A method for data-driven exploration to pinpoint key features in medical data and facilitate expert review

Kristina Juhlin | Kristina Star | G. Niklas Norén

1 Uppsala Monitoring Centre, WHO Collaborating Centre for International Drug Monitoring, Uppsala, Sweden
2 Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

Correspondence
K. Star, Uppsala Monitoring Centre, WHO Collaborating Centre for International Drug Monitoring, Box 1051, S-75140 Uppsala, Sweden.
Email: kristina.star@who-umc.org

Abstract

Purpose: To develop a method for data-driven exploration in pharmacovigilance and illustrate its use by identifying the key features of individual case safety reports related to medication errors.

Methods: We propose vigiPoint, a method that contrasts the relative frequency of covariate values in a data subset of interest to those within one or more comparators, utilizing odds ratios with adaptive statistical shrinkage. Nested analyses identify higher order patterns, and permutation analysis is employed to protect against chance findings. For illustration, a total of 164,000 adverse event reports related to medication errors were characterized and contrasted to the other 7,833,000 reports in VigiBase, the WHO global database of individual case safety reports, as of May 2013. The initial scope included 2000 features, such as patient age groups, reporter qualifications, and countries of origin.

Results: vigiPoint highlighted 109 key features of medication error reports. The most prominent were that the vast majority of medication error reports were from the United States (89% compared with 49% for other reports in VigiBase); that the majority of reports were sent by consumers (53% vs 17% for other reports); that pharmacists (12% vs 5.3%) and lawyers (2.9% vs 1.5%) were overrepresented; and that there were more medication error reports than expected for patients aged 2-11 years (10% vs 5.7%), particularly in Germany (16%).

Conclusions: vigiPoint effectively identified key features of medication error reports in VigiBase. More generally, it reduces lead times for analysis and ensures reproducibility and transparency. An important next step is to evaluate its use in other data.

KEYWORDS
exploratory analysis, individual case safety reports, medication errors, odds ratios, pharmacoepidemiology, statistical shrinkage

1 | INTRODUCTION

Pharmacovigilance is the science and activities relating to the detection, assessment, understanding, and prevention of drug-related problems. Individual reports of suspected harm from drugs remain fundamental to detect new risks, and electronic medical records and social media are currently being explored as complementary sources of information. The challenge is to detect previously unknown patterns without prior hypothesis. Data exploration is at the core of these activities. Traditionally, focus has been on the identification of associations between drugs and adverse drug reactions (ADRs) for which several established methods exist. Beyond that, there is a need to explore large case series in greater

Preliminary versions of this paper have been presented at the annual meeting of the International Society of Pharmacovigilance (ISoP) in October 2013 and 2014.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2017 The Authors. Pharmacoepidemiology & Drug Safety published by John Wiley & Sons Ltd.
depth with the aim of understanding the reporting patterns and of identifying patients at risk. This is growing in importance as the number of contributed reports in many regions has increased substantially in the past 10 years. Similar analyses may also address questions related to data management (“are some ADR terms coded to more often in one region than in others?”) or provide context and points of reference for subsequent signal detection or exploration (“what are the differences between reports for children and adults, in general?”).

A natural first step is to obtain basic descriptive statistics and data visualization. Domain experts familiar with a data source of interest can sometimes directly spot unexpected and important features of a data set from descriptive overviews: a simple example would be an over-represented age group in reports from a certain country compared with the database as a whole. Often, such analyses are performed ad hoc with restricted scope determined by domain experts, because broad, open-ended analyses are challenging to scrutinize. Moreover, manual review may be tedious, yield inconsistent results, and be difficult to reproduce.

Computational identification of key features within large and complex data sets would speed up exploratory analysis, ideally while ensuring transparency and reproducibility. In this context, key success factors are analytics that incorporate and support domain expert interpretation. Individual case reports do not derive from experimental settings, and their collection is not devised to minimize bias and heterogeneity. Complex analytical methods whose output cannot easily be traced back to the underlying data are of little value and potentially harmful, especially considering data quality deficiencies and systematic biases that may undermine correct inference. Simpler analyses on the other hand need to be interpreted with caution because they typically do not account for the influence of other covariates than those of primary interest.

