Drug Safety (2023) 46:857–866
https://doi.org/10.1007/s40264-023-01329-w

ORIGINAL RESEARCH ARTICLE

Mapping Strategies to Assess and Increase the Validity of Published Disproportionality Signals: A Meta-Research Study

Michele Fusaroli[1] · Francesco Salvo[2,3] · Claire Bernardeau[4] · Maryam Idris[2] · Charles Dolladille[5,6] · Antoine Pariente[2,3] · Elisabetta Poluzzi[1] · Emanuel Raschi[1] · Charles Khouri[4,7]

Accepted: 16 June 2023 / Published online: 8 July 2023
© The Author(s) 2023

Abstract

Background and Aim Disproportionality analysis is traditionally used in spontaneous reporting systems to generate working hypotheses about potential adverse drug reactions: the so-called disproportionality signals. We aim to map the methods used by researchers to assess and increase the validity of their published disproportionality signals.

Methods From a systematic literature search of published disproportionality analyses up until 1 January 2020, we randomly selected and analyzed 100 studies. We considered five domains: (1) rationale for the study, (2) design of disproportionality analyses, (3) case-by-case assessment, (4) use of complementary data sources, and (5) contextualization of the results within existing evidence.

Results Among the articles, multiple strategies were adopted to assess and enhance the results validity. The rationale, in 95 articles, was explicitly referred to the accrued evidence, mostly observational data ($n = 46$) and regulatory documents ($n = 45$). A statistical adjustment was performed in 34 studies, and specific strategies to correct for biases were implemented in 33 studies. A case-by-case assessment was complementarily performed in 35 studies, most often by investigating temporal plausibility ($n = 26$). Complementary data sources were used in 25 articles. In 78 articles, results were contextualized using accrued evidence from the literature and regulatory documents, the most important sources being observational ($n = 45$), other disproportionalities ($n = 37$), and case reports ($n = 36$).

Conclusions This meta-research study highlighted the heterogeneity in methods and strategies used by researchers to assess the validity of disproportionality signals. Mapping these strategies is a first step towards testing their utility in different scenarios and developing guidelines for designing future disproportionality analysis.

Emanuel Raschi and Charles Khouri: co-last authors.

✉ Michele Fusaroli
michele.fusaroli2@unibo.it

1 Pharmacology Unit, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

2 Univ. Bordeaux, INSERM, BPH, U1219, Team AHeaD, 33000 Bordeaux, France

3 CHU de Bordeaux, Pôle de Santé Publique, Service de Pharmacologie Médicale, 33000 Bordeaux, France

4 Pharmacovigilance Unit, Grenoble Alpes University Hospital, Grenoble, France

5 UNICAEN, EA4650 SEILIRM, CHU de Caen Normandie, Normandie University, Caen, France

6 Department of Pharmacology, CHU de Caen Normandie, Caen, France

7 Univ. Grenoble Alpes, HP2 Laboratory, Inserm U1300, Grenoble, France

Key Points

Validity of disproportionality is enhanced when combining evidence and robust methods.

No consensus on better approaches exists.

We found high heterogeneity in 100 published studies.

1 Introduction

Spontaneous reporting systems play a key role in monitoring the safety of drugs after they have been approved for the market, complementing drugs' safety profile obtained

from clinical trials [1]. Disproportionality analyses [2] are the most common statistical approach to analyze these reports and identify potential adverse drug reactions. They are used by health agencies, research centers, and pharmaceutical companies [3], for both signal detection (generating new hypotheses about potential adverse drug reactions) and signal refinement (characterizing unknown features of expected adverse drug reactions). Disproportionality analyses look for patterns of disproportional reporting of a specific adverse event in patients who have taken a certain drug, when compared with the overall rates in the database [4]. These statistical disproportions do not imply any kind of causal relationship: the association may also be the result of reverse causality (the drug is taken to treat the event when the event has already been diagnosed [5]), protopathic bias (the drug is taken to treat a prodromic symptom of the event when the event has not yet been diagnosed [6]), indication bias (the reason for using the drug makes the patient more susceptible to the event [7]), notoriety bias (regulatory and media attention increase the reporting likelihood [8]), selective reporting and differences in reporters [9, 10]. In summary, disproportionality signals alone are not sufficient evidence to establish a causal relationship between a drug and an adverse event [2, 11]. They should, instead, be considered as working hypotheses [12] and integrated with evidence from multiple types of studies (e.g., clinical trials, observational studies, pharmacovigilance analyses, and animal experiments), each with its strengths and weaknesses [13]. Healthcare agencies perform this integration to decide, based on the place in therapy of the medicine, whether regulatory decisions are needed [14, 15].

Despite being a starting evidence, disproportionality signals are increasingly published in the literature, with hundreds of safety signals generated each year [2, 13, 16]. These published analyses can be very different in nature: some are hypothesis-based, others purely rely on disproportionality in the reporting (i.e., the agnostic approach). Accordingly, the validity of safety signals may vary greatly [16]. Multiple strategies have been developed to assess and enhance the plausibility of disproportionality signals and to prioritize them [17]. These may involve clinical and pharmacological reasoning, theoretically-based design of disproportionality analyses, implementation and comparison of multiple operative choices (the so-called sensitivity analyses), case-by-case causality assessment, and integration with other data sources (e.g., prescriptions data to approximate incidence, pharmacodynamic data to investigate the mechanism, regulatory documents to distinguish between expected and unexpected reactions) [2, 18, 19]. However, there is no gold standard to guide researchers in the design of disproportionality analysis and a large heterogeneity exists in published studies.

