DIA

Therapeutic Innovation
& Regulatory Science
2018, Vol. 52(3) 294-299

DOI: 10.1177/2168479017728986
tirs.sagepub.com

An Alternative to Disproportionality: A Frequency-Based Method for Pharmacovigilance Data Mining

Jeremy D. Jokinen, PhD[1], Fabio Lievano, MD[1], Linda Scarazzini, MD[1], and Melissa Truffa, RPh[1]

Abstract

Background: Safety surveillance relies on mining of large pharmacovigilance (PV) databases to generate insights regarding the safe use of pharmaceutical products. The predominant approach to PV data mining involves computation of disproportionality scores for drug–adverse event (drug-AE) pairs. However, this approach requires a database to be sufficiently large, sufficiently diverse for the analysis to be reliably sensitive and specific, and fails to consider the particular safety profile of a product. *Objective:* The present study proposes and tests a novel, frequency-based approach to PV data mining that (1) leverages product knowledge and historical drug-AE trends and (2) imposes no requirement for the size and diversity of the database to which it is applied. *Method:* A focus group of physicians and scientists was convened to identify quantitative characteristics of data trends that they consider informative when reviewing counts of adverse events for products under surveillance. Feedback was transferred into a series of decision rules that, when applied to adverse event counts, identifies adverse event trends that are classified as Continuing Trend, Emerging Trend, or No Trend. Regression analyses are completed to verify the presence of a linear trend; and categorical measures of association completed to compare this frequency-based approach to disproportionality scores in a simulated database. *Results:* A significant, positive linear trend is present for the Continuing Trend and Emerging Trend categories ($P < .0001$). There is a significant association between trend categorizations and disproportionality scores ($P < .0001$). *Conclusion:* The proposed alternative frequency-based method for PV data mining would be useful where disproportionalities scores are not appropriate. Additionally, this method may be useful in conjunction with disproportionality scores, where appropriate, highlighting adverse events that are both reported disproportionately and have increasing trends.

Keywords

pharmacovigilance, data mining, signal detection, safety statistics, disproportionality

Introduction

Data mining of longitudinal databases is an analytical task common to disciplines as diverse as public health; engineering; and the biological, political, and social sciences.[1] Within the field of pharmacovigilance (PV), mining of spontaneously reported adverse event databases is a key source of on-market information regarding the safety of medicinal products.[2] Databases with continually accumulating adverse event information are maintained by pharmaceutical companies and regulatory authorities globally. These databases serve as valuable resources to ensure public health; however, the volume of information contained therein can make meaningful use of these data challenging.[3] As a result, methods for mining these databases have been a topic of regulatory, academic, and industry research for over 20 years.[4-6]

The most common approach to mining PV databases is the use of disproportionality measures.[6,7] Disproportionality measures refer to a group of statistical methods that compare the proportion of occurrences of a target adverse event for a given drug (foreground) to the proportion of occurrences of the same adverse event for all other drugs in the database (background). In effect, the background proportion serves as the "expected value" against which the foreground proportion is compared. The computational specifics for disproportionality measures, such as PRR, EBGM, ROR, etc. have been described in numerous publications.[8-10] If the foreground is substantially greater than the expected, background, proportion, the disproportionality measure is substantially greater than, 1.0 indicating disproportionate reporting for the drug-AE pair. Whether the disproportionality is considered a potential safety concern requiring further investigation depends upon the medical and

[1] Abbvie, Inc, North Chicago, IL, USA

Submitted 15-Feb-2017; accepted 7-Aug-2017

Corresponding Author:
Jeremy Jokinen, PhD, Abbvie, Inc, GM60, AP51-2, 1 N Waukegan Rd, North Chicago, IL 60064-1802, USA.
Email: jeremy.jokinen@abbvie.com

safety expertise being applied to the review of the data and the a priori threshold for determining if a disproportionality value is extreme enough to be considered a signal of disproportionate reporting (SDR).

