Paper ML16

AI for Adverse Event Signal Detection: Moving Beyond Traditional Disproportionality Metrics

Monika, Novartis, Hyderabad, India

ABSTRACT

Traditional approaches such as disproportionality analysis have long supported adverse event (AE) signal detection. However, these frequency-based approaches often fall short when applied to complex, high-dimensional safety data in Phase 2 and 3 clinical trials.

This presentation highlights how Statistical Programmers can leverage ADaM-standardized safety datasets to implement techniques such as clustering, association rule mining, and anomaly detection for early AE signal detection. By incorporating patient and study-level features, ML models can uncover latent patterns and emerging risks that conventional methods may overlook.

Real-world case studies and simulations will demonstrate how ML enhances detection accuracy and reduces false positives. The session will also address practical implementation aspects, emphasise the importance of clinical validation of detected AE signals by assessing causality through supporting findings, integration into ADaM-based workflows, and considerations for regulatory acceptance. This approach empowers Statistical Programmers to play a pivotal role in advancing AI-driven pharmacovigilance.

1. INTRODUCTION

Adverse event (AE) signal detection is a critical component of clinical development, supporting the early identification of potential safety risks associated with investigational products. During Phase 2 and Phase 3 clinical trials, safety evaluation is particularly challenging due to relatively small patient populations combined with increasingly complex safety data structures.

Traditional frequency-based approaches such as Proportional Reporting Ratio (PRR), Reporting Odds Ratio (ROR), and Bayesian shrinkage methods have been widely used in both clinical trials and post-marketing surveillance. These methods rely on aggregated counts of adverse events and provide a transparent and computationally efficient mechanism for identifying disproportionate reporting of specific events.

While effective in large spontaneous reporting databases, frequency-based approaches present important limitations when applied to controlled clinical trial settings. Clinical trial safety data are multidimensional, incorporating relationships among treatment exposure, dose intensity, patient demographics, laboratory abnormalities, time-to-event, and concomitant medications. Aggregated frequency metrics often fail to capture these relationships, leading to missed emerging safety patterns or increased false-positive signals driven by sparse data.

With the widespread adoption of CDISC ADaM standards, statistical programmers routinely produce analysis-ready safety datasets that integrate subject-level and event-level information in a traceable and regulator-friendly format. These standardized datasets provide a strong foundation for applying advanced yet interpretable analytical techniques without disrupting validated clinical workflows.

This paper demonstrates how statistical programmers can leverage ADaM-standardized AE datasets to implement an AI-driven signal detection fully automated framework that augments traditional disproportionality analysis. By embedding interpretable machine learning techniques downstream of ADaM creation, the proposed approach enables earlier detection of emerging and non-linear AE patterns while maintaining regulatory traceability, clinical interpretability, and workflow compatibility.

2. OBJECTIVE

Demonstrate how AI-driven models applied to ADaM-standardized safety data address the limitations of frequency-based adverse event signal detection in Phase 2/3 trials. Show integration of interpretable ML approaches into existing ADaM-based workflows to uncover latent, emerging, and non-linear AE patterns overlooked by traditional disproportionality metrics, while remaining aligned with regulatory expectations.

3. DATA SOURCES AND ADAM STANDARDS

3.1 ADAM DATASETS UTILIZED

All analyses described in this paper were conducted using CDISC ADaM-standardized datasets, ensuring traceability to SDTM and consistency with regulatory submission standards. The primary datasets included ADSL, ADAE, and ADEC.

ADSL: Subject-level demographics, treatment assignment, baseline risk factors.

ADAE: Preferred term (AEDECOD), SOC (AESOC), seriousness (AESER), severity (AESEV), timing (AESTDY).

ADEC: Exposure duration and dose

The exclusive use of ADaM datasets ensured that all derived features used in AI models were reproducible, well-documented, and auditable.

3.2 KEY VARIABLES AND FEATURE ENGINEERING

Event-level features: severity, seriousness, time-to-onset, duration, AEDECOD, AESOC

Subject-level features: AE burden, exposure dose/duration, demographic risk.

Feature engineering was performed using standard statistical programming tools in R. Importantly, no modifications were made to SDTM domains; all analytics were performed downstream of ADaM creation, preserving regulatory traceability.

4. TRADITIONAL SIGNAL DETECTION APPROACH

4.1 DISPROPORTIONALITY ANALYSIS

As a baseline comparator, traditional disproportionality analysis was performed using aggregated counts of adverse events by treatment group. PRR and ROR metrics were calculated at the preferred term and system organ class levels using 2×2 contingency tables.