We propose vigiPoint, a simple but versatile method for data exploration that outlines data subsets of interest, pinpoints their key features, and facilitates expert review. It utilizes odds ratios with adaptive statistical shrinkage to balance absolute frequency and strength of association in a single measure of relevance. Individual analyses are univariate, considering one covariate at a time, which places them at the lower end of methodological sophistication, but at the higher end of transparency and robustness to random variability and data quality distortions. We illustrate the use of vigiPoint through a case study of highlighting key features of the 164,000 medication error reports among 8 million reports in the WHO global individual case safety reports database, VigiBase, up until May 31, 2013.

2 | METHODS

2.1 | vigiPoint

We describe vigiPoint, a method for highlighting key features of subsets of data. In this context, a feature can be a specific value (or set of values) of a discrete covariate (eg, country of origin = Tunisia) or a range of values for a numerical covariate (eg, patient age is greater than or equal to 2 years and less than 12 years). It compares a data subset of interest to one or more comparator data sets and identifies substantial differences in the relative frequencies of their respective features. This requires focus on relative differences of great enough magnitude while suppressing highly local patterns. Domain experts and data scientists define subsets of interest, comparators, and features to be analysed, whereas vigiPoint provides a generic analysis method independent of the user-defined scope. The guiding principles for vigiPoint are described in Table 1.

2.2 | Scope

Key design choices for each individual analysis include the subset of interest, the comparator(s), the covariates, and corresponding features for analysis, as illustrated in Figure 1. The subset of interest should encompass the data focus for the analysis. For one-off analyses, this may be any collection of data, but subset definitions based on covariate criteria (for example all records from a given time period or demographic category) ensure reproducibility and facilitate definition of the comparator. An outline for the principles that we employ when defining comparators is provided in Figure 1. To identify key features, the relative frequency of covariate values in the subset of interest is contrasted to those in a comparator data set. The comparator is typically another subset of the same data. In our descriptions below, it is assumed not to include the subset of interest. Several comparators may be used in parallel, in which case only covariate values whose frequencies deviate significantly from those in each comparator will be highlighted.

Table 1 | Guiding principles for vigiPoint

| Description | Guiding principle |
|-------------|------------------|
| Describe broad features | Highlight broad features of the data subset of interest, and suppress highly local patterns that affect very small proportions of the subset. |
| Pinpoint the difference | Highlight features that are substantially more common or substantially rarer in the subset of interest than in the comparator. Focus on relative differences, but ensure applicability for overall common as well as for overall rare covariate values. |
| Facilitate expert review | Support expert manual review by providing a transparent trace to the empirical basis for any highlighted feature, so that the corresponding individual records can be reviewed for relevance. |
| Balance structure with flexibility | Rely on experts to define subsets of interest, comparators, and covariates to be analysed, but provide a generic method for analysis independent of the user-defined scope. |
Shrinkage odds ratios

vigiPoint builds on earlier work where observed-to-expected ratios are used to identify patterns based on strength of association. In vigiPoint, key features are identified using odds ratios subjected to statistical shrinkage. Odds ratios have the advantage over other measures of association such as mutual information, relative risk (including Proportional Reporting Ratios and Information Component values), and cosine similarity, of working well for common as well as for rare covariates. As an example, the log odds ratio for 99% vs 90% has the same absolute value as that for 1% vs 10%, which makes sense because they might describe complementary events (e.g., reports from Canada vs reports from other countries than Canada), and in that sense, reflect the same association.

As vigiPoint focuses on relative frequencies, numerical covariates must be grouped into categories prior to analysis. A key consideration here is granularity: finer categories may enable more specific findings but may also make results difficult to survey and lead to loss of statistical power. An effective compromise may be to display results of finer granularity grouped in broader categories (e.g., analysis by years grouped by decades).

