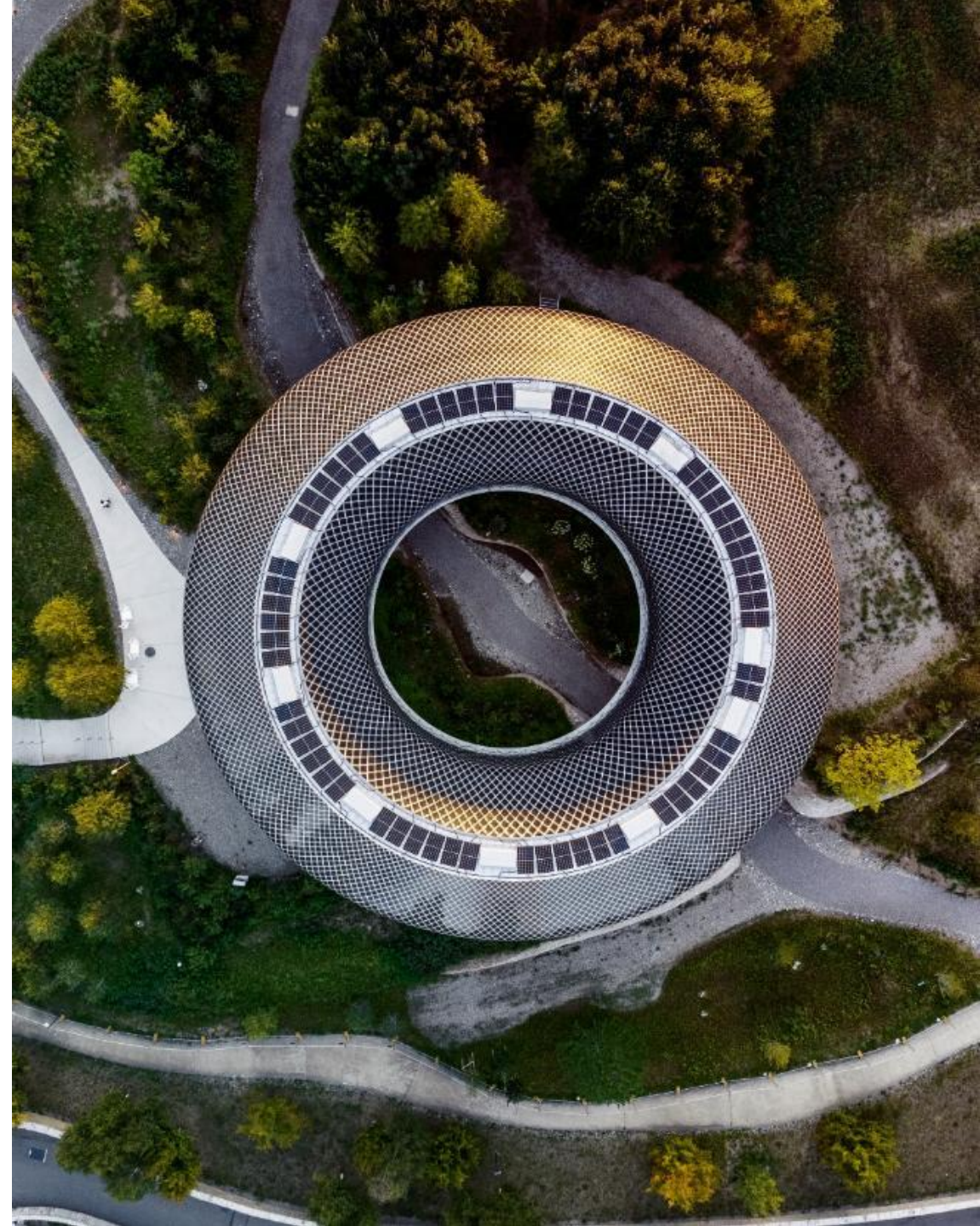


AI for Adverse Event Signal Detection:

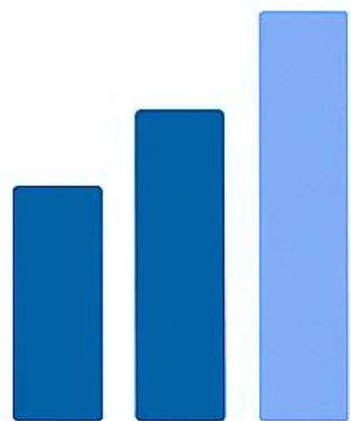
Moving Beyond Traditional Disproportionality Metrics

Phuse APAC - 2026

Monika
Statistical Programmer

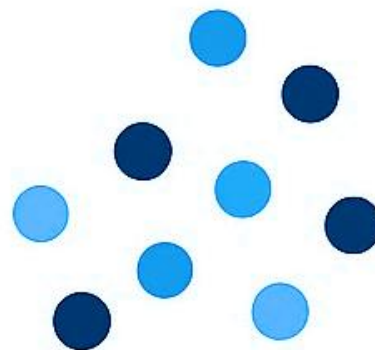
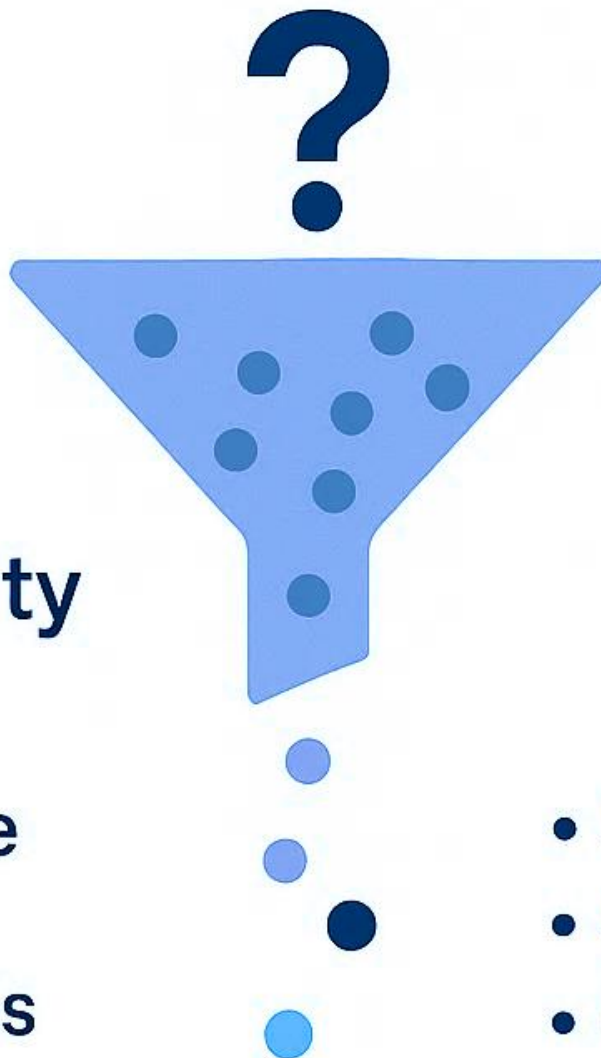


“Why Rethink AE Signal Detection?”