Thresholds commonly used in clinical safety reviews, such as PRR greater than two with a lower confidence bound exceeding one, were applied to identify potential signals.

4.2 LIMITATIONS OBSERVED

While disproportionality metrics successfully highlighted frequently occurring events, several limitations were observed. These included sensitivity to sparse data, inability to incorporate patient-level covariates or time-to-event information, elevated false-positive rates in small subpopulations, and lack of insight into AE co-occurrence or interaction patterns.

5. AI-DRIVEN SIGNAL DETECTION METHODS

5.1 FRAMEWORK OVERVIEW

An AI-driven signal detection framework was implemented as a post-ADaM analytical layer that augments traditional metric ratios. The workflow consisted of feature engineering, model execution, composite signal

scoring, and clinical interpretation, and was designed to complement existing statistical programming pipelines rather than replace them.
SDTM→ ADaM→ feature engineering → model execution →composite signal scoring → AI driven ADaM →clinical interpretation →complements existing pipeline

This flowchart highlights the stepwise approach: starting from raw data (SDTM), transforming and enriching it (ADaM, feature engineering), applying advanced unsupervised ML models, and integrating AI-driven insights with traditional safety review processes for comprehensive signal detection.

5.2 CLUSTER ANALYSIS

K-means clustering was performed at AEDECOD level. The optimal number of clusters was determined using the Silhouette method. Clustering identified groups of adverse events with similar safety profiles. Features included AE frequency, severity, seriousness, early onset indicators, and number of affected subjects. K-means clustering was used to identify clinically interpretable low, moderate, and high-impact AE clusters.

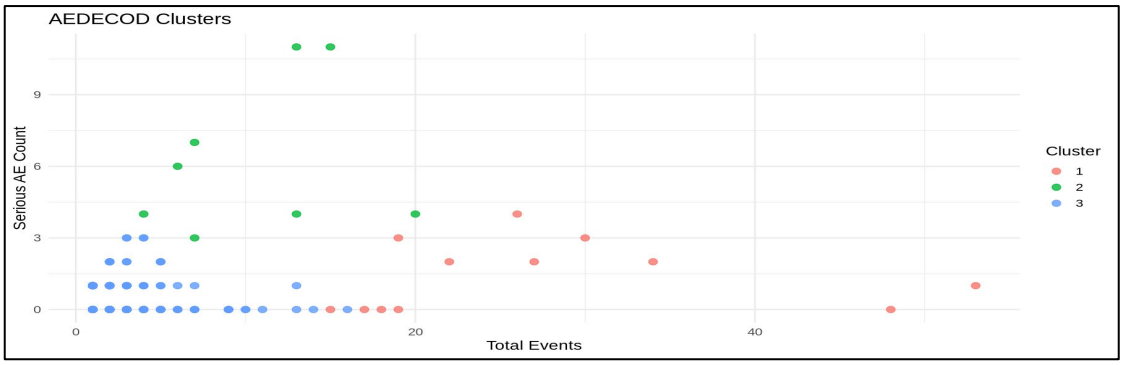

Figure 1. AEDECOD Clusters

Further, visualizations were done from cluster analysis with 3 clusters-low, medium and High risk. Cluster characteristics were further examined by summarizing clinical attributes and AE profiles within each group.