In this meta-research study, we investigated the strategies used to assess and enhance the validity of disproportionality signals. We did so by operationalizing a notion of signals validity, including (a) a priori plausibility of the hypothesis based on already accrued evidence, and (b) methodological robustness, verified as resistance from errors in the results when performing different operative choices and complementary analyses. By creating a mapping of these strategies, the study will provide a foundation for the development of guidelines for assessing the validity of published disproportionality signals.

2 Materials and Methods

2.1 Design of the Study and Article Selection

To obtain a description of strategies used to assess and enhance the validity of safety signals from disproportionality analyses, we conducted a meta-research study based on a previous systematic review [16]. Briefly, a systematic literature search was performed on Medline to identify all published disproportionality analyses since inception up until 1 January 2020, using the search terms "case-non case," "disproportionality analysis," "pharmacovigilance analysis," or "pharmacovigilance study." In the second step, 100 studies were randomly selected for analysis through random selection of article numbers on Excel.

2.2 Development of the Extraction Table

Based on previous research on safety signals prioritization, on Bradford Hill criteria [20], and on the authors' experience in designing disproportionality analyses and assessing safety signals, we identified the variables of interest and designed an extraction table.

Overall, we distinguished five domains in the assessment and enhancing of the validity of a disproportionality safety signal:

1. The explicit rationale for performing the study.
2. The strategies adopted in the design of disproportionality analyses.
3. The strategies adopted in the case-by-case assessment.
4. The implementation of complementary analyses, integrating spontaneous reports with other kinds of data.
5. The contextualization of the results within the existing evidence, including the literature, regulatory documents, and unpublished data.

Domains one and five are strictly related to the a priori plausibility of the hypothesis, given accrued evidence. Domains two, three, and four are related to methodological robustness and the resistance to errors when performing different operative choices.

For each domain, we identified variables of interest and possible associated values (Table S1 in Supplementary Material 1).

2.3 Data Extraction Process and Descriptive Analysis

For each study, two authors extracted the data in parallel (MF, MI, CB, CK, ER). A pilot study was performed on five studies to train the authors in the extraction approach. Disagreements were resolved through discussion and consensus among all the authors. The same procedure—including a phase of extraction led separately by two authors, and a phase of comparison and consensus among all the authors—was repeated for the remaining studies.

A descriptive analysis was performed to summarize the strategies used by the 100 articles to assess and enhance the validity of disproportionality signals.

2.4 Software and Preregistration

R studio (version 4.1.2) was used to process, analyze, and visualize the data. The protocol, together with the original extraction table and the definition of each variable extracted, was preregistered in the OSF platform [21].

3 Results

3.1 Articles Selection and Description

The articles retrieved through the systematic literature search were published between 1983 and 2019. The 100 articles randomly selected for inclusion were published between 1997 and 2019, with 70% of them published after January 2015 (see Supplementary Material 2). Among them, 26% were conceived as brief articles, 40% of the studies were published in specialized clinical journals, and 60% in pharmacological journals, with the most recurrent journals being "*Drug Safety*" (13%), the "*European Journal of Clinical Pharmacology*" (8%), and "*Pharmacotherapy*" (7%). Each study investigated a median of six drugs [interquartile range (IQR): 1–14] and one event (IQR: 1–5), with four articles investigating all drugs found in the database for one or more events, and three articles investigating all the events for one or more drugs. The drugs more commonly analyzed targeted the nervous system (24%), followed by antineoplastic agents (20%) and antiinfective drugs (11%). The most frequently studied adverse events were neurological (11%), cardiac (11%), and vascular disorders (7%) [16]. In 55 articles, a combination of MedDRA®[1] terms was used to retrieve cases.

Among the analyzed articles, the most commonly used spontaneous reporting systems were the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS, $n = 40$), the World Health Organization reporting system (VigiBase, $n = 28$), and the French national database (BNPV, $n = 20$). In 11 articles, more than one database was used. In 78 articles, the entire database was used for disproportionality analysis. Most common restrictions were applied, limiting the analysis to specific demographics ($n = 14$), therapeutic areas ($n = 7$), time ranges ($n = 2$), and serious events only ($n = 2$).