Disproportionality Metrics

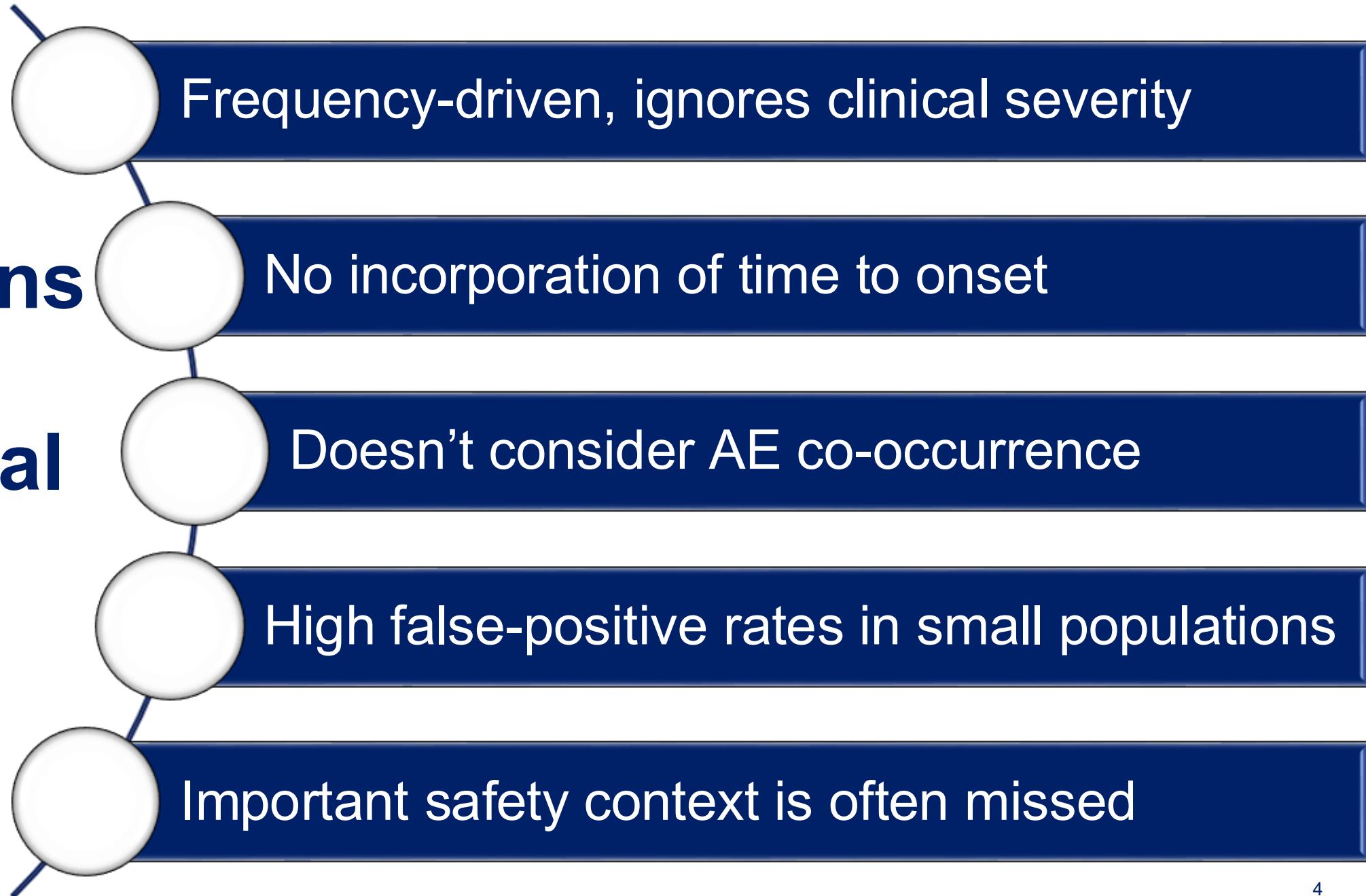
Effective in Large
Spontaneous
Reporting Systems



Phase 2-3 Clinical Trials

- Small Sample Sizes
- Sparse AE Counts
- Complex
Multidimensional
Safety Data

Limitations of Traditional Metrics



PRIMARY OBJECTIVE

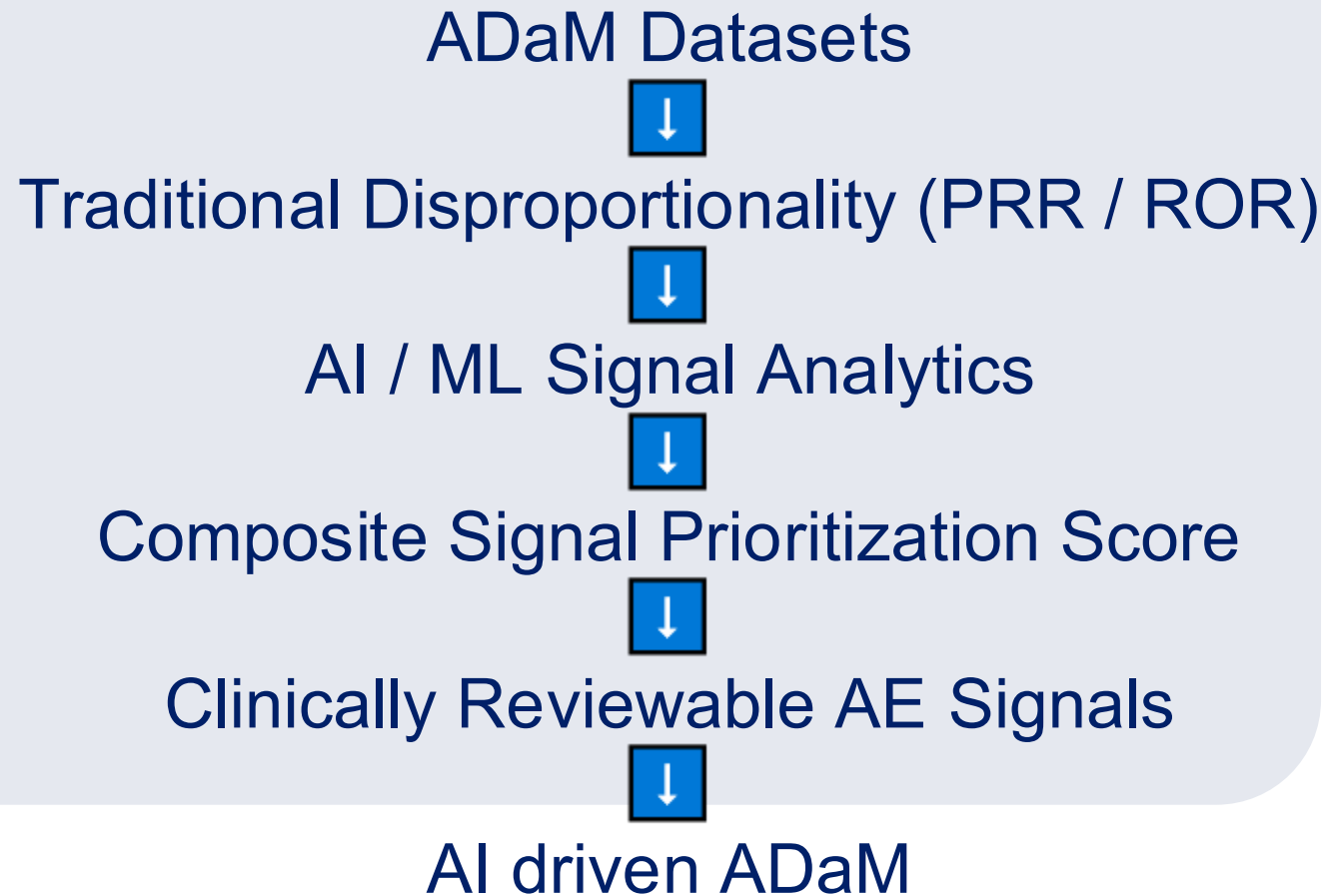
- Demonstrate how **AI-driven methods** can augment traditional AE signal detection
- Leverage **ADaM-standardized datasets**
- Maintain -
 - Interpretability
 - Traceability
 - Regulatory alignment



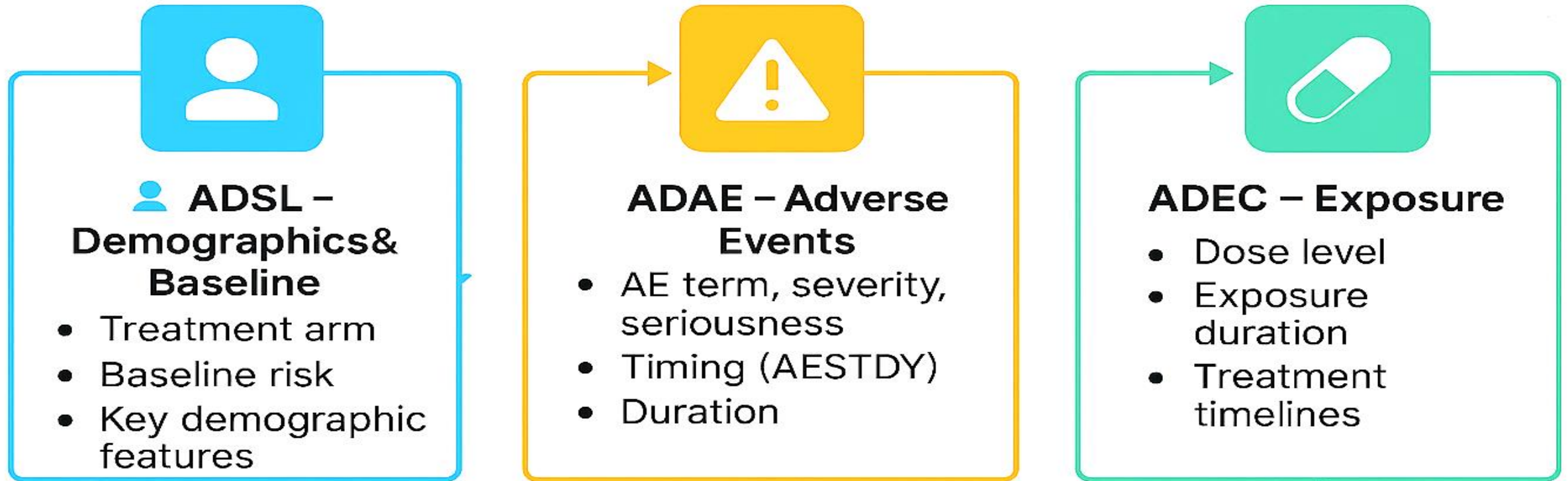
FOCUS

- Phase 2 & 3 clinical trials
- Statistical programmer-driven workflow
- Automated pipeline to detect early signals

Automated AE Detection Pipeline



Where Do the Features Come From?