Table 2 outlines the basic $2 \times 2$ contingency table for the occurrence of a specific covariate value $x$ (e.g., a specific age group) in the subset of interest $(S)$ and the comparator $(C)$, respectively. The odds

![Figure 1](https://wileyonlinelibrary.com)
ratio for the occurrence of $x$ in the subset of interest can be estimated from the contingency table as:

$$OR = \frac{f(x|S)}{f(x|\overline{S})} = \frac{ad}{bc} \quad (1)$$

We may rearrange this ratio and view $O$ as the observed count and $\overline{E} = bc/d$ as the expected count:

$$OR = \frac{O}{\overline{E}} = \frac{a}{bc/d} \quad (2)$$

To protect against spurious associations and to restrict focus to features with strong enough associations and prevalence in the case series, we apply statistical shrinkage to the odds ratio through addition of a constant $k$ to both the observed and the expected counts:

$$\log_2 OR = \log_2 \left( \frac{a + k}{bc/d + k} \right) \quad (3)$$

The shrinkage moderates the odds ratio towards its baseline value of 1, and the log OR towards 0. The moderation is more prominent for smaller observed or expected counts (less than or of the same magnitude as $k$). As the observed and expected counts increase, the shrinkage odds ratio converges to the raw odds ratio.

With the above transformation, the shrinkage odds ratio corresponds to the Bayesian posterior mean of an intensity parameter $\mu$ for a Poisson Po($\mu$) - distributed O with a Gamma prior distribution for $\mu$: $G(k; k)$. The corresponding posterior distribution for $\mu$ is also Gamma (but with parameters $O + k$ and $E + k$). Credibility intervals that indicate a range of values of $\mu$ compatible with data can be calculated by the inverse of the Gamma cumulative distribution function.9

### 2.4 Adaptive shrinkage

The Information Component measure of disproportionality is based on a similar shrinkage observed-to-expected ratio but uses a strength of shrinkage corresponding to $k = 0.5$.9 This is appropriate in the context of ADR surveillance which must be sensitive to rare patterns. The primary focus of vigiPoint on the other hand is to highlight broad patterns that involve a substantial proportion of the subset of interest. For this reason, we use a stronger shrinkage and adapt the strength of the shrinkage to the size of the subset of interest. In our baseline implementation used here, shrinkage is set to $k = 0.01 \cdot N$, where $N$ denotes the size of the subset of interest (except we do not allow $k$ to fall below 0.5, to ensure adequate shrinkage for smaller subsets).

### 2.5 Threshold for key features

To protect against chance findings, we use percentiles of a two-sided Bayesian credibility interval for the shrinkage odds ratio as our basis for pattern discovery rather than the central estimate. The width of the credibility interval is determined by permutation analysis, as described later. As default, we use a 99% two-sided Bayesian credibility interval. Strength of association is measured and presented as the logarithm of the shrinkage odds ratio so that absolute values reflect the strength of association on a common scale regardless of the association direction. In order to only highlight features that deviate substantially from the expected, the threshold for the credibility interval of the log odds ratio is set to 0.5 and $-0.5$, respectively. This corresponds to just over a 40% over-representation or under-representation in the subset of interest in relation to the comparator, thus ensuring that highlighted features are not only statistically significant but relevant also from a practical perspective. Figure 1 illustrates how features are highlighted based on the value of the lower or upper limit of the credibility interval.

### 2.6 Nested analysis

vigilPoint outlines a subset of interest and pinpoints its key features. Often, there will be an interest in following up some of the highlighted key features by nested vigilPoint analysis for a specific subset. For example, if the analysis of reports on a specific drug reveals an unexpectedly large proportion of reports on adolescents, one may want to drill down and characterize these reports separately. Such a nested application of vigilPoint would typically employ (part of) the initial subset of interest, as one of its comparators, alongside a comparator based on the covariate value used to define the nested subset. For the above example, two natural comparators would be reports on the same drug for non-adolescents and reports on other drugs for adolescents. For our general principles of comparator identification, see the Data Subsets section of Figure 1.

These nested analyses would typically be specified by the analyst based on their domain expertise and aim of analysis. Automated nested analysis through recursive application of vigilPoint to subsets defined by highlighted co-variate values in a primary vigilPoint analysis is also possible and an interesting area for future research.