3.2 Strategies to Assess and Increase Validity

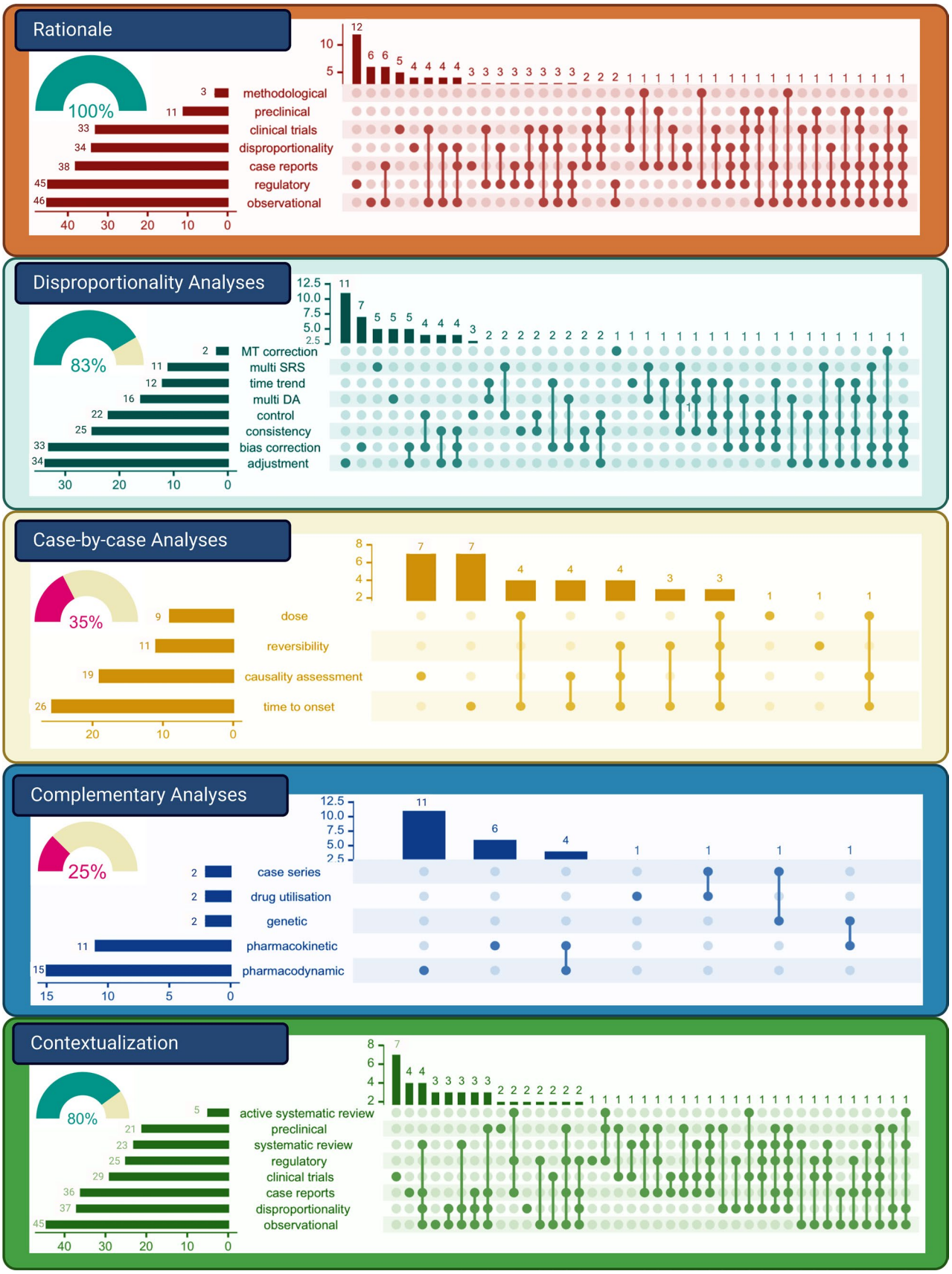
3.2.1 Rationale for the Study

All of the analyzed articles provided a rationale for their study (see Fig. 1). Among the articles, 95 clearly defined the type of evidence over which the study was designed, the most reported ones being observational evidence ($n = 46$) (e.g., pharmacoepidemiological prospective or retrospective studies), regulatory documents ($n = 45$) (e.g., safety warning issued by health authorities, information included in package inserts or summaries of product characteristics), and case reports ($n = 38$), followed by other previous disproportionalities in the same or other spontaneous reporting systems ($n = 36$), clinical trials ($n = 33$), and preclinical studies ($n = 11$). Three studies proposed a methodological rationale (e.g., the implementation of a new technique). Two studies preregistered their protocol.

3.2.2 Design of Disproportionality Analyses

In 83 articles the authors adopted at least one strategy to implement or test the validity of the disproportionality signals; in 48 more than one strategy was implemented (Fig. 1). In the remaining 17 articles, only a raw disproportionality analysis was performed. The most frequently used method was statistical adjustment for general variables not

[1] 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). MedDRA® trademark is owned by the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) on behalf of ICH.