Important

- All features derived **post-ADaM**
- No SDTM modifications
- Fully auditable and reproducible

Feature Engineering

Event-Level Features

AE severity (MILD / MODERATE / SEVERE)

Serious AE flag

Time-to-onset (AESTDY)

AE duration

Subject-Level Features

Total AE burden

Exposure dose & duration

Demographic risk indicators



Enables multidimensional safety analysis

METHODS USED FOR ANALYSIS

Disproportionality
Analysis

Anomaly
Detection

AE Severity
Clustering

Association
Rule Mining

Baseline: Disproportionality Analysis

Traditional Metrics

- Proportional Reporting Ratio (PRR)
- Reporting Odds Ratio (ROR)

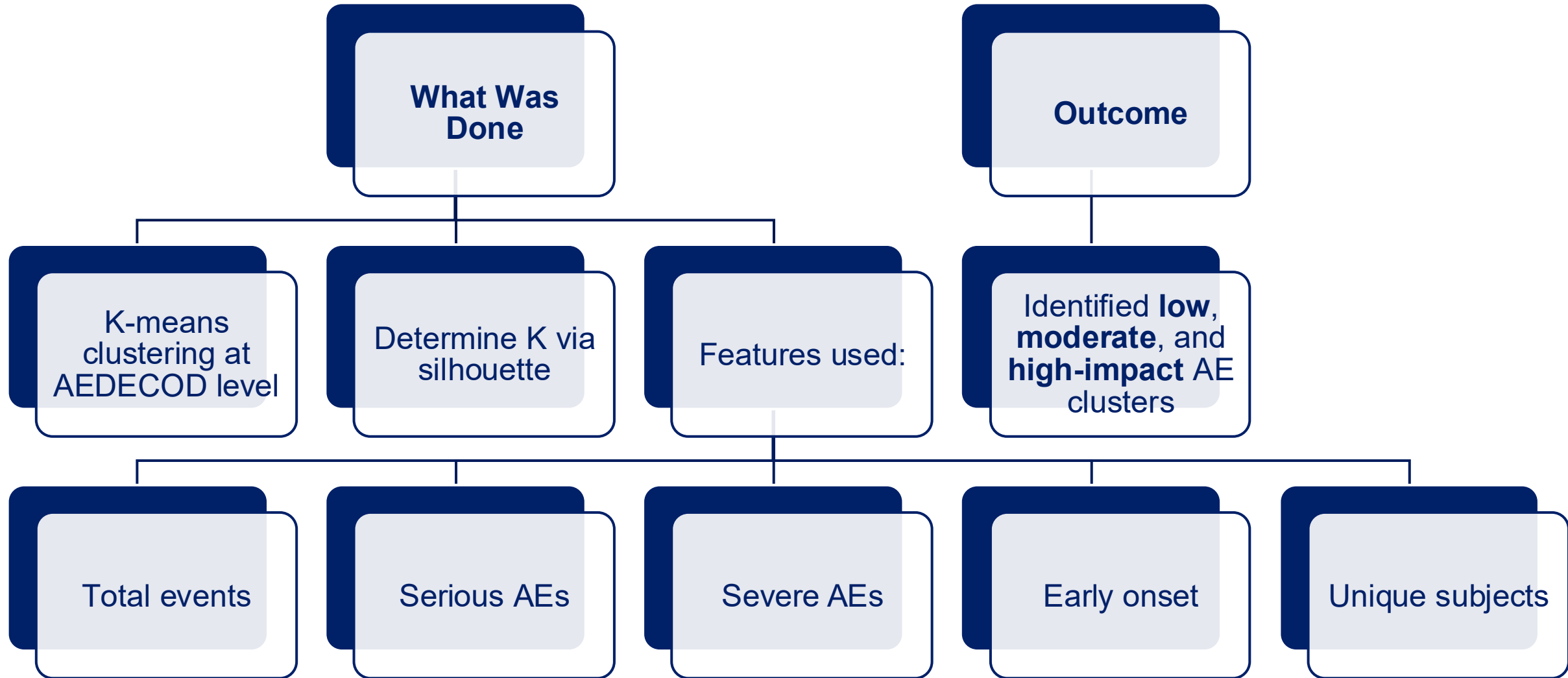
Implementation

- 2×2 contingency tables per AEDECOD
- Continuity correction for sparse counts
- Standard clinical thresholds applied