### 2.7 Permutation analysis

Chance findings are a real concern when executing large numbers of multiple comparisons in parallel.12 To protect against spurious associations from random variability, vigilPoint utilizes statistical shrinkage and Bayesian credibility intervals. Still, its vulnerability to chance findings will vary with the number and granularity of covariates, and their correlation. As a general method to assess this for a specific vigilPoint analysis, we propose a permutation analysis, where subset membership is scrambled, but the size of each subset is intact. We then apply vigilPoint for the same range of covariates. Given the random allocation of reports to the subset of interest and the comparator, any highlighted feature in the permutation analysis is spurious. The proportion of spurious associations can thus be estimated as the number of highlighted features in the permutation analysis compared with that in the original analysis.

### 2.8 Case study

#### 2.8.1 VigiBase

VigiBase is the world’s largest repository of individual case safety reports.10 These are reports of suspected harm from medicines from
health professionals and patients. Reports come from a variety of sources, and the likelihood that a drug caused the suspected adverse reaction varies substantially. In May 2017, VigiBase held around 15 million reports since 1967 from the more than 120 countries that are full members of the WHO Programme for International Drug Monitoring.

2.8.2 Medication errors

Medication errors are mistakes in prescribing, dispensing, or administering medicines. Incidents of medication error can be reported through the pharmacovigilance system in many countries. This study investigated reports of medication errors entered into VigiBase up to May 2013. It exemplifies the use of vigiPoint to explore a data subset of interest to provide points of reference and context for subsequent analyses within the same subset.

For the case study, we defined the subset of interest as reports with at least one medication error related term defined with the MedDRA® High Level Group Term “Medication errors”, except for the subordinated Preferred Terms “Intentional overdose” and “Overdose”. As comparator, we used all other reports in VigiBase.

2.8.3 Scope of analysis for case study

The covariates included in the analysis were year of report entry into VigiBase, country of origin for the report, type of report (eg, spontaneous or report from study), type of notifier (eg, physician or pharmacist), patient age, patient sex, whether the outcome was fatal, reported drugs, and reported adverse events. The analyses of reported drugs were based on the WHO Anatomical Therapeutic Chemical (ATC) classification (level 3) as incorporated in the WHODrug dictionary. The analyses of adverse events were based on MedDRA® High Level Terms.

2.8.4 Nested analysis for case study

Nested vigiPoint analyses were performed for the 5 countries with highest absolute reporting rates of medication error, to explore country-specific patterns. The nested analysis incorporated all covariates in the overall vigiPoint analysis and used two comparators in parallel: all reports from the same country but not reported with medication error, and all reports with medication error from other countries.

3 RESULTS

3.1 Case study

3.1.1 Overall results

A total of 164,000 reports reflecting medication errors were compared with all other reports in VigiBase at the time of this study (n = 7,833,000). The first medication error reports in VigiBase date back to 1994, whilst the first report in VigiBase overall dates back to 1967. The reporting of medication errors has increased over time and was higher than expected in 2012 to 2013 (Figure 2). There were medication error reports from all geographical regions, but North America, and specifically the US, had higher reporting rates than the other regions (Figure 3). Indeed, 145,000 (89%) of the medication error reports were from the US compared with 49% for other reports, and 132,000 (80%) have been submitted since 2008. Medication error was more frequently reported spontaneously and less frequently from studies, clinical trials, and special monitoring. Consumers and non-health professionals (NHP), pharmacists, and lawyers reported medication error to a greater extent than expected, and physicians did so to a lesser extent (Figure 4). Due to differences in report formats, the category “Other” also includes some consumer reports entered in the older INTDIS format. The lower-than-expected rate of medication error for “Other” reflects the overall lower rate of medication error reports historically as well as for countries other than the United States. The apparent under-representation is an example of confounding and underlines the importance of careful interpretation of analysis results.

There were medication error reports for all age groups but higher proportions of reports on 2 to 11 year olds (10% vs 5.7% on reports of other adverse events, Figure 5) and without age specified (Figure 4). The gender distribution was the same for medication error reports as for other reports in VigiBase (M = 39%, F = 61%), but gender was more often missing on medication error reports: 12% vs 6.4% for other reports in VigiBase (Figure 4).

