Drug Safety (2025) 48:923–931
https://doi.org/10.1007/s40264-025-01545-6

ORIGINAL RESEARCH ARTICLE

A Drug Similarity-Based Bayesian Method for Early Adverse Drug Event Detection

Yi Shi[1] · Yuedi Yang[1] · Ruoqi Liu[2] · Anna Sun[1] · Xueqiao Peng[2] · Lang Li[3] · Ping Zhang[2,3] · Pengyue Zhang[1]

Accepted: 17 March 2025 / Published online: 22 April 2025
© The Author(s) 2025

Abstract

Introduction Biochemical drug similarity-based methods demonstrate successes in predicting adverse drug events (ADEs) in preclinical settings and enhancing signals of ADEs in real-world data mining. Despite these successes, drug similarity-based ADE detection shall be expanded with false-positive control and evaluated under a time-to-detection setting.

Methods We tested a drug similarity-based Bayesian method for early ADE detection with false-positive control. Under the tested method, prior distribution of ADE probability of a less frequent drug could be derived from frequent drugs with a high biochemical similarity, and posterior probability of null hypothesis could be used for signal detection and false-positive control. We evaluated the tested and reference methods by mining relatively newer drugs in real-world data (e.g., the US Food and Drug Administration (FDA)'s Adverse Event Reporting System (FAERS) data) and conducting a simulation study.

Results In FAERS analysis, the times to achieve a same probability of detection for drug-labeled ADEs following initial drug reporting were 5 years and ≥ 7 years for the tested method and reference methods, respectively. Additionally, the tested method compared with reference methods had higher AUC values (0.57–0.79 vs. 0.32–0.71), especially within 3 years following initial drug reporting. In a simulation study, the tested method demonstrated proper false-positive control, and had higher probabilities of detection (0.31–0.60 vs. 0.11–0.41) and AUC values (0.88–0.95 vs. 0.69–0.86) compared with reference methods. Additionally, we identified different types of drug similarities had a comparable performance in high-throughput ADE mining.

Conclusion The drug similarity-based Bayesian ADE detection method might be able to accelerate ADE detection while controlling the false-positive rate.

Key Points

Drug similarity might be able to accelerate adverse drug event (ADE) detection.

Posterior probability of the null hypothesis might be able to control for false positives in high-throughput ADE mining.

1 Introduction

Many adverse drug events (ADEs) have been identified in post-marketing pharmacovigilance studies, but not in pre-marketing clinical trials [1]. The US Food and Drug Administration (FDA)'s Adverse Event Reporting System (FAERS) is an important pharmacovigilance database [2]. Commonly used methods to detect signals of ADE include proportional reporting ratio (PRR) [3], reporting odds ratio (ROR) [4], likelihood ratio test [5], multi-item gamma Poisson shrinker (MGPS) or empirical Bayesian geometric mean (EBGM) [6], Bayesian confidence propagation neural network (BCPNN) or information component (IC) [7], logistic regression model-based methods [8, 9], and posterior probability of the null hypothesis under MGPS [10].

Biochemical drug properties are powerful in predicting potential ADEs in preclinical settings [11–14]. Recent studies use biochemical drug property-based drug-similarity networks to enhance signals of ADEs in real-world data mining [14–16], as drugs with similar biochemical properties may also have similar toxicities [17]. Doing so can particularly

Yi Shi, Yuedi Yang and Ruoqi Liu contributed equally.

✉ Ping Zhang
ping.zhang@osumc.edu

✉ Pengyue Zhang
zhangpe@iu.edu

[1] Department of Biostatistics and Health Data Science, Indiana University, Indianapolis, IN, USA

[2] Department of Computer Science and Engineering, The Ohio State University, Columbus, OH, USA

[3] Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA

enhance signals of drugs with a lower frequency such as relatively newer drugs. For instance, the similarity score between two drugs can be defined by the number of common chemical structures over the number of total unique chemical structures [16]. Subsequently, signals (e.g., PRR, ROR, EBGM, and IC) of a relatively newer drug in FAERS can be enhanced by signals of drugs with a higher frequency in FAERS according to the drug similarity network [16]. Besides chemical structures, drug similarity networks can also be derived from drug-targeted genes, proteins, and pathways [11]. Drug similarity-enhanced signals have improved performance on prioritizing true ADEs [15, 16]. However, drug similarity-based early ADE detection has not been expanded with false-positive control in mining all newer drugs and ADEs, nor has this been explicitly evaluated under a time-to-detection setting with respect to different types of drug similarities. In this work, we aimed to test a drug similarity-based Bayesian method for mining signals of ADEs with false-positive control, and to evaluate the performance of different types of drug similarity scores in early ADE detection under a time-to-detection setting.

2 Methods

2.1 Drug Similarity Scores

We computed drug similarity scores based on: (1) chemical substructures (Chem) of drugs, (2) sequences (Seq) of drug-targeted proteins, (3) biological processes (BP) of drug-related genes, (4) molecular functions (MF) of drug-related genes, and (5) cellular components (CC) of drug-related genes.

For the Chem-based similarity score, we used 881 chemical substructures defined in Pubchem [18]. We assigned a binary vector to each drug according to the presence or absence of the corresponding substructures. For a pair of drugs with vectors $\vec{A} = (a_1, a_2, \ldots, a_{881})$ and $\vec{B} = (b_1, b_2, \ldots, b_{881})$ ($a_i = 0/1$ and $b_i = 0/1$), we computed the Chem-based similarity score as $\frac{\sum 1(a_i=1 \text{and} b_i=1)}{\sum 1(a_i=1 \text{or} b_i=1)}$ (e.g., the Jaccard score) [16, 19].