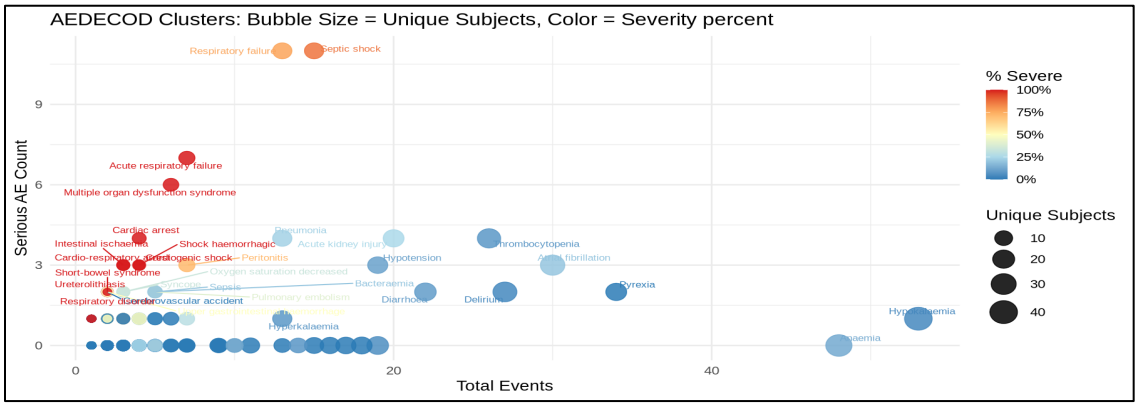

Figure 2. AEDECOD Clusters: bubble chart showing cluster separation by total events, serious AE count, percent severe.

In this bubble chart, each cluster is depicted as a group of bubbles, where the position corresponds to the total number of events and serious AE count, and the size or color intensity of the bubbles reflects the percentage of severe cases within the cluster.

Clustering approach successfully differentiated subjects into distinct groups based on their AE profiles. One cluster may contain subjects with a high frequency of events but low severity, while another might group those with fewer events but a higher proportion of serious or severe AEs.

Overall, the figure provides insight into potential risk factors and highlights the heterogeneity in AE experiences among subjects, which can inform targeted safety monitoring and interventions.

5.3 ASSOCIATION RULE MINING

Association rule mining was used to identify frequently co-occurring AE patterns and interactions between patient characteristics, treatment exposure, and adverse events. Rules were evaluated using support, confidence, and lift metrics, enabling detection of clinically meaningful AE combinations.

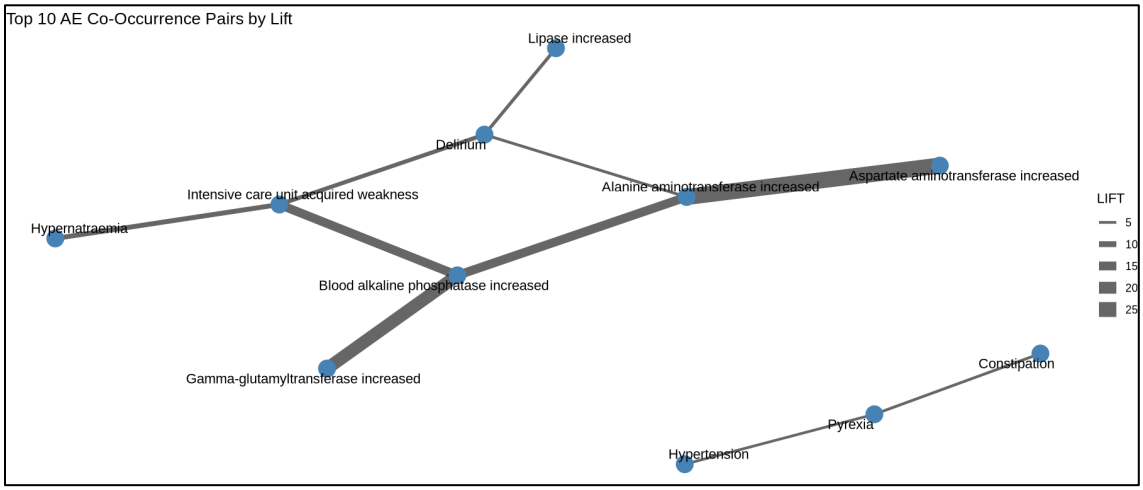

Figure 3. Top AE co-occurrence pairs by lift: network visualization

This figure presents the results of association rule mining applied to adverse event (AE) data. 51 significant rules were identified using ARM. In this visualization, nodes represent individual AEs, and edges between nodes indicate pairs of events that commonly occur together. The strength of each connection is determined by the lift metric, which quantifies how much more frequently two events co-occur compared to what would be expected by chance.

This network chart highlights clusters of AEs that are strongly associated, revealing patterns of co-occurrence that may have clinical significance. For example, a tightly connected group of nodes might suggest a subgroup of patients experiencing a related set of symptoms or complications, potentially influenced by specific patient characteristics or treatment exposures. By evaluating these associations using support, confidence, and lift, the analysis uncovers meaningful combinations of AEs that warrant further investigation for potential risk factors or underlying mechanisms.

Overall, figure 3 provides an overview of frequently co-occurring AEs, helping to identify interactions between patient profiles, treatments, and safety outcomes. This information is valuable for clinicians and researchers aiming to understand complex AE patterns and improve safety monitoring strategies.