Greater than expected reporting rates of medication errors were found for drugs used for respiratory and nervous systems disorders, as well as drugs used for alimentary tract and metabolism disorders (eg, drugs used for gastro-oesophageal reflux, propulsives, and insulins). A number of MedDRA High-Level Terms found in the MedDRA system organ classes “Social circumstances,” “Surgical and
medical procedures,” and “Psychiatric disorders” were over-represented among the medication error reports (Figure 6).

3.1.2 Nesting analysis results

A nested analysis was performed for the 5 countries with highest absolute reporting rates of medication error: United States, Canada, Germany, United Kingdom, and Republic of Korea contrasted with the reports of other adverse events in the respective countries and with medication error reports for other countries (Figures 2, 4–6).

A greater than expected proportion of medication error reports from consumers and non-health professionals were found in 2 countries: for the United States, 59% vs 31% of the other reports from the same country and vs. 19% of reports on medication error from other countries; for Republic of Korea, 87% vs 15% of the other reports from the same country and vs 53% of reports on medication error from other countries. In the United Kingdom, medication error was more often than expected reported by pharmacists: 25% vs 9.0% of other reports from the same country and vs. 12% of reports on medication error from other countries. For Canada and Germany, literature reports were over-represented among reports on medication error: for Canada 4.0% vs 1.4% of other reports from the same country and vs. 0.51% of reports on medication error from other countries; for Germany, 5.8% vs 2.9% of other reports from the same country and vs 0.46% of reports on medication error from other countries (Figure 4).

The previously observed over-representation of reports for children aged 2-11 years with medication error was even more pronounced in Germany: 16% vs 3.1% of other reports from the same country and vs 9.9% of reports on medication error from other countries (Figure 5). Medication error reports from the United States and Republic of Korea lacked information on patient age more often than expected based on the comparator datasets (Figure 4).

3.1.3 Permutation analysis results

The permutation analysis for the primary and nested VigiPoint analyses did not highlight any key features at the suggested thresholds, OR_{905} > 0.5 or OR_{995} < -0.5. This can be compared with 109 key features highlighted in the primary analysis for the case study, which suggests that the expected proportion of chance findings in our analysis is very low.

4 DISCUSSION

VigiPoint provides a method to quickly highlight key features that distinguish a data subset of interest from one or more comparators. In our case study, it identified the large proportions of medication error reports that come from the United States and from patients, but also drew to light subtler patterns such as the over-representation.
of medication error reports for children in Germany, and from pharmacists in the United Kingdom. The over-representation of medication error reports from the United States is obvious and would easily have been detected in a descriptive analysis. However, to identify subtler reporting patterns, the expert would have needed prior knowledge on the general reporting patterns in the data to pinpoint key features. The context and knowledge of the dataset under scrutiny provided by vigiPoint can guide the expert in their consideration of further investigations.

Speed is of the essence, and vigiPoint limits the number of steps to set up each individual analysis, thus reducing the lead times, while ensuring reproducibility and consistency. Once the characterization of medication error reports overall had been performed, the nested analysis for each country was straightforward: all it required was a definition of the subset of interest (eg, all medication error reports from Germany) and its comparators (eg, all medication error reports from other countries and all other reports from Germany, respectively).

vigiPoint allows for the consideration of more covariates, and combinations of covariates, than would have been possible to review manually, thus enabling a wider exploration of potential patterns. In our analysis of medication error reports, the inclusion of a very large number of countries, drug classes, and combinations thereof, led to the highlighting of a disproportionate number of medication error reports for opioids and hormonal contraceptives in Republic of Korea and of antipsychotic medicines in Germany and the United Kingdom.

Each individual vigiPoint analysis is essentially univariate: its identification of key features considers each covariate separately. This ensures transparency between highlighted features and the underlying descriptive statistics. As an example, analysts can directly see why direct patient reports are highlighted as a key feature of medication error reports overall in VigiBase: more than half of the medication error reports are from patients compared with 1 in 6 reports overall. The lower flexibility of univariate analyses should make them less prone to over fit to peculiarities of the data set at hand, and so less vulnerable

**FIGURE 4** Distribution of report type (top panel), type of reporter (panel 2), fatal outcome (panel 3), gender (panel 4), and amount of missing data (bottom panel) for medication error reports overall as well as in each of the 5 countries examined in-depth. The bars show the proportion of the covariate in the subset of interest, while the lines show the proportion(s) in the comparator(s). Black lines denote reports without medication errors and grey lines denote medication error reports from other countries. Key features are marked with yellow (over-represented) or blue (under-represented) with stated percentages for the subset of interest and for the comparator group closest in relative frequency [Colour figure can be viewed at wileyonlinelibrary.com]
to random variability and data quality issues. Moreover, the possibility to associate the key feature with individual records enables further follow-up and action related to these.