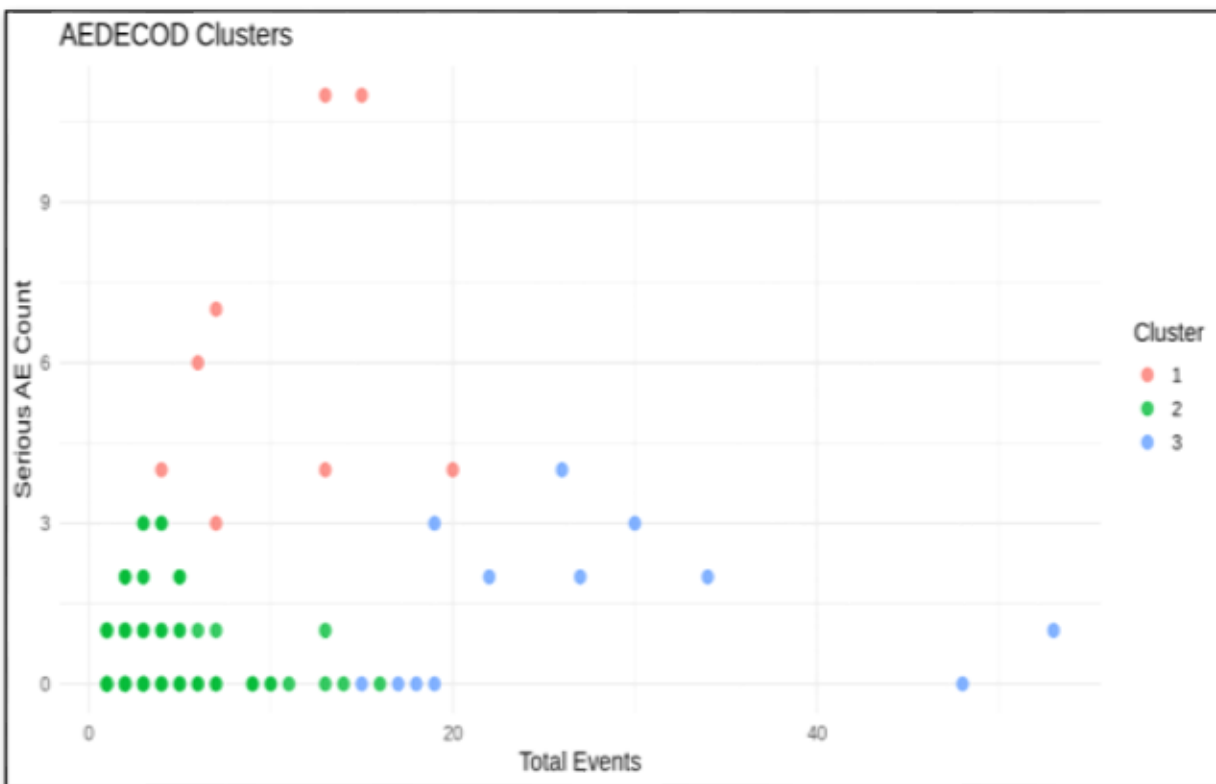
Role in Pipeline

- Serves as **baseline comparator**

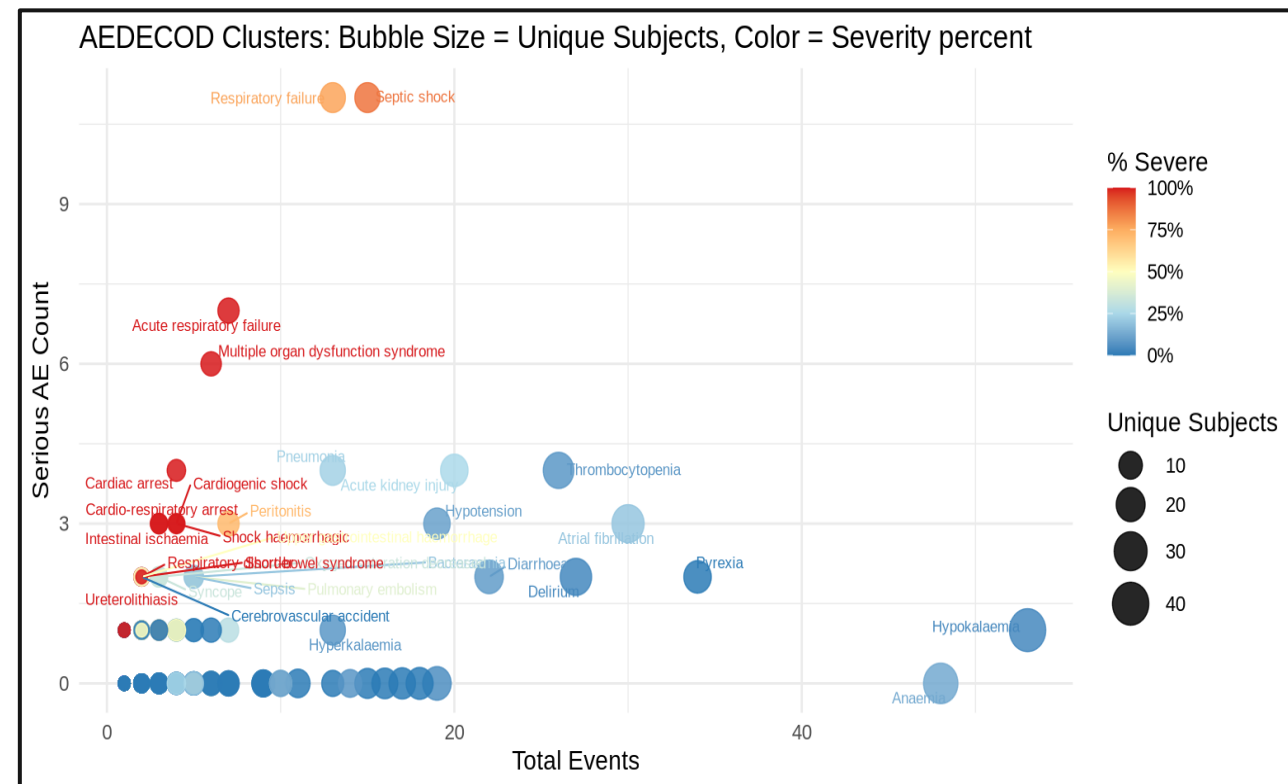
ML Layer 1: AE Severity Clustering



Visualizations based on Clustering Analysis



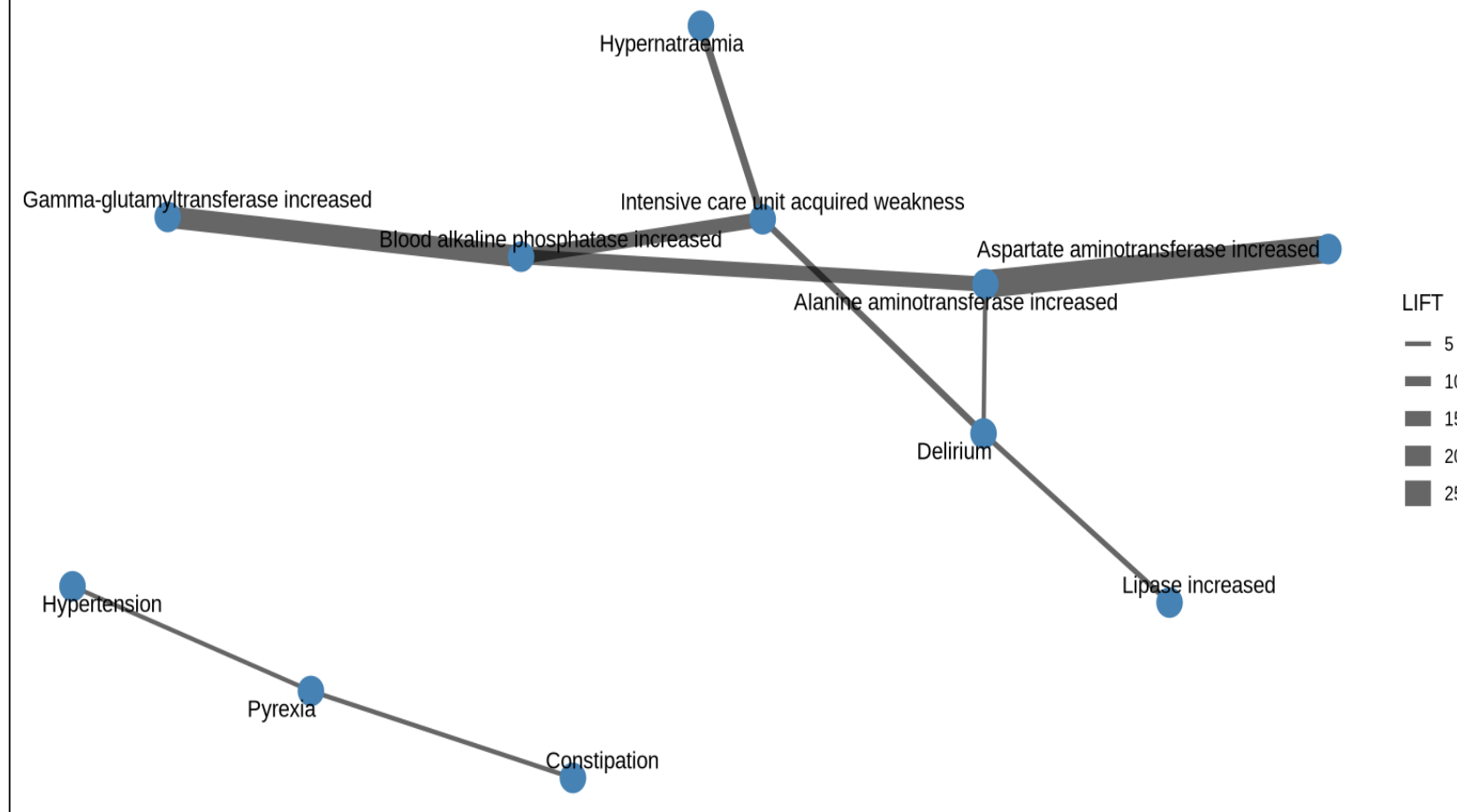
- Three clusters based on AEDECOD identified.
- Improves clinical prioritization



- Separates:
 - High-frequency but low-severity AEs
 - Low-frequency but high-severity AEs
- Highlights AEs with early onset or severe outcomes

ML Layer 2: AE Co-Occurrence Analysis

Top 10 AE Co-Occurrence Pairs by Lift



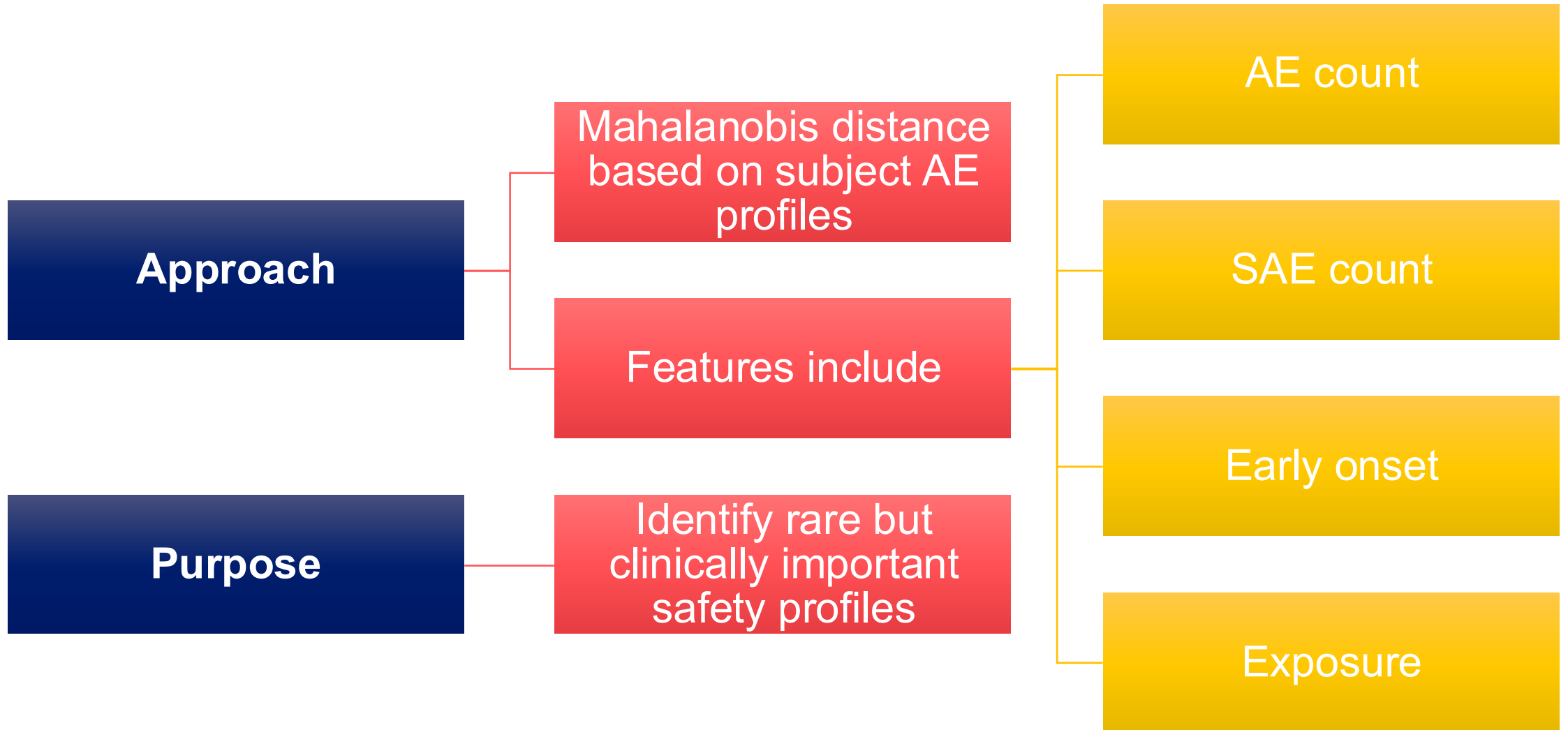
Method

- Association rule mining
- AE “basket” created per subject
- Metrics:
 - Support
 - Confidence
 - Lift