Rationale
100%
3 methodological
11 preclinical
33 clinical trials
34 disproportionality
38 case reports
45 regulatory
46 observational
40 30 20 10 0
12 6 6 5 4 4 4 4 3 3 3 3 3 3 3 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Disproportionality Analyses
83%
2 MT correction
11 multi SRS
12 time trend
16 multi DA
22 control
25 consistency
33 bias correction
34 adjustment
30 20 10 0
11 7 5 5 5 4 4 4 3 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Case-by-case Analyses
35%
9 dose
11 reversibility
19 causality assessment
26 time to onset
20 10 0
7 7 4 4 4 3 3 1 1 1
Complementary Analyses
25%
2 case series
2 drug utilisation
2 genetic
11 pharmacokinetic
15 pharmacodynamic
15 10 5 0
11 6 4 1 1 1 1
Contextualization
80%
5 active systematic review
21 preclinical
23 systematic review
25 regulatory
29 clinical trials
36 case reports
37 disproportionality
45 observational
40 30 20 10 0
7 4 4 3 3 3 3 3 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

◂**Fig. 1** Use of strategies to assess/enhance validity in the 100 disproportionality studies randomly selected, grouped by domain. The five colored bars represent the five domains. Each domain has a gauge plot on the left, showing the percentage of studies using at least a technique of the domain, and an UpSet plot on the right, showing the frequency of use of each technique (side bar plot) and of each combination (bar plot on the top). In some cases, a rationale or a literature support was provided based on preexisting works, for which the underlying evidence was not clearly specified, as it came from narrative reviews, opinions, and commentaries. Articles with underspecified rationale or literature support were counted in the gauge plot, but not reported in the UpSet plot. Legend: MT correction, correction for multiple testing; multi DA, multiple disproportionality analysis; multi SRS, multiple spontaneous reporting system. Created with biorender.com

explicitly assumed to be effect modifiers or confounders ($n = 34$), particularly for age ($n = 26$), gender ($n = 23$), and comedications ($n = 21$). Moreover, specific strategies to correct for anticipated or identified biases were implemented in 33 articles, the most frequent being indication bias ($n = 11$), comedication bias ($n = 9$), and masking bias ($n = 7$). In 25 studies, the signal stability (i.e., consistency) was tested on multiple subpopulations, particularly on different age groups ($n = 8$) and indications for use ($n = 5$). In 22 articles, control was implemented: 9 were only positive, 1 was only negative, and 12 were both. In 16 studies, multiple disproportionality methods were used: in 13 cases a Bayesian approach combined with a frequentist approach was used to maximize specificity and sensitivity, and in 3 cases multiple frequentist methods were used. In 12 articles, the time trend of the signal was investigated, and 11 articles accessed multiple spontaneous reporting systems. Two studies corrected the statistical significance for multiple testing.

3.2.3 Case-by-Case Causality Assessment

Together with the disproportionality analysis, a case-by-case assessment was also performed in 35 studies (Fig. 1). The most common form was investigating time to onset and temporal plausibility ($n = 26$). Among the 100 articles revised, 19 performed some kind of causality assessment. In 14 articles this consisted only of a differential diagnosis excluding alternative causes, such as comorbidities [e.g., human immunodefificency virus (HIV) is a confounder in the investigation of TNF-α inhibitors-related Kaposi sarcoma [22]] and concomitants (e.g., LiverTox and AZCERT lists of hepatotoxic and torsadogenic drugs [23, 24]). Five articles also adopted a validated causality algorithm (two adopted the WHO algorithm [25], one adopted the French Agency BNPV algorithm [26], and two adopted both of these algorithms). The reversibility (dechallenge/rechallenge) was investigated in 11 articles, and the dose gradient in 9 articles.

Among the 17 articles performing only a raw disproportionality analysis, 4 also performed a case-by-case assessment.

3.2.4 Complementary Analyses

In 25 articles, the authors complemented the analyses on spontaneous reporting data with information from other data sources (Fig. 1). In 23 cases, the biological plausibility was investigated by linking spontaneous reports with pharmacometrics (pharmacodynamics in 15 articles, pharmacokinetics in 11 articles, genetic data in two articles). In two articles, the authors linked drug utilization data to overcome the lack of exposure data, in the attempt to estimate reporting incidence measures. In two studies, one or more additional case reports were published and discussed together with the disproportionality analysis.

Among the 17 articles performing only a raw disproportionality analysis, 5 also performed a pharmacometric evaluation.

3.2.5 Contextualization Within Existing Evidence

In 78 articles, the results of the disproportionality analyses were contextualized within accrued evidence from the literature and regulatory documents (Fig. 1), the most important sources being observational ($n = 45$), other disproportionalities ($n = 37$), and case reports ($n = 36$), followed by clinical trials ($n = 29$), regulatory documents ($n = 25$), systematic reviews ($n = 23$), and preclinical data ($n = 21$). Five articles reported a formal active systematic review or a meta-analysis of the evidence for and against the investigated hypothesis.

4 Discussion

Our meta-research study mapped strategies used in the assessment and enhancement of the validity of safety signals from disproportionality analyses. We here present and discuss our findings, contextualizing them within methodological debates about disproportionality analyses from the literature.

Among the 100 selected articles, we observed a large heterogeneity in the approaches to assess and enhance the validity of disproportionality signals.

The heterogeneity of study designs in disproportionality analysis should not always be seen as a threat to study validity. It is important to acknowledge that different research questions necessitate different study designs. Some studies aim to uncover general associations between a drug and an adverse event, while others focus on specific subpopulations,

explore interactions with concurrent factors, or investigate variations among countries. This diversity of study designs allows researchers to effectively address their specific research objectives and specific bias. However, the omission of a case-by-case assessment can jeopardize the validity of the findings, as it plays a crucial role in accurately defining a signal. On the other hand, the assessment of reversibility, which is rarely irrelevant to the research question (e.g., for fatal events, pregnancy-related outcomes, or situations where drug suspension is not feasible), is often neglected. The absence of standardized guidelines frequently leaves researchers to make subjective decisions on study design, leading to a broader range of approaches and methodologies.