On the other hand, each univariate analysis on its own is vulnerable to confounding, where apparent associations result from imbalances in other covariates. For example, the higher rate of medication error reports for patients might have simply reflected the higher rates of both direct patient reports and medication errors from the United States; in theory, medication errors could have been less common on direct patient reports within every single country and still more common for patient reports overall, because so many of the medication error reports are from the United States where a larger proportion of reports are from patients. This would have been an example of Simpson’s paradox, where it would be incorrect to infer that patients are more likely than others to report medication error. In reality, medication errors were more common on direct patient reports than on other reports for each of the 5 countries in our nested analysis.

An additional challenge with univariate analyses is that they do not encompass higher order associations. There may for example be countries with high rates of medication error reports from physicians, even though the overall rate of medication error reports from physicians is low. This can to some extent be handled by nested vigiPoint analyses: in our study, such patterns would have been highlighted for the 5 countries analysed in closer detail, but not for any of the other countries.

Our use of univariate analysis over regression for primary analyses in vigiPoint derives from the ambition to first describe and then interpret, and for vigiPoint to serve as a guide rather than a final answer when screening for data patterns. This transparency also makes it easier to detect data quality issues and implementation errors, because unexpected patterns cannot be attributed to sophisticated analytics. An interesting area of future research is to explore whether performing regression in parallel with univariate analyses can support inference. Another possible area of future research is automated nested analysis through recursive application of vigiPoint to subsets defined by highlighted covariate values in a primary vigiPoint analysis.

vigiPoint projects numerical covariates into categories prior to analysis, which may limit resolution and dilute the results. The age group 2-11 year olds was highlighted as a key feature of medication error reports, but closer examination reveals that the over-representation is most pronounced for the 2-3 year olds, whereas the deviation for the 4-11 year olds is in the same direction but not statistically significant. One possible extension of the methodology would be to explore the comparison of distributions of numerical variables using, for example, Kullback-Leibler divergence.

The value of pattern discovery ventures resides to some extent in the eye of the beholder: are the highlighted features of interest to domain experts and are all interesting features according to domain experts highlighted by the method? Gold standards for this are difficult to obtain because subjective judgement is fraught with inter-rater and intra-rater variability. However, the permutation analysis integral to vigiPoint ensures that highlighted patterns most likely reflect real differences and not random fluctuations, which is an important first step. Of course, these patterns reflect variations in data and not necessarily in the underlying population: our case study has highlighted differences in the reporting of medication errors that may not reflect variations in its occurrence. As always, it is important to bear in mind the nature of the underlying data, in drawing conclusions based on the analysis.

As regards the relevance to domain experts, vigiPoint (or earlier incarnations thereof) has been used and found valuable in a number of applied analyses of VigiBase, ranging from characterizations of
FIGURE 6  Distribution of MedDRA high-level term (top panel) and ATC group (level 3) (bottom panel) for medication error reports overall as well as in each of the 5 countries examined in-depth. The bars show the proportion of the covariate in the subset of interest, while the lines show the proportion(s) in the comparator(s). Black lines denote reports without medication errors, and grey lines denote medication error reports from other countries. Key features are marked with yellow (over-represented) with stated percentages for the subset of interest and for the comparator group closest in relative frequency. Due to limited space, only features that are over-represented in at least one of the subsets are included [Colour figure can be viewed at wileyonlinelibrary.com]
reports on drugs for cardiometabolic disease in Sub-Saharan Africa,\textsuperscript{15} of well-documented reports,\textsuperscript{16} of suspected duplicate reports,\textsuperscript{17} and of reports for children.\textsuperscript{18} An informal evaluation of vigiPoint against an already published characterization of paediatric ADR reports gave encouraging results: vigiPoint reproduced the findings of the published study and additionally highlighted an over-representation of severe skin reactions (erythema multiforme)\textsuperscript{18} that was of interest to domain experts but had not been identified in the original study.