A major concern regarding disproportionality measures is the appropriateness of the comparison of the foreground proportion to the background proportion. One needs to consider if the drug-AE in the foreground is reasonable to compare to the background drugs and adverse events. It is generally believed that disproportionality measures are acceptable data mining analyses if the database is suitably large and suitably diverse.[8,9] If so, the background proportion can be considered an estimate of the occurrence of this AE in the general population, and therefore the foreground is a comparison to an estimate of the proportion experiencing an AE in the general population. This requirement for size and diversity may be met with large regulatory databases, such as FDA FAERS or WHO VigiBase, but it is unlikely to be met with data contained in a company-owned, proprietary, safety database. A company-owned database likely contains data for a limited number of products intended to treat patients from a limited subset of therapeutic areas, which can cause disproportionality measures to produce unpredictable results (J. D. Jokinen and M. Henderson, unpublished data, 2016). For example, a consumer product safety database might contain a large amount of safety data for an over-the-counter pain medication, but the background in this database might be composed of allergy medications, cosmetics, or topical sunscreens. The potential SDR for the pain reliever is dependent upon sunscreen safety reports and allergies and, more generally, dependent upon seasonal fluctuations in reporting of these background products and events. In these databases, one needs to carefully consider whether it is reasonable to compare foreground to background using disproportionality measures.

The present study proposes a data mining algorithm that is not dependent upon this foreground-to-background comparison. The approach results in an algorithm based on the frequency of reports over time and leverages insights from the medical and scientific subject matter experts (SMEs). In this frequency-based approach, the comparison is not foreground to background proportion, but rather the count of adverse events for a product in the current time period relative to previous time periods. A focus group of medical and scientific safety experts were convened to identify quantitative characteristics that are of potential clinical significance when reviewing safety data. The decision was made to focus on quantitative aspects, for example, changes in the frequency of adverse event reports, as opposed to qualitative aspects, such as adverse event seriousness or source of adverse event report. Some authors have suggested that quantitative and qualitative aspects play complementary roles in safety surveillance.[11] Therefore, we approached these topics independently in order to provide flexibility in combining quantitative and qualitative aspects across the entire surveillance ecosystem in the future.

The following describes our approach to build a judgment-based algorithmic approach, a sieve analysis, against which data from a spontaneous report database are evaluated. Research exists demonstrating the value of statistical approaches that prioritize potential signals for further evaluation of medical SMEs.[12,13] Our approach is to extend this thinking by making the prioritization criteria the analysis itself, by translating SME judgment into a series of filters, which results in a hierarchy of data for review. This analysis is similar to the geological practice of applying sieves to separate finer and finer layers of rock and soil.[14] The process, criteria for prioritization, and results when applied to a simulated subset of data are reviewed. Finally, the correspondence between this approach and proportional reporting ratios (PRRs), a standard disproportionality approach, is examined.

Method

Gather SME Input

Safety physicians and scientists for a major pharmaceutical company were engaged in a focus group regarding data they review periodically to ensure patient safety. The focus group was conducted by a doctoral-level psychologist with experience conducting qualitative data-collection interviews. The intention of the focus group is to identify characteristics within data that are significant to this group of SMEs and may potentially indicate a safety signal. Feedback was recorded as verbatim text, common themes were identified and summarized. For example, SMEs routinely removed from consideration any adverse events that have 0 reports for the current period of evaluation. This may seem self-evident, but it is not uncommon for a disproportionality analysis to flag for further review adverse events with no new information for the current evaluation period. Additionally, SMEs regard adverse events in the current period of evaluation as more informative or interesting if the number of events for the current quarter is large relative to the cumulative number of reports for a given adverse event. That is, the current data represent a large proportion of all data in the database for a given adverse event. Finally, if this reporting period included data for an adverse event that had not been reviewed previously, that adverse event would be interesting to review. From this discussion, the filters for the sieve analysis were developed and programmed to sort data into a series of priority levels. The filters were chosen such that the criteria are increasingly restrictive and only the most significant "must see" data flow to the far end of the sieve requiring priority review. For example, the requirement that there be at least one report in this reporting period is least restrictive, and many PTs meet this requirement, but the first time a PT has been observed for this drug is most restrictive, very few PTs meet this requirement, and it results in the highest priority.