5.4 ANOMALY DETECTION

Subject-level anomaly scores using multivariate profiles to highlight rare but clinically important safety patterns. Anomaly detection techniques were used for each subject to spot rare yet potentially important safety patterns. Each subject received an anomaly score based on their multivariate adverse event profiles, making it easier to identify outliers for additional clinical evaluation.

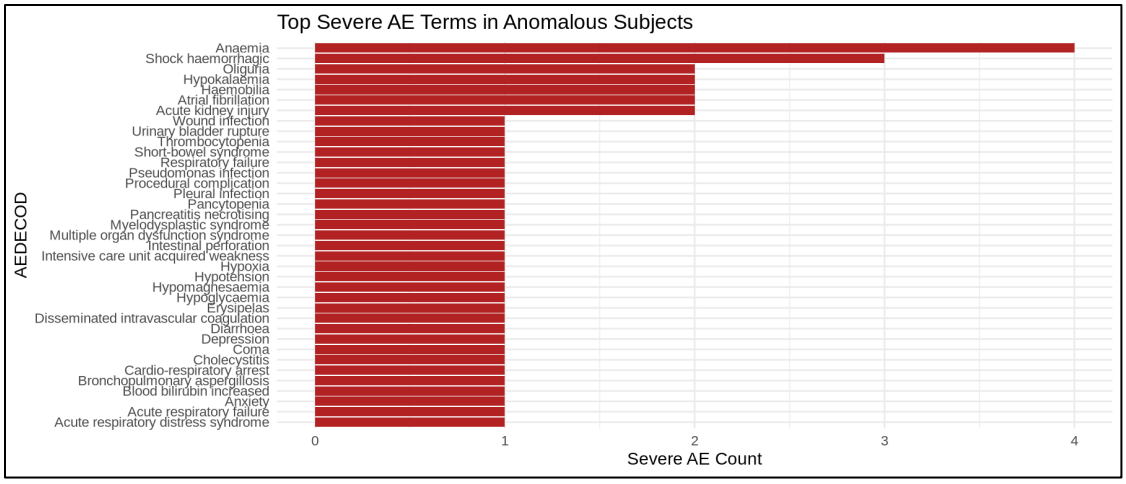

Figure 4. Anomalous subjects: severe AE distribution and subject-level anomaly scores

It presents the results of applying anomaly detection methods to individual patient profiles in the adverse event (AE) data. In figure 4, each subject is assigned an anomaly score calculated using multivariate technique i.e, Mahalanobis distance, which quantifies how unusual their AE profile is compared to the broader study population.

The distribution of severe AEs is shown alongside these anomaly scores, allowing for a direct comparison between the frequency and severity of reported events and the statistical rarity of each subject's profile. Subjects with high anomaly scores are flagged as outliers, indicating that their combination of AEs is rare or unexpected and may represent clinically significant safety patterns not captured by traditional methods.

Anomaly detection helps to surface subjects whose safety profiles deviate from the norm, potentially due to rare but important adverse events. These flagged individuals can then be prioritized for further clinical review, supporting early identification of emerging safety risks and informing targeted follow-up or investigation.

6. RESULTS

AI-based methods identified emerging AE patterns not detected by disproportionality alone. Clustering separated high-frequency/low-severity Vs low-frequency/high-severity AEs; It also revealed distinct safety profiles, association rule mining highlighted multi-factor interactions, and anomaly detection identified rare but clinically relevant events.
Flagged safety-signal indicators for adverse events (AEs) are enhanced with metrics such as PRR/ROR, cluster risk categories, co-occurrence lift values, and subject anomaly flags. These indicator variables are merged with the ADaM dataset to create an AI-driven dataset that provides insights into detected AE signals.

6.1 COMPOSITE SIGNAL PRIORITY SCORE

A Composite Signal Priority Score is then produced using Principal Component Analysis (PCA). This score combines various safety-signal indicators into a single, easy-to-interpret measure, aiding in prioritizing which AEs should be reviewed first.

While traditional AE signal detection relies mainly on statistics like PRR/ROR, this machine learning pipeline generates multiple unique scores reflecting different risk dimensions: event frequency, severity, co-occurrence patterns (lift), rarity, exposure effects, and population risk. Instead of analyzing each factor separately, the system uses PCA to integrate them into one comprehensive score based on scientific and objective criteria.