For the Seq-based similarity score, we collected drug-targeted proteins from DrugBank [20]. We used the R package *Biostrings* to compute similarities between drug-targeted proteins with substitutionMatrix = BLOSUM62, gapOpening = 10, and gapExtension = 1 [21–23]. Where P denotes a protein, ρ denotes similarity between two proteins, $\boldsymbol{A}$ and $\boldsymbol{B}$ denote sets including drug-targeted proteins for two drugs, and 1[•] denotes the indicator function, for a pair of drugs with drug-targeted protein sets $\boldsymbol{A}$ and $\boldsymbol{B}$, we computed the Seq-based similarity score as:

$$\frac{\sum_{P_i \in \boldsymbol{A}} \sum_{P_j \in \boldsymbol{B}} \left\{ 1\left[P_i = P_j\right] + \rho_{ij} \times 1\left[\left(P_i, P_j\right) \notin (\boldsymbol{A} \cap \boldsymbol{B})\right] \right\}}{\sum_{P_i \in \boldsymbol{A}} \sum_{P_j \in \boldsymbol{B}} \left\{ 1\left[P_i = P_j\right] + 1\left[\left(P_i, P_j\right) \notin (\boldsymbol{A} \cap \boldsymbol{B})\right] \right\}} \quad (1)$$

The similarity score in Eq. (1) was derived from the total number of overlapped proteins (i.e., the function $1\left[P_i = P_j\right]$) and average similarity score between non-overlapped proteins (i.e., the function $1\left[\left(P_i, P_j\right) \notin (\boldsymbol{A} \cap \boldsymbol{B})\right]$), which ensured two drugs with complete overlapped proteins to have a similarity score = 1.

For BP-, MF-, and CC-based similarity scores, we collected drug-targeted genes and their ontologies from Gene Ontology (GO) [24]. For each type of ontology (e.g., BP, MF, or CC), we used R package *GOSemSim* to compute similarities between drug-targeted genes with measure = Wang [25, 26]. For a pair of drugs, we computed BP-, MF-, and CC-based similarity scores by using equation (1), in which ρ_{ij}, $\boldsymbol{A}$, and $\boldsymbol{B}$ in equation (1) refer to genes.

2.2 US Food and Drug Administration (FDA)'s Adverse Event Reporting System (FAERS) Data Preparation

We followed Banda et al. and Wang et al. to process FAERS data (2004Q1–2022Q2) [27, 28]. We removed duplicates, and normalized drug names to drug ingredient names according to RxNorm and DrugBank [20, 29]. We used MedDRA® Preferred Terms (PT) as ADE names [30]. MedDRA® (the Medical Dictionary for Regulatory Activities terminology) is the international medical terminology developed under the auspices of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). The MedDRA® trademark is owned by the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) on behalf of the ICH. We excluded reports with > 10 drug ingredient names due to a high uncertainty in causal relationships involving > 10 ingredients. We included 12 million reports with ≤ 10 drug ingredient names. In our processed data, Chem-, Seq-, BP-, MF-, and CC-based similarity scores included all two-drug pairs among 2,659 drugs (i.e., Chem), 1,659 drugs (i.e., Seq), 692 drugs (i.e., BP), 692 drugs (i.e., MF), and 692 drugs (i.e., CC), respectively.

We included 445 relatively newer drugs with: (1) total reports ≤ 50 between 2004Q1 and 2008Q4, and ≥ 500 in 2022Q2; and (2) similarity scores with other drugs. We derived matched datasets for newer drugs to control for a confounding effect. For a newer drug, reports with and without the drug were matched on year/quarter (exact match), number of drugs in the report (exact match), and concomitant drugs in the report (best match). For instance, a report including a newer drug and two concomitant drugs (e.g.,

new drug + A + B) could be matched with other reports including three drugs (e.g., X + A + B, X + Y + B, X + Y + A, or X + Y + Z), in which "X + A + B" was defined as the best match. As we focused on early ADE detection, we followed newer drugs before their total reports exceeded 5,000. We selected the matching ratio as 1:100 for reports with versus without a newer drug. We computed percentages of matched concomitant medications.

2.3 Bayesian Drug Similarity-Based Adverse Drug Event (ADE) Detection

We used a Bayesian approach, as the posterior probability of the null hypothesis provided false-positive control in high-throughput mining [31, 32], especially for investigating a large amount of drug-ADE pairs [10]. We used the following notations. For a drug and an ADE, let N and X be the frequencies of the drug and the drug-ADE pair, respectively. Let θ be the probability of ADE (e.g., θ = Prob[ADE | drug]). Let θ^* be the reference value for testing θ, where θ^* could be determined by subject matter expert or in a data driven manner (e.g., θ^* = Prob[ADE | no drug]). Let $\mathrm{Bin}(X, N; \theta)$ denote binomial distribution, $\mathrm{Beta}(\theta; \alpha, \beta)$ denote beta distribution, $\mathrm{IB}(x; \alpha, \beta) = \int_0^x \mathrm{Beta}(\theta; \alpha, \beta) d\theta$ denote incomplete beta distribution, and $\mathrm{BB}(X, N; \alpha, \beta)$ denote beta-binomial distribution. Let subscript i indicate the i th drug, subscript j indicate the j th ADE, and subscript (k) indicate the k th time point (e.g., quarters after initial drug reporting in FAERS data).