Transferring SME Input to Data Filters—Sieve Analysis

Feedback from SMEs was translated by data scientists into a sequence of data filters. The data filters are specifically designed to (1) closely approximate the verbal descriptors

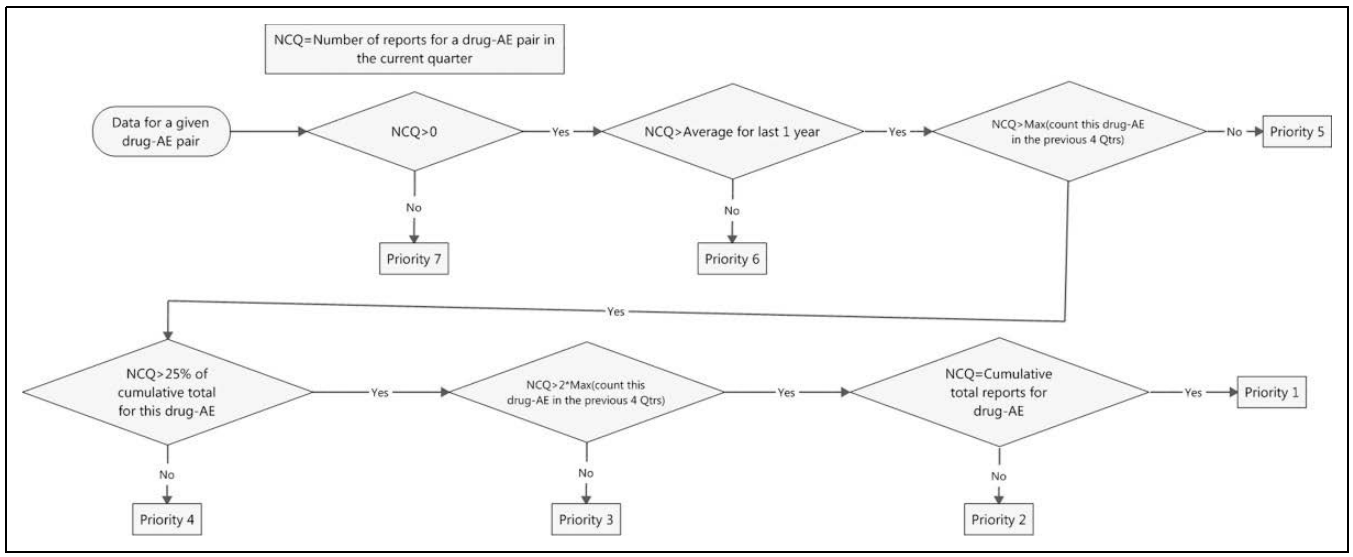

Figure 1. Prioritization according to the sieve analysis decision tree.

provided by the SMEs and (2) be ordered relative to the "importance" or "informativeness" of the data to the SMEs. In the present example, as data proceed from the top filters (Priority 7) to the bottom of the sieve (Priority 1), the drug-AE pair data are more likely to be of interest to the SME reviewers responsible for product safety as the data are more likely to represent an upward trend. Priority is assigned to the data according to the number of filters in the sieve that data points flow through. Figure 1 shows the resulting sieve to be applied to the adverse event data.

