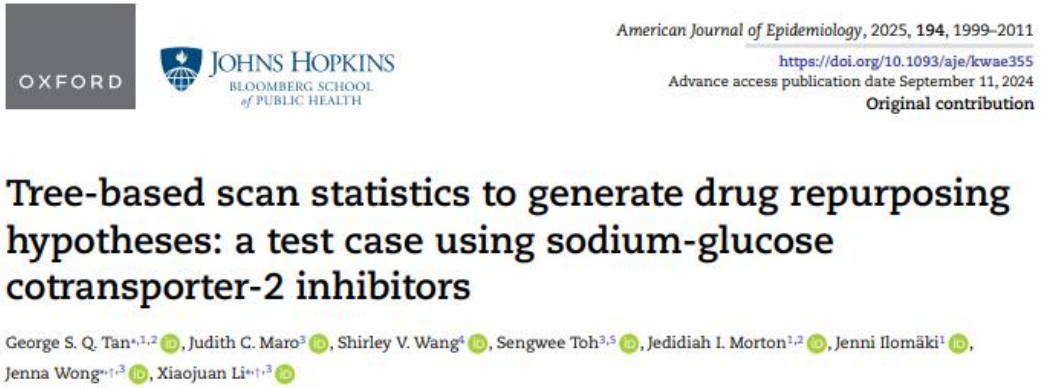
TABLE 3 Signals for potential adverse events based on any diagnoses among canagliflozin initiators vs. initiators of other diabetes drugs in two pairwise propensity score-matched cohorts

		Canagliflozin vs. DPP4i						Canagliflozin vs. GLP-1RA					
Potential adverse event (ICD-9 code)[a]	Tree level	N events Canagliflozin	N events DPP4i	RR	RD[b]	LLR	*p*	N events Canagliflozin	N events GLP-1RA	RR	RD[b]	LLR	*p*
Mycoses (110–118)	2	1,861	2,741	1.4	33.4	44.7	.0001	2,217	2,041	1.4	38.6	63.9	.0001
Candidiasis (112.xx)	3	872	635	2.7	37.7	137.8	.0001	1,047	525	2.6	38.3	172.9	.0001
Candidiasis of vulva and vagina (112.1)	4	498	254	3.9	25.2	137.8	.0001	606	232	3.4	25.5	143.1	.0001
Candidiasis of other urogenital sites (112.2)	4	61	34	3.6	3.0	16.8	.0001	60	29	2.7	2.2	8.4	.0681
Candidiasis of unspecified site (112.9)	4	163	86	3.8	8.2	40.8	.0001	198	86	3.0	7.9	41.2	.0001
Balanoposthitis (607.1)	4	88	70	2.5	3.6	15.8	.0001	83	35	3.1	3.3	14.4	.0002
Inflammatory disease of female pelvic organs (614–616)	2	640	544	2.4	25.0	81.8	.0001	785	489	2.1	24.4	106.1	.0001
Inflammatory disease of cervix vagina and vulva (616.x)	3	606	492	2.5	24.5	87.2	.0001	751	422	2.3	25.4	119.6	.0001
Vaginitis and vulvovaginitis (616.1x)	4	519	363	2.9	23.0	98.8	.0001	661	308	2.8	25.3	135.2	.0001

Fralick M, et al. Endocrinol Diab Metab 2021; 4e00237.

Approaches for Pharmacovigilance

(2) Beyond PV – TreeScan as a tool for drug repurposing



OXFORD

JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

American Journal of Epidemiology, 2025, 194, 1999–2011
https://doi.org/10.1093/aje/kwae355
Advance access publication date September 11, 2024
Original contribution

Tree-based scan statistics to generate drug repurposing hypotheses: a test case using sodium-glucose cotransporter-2 inhibitors

George S. Q. Tan[*,1,2], Judith C. Maro[3], Shirley V. Wang[4], Sengwee Toh[3,5], Jedidiah I. Morton[1,2], Jenni Ilomäki[1], Jenna Wong[*,†,3], Xiaojuan Li[*,†,3]

- **Proof-of-concept of TreeScan for identifying "repurposing" signals**
  - *"…to demonstrate how TBSS can be used to generate new drug repurposing hypotheses from RWD."*
  - *"…**inverse association** between drug exposure and a health outcome identified by the scan statistics may suggest a potential repurposing signal relating to the outcome."*