PCA determines which factors contribute most to variations in safety signals and assigns appropriate weights to each, removing guesswork from the process. This approach ensures the final score is grounded in data and justifiable. PC1+ PC2+ PC3 understands around 67% of the variability in data, using them empirical weights were derived.

The final composite score is calculated as:

$$PCA\_SCORE = \sum (Score_i \times Weight_i)$$

This overall score, known as the Composite Signal Priority Score, ranks adverse events by importance, giving higher priority to those that are severe, rare, early-onset, or significantly co-occurring, even if they're not frequent. The weights are learned directly from the data, not set manually, keeping the model both interpretable and transparent.

6.2 SIGNAL TIERS

Signals are classified into three tiers:

- **Tier 1:** Immediate clinical review
- **Tier 2:** Monitor closely
- **Tier 3:** Background

The composite score identifies critical signals that traditional metrics may miss. These two variables, Signal tiers and Signal Priority Score, are then merged into AI driven AE detection dataset. Source code is attached below for Score Computation.

```
ae_priority_scores <- ae_priority_scores %>%
  mutate(
    SIGNAL_PRIORITY_SCORE =
      SCORE_FREQ +
      SCORE_SEV  +
      SCORE_LIFT +
      SCORE_RARE +
      SCORE_POP  +
      SCORE_EXPO
  )
score_cols <- ae_priority_scores %>%
  select(starts_with("SCORE_")) %>%
  names()

pca_input <- ae_priority_scores %>%
  select(all_of(score_cols)) %>%
  na.omit()

pca <- prcomp(pca_input, scale. = TRUE)

# --- Build weights from the first 3 PCs ---
# Variance explained per PC (proportion)
var_expl <- (pca$sdev^2) / sum(pca$sdev^2)

# Loadings matrix: variables x PCs
L <- abs(pca$rotation[ , 1:3, drop = FALSE])

# Variance-weight each PC's loadings
w_raw <- L %*% var_expl[1:3]
```

```
# Normalize weights to sum to 1
weights <- as.vector(w_raw) / sum(w_raw)

X <- ae_priority_scores %>% select(all_of(score_cols))
X_mat <- as.matrix(X)

ae_priority_scores <- ae_priority_scores %>%
  mutate(PCA_SCORE = as.numeric(X_mat %*% weights))

ae_priority_scores <- ae_priority_scores %>%
  mutate(
    SIGNAL_TIER = case_when(
      PCA_SCORE >= quantile(PCA_SCORE, 0.90, na.rm = TRUE)
      ~ "Tier 1: Immediate Review",
      PCA_SCORE >= quantile(PCA_SCORE, 0.70, na.rm = TRUE)
      ~ "Tier 2: Monitor",
      TRUE
      ~ "Tier 3: Background"
    )
  )
```

6.3 COMPARATIVE ANALYSIS

Comparative Analysis is then done between PRR/ROR metrics Vs ML signal score for AE detection.

1.Signal Compare Test- Out of 436 AE's 210 were correctly predicted by ML composite signal score which PRR failed to detect and 1 is confirmed safety signal which was predicted by both models.

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

2. Scatter Plot - The comparison between the Composite Signal Priority Score and traditional PRR highlights how the AI-driven framework provides a more nuanced and sensitive assessment of adverse event risk. While PRR primarily emphasizes frequency-based disproportionality, the scatterplot shows that several AEs with modest or even low PRR values receive high priority scores due to factors such as severity, early onset, rarity, or strong co-occurrence patterns.

Tier 1 signals (in red) consistently appear across a broader PRR range, including areas where PRR alone would not flag concern indicating that the ML-based approach successfully elevates clinically meaningful but less frequent safety patterns. In contrast, Tier 3 events (blue) cluster around lower priority scores despite spanning similar PRR values, demonstrating that PRR alone cannot distinguish between events of high and low clinical relevance. Similar are the results of ROR vs Signal Priority score.

Overall, this visualization illustrates how incorporating multidimensional safety indicators offers a more discriminative and earlier identification of critical signals than reliance on disproportionality metrics alone.