We assumed $\theta_{ij(k)} \leq \theta^*_{ij(k)}$ under the null hypothesis (e.g., H_0 : Prob[ADE | drug] $\leq$ Prob[ADE | no drug]). We assumed $X_{ij(k)} \sim \mathrm{Bin}\left(X_{ij(k)}, N_{ij(k)}; \theta_{ij(k)}\right)$ and $\theta_{ij(k)} \sim \mathrm{Beta}\left(\theta_{ij(k)}; \alpha_{ij(k)}, \beta_{ij(k)}\right)$ (i.e., the prior distribution). For a new drug with similarity scores, let $X_{h(k)}$ and $N_{h(k)}$ be the corresponding frequencies of the similar drugs. Parameters in the prior distribution could be estimated by optimizing the objective function based on beta-binomial distributions (Eq. 2), where box-constrained optimization could be used to control the influence of prior distribution or prevent potential overfitting in automatic high-throughput mining.

$$\hat{\alpha}_{ij(k)}, \hat{\beta}_{ij(k)} = \underset{[\alpha_{ij(k)}, \beta_{ij(k)}] \in \mathbb{C}}{\arg\max} \sum_h \log \mathrm{BB}\left(X_{h(k)}, N_{h(k)}; \alpha_{ij(k)}, \beta_{ij(k)}\right) \tag{2}$$

We defined the posterior probability of the null hypothesis as:

$$\begin{aligned} &\mathrm{Prob}\left(\theta_{ij(k)} \leq \theta^*_{ij(k)} | X_{ij(k)}, N_{ij(k)}\right) = q_{ij(k)} \\ &= \mathrm{IB}\left(\theta^*_{ij(k)}; \hat{\alpha}_{ij(k)} + X_{ij(k)}, \hat{\beta}_{ij(k)} + N_{ij(k)} - X_{ij(k)}\right) \end{aligned} \tag{3}$$

We defined the signal as $q_{ij(k)}$ <0.05 and $X_{ij(k)}$ >0. The average of $q_{ij(k)}$s at times of detection estimated the false discovery rate (FDR) for mining all newer drugs and ADEs [10].

2.4 Performance Evaluation

We constructed a benchmark including 170,386 drug-labeled drug-ADE pairs according to SIDER and Demner-Fushman et al. [33, 34], where PT terms were available as ADE names. In real-world data-based performance evaluation, we used benchmark drug-ADE pairs as true positives, and non-benchmark drug-ADE pairs as true negatives. We summarize the performance evaluation analyses in the rest of this section, and present details (e.g., reference methods, an additional real-world dataset, and simulation procedures) in the Online Supplemental Material, S1.

First, we analyzed the processed FAERS data. The tested method was drug similarity-based posterior probability of the null hypothesis (equation 3) derived from top-25 similar drugs. The reference methods (Online Supplemental Material, S1.1) included empirical Bayes geometrical mean under multi-item gamma Poisson shrinker (EBGM(MGPS)) [6], the information component under a Bayesian confidence propagation neural network (IC(BCPNN)) [7], positive false discovery rate under MGPS (pFDR(MGPS)) [35], and posterior probability of the null hypothesis under MGPS (PP(MGPS)) [10]. As only the tested method, pFDR(MGPS), and PP(MGPS) provided false-positive control [10, 35], we detected ADE signals with FDR < 0.05 using these methods, and computed times to detection. We compared the tested method, pFDR(MGPS) and PP(MGPS) on time-to-detection. We also compared all methods on prioritizing the benchmark ADEs (i.e., area under the curve (AUC)).

Second, we analyzed ~1.5 million potential ADE cases derived from Optum's de-identified Clinformatics® Data Mart Database (Online Supplemental Material, S1.2) [36]. The data included four types of ADEs (e.g., acute myocardial infarction, acute kidney injury, fall, and hypoglycemia) and 358 relative newer drugs in January 2008. The tested method was drug similarity-based posterior probability of the null hypothesis (equation 3) derived from the top-25 similar drugs. The reference methods (Online Supplemental Material, S1.2) included the positive false discovery rate under the precision mixture risk model (pFDR(PMRM)) [35], and posterior probability of null hypothesis under PMRM (PP(PMRM)) [36]. We detected ADE signals with FDR < 0.05 under all approaches, and computed times to detection. We compared the tested method with the reference methods on time-to-detection and prioritizing the benchmark ADEs (i.e., area under the curve (AUC)).

Third, we conducted simulation studies. The simulation procedures are described in Online Supplemental Material S1.3. In short, we used the estimated prior distributions of ADE probabilities and the marginal reporting frequencies in FAERS data analysis for simulation probabilities of ADEs (i.e., $\theta_{ij(k)} \sim \text{Beta}\left(\theta_{ij(k)};\widehat{\alpha}_{ij(k)},\widehat{\beta}_{ij(k)}\right)$ and frequencies of drug-ADE pairs $X_{ij(k)} \sim \text{Bin}\left(X_{ij(k)},N_{ij(k)};\theta_{ij(k)}\right)$). We determined condition positives and negatives by comparing the simulated ADE probabilities with the corresponding ADE reporting rates in FAERS data (i.e., the null hypothesis). As similarity-based prior risk distributions could be mis-specified, we further assumed 20% of the prior risk distributions could be mis-specified. For the tested method, we used true parameters (i.e., correctly specified prior risk distributions) and randomly selected parameters (e.g., mis-specified prior risk distributions) to compute the drug similarity-based posterior probability of the null hypothesis depending on randomly simulated binary indicators with a probability of success = 0.8. For reference methods, we computed pFDR(MGPS) and PP(MGPS). We detected signals with FDR < 0.05 with all methods (i.e., predicted positives). We conducted 10,000 simulations. We computed the averages of FDR, false omission rate, probability of detection, specificity, and AUC.