In this study, we grouped these strategies in five domains: (1) the rationale for the study, (2) the design of disproportionality analyses, (3) the case-by-case assessment, (4) the use of complementary data sources, and (5) the contextualization of the results within existing evidence. Domains one and five were related to the a priori plausibility of the hypothesis. Domains two, three, and four were, instead, more strictly related to methodological robustness and inferential validity of the quantitative and qualitative methods. In the following paragraphs, we will discuss separately each of these two components of signals validity.

4.1 A Priori Evidence

Health agencies like the European Medicines Agency (EMA) or the FDA mainly use disproportionality analysis in the agnostic (i.e., non-targeted) routine detection of new signals, with no prespecified hypothesis [3, 27]. Most of the published disproportionality analyses are conceived and conducted to add further knowledge on an already identified safety signal or hypothesis, which is usually emerging from other observational studies, clinical trials, or case reports. Researchers' disproportionality analyses, therefore, complement regulatory signal detection with a more targeted signal refinement.

First, starting from prespecified hypotheses, it is essential to contextualize the study within accrued evidence, to assess the a priori plausibility of the hypothesis. Ideally, this may consist in a systematic presentation of clinical and pre-clinical evidence in favor and against the hypothesis of a causal link between the drug(s) and the adverse event(s) of interest [28], taking into account the external validity of preclinical findings (relevance of model, surrogate outcome, exposure, route of administration, dose, and drug) [29]. In our sample, only five studies performed a formal systematic review of the evidence, rarely together with a proper meta-analysis [30]. While the timeliness of signal detection in disproportionality analysis is important, it should not compromise the thoroughness of signal refinement. Conducting a systematic review is time-consuming, and it is crucial to strike a balance between two essential factors: the need to provide a comprehensive synthesis of existing evidence alongside the identified signal, and the potential risk of delaying necessary regulatory actions. In fact, more research is warranted to study the timeliness of the publication of disproportionality analyses in the evolution of a safety signal, from the first warning to the potential final validation or refutation [31].

Second, starting from prespecified hypotheses, it is essential to justify the study, given already accrued evidence. For example, if the study is conceived on results from an observational study, which, by definition, should be used to confirm or refute disproportionality signals from spontaneous reporting systems, the relevance of disproportionality analyses is questionable [32] and the added value should be clearly specified. In this case, a disproportionality analysis may be justified by the focus on populations not previously addressed by observational studies, or by obsolete results due to significant changes in clinical practice (e.g., changes in the dose administered).

4.2 Methodological Robustness of the Qualitative and Quantitative Analyses

4.2.1 Design of Disproportionality Analysis

To achieve methodological robustness, disproportionality analysis should account for potential confounders and address the inherent variability of results stemming from subjective design choices. These operative choices involve discretionary decisions made during the analysis, such as the choice of background population, statistical methodologies employed, and the determination of signal detection thresholds. It is important to justify and report these subjective choices to enhance the transparency and reproducibility of the analysis. Additionally, different operative choices can be simultaneously employed to explore the robustness and consistency of findings.

As previously emphasized [9], we found a high amount of heterogeneity in statistical methods used to calculate disproportionality.

The most commonly used technique was to statistically adjust or control for one or more variables. This adjustment was often made on general variables, such as sex and age, without explicitly assuming that these variables would affect the outcome in any specific way. While sex and age may affect the outcomes, any adjustment for covariates should be supported by a transparent a priori defined causal model, otherwise it may even introduce new biases [33]. Furthermore, because non-complete fields in spontaneous reporting are common, the practice of adjusting may lead to spurious and hardly interpretable results and should instead be limited to expected confounding factors (a targeted approach that we referred to, in the extraction table, as controlling for

bias). Conversely, searching for consistency in subpopulations (i.e., subgroup or stratified analyses), which identified the lack or the presence of a specific confounder, is recommended [34].

The adoption of multiple frequentist disproportionality estimates (e.g., RRR, ROR, PRR) is also debated since their results overlap in big databases [35]. The adoption of a frequentist and a Bayesian approach instead may help to prioritize among signals, integrating the higher sensitivity of the former and the highest specificity of the latter [36]. Furthermore, running the analyses on multiple spontaneous reporting systems usually has consistent results [35] and may not be necessary apart from specific drug-event pairs, depending on the database-specific pattern of use [37].

Finally, when agnostic signal detection is performed, strategies to control for false discovery rates and to prioritize signals can also reduce spurious signal generation [32, 38], but there is still no consensus on criteria and thresholds [28, 39, 40].

