AI-Assisted Prioritization of VAERS Signal Candidates

A retrospective study of next-quarter threshold crossings for expert review

Natanael Alpay | nalpay@uci.edu | University of California, Irvine

UC Irvine
School of Physical Sciences

47.4%
fewer reviews at 50% capture

806 vs 1,531
gradient boosting vs threshold proximity

32 / 100
top-ranked pairs crossed next quarter

01 STUDY DESIGN

Question and data

Expert review capacity is limited. Can recent reporting patterns identify which currently non-strong vaccine–symptom pairs will cross a prespecified rule next quarter?

U.S. VAERS, 2022Q1–2026Q2.

- Primary report sequences (`ORDER=1`); tables linked by report ID and `ORDER`.
- Unit: **vaccine type** × **MedDRA term** × **receipt quarter**.
- **448,942** eligible reports and **200,220** candidate-quarter rows.

Receipt dates define the time axis. The incomplete 2026Q3 data in the supplied snapshot are excluded.

Outcome and chronological split

Prespecified strong-signal proxy

$$a \geq 3, \qquad \mathrm{PRR} \geq 2, \qquad \mathrm{LCB}_{95}(\mathrm{PRR}) > 1.$$

a is the number of reports containing both the vaccine type and term. PRR is the proportional reporting ratio; LCB is its lower confidence bound. Ratio calculations use a 0.5 continuity correction.

- At quarter t: observed and below the proxy rule.
- Outcome: meets the rule in **exact quarter** $t+1$.
- Absent next quarter: retained as a negative outcome.

Final retrospective temporal test

Fit	Candidate quarters through 2025Q3
Calibrate	2025Q4 candidates
Rank	2026Q1 candidates
Evaluate	Strong-signal status in 2026Q2

Predictions are made after the candidate quarter ends.

Models and evaluation

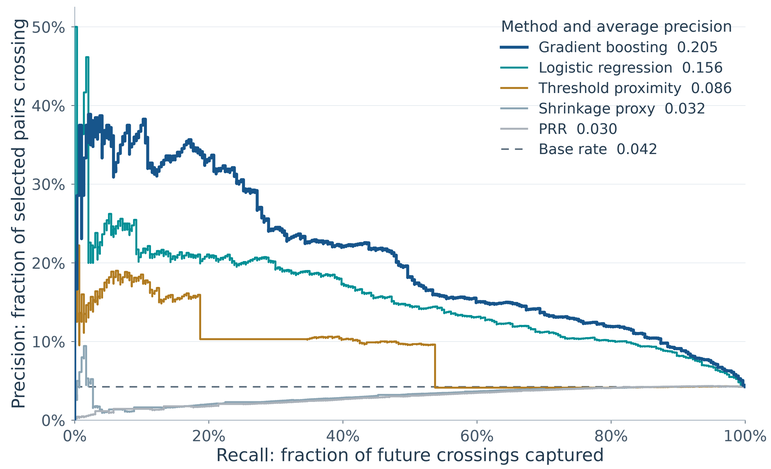
- **31 features**: counts, classical scores, report context, and exact one- and two-quarter history.
- Class-weighted logistic regression and histogram gradient boosting; sigmoid calibration on later data.
- **Training-only age imputation**, learned separately within each fitting window.
- Baselines: PRR, a Gamma–Poisson-style shrinkage proxy, and proximity to the three rule thresholds.
- Average precision (AP), top-100 precision, ranked-pair counts at fixed capture, and nine rolling test quarters.

The shrinkage score is a simple proxy, not the full MGPS/EBGM method.