The decision points in the sieve analysis cover a variety of quantitative aspects of the data that medical SMEs have identified as being of keen interest. For example, the criterion "NCQ >25% of the cumulative total for this drug-AE" is intended to prioritize trends in data for drug-AE combinations where little historical data exist. If the number of reports in this quarter comprises more than 25% of the historical total in the database, this is a drug-AE combination for which a substantial proportion of the reporting has occurred in the recent time frame and therefore receives a higher prioritization. Twenty-five percent was selected arbitrarily based on discussion with the SMEs regarding how they would quantify the amount of evidence this quarter relative to the total ever collected that would "catch their attention." Similarly, if the current quarter is the first time this drug-AE combination has appeared in the data, the criterion "NCQ=Cumulative total reports for this drug-AE" results in a priority of 1 assigned to this pair. This filter was assigned the highest priority as all SMEs agreed that they would want to see the first report of an AE, even if ultimately the case was considered uninformative. Additionally, frequency analyses like those carried out in the current study might identify trends associated with increased volume of drug sales and patient exposure, rather than increasing drug-AE trends. Therefore, priority filters requiring a frequency greater than the average for the last year and greater than the maximum for the last 4 quarters are intended to prioritize drug-AE combinations that are increasing relative to any underlying increasing trend. The result of this analysis is fewer and fewer data points assigned to each subsequent priority level. Additionally, the relevance of those data points to the decision task—determination of an upward trend within the spontaneous report database—increases.

Following calculation of the 7 possible priority scores for each drug-AE pair for the current time period and the previous time period, the priorities are compared and the number of priorities simplified to aid interpretation of the results and action to be taken by safety review teams. The result is a 3-category system. The first priority, "Continuing Trend," contains drug-AE pairs that were scored a priority 1, 2, or 3 for both the current time period and the previous time period. In other words, these are drug-AE pairs that had an upward trend in the previous time period and the current one. "Emerging Trend" is assigned to drug-adverse pairs that were priority 4, 5, 6, or 7 in the previous time period and priority 1, 2, or 3 for the current time period. This trend is newly emerging as the previous time period data indicated little or no trend, but the current quarter does indicate an upward trend. All other combinations are indicated as "No Trend."

Data Simulation

A random data set of counts of adverse events from a live, spontaneous report database was extracted without adverse event or drug names included. The data set contained 80,000 rows of 8 quarter counts of the number of reports contained in the database, for a total of 640,000 quarter-count data points. Bootstrap resampling of these quarter counts was used to create a simulated database of 2 million rows of 8 quarters of data that resemble spontaneous report data contained within the original

Table 1. Percentage of Drug–Adverse Event Pairs by Priority.

Priority	Percent
Continuing trend	0.53
Emerging trend	4.81
No trend	94.66

Table 2. Summary of Linear Regression of Adverse Event Counts on Time Period.

Priority	Time Trend Parameter Value	*P* Value
Continuing trend	14.58	<.0001
Emerging trend	0.42	<.0001
No trend	0.06	<.0001

database. Finally, 100 mock drug names and 1000 mock adverse events were randomly assigned to the data rows.

Results

Descriptive analyses, counts and percentages, were examined to verify that the 3 priorities resulted in a smaller number of items for review. If the prioritization scheme failed to identify a subset for prioritized review, the approach would not be tenable for a surveillance program with limited resources to review data. Table 1 summarizes the percentage of the 2 million data rows that are categorized in each of the 3 priorities.

The prioritization reduced the number of drug-AE pairs for review to a subset of the original. The substantial majority of all data points are categorized as No Trend. However, additional analyses are required to verify that linear trends are present in the first 2 priority categories and absent in the last.

Linear regression analyses were conducted for each priority category to examine the extent to which the categorizations correctly identify an upward trend present in the data. For each analysis, adverse event counts were regressed on the time period. Because of the large number of data points, it is anticipated that all parameters would be statistically significant. Examination of the parameter for the time component provides information regarding the trend present in the data. Table 2 below summarizes the results of the regression analysis.

The greatest value for the time parameter is in the Continuing Trend subset. For every 1-unit increase in the time parameter, there is an average increase of 14.58 adverse events. For the No Trend subset, the time parameter is nearly 0. For every 1-unit increase in time, there is an average increase of 0.06 adverse events. The Emerging Trend results are between the 2 extremes.