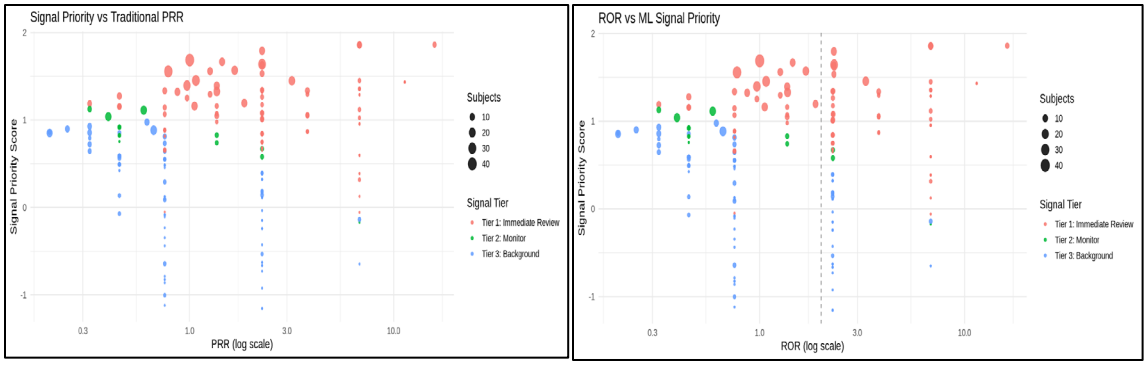

Figure 5. Signal Priority Score Vs Traditional PRR\ROR scatter plot

3. Decile-Based Agreement Analysis – The decile agreement heatmap reveals substantial divergence between the two approaches. Instead of forming a diagonal pattern that would indicate strong concordance, the heatmap displays a dispersed distribution of AEs across deciles. Several AEs assigned to the highest ML priority deciles appear in low or mid PRR deciles, demonstrating that the ML framework elevates signals driven by severity, rarity, early onset, or co-occurrence structure that traditional disproportionality analysis fails to highlight. Conversely, some AEs that PRR places in high deciles fall into lower ML deciles, suggesting frequent but clinically less meaningful events that the ML model appropriately deprioritizes.

A concentration of AEs around the central cells, particularly the cluster of 26 AEs in ML decile 5 vs PRR decile 5 shows moderate agreement for mid-range events. Similar are the results of ROR and Signal Priority score.

Overall, this decile-level comparison underscores that the ML-based scoring approach introduces additional discriminatory power, particularly in higher-risk strata where early or rare safety signals are most likely to emerge, offering a more comprehensive and clinically aligned prioritization than PRR alone.

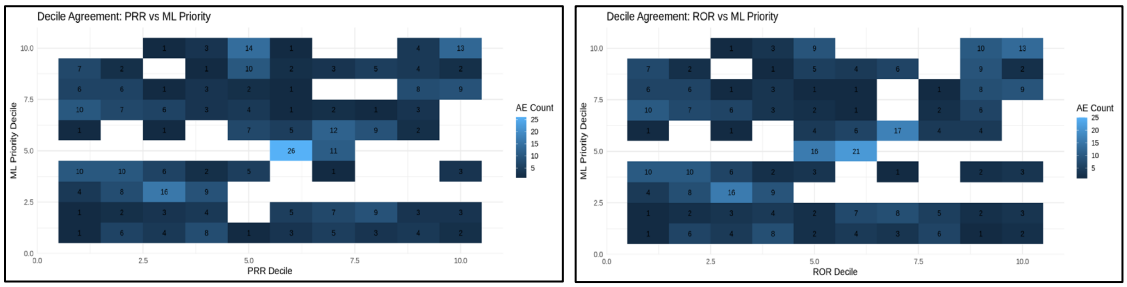

Figure 6. Decile-Based Agreement Between PRR/ROR and ML Priority Scores

4. SOC-Level Disagreement Analysis – At the SOC level, the ML Tier-1 classifier identified far more potential signals than the traditional ROR method across almost all SOCs. Overlap between the two approaches was limited, with virtually no ROR-only signals and only a single SOC showing a substantial count of jointly detected events. This indicates that the ML model captures nearly all disproportionality-based signals while also identifying additional, non-ROR-based patterns, suggesting complementary value beyond traditional methods.