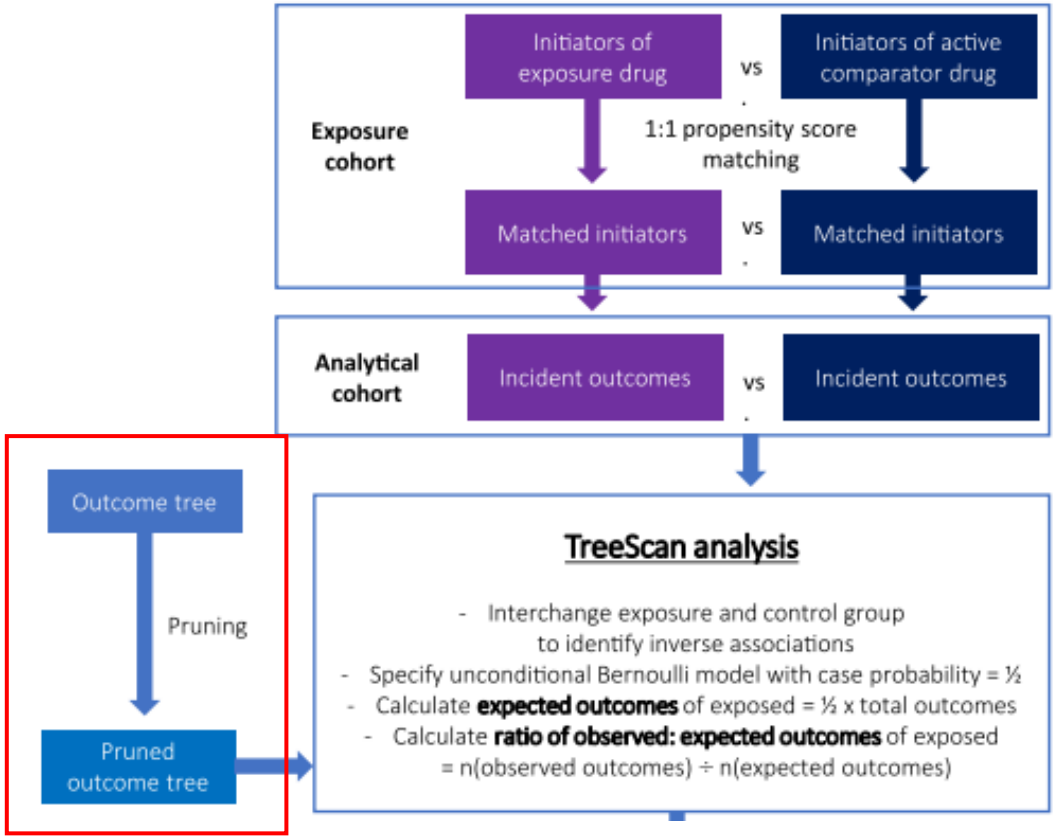
Tan GSQ, et al. Am J Epidemiol 2025; 194:1999-2011.

Approaches for Pharmacovigilance

(2) Beyond PV – TreeScan as a tool for drug repurposing

- Pruning ICD-10 code branches with diagnoses that are less plausible as a “repurposing” candidates

Tree	Description
Original pruned tree	All ICD-10-CM codes Exc. • V00-Y99 (External causes of morbidity and mortality) • Z00-Z99 (Factors influencing health status and contact with health services) • P00-P96 (Conditions originating in the perinatal period) • O00-O9A (Pregnancy, childbirth and the puerperium)