In order to provide additional context for this novel analytics approach, the data were also analyzed using the typical PRR approach. Signals of disproportionate reporting (SDR) were identified using Evan's criteria.[15] PRR resulting in SDR were compared to the 3 priority categories to examine the correspondence between the sieve, frequency-based analysis, and disproportionality measures for each drug-AE pair. Table 3 summarizes the percentage of data in each category.

Table 3. Percentage of Each Priority Category that are PRR SDR/NoSDR.

	Percent	
Priority	No SDR	SDR
Continuing trend	36.32	63.68
Emerging trend	81.01	18.99
No trend	92.44	7.56

Table 3 demonstrates a strong correspondence between the trend categories of the sieve analysis and signals of disproportionate reporting as identified by PRR. Of the drug-AE pairs considered Continuing Trend, 63.68% were also identified as a signal of disproportionate reporting. Of the No Trend drug-AE pairs, only 7.56% were identified as signals of disproportionate reporting. Results of the Cochran-Armitage trend test are consistent with the observations of the percentage scores. The Cochran-Armitage trend test is significant ($Z = 97.18$, $P < .0001$), indicating a statistically significant association between the priority categories and the SDR/No SDR disproportionality results.

Conclusion and Discussion

Initial development of the frequency-based sieve analysis is promising. Results indicate that a small proportion of drug-AE pairs in large simulated databases is prioritized for review. Additionally, those that were prioritized demonstrated a significant, upward time trend in adverse event counts. Finally, the results of the prioritization are positively correlated with the results of the well-established PRR method for identifying signals of disproportionate reporting.

The sieve analysis methodology offered many additional benefits. First, the algorithm is easily implemented with minimal statistical and data science knowledge. Second, the results do not depend upon assumptions regarding the size and diversity of the database being analyzed. Finally, the authors found the process of working with medical and scientific SMEs to generate levels of the sieve analysis promoted necessary discussions regarding outcomes and characteristics of data that were of greatest interest. This discussion facilitated both the initial development and future refinements to the algorithm that are necessary for ongoing monitoring of spontaneous report data.[16] Refinements may include incorporating both quantitative and qualitative algorithms, which Caster et al have argued may improve overall signal detection.[11] Inclusion of qualitative measures may improve the positive predictive value of data mining efforts, and this statistical/medical SME collaboration may result in a more robust PV system overall.

This sieve analysis is a frequency-based approach to data mining. As such, the analysis is impacted by all factors that impact reporting frequency of adverse events to spontaneous

report databases. Factors such as product sales and stimulated reporting due to new product launches or other publicity effects can result in upward trends in spontaneous reports of adverse events. These upward trends might be signaled by frequency analyses as high-priority drug-AE pairs that should be evaluated as safety signals. It is possible that a combination of frequency-based analyses and other methods, such as disproportionality measures, could reduce false positives that may result from stimulated reporting. Future research should investigate the potential value of signals of disproportionality that also show increasing trends as identified by frequency-based methods.

Alternatives to this sieve approach, such as disproportionality measures for data mining, are successfully applied but should be interpreted with caution. Interpretation of disproportionality measures requires consideration of four major components: the disproportionality measure itself, the threshold for determination of a signal of disproportionate reporting, the foreground proportion, and the background proportion. In a form of denominator neglect, often the background is not thoroughly considered in the interpretation of disproportionalities.[17] SDR could result from changes in the foreground proportion, but could also result from changes in the background proportion. Changes in the background are not obvious in disproportionality measures and require careful consideration of the data overall. The sieve analysis approach, however, relies only on the time course of a specific adverse event for a given drug and not the background reporting of the adverse event of interest for other drugs.

