

BACKGROUND

- Tree-based scan statistics allow simultaneous surveillance across thousands of hierarchical diagnosis codes and code groupings. The method inherently adjusts for multiple comparisons to enable a general surveillance safety net without a high false positive rate.
- However, this adjustment raises the numerical threshold for alerting and thus may miss clinically important safety signals.
- To focus surveillance efforts in adverse event reporting systems, regulatory agencies often use lists of 'Designated Medical Events' (DMEs) that represent outcomes of particular interest. A DME-based tree contains fewer outcomes and therefore decreases the sample size required for multiple testing. However, these lists have not been systematically adapted for active drug safety surveillance in claims-based systems.

OBJECTIVES

- We aimed to construct targeted outcome trees based on regulatory DME lists and compare their performance to routinely used untargeted trees by reanalyzing a prior signal identification analysis of erenumab in the U.S. Sentinel System.¹

METHODS

- We compiled the DME lists of the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), along with additional outcomes of interest (e.g., stroke, myocardial infarction).
- MedDRA terms were mapped to ICD-10-CM codes using standard crosswalks² with additional curation. We constructed three corresponding targeted trees, two were restricted to each agency's DME list and one combined them (DME+) (Figure 1).
- We reanalyzed a cohort used in a prior Sentinel self-controlled signal identification analysis (May 2018–Jun 2022) of incident adult erenumab users and employed identical tree temporal scan statistics to detect outcomes 1–183 days post-index in the inpatient/emergency department setting (primary) and all care settings (sensitivity). Statistical alert thresholds (log likelihood ratios (LLRs) with $p \leq 0.05$) were compared between the original untargeted tree and three targeted trees.

RESULTS

- A total of 68,345 nodes were evaluated in the untargeted tree, compared with 1,984, 751, and 3,082 nodes in the FDA, EMA, and DME+ trees, respectively.
- The overall distribution of calculated LLRs from the Monte-Carlo simulation is substantially lower for the EMA, FDA and DME+ trees than for the untargeted tree (Figure 2).
- The alert threshold in the primary analysis corresponded to lower test statistics by 20% (DME+), 26% (FDA), and 29% (EMA). These findings remained robust in the sensitivity analysis, which showed comparable reductions of 22% (DME+), 27% (FDA), and 28% (EMA) (Table 1).

CONCLUSION

- Targeted trees may serve as a valuable complementary screening tool to optimize the balance between coverage and statistical efficiency in safety signal detection.
- Performance also varied across DME tree versions despite relatively small differences in outcome selection, highlighting opportunities to further refine targeted trees for longitudinal claims-based active surveillance.