Protocols are a relatively new development, which has still not gained widespread adoption in pharmacovigilance disproportionality analyses. Given the retrospective nature of spontaneous reporting data and the absence of preregistered protocols, there is risk for publication bias, selective outcome, and selective analysis reporting. Therefore, the results of only one analysis, in particular when implemented with unusual thresholds and comparator groups, need to be interpreted very cautiously. Instead, we advocate for the presentation of a set of motivated analyses (e.g., signal detection thresholds, comparator groups, adjustment strategies, and population subgrouping), ideally prespecified in a protocol, and an assessment of the consistency of the estimates presented alongside crude results [9, 16]. The assessment of the variability of the results according to event and drug terms selection is also an important point, given their impact on the results [34, 41, 42].

4.2.2 Case-by-Case Assessment

Case-by-case analysis and causality assessment, including the evaluation of temporal and pharmacological plausibility, as well as a differential diagnosis to exclude alternative explanations, is an important step in the evaluation of safety signals by drug agencies. Even if case narratives are not always available in international databases due to data privacy policies, other data useful for causality assessment are often reported (time to onset, drug dose, action taken, and evolution). We recommend, if appropriate and whenever possible, the implementation of a causality assessment, especially for designated medical events such as Torsade de Pointes, drug-induced liver injury, and severe skin reactions, for example using the WHO causality assessment method [25] or Bradford Hill criteria [20]. Even if these methods cannot be fully implemented, researchers should still attempt to apply these criteria or their adapted form for spontaneous reporting systems [43, 44] to the reports inherent to the investigated signal. Doing so may reveal previously undetected duplicates and non-predicted confounders.

4.2.3 Complementary Analyses

New clinical data are sometimes published together with disproportionality analyses, presented as the evidence driving the disproportionality (e.g., case reports) or as a part of the study (e.g., considering prescription flows). Combining multiple data sources, albeit challenging, may provide a broader perspective and a more precise evaluation of the safety profile of a given drug [45, 46].

The integration with drug utilization data, a practice emerging thanks to the tracking of all the coronavirus disease 2019 (COVID-19) vaccines administered, is a way to partly map the exposure to the drug and try to estimate reporting rate measures [47]. Drug utilization data can be also useful as a tool for signal prioritization or to identify drug- and country-specific scenarios [48].

Correlating spontaneous reporting data with other pharmacometrics data, performed at least in a speculative manner in 22 of the 100 investigated studies, is a new and promising method to explore the underlying pharmacological basis. Nonetheless, a consensus on this approach has still not been reached [49–51].

4.3 Limitations

We investigated techniques used by researchers to assess or enhance the validity of their disproportionality signals, without checking whether the signal was found to be valid. Also, we did not aim to assess the comparative utility of the different techniques. Furthermore, our findings should be considered in light of several limitations. Because some of the selected articles were conducted in the late 1990s, preregistration of protocols would not have been possible given the non-existence of such registries. Because of the sampling, we may have missed some less-used or more recent techniques. Although we tried to minimize data-collection errors by performing extraction in parallel by two authors, sometimes the lack of transparency and the fact that we did not contact authors for clarification may have introduced minor misinterpretation of methods.

However, this meta-research is an initial attempt to formally and comprehensively characterize criteria to define the validity of published disproportionality analyses, in the promotion of high-quality research in pharmacovigilance, namely conception, conduction, and reporting of disproportionality analyses. We strongly believe that harmonizing these approaches will finally increase the transferability

of results from spontaneous reporting systems in clinical practice.

5 Conclusions

This meta-research study highlights the heterogeneity in methods and strategies used by researchers to assess and increase the validity of disproportionality signals. This validity is built on both plausibility (derived from existing evidence) and methodological robustness. Mapping available strategies is the first step toward validation studies and formal pharmacovigilance expert consensus on guidelines for designing disproportionality analyses and for assessing the validity of a published disproportionality analysis. Standardization of items and the development of a checklist will allow researchers to better design and transparently present the strengths and weaknesses of their study. Standardization of items would also allow readers, editors, and clinicians to assess the validity of disproportionality signals.

A more valid signal refinement activity, providing not only a transparent report of a methodologically robust disproportion pattern, but also its assessment and contextualization within existing evidence, would support regulatory agencies in more easily prioritizing and getting the most from published results, thus avoiding the accumulation of hard-to-handle disproportionality noise. We encourage more complete, valid, and transparent—even if a bit delayed—disproportionality studies to better manage safety signals by regulatory agencies and increase clinical transferability [52].

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40264-023-01329-w.

Acknowledgements Part of the results has been presented at the ISOP (International Society of Pharmacovigilance) 2022, held in Verona on the 20–23 of September 2022. ER and EP were supported by institutional research funds (Ricerca Fondamentale Orientata).

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

Declarations

Funding No specific funding supported this research.

Conflict of interest The authors declare no conflict of interest specific for this research.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Availability of data and material Our data consists of articles retrieved from Medline.

Code availability The code used for descriptive analysis and visualization is available upon request to the corresponding author.

Protocol registration The study was preregistered, and the protocol was made available at: https://osf.io/4exnh.

Author contributions MF, CK, ER, and FS conceptualized and designed the study and developed the methodology. MF, ER, CK, CB, and MI extracted data from the articles. MF and CK performed the visualization. MF and CK wrote the original draft. All the authors strongly contributed to the interpretation of results, and to the review and editing of the draft. All the authors read and approved the final version.