3 Results

3.1 FAERS Data

The FAERS dataset included 445 relatively newer drugs. For these drugs, the medians and interquartile ranges (IQRs) were: (1) 2363 (1028–4614) for the total number of reports at the end of follow-up, and (2) 40% (6–62%) for reports including only one drug. For reports including new drugs and comedication(s), the median of matched concomitant medications in the matched dataset was 33% (IQR: 22–40%). The tested chemical substructure-based posterior probability of the null hypothesis (PP_Chem) identified 3.7% of all investigated drug-ADE pairs as signals. For reference methods, the MGPS-based posterior probability of the null hypothesis (PP(MGPS)) and the positive false discovery rate (pFDR(MGPS)) identified 1.5% and 0.6% of all investigated drug-ADE pairs as signals, respectively. The expected prior ADE probability based on chemical similarity to the observed ADE probability in FAERS dataset ratio had medians = 1.5 (IQR: 0.9–2.4) and 0.8 (IQR: 0.2–2.7) for benchmark and non-benchmark drug-ADE pairs, respectively (Mann-Whitney test $p < 0.01$). For other tested types of drug similarity scores, the percentages of signals among all investigated drug-ADE pairs were 3.8–4.0% for the posterior probability of the null hypothesis based on protein sequences (PP_Seq), biological processes (PP_Seq), molecular functions (PP_BP), and cellular components (PP_MF). Full results of ADE signals and their detection times for all methods are presented in Table S1 (Online Supplemental Material).

3.2 Time to ADE Detection

We evaluated the time-to-detection for ADEs defined by drug labels [33, 34]. Figure 1 shows the cumulative incidence of detection following initial drug reporting in the FAERS data analysis. We identified the tested PP_Chem had a faster time-to-detection compared with reference methods [e.g., PP(MGPS) and pFDR(MGPS)] (Fig. 1). For instance, the tested PP_Chem and the reference PP(MGPS) used 5 and 7 years to achieve the same probability of detection for drug-labeled ADEs (e.g., a 2-year reduction; Fig. 1). We used two drug-labeled ADEs (e.g., cabazitaxel-induced leukopenia and iloperidone-induced urinary incontinence) to illustrate the use of the drug similarity-based Bayesian posterior probability of the null hypothesis for monitoring signals of ADE sequentially in FAERS data [37, 38]. The tested PP_Chem identified cabazitaxel-induced leukopenia and iloperidone-induced urinary incontinence at 12 and 3 quarters after initial drug reporting (Fig. 2), respectively. The reference methods (e.g., PP(MGPS) and LFDR(MGPS)) compared with the tested approach used ≥ 3 quarters to identify these ADEs (Fig. 2). In Clinformatics® data analysis, the tested PP_Chem compared with reference methods also used a shorter time to achieve the same probability of detection for drug-labeled ADEs (Fig. S2A).

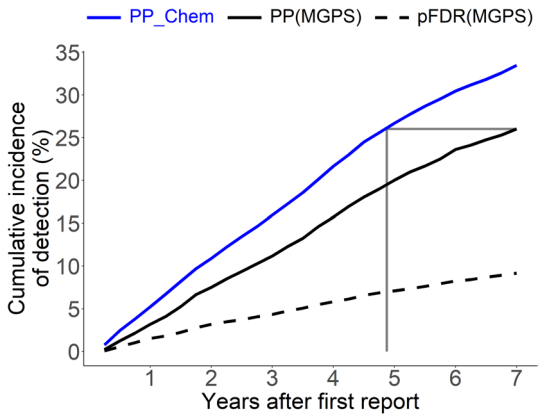


Fig. 1 Cumulative incidence from initial drug reporting to detection for adverse drug events defined by drug labels (*PP_Chem* chemical similarity-based posterior probability of null hypothesis, *PP(MGPS)* posterior probability of null hypothesis under multi-item gamma Poisson shrinker, *pFDR(MGPS)* positive false discovery rate under multi-item gamma Poisson shrinker)

3.3 Additional Performance Evaluation Analyses

In FAERS analysis, the tested PP_Chem had higher AUC values compared with reference methods (e.g., AUC: 0.57–0.79 vs. 0.32–0.71; Fig. 3A), especially in the first 3 years following initial reporting. All types of drug similarity-based posterior probabilities had comparable AUC values over time (Fig. 3B). In Clinformatics® data analysis, the tested PP_Chem also had higher AUC values compared with reference methods (Fig. S2B (Online Supplemental Material)).