An important next step is to evaluate its use in other domains, to ensure that the chosen analytical approach and design choices generalize to other settings. Design choices that may need re-consideration in other domains include the strength of shrinkage, the confidence interval percentiles for the log-odds ratio to use in screening, and the associated threshold.

ETHICS STATEMENT
The authors state that no ethical approval was needed.

ACKNOWLEDGEMENTS
The authors are indebted to all the national centres that contribute reports to VigiBase. However, the opinions and conclusions in this study are not necessarily those of the various national centres, nor of the WHO. Kristina Juhlin is a former employee of the Uppsala Monitoring Centre. We would also like to thank Johan Hopstadius for contributions in the early phase of this project.

No funding was received for this study.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ORCID
Kristina Star \url{http://orcid.org/0000-0002-5384-8609}

REFERENCES
1. World Health Organization. The importance of pharmacovigilance. Safety monitoring of medicinal products. 2002.
2. Vandenbroucke JP. In defense of case reports and case series. Ann Intern Med. 2001;134:330-334.
3. McNaughton R, Huet G, Shakir S. An investigation into drug products withdrawn from the EU market between 2002 and 2011 for safety reasons and the evidence used to support the decision-making. BMJ Open. 2014;4. https://doi.org/10.1136/bmjopen-2013-004221
4. Norén GN, Hopstadius J, Bate A, Star K, Edwards IR. Temporal pattern discovery in longitudinal electronic patient records. Data Mining and Knowledge Discovery. 2010;20:361-387. https://doi.org/10.1007/s10618-009-0152-3
5. Powell GE, Seifert HA, Reblin T, et al. Social media listening for routine post-marketing safety surveillance. Drug Saf. 2016;39:443-454. https://doi.org/10.1007/s40264-015-0385-6
6. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Pharmacol. 1998;54:315-321.
7. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. American Statistician. 1999;53:177-190.
8. Evans SJW, 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.
9. Norén GN, Hopstadius J, Bate A. Shrinkage observed-to-expected ratios for robust and transparent large-scale pattern discovery. Stat Methods Med Res. 2013;22:57-69.
10. Lindquist M. VigiBase, the WHO Global ICSR Database System: basic facts. Drug Inf J. 2008;42:409-419. https://doi.org/10.1177/009286150804200501
11. Tan P-N, Kumar V, Srivastava J. Selecting the right interestingness measure for association patterns. Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining; Edmonton, Alberta, Canada. 775053: ACM; 2002. p. 32–41.
12. Webb G. Discovering significant patterns. Machine Learning. 2007;68:1-33. https://doi.org/10.1007/s10994-007-5006-x
13. Simpson EH. The interpretation of interaction in contingency tables. JR Stat Soc B Methodol. 1951:13:238-241.
14. Kullback S, Leibler RA. On information and sufficiency. The Annals of Mathematical Statistics. 1951:79-86.
15. Bergvall T, Norén GN, Lindquist M. vigiGrade: a tool to identify case safety reports. Eur J Intern Med. 2013;24:69.
16. Bengoa E, Norén GN, Kumar V, Srivastava J. Selecting the right interestingness measure for association patterns. Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining; Edmonton, Alberta, Canada. 775053: ACM; 2002. p. 32–41.
17. Tregunno P, Fink D, Fernandez-Fernandez C, Lázaro-Bengoa E, Norén GN. Performance of probabilistic method to detect duplicate individual case safety reports. Drug Saf. 2014;37:65-77.
18. Juhlin K, Star K, Norén GN. Pinpointing key features of case series in pharmacovigilance—a novel method. International Society of Pharmacovigilance Annual Meeting; Pisa. Drug Saf. 2013;36:912-913. https://doi.org/10.1007/s40264-013-0087-x

How to cite this article: Juhlin K, Star K, Norén GN. A method for data-driven exploration to pinpoint key features in medical data and facilitate expert review. Pharmacoepidemiol Drug Saf. 2017;26:1256–1265. https://doi.org/10.1002/9781119123850.ch39