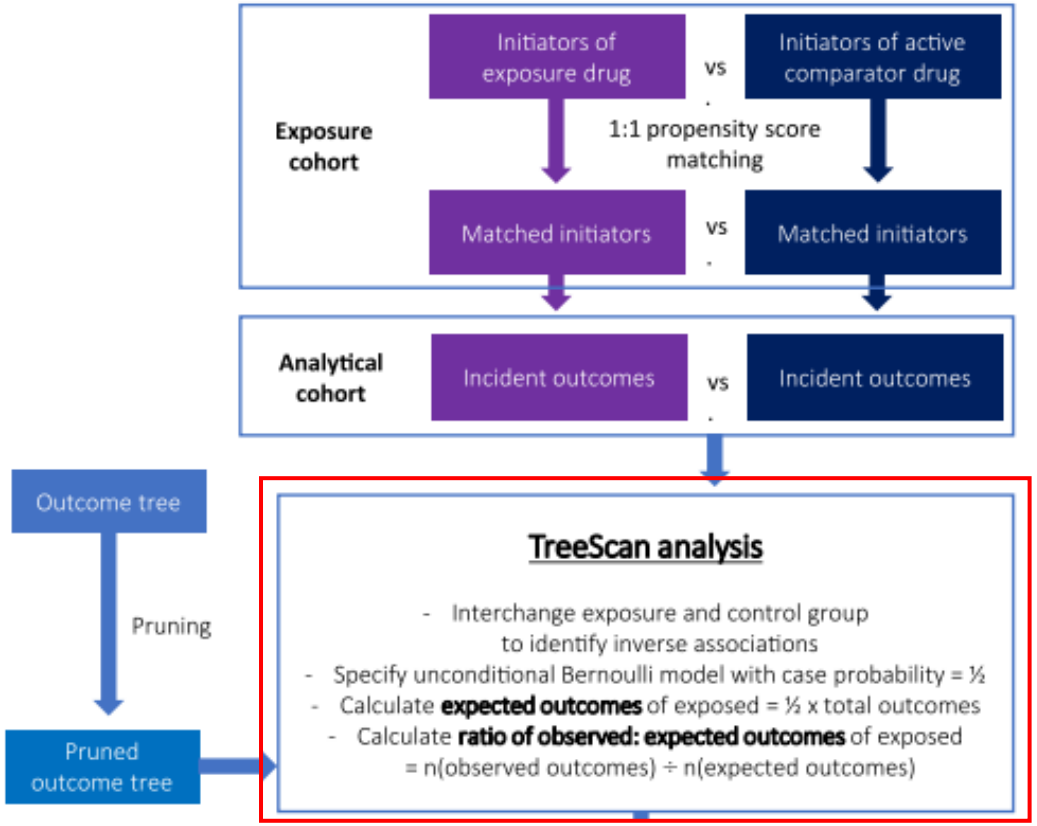


Tan GSQ, et al. Am J Epidemiol 2025; 194:1999-2011.

Approaches for Pharmacovigilance

(2) Beyond PV – TreeScan as a tool for drug repurposing

- TreeScan operates in a way that it identifies clusters of signals with higher-than-expected (i.e., positive association between exposure & outcome)
- For the purpose of identifying the “repurposing” signals… (i.e., Lower-than-expected; inverse association)
- ***“…implemented by Interchanging the exposure and comparator groups…”***



Tan GSQ, et al. Am J Epidemiol 2025; 194:1999-2011.

Approaches for Pharmacovigilance

(2) Beyond PV – TreeScan as a tool for drug repurposing

Table 2. Tree-based scan statistics for associations between SGLT2i vs DPP4i and outcomes (only for associations with $P < .1$).[a]