In the simulation study, the tested similarity-based PP (PP_sim), and the reference methods (e.g., PP(MGPS) and pFDR(MGPS)) were all able to control the FDR (Fig. 4A). Specifically, the empirical values of FDR for PP_sim were 0.01% for an average model-based FDR = 0.05% (e.g., model-based FDR < 0.1%), 0.4% for an average model-based FDR = 0.6% (e.g., model-based FDR in 0.1–1%), and 2.4% for an average model-based FDR = 3% (e.g., model-based FDR in 1–5%). Additionally, the tested method compared to reference methods had better values on the probability of detection (0.31–0.60 vs. 0.11–0.41), false omission rate FOR; 0.08–0.11 vs. 0.12–0.15) and AUC [0.88–0.95 vs. 0.69–0.86 (Fig. 4B–D)]. All methods had specificity ≥ 0.99.

4 Discussion

In this work, we tested a drug similarity-based Bayesian method for early ADE detection. Under the tested method, prior distribution of ADE probability of a less frequently used drug could be derived from biochemically similar and more frequently used drugs, and the posterior probability of the null hypothesis (PP) could be used for detecting ADEs and controlling for false-positive rate. We analyzed relatively newer drugs in real-world data (e.g., FAERS data), as well as conducted performance evaluation analyses. The tested drug similarity-based PP might be able to accelerate ADE detection with false-positive control.

We identified signals of ADEs for newer drugs with a lower frequency in FAERS data. Similarity-based PP detects more ADE signals with false-positive control [10, 31, 32]. Additionally, similarity-based PP can accelerate ADE detection (Figs. 1 and 2). For instance, the similarity-based PP might be able to detect cabazitaxel-induced leukopenia and iloperidone-induced urinary incontinence (e.g., drug-labeled ADEs) within 3 years after initial drug reporting [37, 38], while the non-drug-similarity-based methods compared with a similarity-based method uses ≥ 3 quarters to identify these ADEs (Fig. 2).

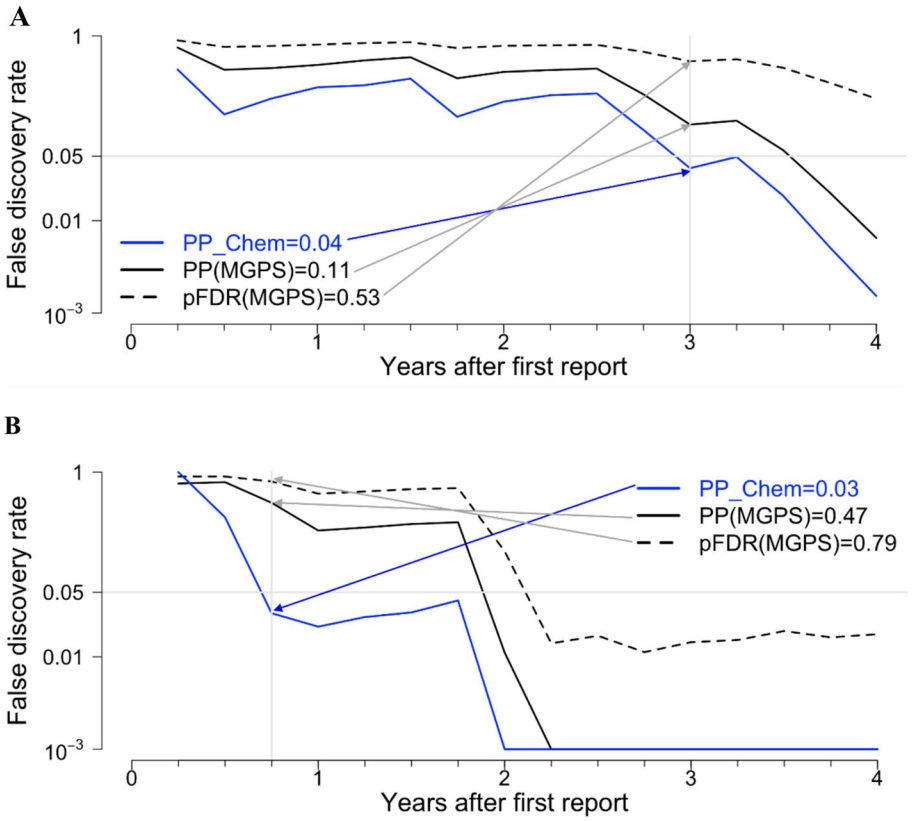


Fig. 2 Case studies: **A** Cabazitaxel and leukopenia. **B** Iloperidone and urinary incontinence

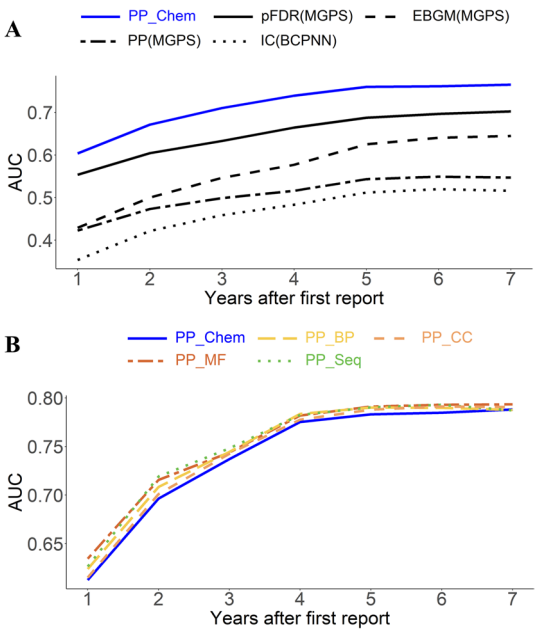


Fig. 3 Area under the ROC curve (AUC). **A** Evaluation of chemical similarity-based posterior probability (PP_Chem); multi-item gamma Poisson-based (MGPS-based) posterior probabilities (PP(MGPS)), positive false discovery rate (pFDR(MGPS)) and empirical Bayes geometric mean (EBGM(MGPS)); and Bayesian confidence propagation neural network-based (BCPNN-based) information component (IC(BCPNN)). **B** Evaluation of PP_Chem, protein sequences-based PP (PP_Seq), biological processes-based PP (PP_BP), molecular functions-based PP (PP_MF), and cellular components-based PP (PP_CC)