What It Reveals

- AE interaction patterns
- AEs often occur **together**, not independently
- Syndromic or multi-system toxicity signals

ML Layer 3: Subject-Level Anomaly Detection

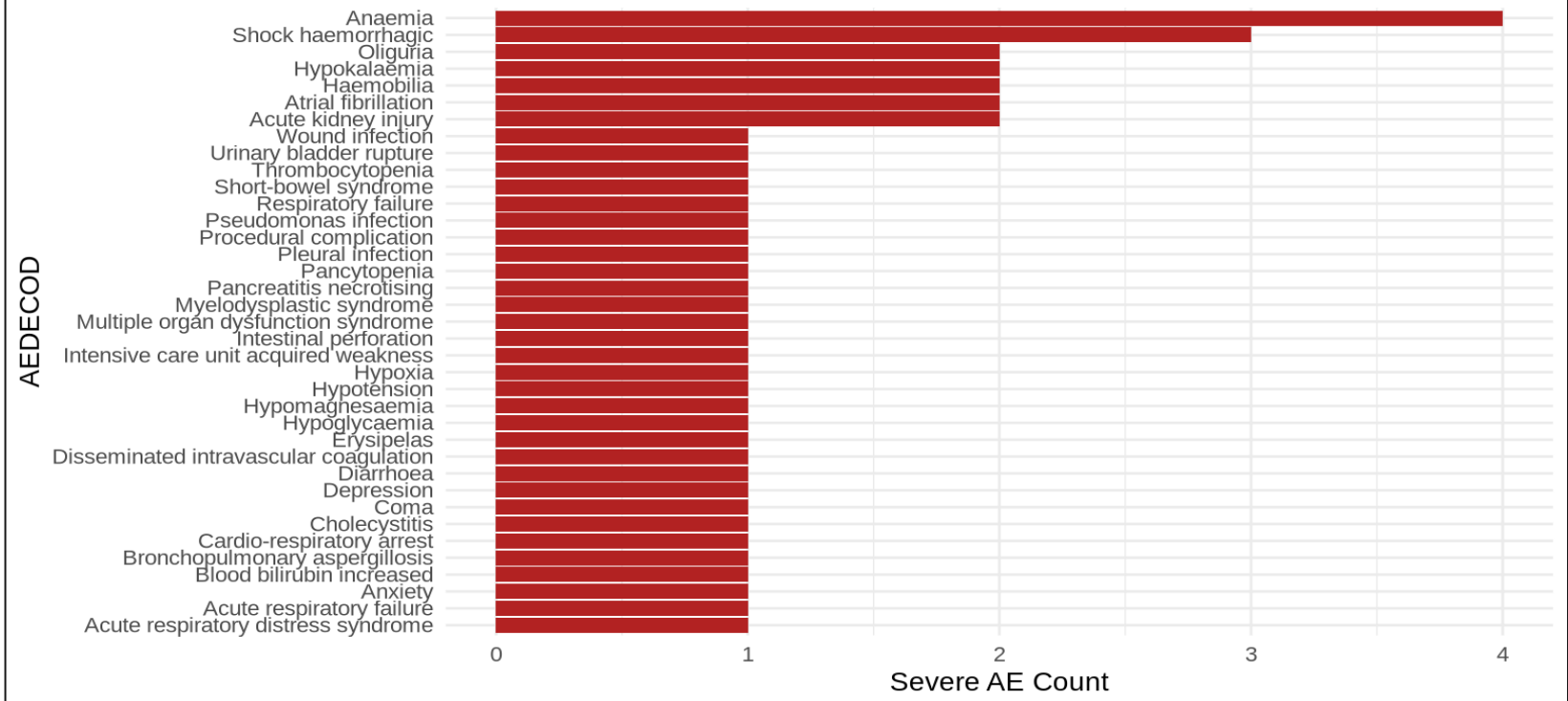


AE_COUNT	SAE_COUNT	SEVERE_CNT	EARLY_CNT	EXPO_DAYS	EXPO_DOSE	ANOM_SCORE	ANOM_FLAG
<int>	<int>	<int>	<int>	<dbl>	<dbl>	<dbl>	<lgl>
1	0	1	1	0	500	183.	TRUE
26	3	14	20	0	200	74.3	TRUE
53	2	3	36	0	200	67.4	TRUE
34	0	0	31	0	200	41.7	TRUE
18	3	7	14	2	200	31.2	TRUE
1	1	1	1	0	80	31.1	TRUE
12	3	1	11	2	200	26.9	TRUE
18	2	8	18	0	200	26.8	TRUE
22	1	5	12	0	200	25.3	TRUE
19	5	6	19	0	200	23.4	TRUE
32	1	1	24	0	200	21.1	TRUE

Why Anomaly Detection Matters

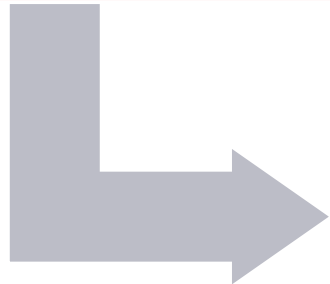
- It Flags:
 - Rare serious events
 - Early-onset toxicities
- Prioritizes **review-worthy subjects**

Top Severe AE Terms in Anomalous Subjects



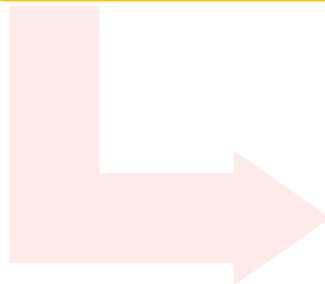
Integrating Results: AI Driven Dataset

**AI-
Enriched
AE Dataset**



**Each AE
enriched
with:**

- PRR / ROR metrics
- Cluster risk category
- Co-occurrence lift
- Subject anomaly flags