02 MAIN RESULTS

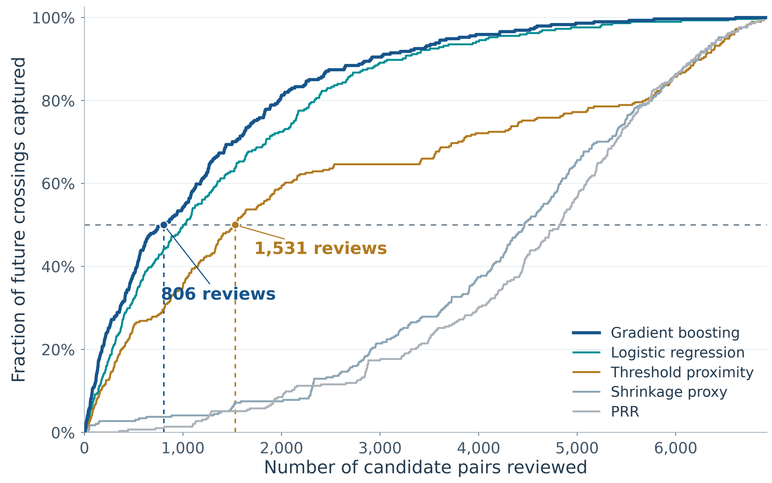
Ranking quality on the final test

6,926 candidates | 294 crossings | 4.24% base rate



Gradient boosting AP: **0.205**. Descriptive 95% row-bootstrap interval: **0.169–0.249** (200 resamples).

Fewer reviews for the same capture



Target captured	**50%**	**80%**
Gradient boosting	806	1,980
Threshold proximity	1,531	5,696
Reduction in reviews	**47.4%**	**65.2%**

725 fewer pairs need review to capture half of the eligible future crossings.

Counts assume equal review effort per pair. They do not measure staff hours, cost savings, or clinical benefit.

03 INTERPRETATION

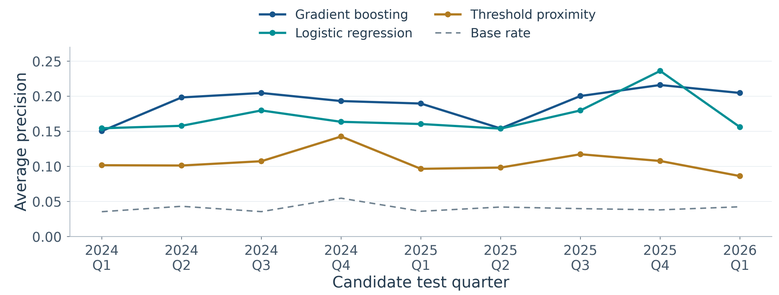
Most crossings are returning signals

245 / 294
83.3% had been strong in an earlier quarter.

49
had no earlier strong quarter in the available history.

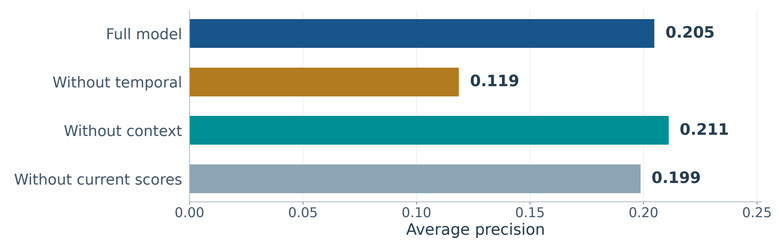
History starts in 2022. Eligibility excludes pairs strong *now*; earlier strong quarters are allowed. The endpoint includes recurrence.

Performance across nine test quarters



Both fitted models exceeded threshold proximity in all nine quarters. Logistic regression led in 2024Q1 and 2025Q4.

Recent history adds predictive value



Removing temporal features lowered AP to **0.119**. Omitting context did not lower AP.

Interpretation and limits

The ranking enriches a queue for expert review. Outcomes are statistical proxy crossings. Passive VAERS reports establish neither incidence nor vaccine causality.

Public-data availability lags limit real-time claims. Pairs share reports, so row-bootstrap intervals are descriptive. Clinical usefulness requires expert validation.

AIPM 2026 Data: U.S. VAERS public files, 2022–2026. Analysis restricted to receipt quarters through 2026Q2.

References: Evans et al. *Pharmacoepidemiol Drug Saf.* 2001;10:483–486. doi:10.1002/pds.677. Bate et al. *Eur J Clin Pharmacol.* 1998;54:315–321. doi:10.1007/s002280050466. VAERS data guidance.

Contact
nalpay@uci.edu
Website / QR