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. Raschi E, Moretti U, Salvo F, Pariente A, Antonazzo IC, Ponti FD, et al. Evolving roles of spontaneous reporting systems to assess and monitor drug safety. In: Kothari CS, Shah M, Manthan Patel R, editors. Pharmacovigilance. IntechOpen; 2018; https://www.intechopen.com/online-first/evolving-roles-of-spontaneous-reporting-systems-to-assess-and-monitor-drug-safety. Accessed 3 Feb 2019.
2. Sartori D, Aronson JK, Norén GN, Onakpoya IJ. Signals of adverse drug reactions communicated by pharmacovigilance stakeholders: a scoping review of the global literature. Drug Saf. 2022. https://doi.org/10.1007/s40264-022-01258-0.
3. Dhodapkar MM, Shi X, Ramachandran R, Chen EM, Wallach JD, Ross JS. Characterization and corroboration of safety signals identified from the US Food and Drug Administration Adverse Event Reporting System, 2008–19: cross sectional study. BMJ. 2022. https://doi.org/10.1136/bmj-2022-071752.
4. Faillie J-L. Case-non-case studies: principle, methods, bias and interpretation. Therapies. 2019;74:225–32. https://doi.org/10.1016/j.therap.2019.01.006.
5. Fusaroli M, Simonsen A, Borrie SA, Low DM, Parola A, Raschi E, et al. Identifying medications underlying communication atypicalities in psychotic and affective disorders: a pharmacovigilance study within the FDA Adverse Event Reporting System. J Speech Lang Hear Res. https://doi.org/10.1101/2022.09.05.22279609. (**in press**).
6. Gerhard T. Bias: considerations for research practice. Am J Health Syst Pharm. 2008;65:2159–68.