Our performance evaluation analysis shows that similarity-based PP has better AUC values, especially in the first few years after initial drug reporting in FAERS (Figs. 3A), as well as proper false-positive control (Fig. 4A). Thus, similarity-based PP has better performance in mining all newer drugs and all ADEs while controlling the false-positive rate. Similarity-based PP has a better performance for a lower drug reporting frequency, as it leverages drug similarity-based prior distributions. Our findings agree with the existing literature that drug similarity-enhanced signals have an improved performance [15, 16]. However, drug similarity-based PP compared with drug similarity-enhanced signal provides the much-desired false-positive control [10, 31, 32]. Both drug similarity-based PP and non-drug similarity-based PP (e.g., PP under multi-item gamma Poisson shrinker) provide theoretical false-positive control under the Bayesian framework, and can reduce potential false positives by penalizing signals with a lower frequency [10]. However, drug similarity-informed prior distribution might be able to accelerate ADE detection, as drug similarity contributes relevant information to ADE detection [15, 16]. In addition, compared with fiscal year-based evaluation in existing literature [15, 16], our time-to-event-based evaluation provides a more comprehensive and explicit longitudinal relationship between time after initial observation and performance of ADE detection (Figs. 1 and 2). Moreover, similarity-based PP has a better probability of detection (Fig. 4B), which provides reassurance that similarity-based PP identifies more signals than reference methods (Fig. 1). We also find that

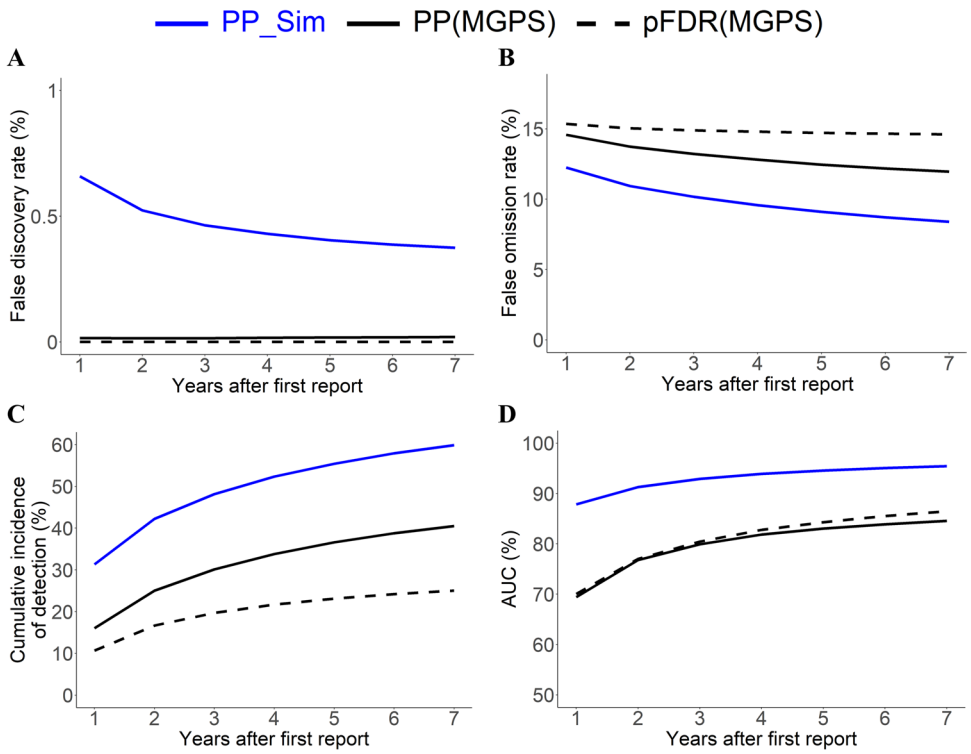


Fig. 4 Simulation results for similarity-based posterior probability (PP_Sim), and multi-item gamma Poisson shrinker-based (MGPS-based) posterior probabilities (PP(MGPS)) and positive false discovery rate (pFDR(MGPS)). **A** False discovery rate. **B** False omission rate. **C** Probability of detection. **D** Area under the curve

different types of drug similarities have a comparable performance (Fig. 3C), which implies drug-related chemical structures, proteins, and genes may have a similar power for predicting ADEs in a high-throughput setting.