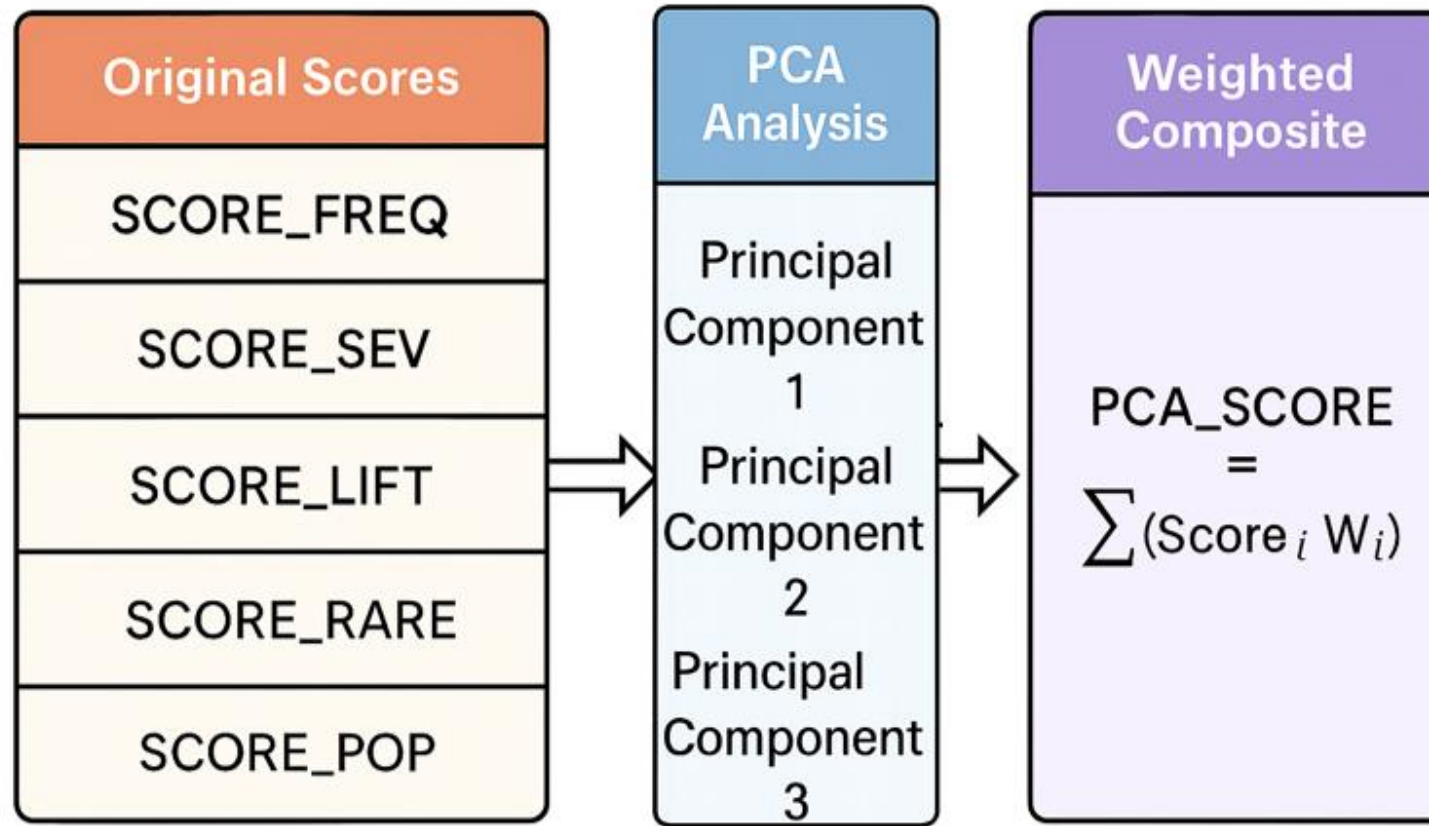


Result

Single analysis-ready
safety dataset

CLUSTER_RISK	PRR_FLAG	ROR_FLAG	ML_LIFT_FLAG	ML_CLUSTER_FLAG	ML_ANOM_FLAG
Moderate	0	0	1	0	0
Moderate	0	0	0	0	0
Moderate	0	0	0	0	0
Moderate	0	0	1	0	0
Moderate	0	0	0	0	0
Moderate	0	0	0	0	0
Low	0	0	1	0	0
Low	0	0	1	0	0
Moderate	0	0	1	0	0

What is Composite Signal Priority Score?

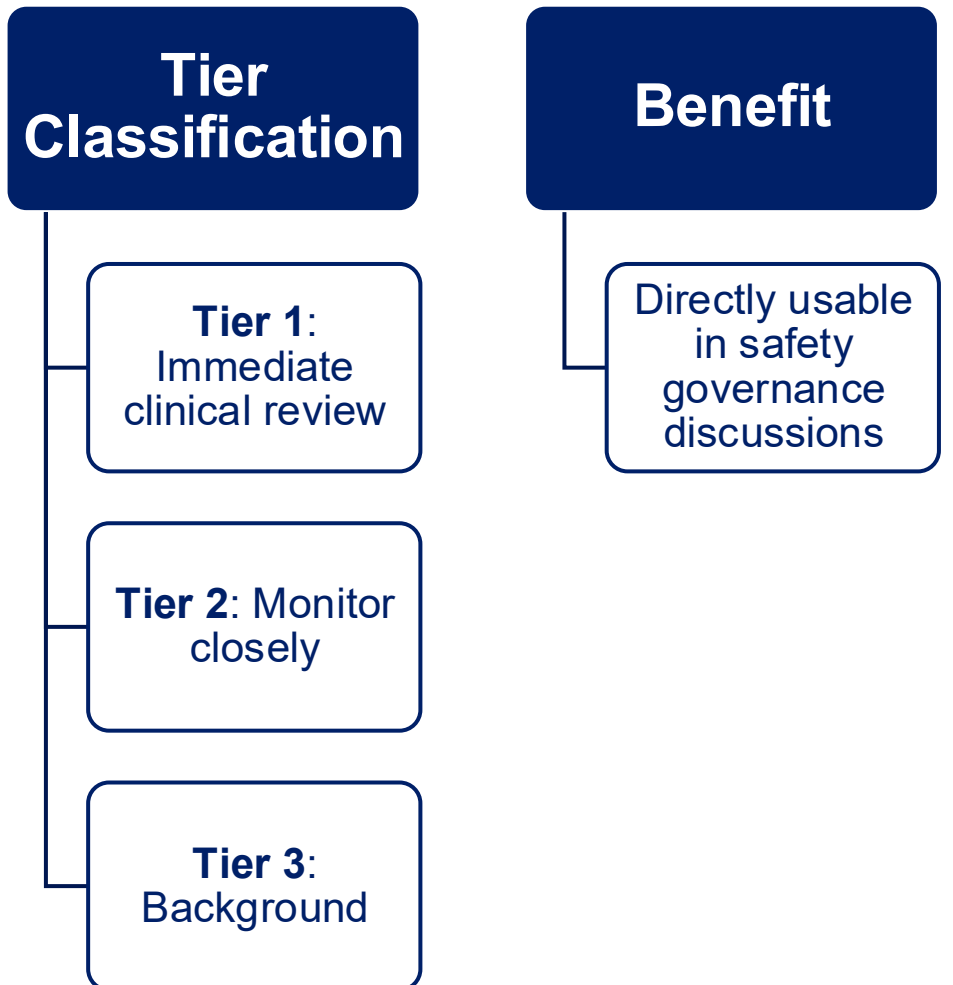


Why PCA?

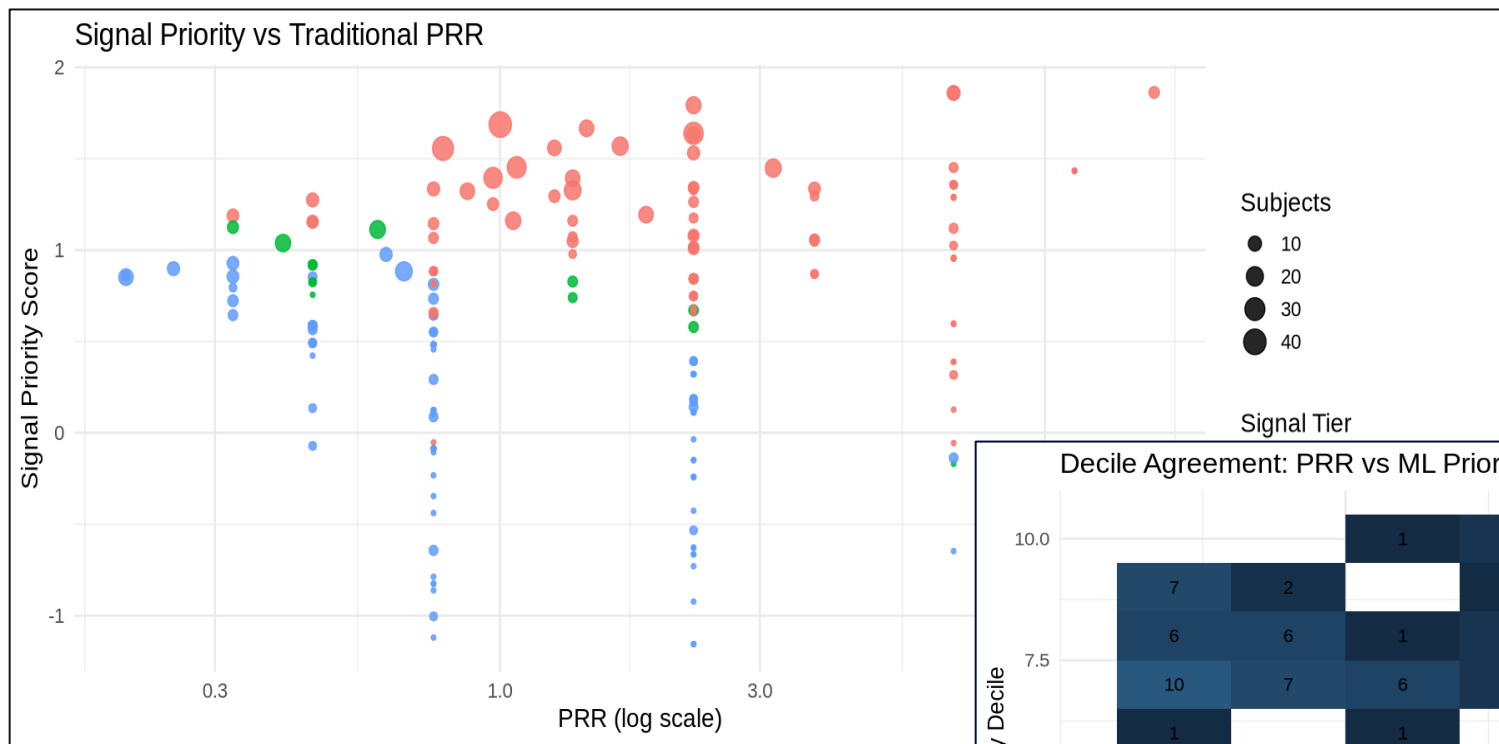
Applied Principal Component Analysis (PCA) to derive empirical weights.