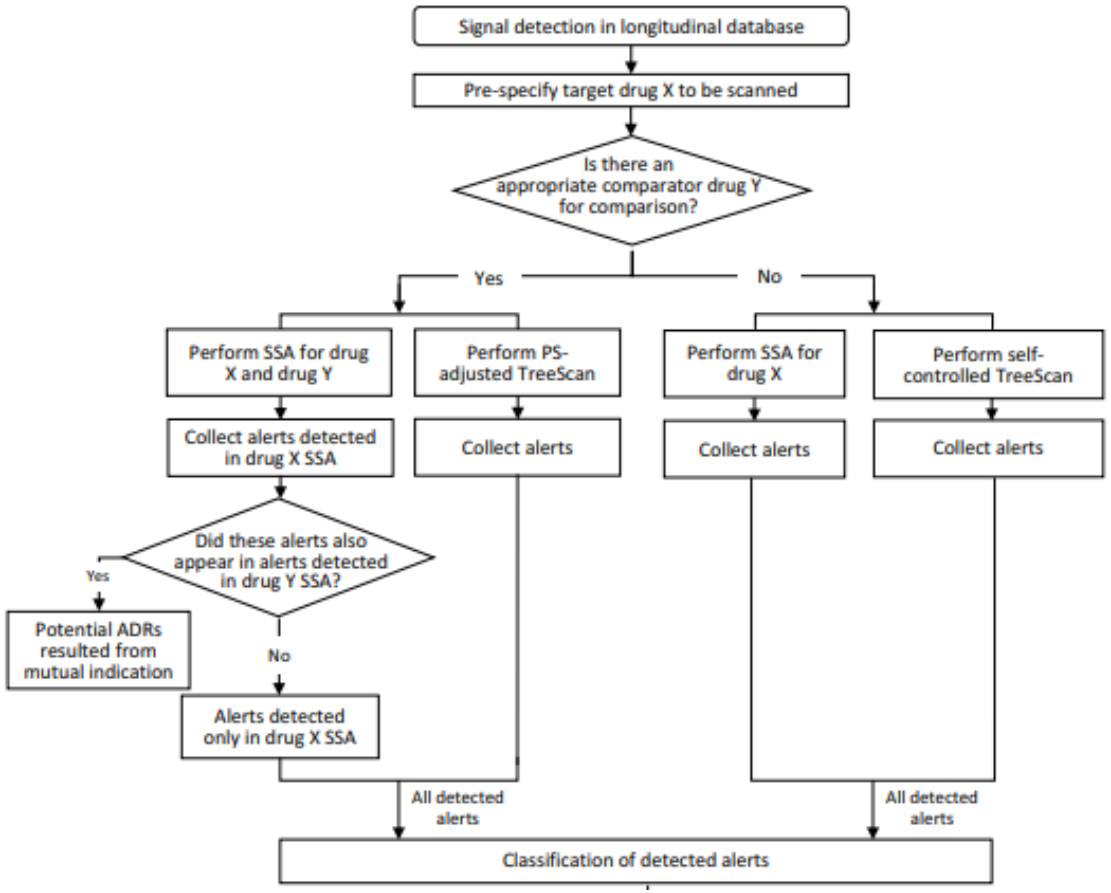
Node	Description	Total outcomes	Observed outcomes (SGLT2i)	Expected putcomes[b] (SGLT2i)	Observed: expected outcomes (SGLT2i)	Log likelihood ratio (scan statistic)	P value
Associations with statistical alert ($P \leq .01$)							
N18	Chronic kidney disease (CKD)	1470	594	735	0.81	27.21738	0.0001
N18.3	Chronic kidney disease, stage 3 (moderate)	722	270	361	0.75	23.18839	0.0001
D64	Other anemias	1415	581	707.5	0.82	22.7401	0.0001
D64.9	Anemia, unspecified	1356	556	678	0.82	22.07283	0.0001
R60.0	Localized edema	941	371	470.5	0.79	21.20169	0.0001
R60	Edema, not elsewhere classified	1564	656	782	0.84	20.39056	0.0001
I12	Hypertensive chronic kidney disease	833	333	416.5	0.8	16.85408	0.0001
E83.4	Disorders of magnesium metabolism	307	106	153.5	0.69	14.94276	0.0002
I12.9	Hypertensive chronic kidney disease with stage 1-4 or unspecified chronic kidney disease	802	324	401	0.81	14.87777	0.0002
⋮							
Nonsignificant associations ($P > 0.01$)[a]							
I50	Heart failure	846	357	423	0.84	10.34007	0.0167
R14.2	Eructation	42	7	21	0.33	10.18861	0.018
N25	Disorders resulting from impaired renal tubular function	95	26	47.5	0.55	10.09454	0.0198
D63	Anemia in chronic diseases classified elsewhere	217	76	108.5	0.7	9.886091	0.0236
N25.81	Secondary hyperparathyroidism of renal origin	70	17	35	0.49	9.715734	0.0266
R09	Other symptoms and signs involving the circulatory and respiratory system	1967	887	983.5	0.9	9.483731	0.0372
N13.2	Hydronephrosis with renal and ureteral calculous obstruction	127	40	63.5	0.63	8.907116	0.0713
N19	Unspecified kidney failure	119	37	59.5	0.62	8.72376	0.0784