The drug similarity-based Bayesian method for ADE detection has limitations and challenges. First, our FAERS analysis is subject to biases due to under-reporting, selective reporting, and unfair sampling. As administrative claims data and electronic health records (EHR) data are less subject to reporting bias, mining of these data sources promises generation of high-confidence signals. While the drug similarity-based ADE detection approach can be used to mine claims and EHR data, ADE phenotypes are not well defined in these data sources compared with FAERS data. We expect more claim- and EHR-based ADE phenotyping algorithms to be developed. Second, our FAERS analysis is based on MedDRA® PT terms. Using PT compared with High Level Terms (HLTs) provides a higher resolution, but may compromise the clinical relevance and/or statistical power for ADE detection. Third, suboptimal prior distributions and comparisons and prior-data conflicts could occur in automatic high-throughput mining of large-scale real-world data. Reasons for these challenges include lack of similarity, imperfect selections of similar drugs and null hypotheses, and model underfitting and overfitting. Unlike automatic high-throughput data mining, the second and third limitations can be addressed in carefully designed clinical and real-world data-based studies with a reasonable number of hypotheses (e.g., testing drug-ADE pairs), in which all of the prior distributions, ADE outcomes, and comparisons can be delicately derived and rigorously adjudicated [39–42]. Additionally, we anticipate that new research will refine ADE outcomes, comparisons, and drug similarity-based prior distributions and their relationships between observed data in a high-throughput mining setting. Fourth, the ADE benchmark derived from SIDER and Demner-Fushman et al. might not be comprehensive [33, 34]. We use SIDER and Demner-Fushman et al. as these databases include a significant number of drug-ADE pairs [33, 34], and have been previously used to evaluate drug similarity-enhanced signals [15, 16]. Alternatively, other gold standards either only include a small number of drug-ADE pairs and/or are subject to mislabel [43, 44], especially for relative newer drugs. We hope more comprehensive benchmarking for pharmacovigilance studies to be generated in future. Fifth, our simulation settings could deviate from real clinical scenarios. Given these limitations, the signals generated in our study should be further validated and adjudicated, and should not be considered as true adverse drug reactions. Finally, we would like to point out that integrated mining of non-clinical drug features, food features, and clinical outcomes is an important direction for future research, as recent works demonstrate machine learning of non-clinical drug and food features can identify adverse drug-drug and drug-food interactions [12, 13].

5 Conclusion

A drug-similarity based Bayesian method might be able to accelerate ADE detection while controlling the false-positive rate.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40264-025-01545-6.

Declarations

Funding This work was supported by NIH (R01GM141279).

Conflicts of Interest The authors declare no conflict of interest.

Ethics Approval Not applicable. No human data were involved in this study. The Indiana University Institutional Review Board (IRB) designated this study as exempt.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Availability of Data and Materials FAERS data are available from the US Food and Drug Administration (FDA, https://www.fda.gov/drugs/fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard). Optum's de-identified Clinformatics® Data Mart Database is available from Optum (optum.com). SIDER data is available from SIDER 4.1: Side Effect Resource (http://sideeffects.embl.de/). Data from Demner-Fushman et al. is available form doi: https://doi.org/10.1038/sdata.2018.1. Summary statistics generated during this study are included in this published article and its supplementary files.

Code Availability Model codes are provided in the Online Supplementary files and GitHub (https://github.com/PengyueLab/Drug_Sim_ADE_Mining).

Author Contributions PeZ and PiZ conceived the study; YS, YY, RL, AS and XP conducted data analysis; YS, YY, PeZ and PiZ drafted the manuscript. All authors interpreted the results, critically revised, and gave final approval of the manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