Signal Tiers

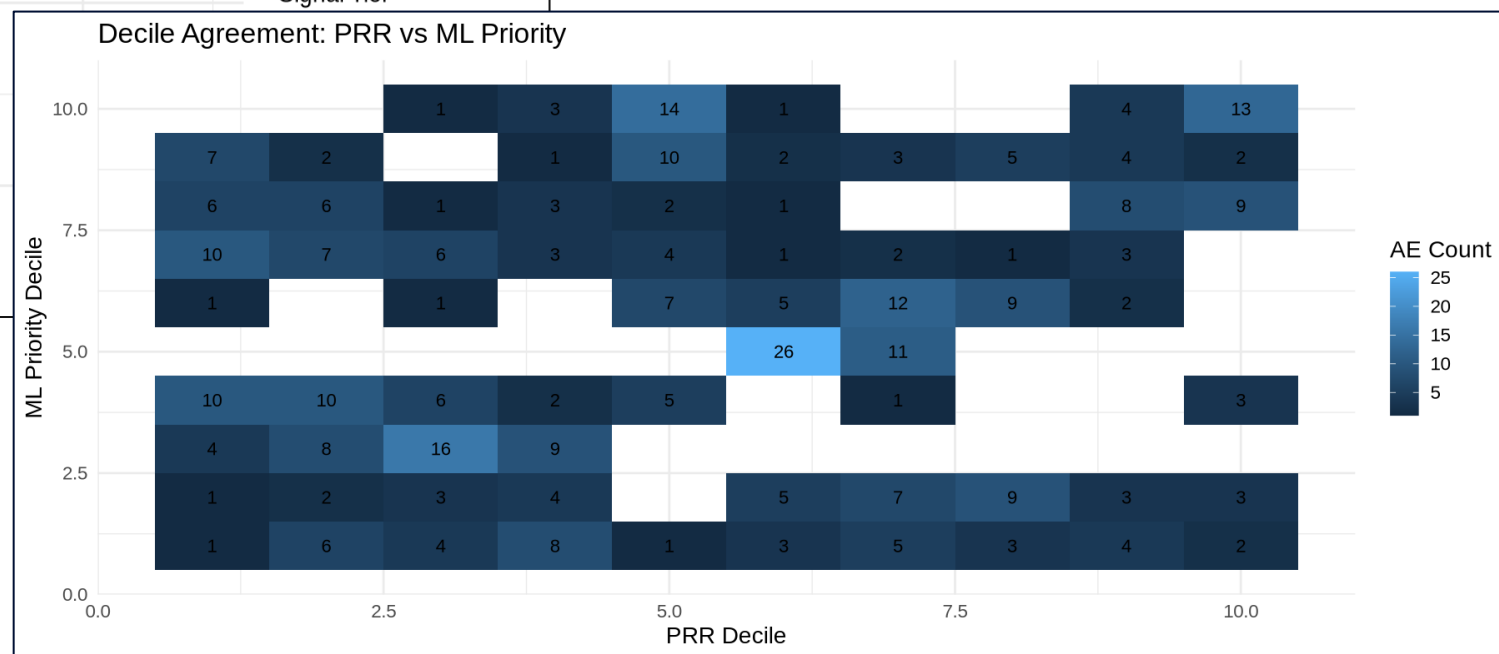
SCORE_FREQ	SCORE_SEV	SCORE_LIFT	SCORE_RARE	SCORE_POP	SCORE_EXPO	SIGNAL_PRIORITY_SCORE	PCA_SCORE[, 1]	SIGNAL_TIER
0.3017330	0.5	0.9353846	0	-0.5907677	-0.05614805	1.090201806	0.36748978	Tier 2: Monitor
0.3017330	0.5	0.2056753	0	0.1322913	-0.05614805	1.083551549	0.41750849	Tier 2: Monitor
0.3017330	0.5	0.9353846	0	-0.5907677	-0.05614805	1.090201806	0.36748978	Tier 2: Monitor
-1.2358464	0.5	0.2056753	0	0.1322913	-0.05614805	-0.454027880	0.09235752	Tier 3: Background
-0.0867002	0.0	-1.9917977	1	3.3275057	0.20784159	2.456849301	1.36110694	Tier 1: Immediate Review
0.3017330	0.5	0.2056753	0	0.1322913	-0.05614805	1.083551549	0.41750849	Tier 2: Monitor
0.3017330	0.5	0.9353846	0	-0.5907677	-0.05614805	1.090201806	0.36748978	Tier 2: Monitor
-0.8771709	0.5	0.2056753	1	0.1322913	-0.05614805	0.904647701	0.55251668	Tier 2: Monitor
1.0235123	0.0	-1.0979426	1	1.8813875	-0.05614805	2.750809195	1.17609105	Tier 1: Immediate Review
-0.8771709	0.5	0.2056753	0	0.1322913	-0.05614805	-0.095352299	0.16820642	Tier 3: Background
-0.8771709	0.5	0.2056753	0	0.1322913	-0.05614805	-0.095352299	0.16820642	Tier 3: Background
-1.2358464	0.5	-1.2347145	1	2.0914280	-0.05614805	1.064719086	0.91127847	Tier 2: Monitor
-0.8771709	0.5	-0.2191773	0	0.6453099	-0.05614805	-0.007186328	0.25545019	Tier 3: Background
-0.8441106	1.0	-2.6517475	1	4.5133466	-0.05614805	2.961340477	1.89053088	Tier 1: Immediate Review
1.9539934	0.5	0.9353846	0	-0.5907677	-0.05614805	2.742462246	0.71689228	Tier 1: Immediate Review
1.9539934	0.5	0.9353846	0	-0.5907677	-0.05614805	2.742462246	0.71689228	Tier 1: Immediate Review