A reasonable criticism of the sieve approach is a criticism common to all data mining efforts applied to spontaneous report databases—the sieve analysis will never eliminate the need for SME review in judgments of safety and benefit risk. It is important to remember that frequency-based analyses, and PV data mining in general, are statistical approaches to what is inherently a medical question: Is there a safety issue present in the data? Ultimately, human expertise is required to review complex issues such as potential drug interactions or the biological plausibility of adverse events for a particular drug. However, the sieve analysis provides a prioritization that can triage the most critical data and allow experts to budget time and resources accordingly.[12,13]

The proposed method for frequency-based data mining may play an important role in a company's safety surveillance ecosystem. The analyses presented above demonstrate that the results are consistent with, but complementary to, traditional disproportionality measures. These 2 approaches combined may yield useful information regarding adverse events that are being reported both disproportionately and with increasing frequency. Additionally, augmenting these quantitative approaches with a qualitative approach such as the vigiRank approach may start to create an algorithmic approach that more closely aligns with human judgment which integrates both the quantitative and the qualitative aspects of a signal.[11,18]

Declaration of Conflicting Interests

All authors are employees of AbbVie, Inc.

Funding

No financial support of the research, authorship, and/or publication of this article was declared.

References

1. Frisén M. Methods and evaluations for surveillance in industry, business, finance, and public health. *Q Reliabil Eng Int*. 2011;27:611-621.
2. Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. *Clin Pharmacol Ther*. 2013;93:539-546.
3. Hauben M, Madigan D, Gerrits CM, Walsh L, Van Puijenbroek EP. The role of data mining in pharmacovigilance. *Expert Opin Drug Saf*. 2005;4:929-948.
4. Chan KA, Hauben M. Signal detection in pharmacovigilance: empirical evaluation of data mining tools. *Pharmacoepidemiol Drug Saf*. 2005;14:597-599.
5. Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiol Drug Saf*. 2009;18:427-436.
6. Quattrini G, Zambon A, Simoni L, Fiori G. Disproportionality measures used in signal detection: an assessment on pharmacovigilance adverse event reporting system data. *Value Health*. 2015;18:A720.
7. Hauben M, Zhou X. Quantitative methods in pharmacovigilance. *Drug Saf*. 2003;26:159-186.
8. Almenoff JS, LaCroix KK, Yuen NA, Fram D, DuMouchel W. Comparative performance of two quantitative safety signalling methods: implications for use in a pharmacovigilance department. *Drug Saf*. 2006;29:875-887.
9. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. *Drug Saf*. 2003;26:159-186.
10. Jokinen JD. Determination of change in online monitoring of longitudinal data: An evaluation of methodologies [dissertation]. Athens, OH: Ohio University; 2015.
11. Caster O, Juhlin K, Watson S, Noren GN. Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank. *Drug Saf*. 2014;37:617-628.
12. Levitan B, Yee C, Russo L, Bayney R, Thomas A, Klincewicz S. A model for decision support in signal triage. *Drug Saf*. 2008;31:727-735.
13. Stahl M, Lindquist M, Edwards IR, Brown EG. Introducing triage logic as a new strategy for the detection of signals in the WHO Drug Monitoring Database. *Pharmacoepidemiol Drug Saf*. 2004;13:355-363.
14. McCave I, Syvitski JP. Principles and methods of geological particle size analysis. In: Syvitski JP, ed. *Principles, Methods and Application of Particle Size Analysis*. Cambridge: Cambridge University Press; 1991:3-21.

15. 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;10:483-486.
16. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) Module IX Addendum I—methodological aspects of signal detection from spontaneous reports of suspected adverse reactions 2016. Updated August 4, 2016. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2016/08/WC500211715.pdf. Accessed September 22, 2016.
17. Reyna VF, Brainerd CJ. Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learn Indiv Differ*. 2008;18:89-107.
18. Bergvall T, Noren GN, Lindquist M. vigiGrade: a tool to identify well-documented individual case reports and highlight systematic data quality issues. *Drug Saf*. 2014;37:65-77.