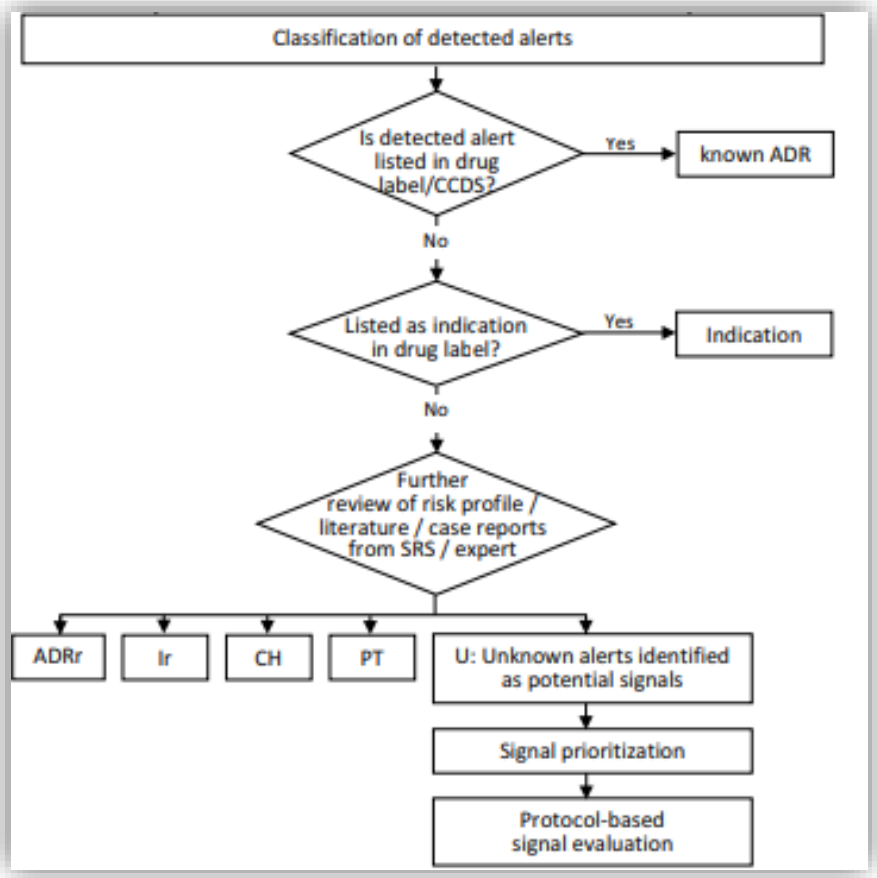
[a]Refer to table in Table S6 for associations with $P > .1$.
[b]The number of expected outcomes at a node was calculated as half of the total number of outcomes from both exposure and comparator group.

SQ, et al. Am J Epidemiol 2025; 194:1999-2011.

Approaches for Pharmacovigilance

(3) Steps to be undertaken after signal detection





Clinical Epidemiology
Dovepress
open access to scientific and medical research
Open Access Full Text Article
ORIGINAL RESEARCH

A New Drug Safety Signal Detection and Triage System Integrating Sequence Symmetry Analysis and Tree-Based Scan Statistics with Longitudinal Data

Miyuki Hsing-Chun Hsieh[1,2,*], Hsun-Yin Liang[2,*], Chih-Ying Tsai[2], Yu-Ting Tseng[2], Pi-Hui Chao[2], Wei-I Huang[2], Wen-Wen Chen[2], Swu-Jane Lin[3], Edward Chia-Cheng Lai[1]

Hsieh MH, et al. Clinical Epidemiology 2023; 15:91-107.

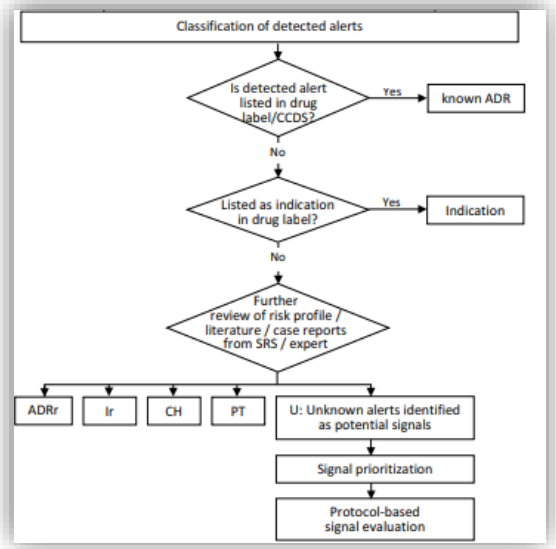
Approaches for Pharmacovigilance

- In summary…

1) Pharmacoepidemiologic designs (e.g., case-only study, cohort study)
    - More suitable for signal evaluation or hypothesis testing
2) Sequence symmetry analysis (SSA)
    - Rely on an assumption of “event onset influences drug initiation”
3) Supervised machine learning (i.e., Gradient boosting machine, random forest)
    - No established guides on constructing input dataset (positive and negative label) for training and validation
4) Tree-based scan statistic
    - Ideal candidate for outcome surveillance in RWD
    - Combined with pharmacoepi approaches (i.e., new-user, active comparator design, PS-based methods)

Multiple data sources

- **Federated network approach and multi-database study using TreeScan**
  - Select drug(s) of interest and comparator drug(s)
  - Apply TreeScan to establish outcome candidates from each database
  - Triage system (Hsieh, et al.) for shortlisting outcome candidates
  - Protocol development and execution in each database



- **Considerations and discussion points**
  - Common hierarchical outcome tree that can be implemented across different databases
  - How to handle the heterogeneity in detected signals across databases
  - For each outcome candidates, further evaluation needed for feasibility (i.e., outcome definition)
  - Covariates for PS that can be used generically for all outcomes

Thank you for your attention

napa928@hotmail.com