References

1. Santoro A, Genov G, Spooner A, Raine J, Arlett P. Promoting and protecting public health: how the European Union Pharmacovigilance System Works. Drug Saf. 2017;40(10):855–69.
2. Food and Drug Administration (FDA). FDA Adverse Events Reporting System (FAERS) Public Dashboard. 2022 [cited 2022/07/05; Available from: https://fis.fda.gov/sense/app/95239e26-e0be-42d9-a960-9a5f7f1c25ee/sheet/7a47a261-d58b-4203-a8aa-6d3021737452/state/analysis. Accessed 12 Jan 2024.
3. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf. 2001;10(6):483–6.
4. van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf. 2002;11(1):3–10.
5. Huang L, Zalkikar J, Tiwari RC. Likelihood ratio test-based method for signal detection in drug classes using FDA's AERS database. J Biopharm Stat. 2013;23(1):178–200.
6. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. Am Stat. 1999;53(3).
7. Bate A. Bayesian confidence propagation neural network. Drug Saf. 2007;30(7):623–5.
8. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med. 2012;4(125):125ra31.
9. Zhang P, Li M, Chiang CW, Wang L, Xiang Y, Cheng L, et al. Three-component mixture model-based adverse drug event signal detection for the adverse event reporting system. CPT Pharmacomet Syst Pharmacol. 2018;7(8):499–506.
10. Ahmed I, Haramburu F, Fourrier-Reglat A, Thiessard F, Kreft-Jais C, Miremont-Salame G, et al. Bayesian pharmacovigilance signal detection methods revisited in a multiple comparison setting. Stat Med. 2009;28(13):1774–92.
11. Cheng F, Li W, Wang X, Zhou Y, Wu Z, Shen J, et al. Adverse drug events: database construction and in silico prediction. J Chem Inf Model. 2013;53(4):744–52.
12. Kha QH, Le VH, Hung TNK, Nguyen NTK, Le NQK. Development and validation of an explainable machine learning-based prediction model for drug-food interactions from chemical structures. Sensors (Basel). 2023;23(8).
13. Le NQK. Predicting emerging drug interactions using GNNs. Nat Comput Sci. 2023;3(12):1007–8.
14. Vilar S, Tatonetti NP, Hripcsak G. 3D pharmacophoric similarity improves multi adverse drug event identification in pharmacovigilance. Sci Rep. 2015;6(5):8809.
15. Ji X, Cui G, Xu C, Hou J, Zhang Y, Ren Y. Combining a pharmacological network model with a Bayesian signal detection algorithm to improve the detection of adverse drug events. Front Pharmacol. 2021;12: 773135.
16. Liu R, Zhang P. Towards early detection of adverse drug reactions: combining pre-clinical drug structures and post-market safety reports. BMC Med Inform Decis Mak. 2019;19(1):279.
17. Raies AB, Bajic VB. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Comput Mol Sci. 2016;6(2):147–72.
18. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. Nucleic Acids Res. 2016;44(D1):D1202–13.
19. Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bulletin de la Société vaudoise des sciences naturelles. 1901;37:547–79.
20 Knox C, Wilson M, Klinger CM, Franklin M, Oler E, Wilson A, et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res. 2024;52(D1):D1265–75.
21 Kitsiranuwat S, Suratanee A, Plaimas K. Integration of various protein similarities using random forest technique to infer augmented drug-protein matrix for enhancing drug-disease association prediction. Sci Prog. 2022;105(3):368504221109215.
22. Pagès H, Aboyoun P, Gentleman R, DebRoy S. Biostrings: Efficient manipulation of biological strings. 2024 [cited; Available from: https://bioconductor.org/packages/Biostrings. Accessed 12 Jan 2024.
23 Pearson WR. Selecting the Right Similarity-Scoring Matrix. Curr Protoc Bioinformatics. 2013;43:351–9.
24. Gene Ontology C, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The Gene Ontology knowledgebase in 2023. Genetics. 2023;224(1).
25. Wang JZ, Du Z, Payattakool R, Yu PS, Chen CF. A new method to measure the semantic similarity of GO terms. Bioinformatics. 2007;23(10):1274–81.
26. Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics. 2010;26(7):976–8.
27. Banda JM, Evans L, Vanguri RS, Tatonetti NP, Ryan PB, Shah NH. A curated and standardized adverse drug event resource to accelerate drug safety research. Sci Data. 2016;10(3): 160026.
28. Wang L, Shendre A, Chiang CW, Cao W, Ning X, Zhang P, et al. A pharmacovigilance study of pharmacokinetic drug interactions using a translational informatics discovery approach. Br J Clin Pharmacol. 2022;88(4):1471–81.
29 Nelson SJ, Zeng K, Kilbourne J, Powell T, Moore R. Normalized names for clinical drugs: RxNorm at 6 years. J Am Med Inform Assoc. 2011;18(4):441–8.
30 Lindstrom-Gommers L, Mullin T. International Conference on harmonization: recent reforms as a driver of global regulatory harmonization and innovation in medical products. Clin Pharmacol Ther. 2019;105(4):926–31.
31. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst. 2004;96(6):434–42.
32. Wakefield J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am J Hum Genet. 2007;81(2):208–27.
33. Side Effect Resource (SIDER 4.1). [cited; Available from: http://sideeffects.embl.de/. Accessed 12 Jan 2024.
34. Demner-Fushman D, Shooshan SE, Rodriguez L, Aronson AR, Lang F, Rogers W, et al. A dataset of 200 structured product labels annotated for adverse drug reactions. Sci Data. 2018;30(5):180001.
35. Storey JD. The positive false discovery rate: a Bayesian interpretation and the q-value. Ann Stat. 2003;31(6):2013–35.
36. Shi Y, Eadon MT, Chen Y, Sun A, Yang Y, Chiang C, et al. A precision mixture risk model to identify adverse drug events in subpopulations using a case-crossover design. Stat Med. 2024;43(27):5088–99.
37. Food and Drug Administration (FDA). Drug label of cabazitaxel. [cited; Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.accessdata.fda.gov/drugsatfda_docs/label/2010/201023lbl.pdf. Accessed 12 Jan 2024.
38. Food and Drug Administration (FDA). Drug label of iloperidone. [cited; Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/022192lbl.pdf. Accessed 12 Jan 2024.
39. Goligher EC, Tomlinson G, Hajage D, Wijeysundera DN, Fan E, Juni P, et al. Extracorporeal membrane oxygenation for severe

acute respiratory distress syndrome and posterior probability of mortality benefit in a post hoc bayesian analysis of a randomized clinical trial. JAMA. 2018;320(21):2251–9.
40. Muehlemann N, Zhou T, Mukherjee R, Hossain MI, Roychoudhury S, Russek-Cohen E. A tutorial on modern Bayesian methods in clinical trials. Ther Innov Regul Sci. 2023;57(3):402–16.
41. Stefan AM, Gronau QF, Schonbrodt FD, Wagenmakers EJ. A tutorial on Bayes Factor Design Analysis using an informed prior. Behav Res Methods. 2019;51(3):1042–58.
42. van de Schoot R, Broere JJ, Perryck KH, Zondervan-Zwijnenburg M, van Loey NE. Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. Eur J Psychotraumatol. 2015;6:25216.
43. Hauben M, Aronson JK, Ferner RE. Evidence of misclassification of drug-event associations classified as gold standard "negative controls" by the observational medical outcomes partnership (OMOP). Drug Saf. 2016;39(5):421–32.
44. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. Drug Saf. 2013;36(Suppl 1):S33-47.