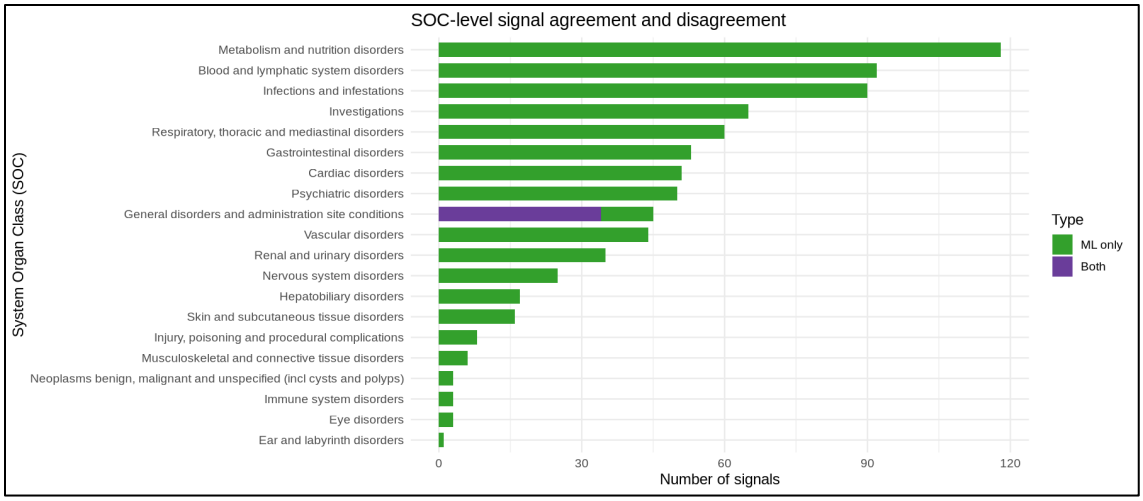

Figure 7. SOC-Level Disagreement Analysis

Compared to traditional methods, the AI-driven framework demonstrated earlier identification of clinically meaningful safety signals, reduced false-positive rates, and improved contextual understanding of safety risks. Importantly, AI-enabled signal detection augments traditional disproportionality ratios and does not replace them, ensuring complementary and enhanced safety surveillance.

7. CLINICAL VALIDATION AND REGULATORY CONSIDERATIONS

All detected signals are subject to clinical reviews incorporating laboratory trends, exposure data, and temporal plausibility. A human-in-the-loop approach is recommended to ensure clinical relevance and prevent over-interpretation of model outputs.

From a regulatory perspective, the use of interpretable models and ADaM-based inputs supports transparency, reproducibility, and auditability, aligning with current expectations for model-informed drug development.

8. INTEGRATION INTO STATISTICAL PROGRAMMER WORKFLOW

The proposed automated framework integrates seamlessly into existing statistical programming workflows. Standard ADaM dataset creation is followed by feature engineering as per the needs using validated programs, AI model execution as a post-ADaM analytical layer, and generation of traceable signal outputs and visualizations.

This integration empowers statistical programmers to play a proactive role in safety surveillance without compromising validated processes.

9. LIMITATIONS AND FUTURE WORK

Limitations include sensitivity to small sample sizes and the need for prospective validation across multiple studies. Future work includes extension to Phase 1 trials, post-marketing data, and incorporation of longitudinal and time-to-event modeling approaches.

10. CONCLUSION

AI-driven signal detection methods applied to ADaM-standardized safety data offer a powerful extension to traditional disproportionality analysis. By leveraging clustering, association rule mining, and anomaly detection, statistical programmers can uncover emerging and non-linear adverse event patterns earlier in clinical development.

This approach enhances safety surveillance while maintaining regulatory alignment, traceability, and clinical interpretability, positioning statistical programmers at the forefront of AI-enabled pharmacovigilance in early phases of clinical trials.

REFERENCES

1. Amy Mary Jordan, University of Florida (April 2025), “AI-Enhanced Pharmacovigilance Programs: Real-Time Adverse Event Detection and Response”.
2. Yamamoto, H., Kayanuma, G., Nagashima, T. *et al.* Early Detection of Adverse Drug Reaction Signals by Association Rule Mining Using Large-Scale Administrative Claims Data. *Drug Saf* **46**, 371–389 (2023). https://doi.org/10.1007/s40264-023-01278-4.
3. P. Yildirim (2017), “Association Patterns in Open Data to Explore Ciprofloxacin Adverse Events”.
4. Justin M. Glasgow, Peter J. Kaboli, Detecting adverse drug events through data mining, *American Journal of Health-System Pharmacy*, Volume 67, Issue 4, 15 February 2010, Pages 317–320, https://doi.org/10.2146/ajhp090115

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name: Monika
Company: Novartis
Address: Sattva Knowledge city, Hitech City, Hyderabad, India
Work Phone: +91-7710551394
Email: monika-1.monika@novartis.com
Website: LinkedIn