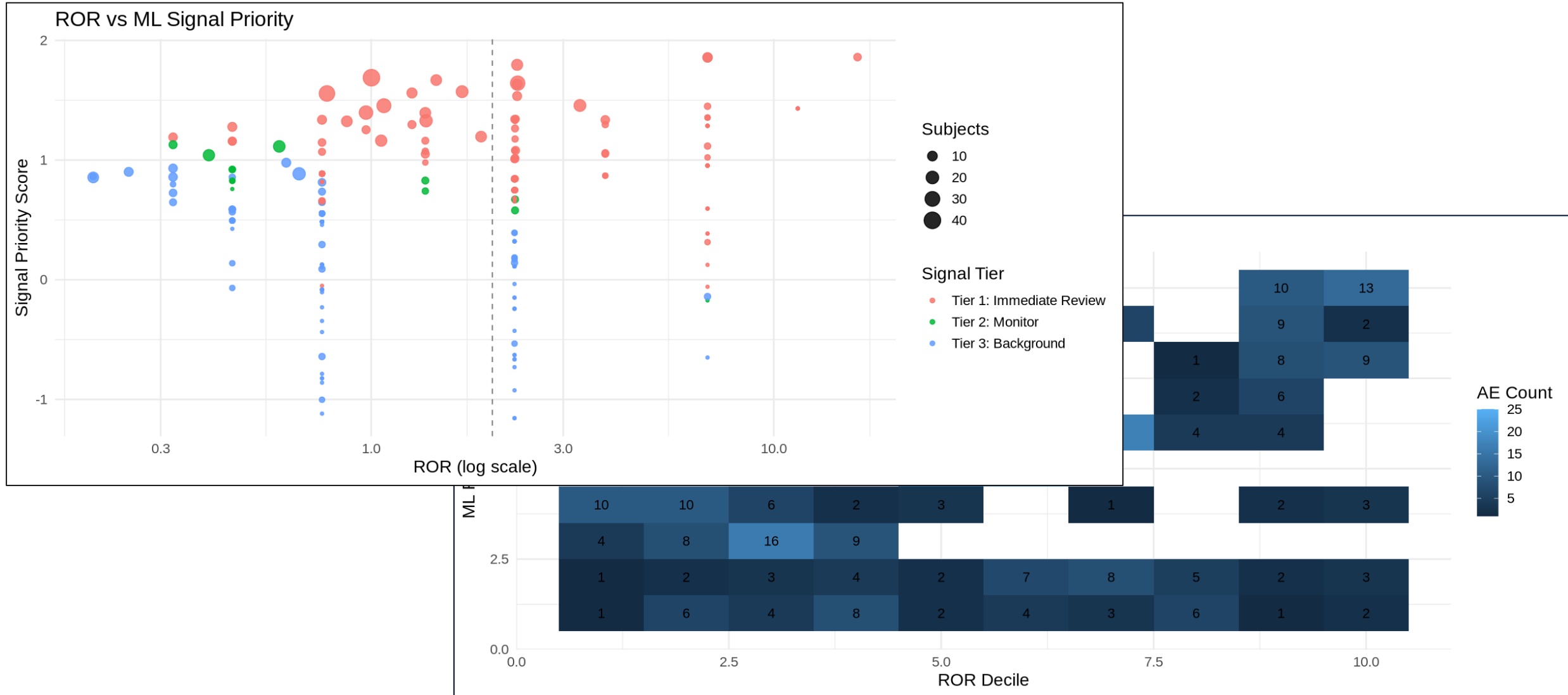
Comparative Results: PRR Vs AI

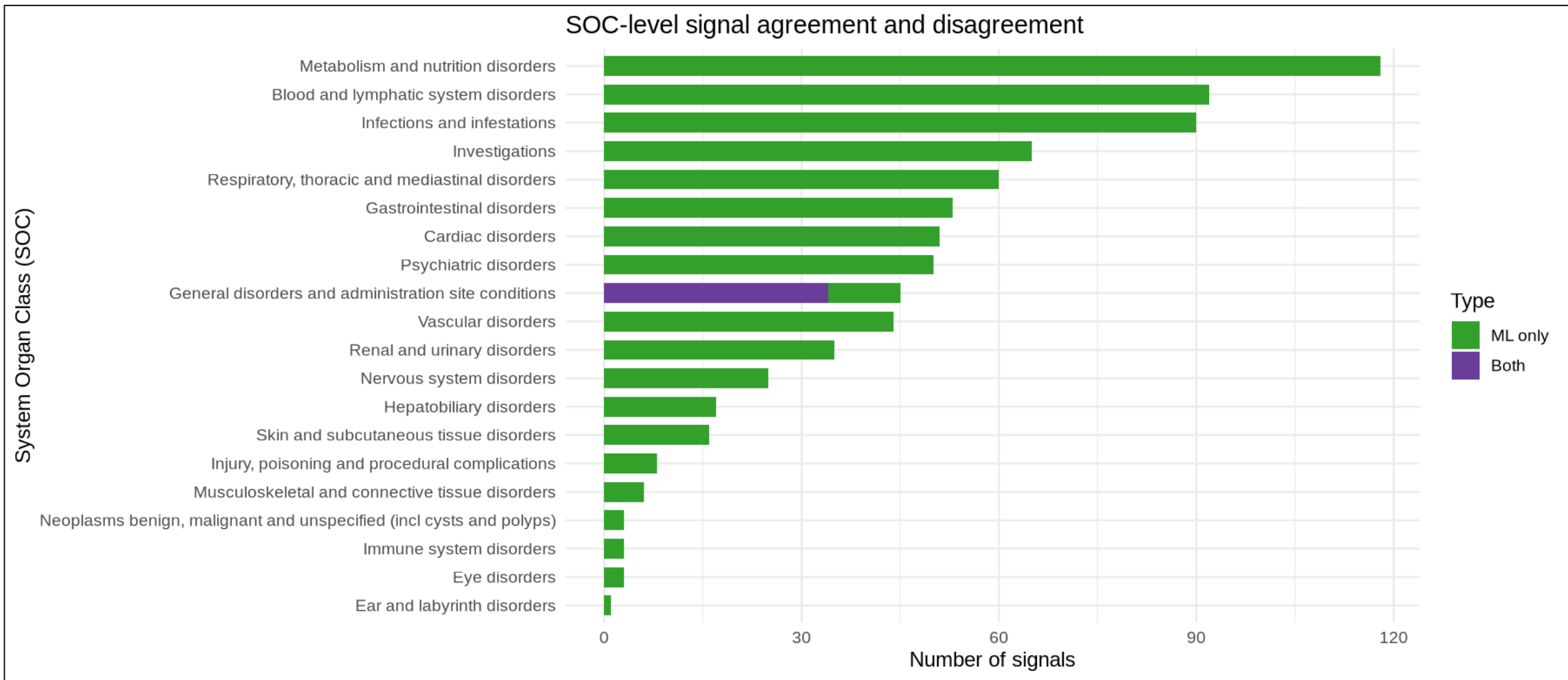


PRR_SIGNAL	PRIORITY_SIGNAL	n
FALSE	FALSE	225
FALSE	TRUE	210
TRUE	TRUE	1

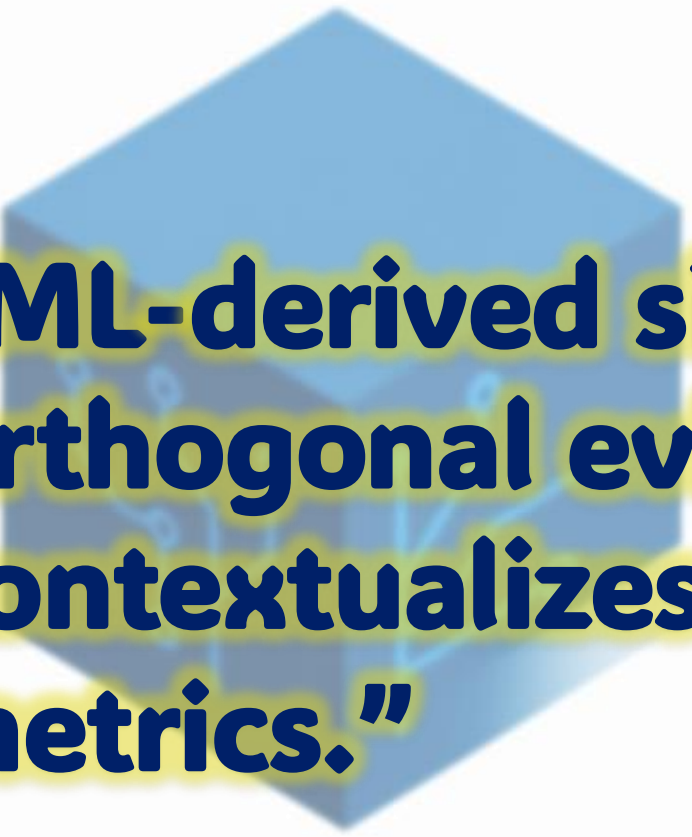


Comparative Results: ROR Vs AI





This suggests Composite score provided broader coverage and may capture clinically relevant signals earlier.



“ML-derived signal features provide orthogonal evidence that contextualizes frequency-based metrics.”

ML signal



Frequency

Regulatory & Workflow Considerations

Regulatory Alignment

- Interpretable models
- ADaM-based traceability
- Human in the loop review

Workflow Fit

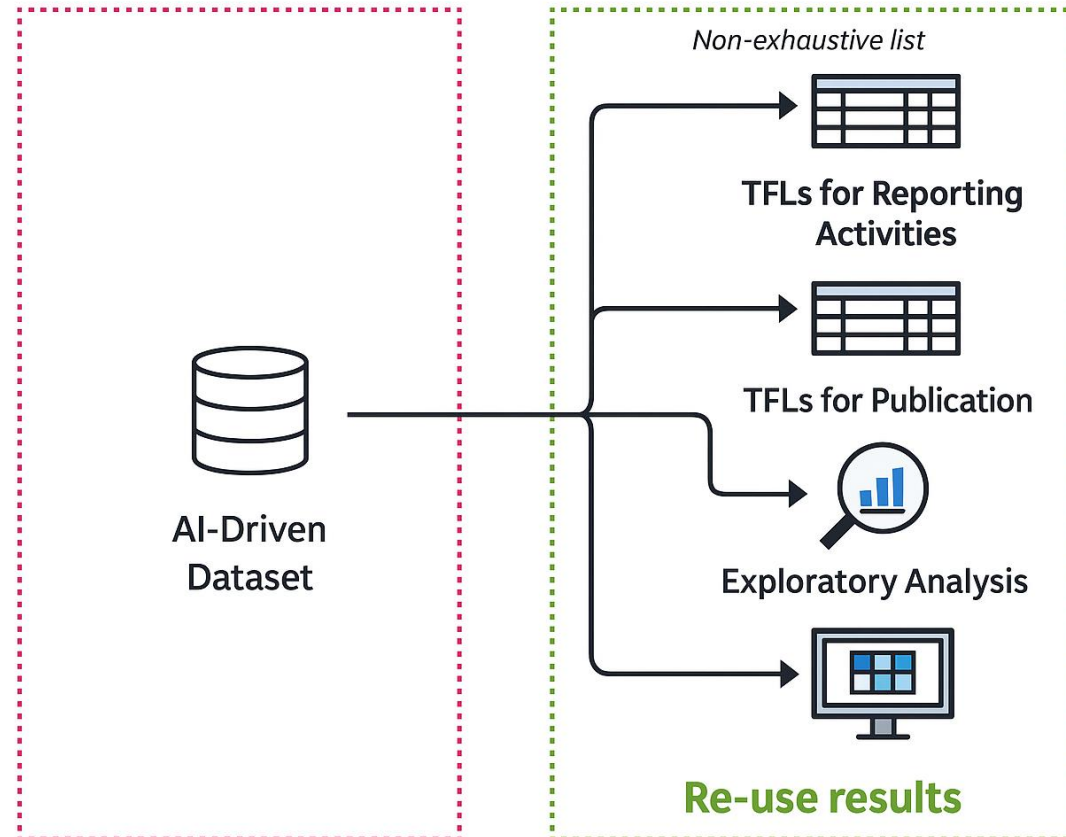
- Seamless integration into SP workflow pipelines
- No disruption to validated processes

Key Takeaways



PRR/ROR alone are insufficient in Phase 2 and 3 trials

Statistical programmers can lead early AI-driven safety innovation



Which method according to you helps detect early-onset toxicity patterns most?

If you were to prioritize safety signals tomorrow, what extra dimension would you want AI model to consider?

Acknowledgement and Contact Info

Monika

Statistical Programmer

Novartis

Email- monika-1.monika@novartis.com

LinkedIn- <https://www.linkedin.com/in/monika-42a471222/>