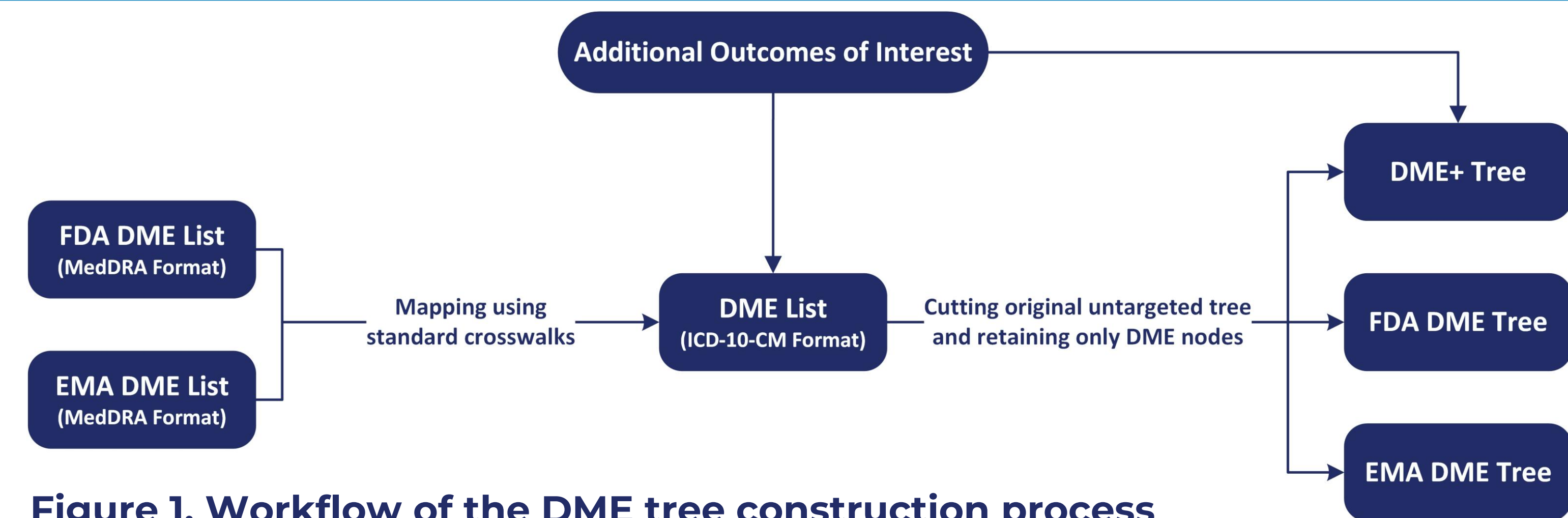


Figure 1. Workflow of the DME tree construction process

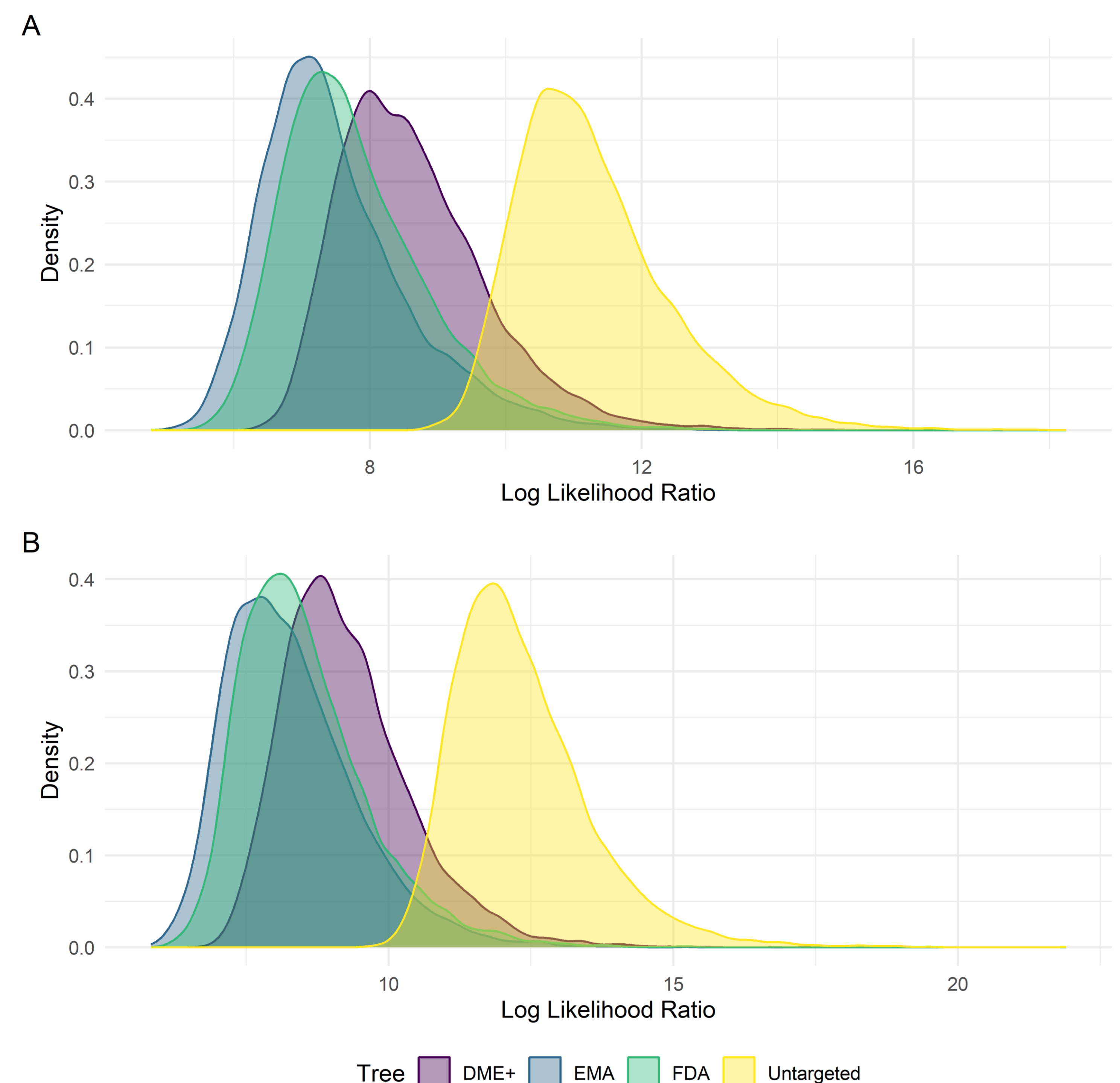


Figure 2. Density plots of log likelihood ratios generated in the primary (A) and sensitivity (B) analysis with the untargeted and three DME trees

Table 1. Comparison of scan statistics between untargeted tree and DME trees

Metrics	Untargeted Tree		FDA Tree		EMA Tree		DME+ Tree	
	N		N	Δ vs Untargeted	N	Δ vs Untargeted	N	Δ vs Untargeted
Primary analysis								
Total cases	83585	2690		-96.78%	2216	-97.35%	4823	-94.23%
Number of nodes	152036	3855		-97.46%	1123	-99.26%	5535	-96.36%
Number of nodes evaluated	68345	1984		-97.10%	751	-98.90%	3082	-95.49%
Number of most likely cuts	7	6		-14.29%	14	100.00%	10	42.86%
Number of significant alerts	0	0		0 (abs)	0	0 (abs)	1	+1 (abs)
LLR at 0.05 sig level	13.40	9.86		-26.46%	9.57	-28.59%	10.67	-20.39%
Sensitivity analysis								
Total cases	485355	8689		-98.21%	7080	-98.54%	13843	-97.15%
Number of nodes	152036	3855		-97.46%	1123	-99.26%	5535	-96.36%
Number of nodes evaluated	68345	1984		-97.10%	751	-98.90%	3082	-95.49%
Number of most likely cuts	30	15		-50.00%	9	-70.00%	16	-46.67%
Number of significant alerts	7	0		-100.00%	0	-100.00%	0	-100.00%
LLR at 0.05 sig level	14.55	10.67		-26.65%	10.40	-28.49%	11.37	-21.84%

LIMITATIONS

- The inclusion of additional clinical outcomes in the DME+ tree carries inherent subjectivity regarding event selection.
- A DME targeted tree may miss signals outside the curated tree.
- Evaluating this framework using a single empirical case study (erenumab) introduces potential randomness, which may limit the generalizability of our conclusions.

ACKNOWLEDGEMENTS/DISCLOSURES

- The contents are those of the authors, and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS or the U.S. Government.
- T.T. and J.C.M. are employees of HPHCI, an organization which conducts work for government and private organizations, including pharmaceutical companies.
- Many thanks are due to the Sentinel Data Partners who provided data used in this analysis.

¹Outcome Monitoring Following Erenumab Use: A Signal Identification Analysis. Sentinel Initiative; 2023. <https://www.sentinelinitiative.org/studies/drugs/individual-drug-analyses/outcome-monitoring-following-erenumab-use-signal>

²The ICD-10 to MedDRA mapping (3rd release) (<https://www.meddra.org/mapping>), Sentinel's Code Lookup Tool, and the Clinical Classifications Software Refined (CCSR) for ICD-10-CM Diagnoses (v2025.1) (https://hcup-us.ahrq.gov/toolssoftware/ccsr/ccsr_refined.jsp)