7. Raschi E, Poluzzi E, Salvo F, Pariente A, De Ponti F, Marchesini G, et al. Pharmacovigilance of sodium-glucose co-transporter-2 inhibitors: what a clinician should know on disproportionality analysis of spontaneous reporting systems. Nutr Metab Cardiovasc Dis. 2018;28:533–42.
8. Pariente A, Gregoire F, Fourrier-Reglat A, Haramburu F, Moore N. Impact of safety alerts on measures of disproportionality in spontaneous reporting databases: the notoriety bias. Drug Saf. 2007;30:891–8.
9. Khouri C, Nguyen T, Revol B, Lepelley M, Pariente A, Roustit M, et al. Leveraging the variability of pharmacovigilance disproportionality analyses to improve signal detection performances. Front Pharmacol. 2021;12: 668765.
10. Khouri C, Roustit M, Cracowski J-L. Caution in interpreting facial paralysis data to understand COVID-19 vaccination risks—reply. JAMA Intern Med. 2021;181:1420–1.
11. Dias P, Penedones A, Alves C, Ribeiro CF, Marques FB. The role of disproportionality analysis of pharmacovigilance databases in safety regulatory actions: a systematic review. Curr Drug Saf. 2015;10:234–50.
12. Hauben M, Aronson JK. Defining "signal" and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug Saf. 2009;32:99–110.
13. Raschi E, Gatti M, Gelsomino F, Ardizzoni A, Poluzzi E, De Ponti F. Lessons to be learnt from real-world studies on immune-related adverse events with checkpoint inhibitors: a clinical perspective from pharmacovigilance. Target Oncol. 2020;15:449–66.
14. Farcaş A, Măhălean A, Bulik NB, Leucuta D, Mogoşan C. New safety signals assessed by the Pharmacovigilance Risk Assessment Committee at EU level in 2014–2017. Expert Rev Clin Pharm. 2018;11:1045–51.
15. Insani WN, Pacurariu AC, Mantel-Teeuwisse AK, Gross-Martirosyan L. Characteristics of drugs safety signals that predict safety related product information update. Pharmacoepidemiol Drug Saf. 2018;27:789–96.
16. Khouri C, Revol B, Lepelley M, Mouffak A, Bernardeau C, Salvo F, et al. A meta-epidemiological study found lack of transparency and poor reporting of disproportionality analyses for signal detection in pharmacovigilance databases. J Clin Epidemiol. 2021;139:191–8.
17. Pacurariu AC, Coloma PM, Gross-Martirosyan L, Sturkenboom MC, Straus SM. Decision making in drug safety—a literature review of criteria used to prioritize newly detected safety issues. Pharmacoepidemiol Drug Saf. 2017;26:327–34.
18. Croteau D, Pinnow E, Wu E, Muñoz M, Bulatao I, Dal Pan G. Sources of evidence triggering and supporting safety-related labeling changes: a 10-year longitudinal assessment of 22 new molecular entities approved in 2008 by the US Food and Drug Administration. Drug Saf. 2022;45:169–80.
19. Tau N, Shochat T, Gafter-Gvili A, Tibau A, Amir E, Shepshelovich D. Association between data sources and us food and drug administration drug safety communications. JAMA Intern Med. 2019;179:1590–2.
20. Hill AB. The Environment and disease: association or causation? Proc R Soc Med. 1965;58:295–300.
21. Fusaroli M, Raschi E, Khouri C, Idris M, Salvo F. The approaches to assess robustness of disproportionality signals in pharmacovigilance. A meta-epidemiological research. OSF; 2022 [cited 2022 Aug 1]. Available from: https://osf.io/4exnh.
22. Sáinz-Gil M, Rodríguez-Jiménez P, Cárdaba ME, Martín Arias LH. Kaposi sarcoma in patients on tumour necrosis factor alpha inhibitors: a case–noncase study. J Eur Acad Dermatol Venereol. 2020;34:e156–8.
23. Antonazzo IC, Poluzzi E, Forcesi E, Riise T, Bjornevik K, Baldin E, et al. Liver injury with drugs used for multiple sclerosis: a contemporary analysis of the FDA Adverse Event Reporting System. Mult Scler. 2019;25:1633–40.
24. Poluzzi E, Raschi E, Koci A, Moretti U, Spina E, Behr ER, et al. Antipsychotics and torsadogenic risk: signals emerging from the US FDA Adverse Event Reporting System database. Drug Saf. 2013;36:467–79.
25. WHO—Uppsala Monitoring Centre. The use of the WHO-UMC system for standardized case causality assessment. 2018 [cited 2022 Dec 18]. Available from: http://www.who-umc.org/Graphics/24734.pdf.
26. Miremont-Salamé G, Théophile H, Haramburu F, Bégaud B. Causality assessment in pharmacovigilance: the French method and its successive updates. Therapie. 2016;71:179–86.
27. European Medicines Agency. 2020 Annual Report on EudraVigilance for the European Parliament, the Council and the Commission. 2021. Available from: https://www.ema.europa.eu/en/documents/report/2020-annual-report-eudravigilance-european-parliament-council-commission_en.pdf.
28. Hlavaty A, Roustit M, Montani D, Chaumais M-C, Guignabert C, Humbert M, et al. Identifying new drugs associated with pulmonary arterial hypertension: a WHO pharmacovigilance database disproportionality analysis. Br J Clin Pharmacol. 2022;88:5227–37.
29. Whaley P, Piggott T, Morgan RL, Hoffmann S, Tsaioun K, Schwingshackl L, et al. Biological plausibility in environmental health systematic reviews: a GRADE concept paper. Environ Int. 2022;162:107109.
30. Khouri C, Lepelley M, Roustit M, Montastruc F, Humbert M, Cracowski JL. Comparative safety of drugs targeting the nitric oxide pathway in pulmonary hypertension: a mixed approach combining a meta-analysis of clinical trials and a disproportionality analysis from the World Health Organization pharmacovigilance database. Chest. 2018;154(1):136–47.
31. Faillie J-L, Petit P, Montastruc J-L, Hillaire-Buys D. Scientific evidence and controversies about pioglitazone and bladder cancer: which lessons can be drawn? Drug Saf. 2013;36:693–707.
32. Raschi E, Salvo F, Khouri C. Conceiving, conducting, reporting, interpreting, and publishing disproportionality analyses: a call to action. Br J Clin Pharmacol. 2022;88:3535–6.
33. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021;50:620–32.
34. Wisniewski AFZ, Bate A, Bousquet C, Brueckner A, Candore G, Juhlin K, et al. Good signal detection practices: evidence from IMI PROTECT. Drug Saf. 2016;39:469–90.
35. Vogel U, van Stekelenborg J, Dreyfus B, Garg A, Habib M, Hosain R, et al. Investigating overlap in signals from EVDAS, FAERS, and VigiBase®. Drug Saf. 2020;43:351–62.
36. Noguchi Y, Tachi T, Teramachi H. Detection algorithms and attentive points of safety signal using spontaneous reporting systems as a clinical data source. Brief Bioinform. 2021;22:bbab347.
37. Raschi E, Poluzzi E, Salvo F, Koci A, Suling M, Antoniazzi S, et al. The contribution of national spontaneous reporting systems to detect signals of torsadogenicity: issues emerging from the ARITMO project. Drug Saf. 2016;39:59–68.
38. Ahmed I, Thiessard F, Miremont-Salamé G, Haramburu F, Kreft-Jais C, Bégaud B, et al. Early detection of pharmacovigilance signals with automated methods based on false discovery rates: a comparative study. Drug Saf. 2012;35:495–506.
39. Gastaldon C, Schoretsanitis G, Arzenton E, Raschi E, Papola D, Ostuzzi G, et al. Withdrawal syndrome following discontinuation of 28 antidepressants: pharmacovigilance analysis of 31,688 reports from the WHO spontaneous reporting database. Drug Saf. 2022;45:1539–49.

40. Gatti M, Antonazzo IC, Diemberger I, De Ponti F, Raschi E. Adverse events with sacubitril/valsartan in the real world: emerging signals to target preventive strategies from the FDA adverse event reporting system. Eur J Prev Cardiol. 2021;28:983–9.
41. Pearson RK, Hauben M, Goldsmith DI, Gould AL, Madigan D, O'Hara DJ, et al. Influence of the MedDRA hierarchy on pharmacovigilance data mining results. Int J Med Inform. 2009;78:e97-103.
42. Reich CG, Ryan PB, Schuemie MJ. Alternative outcome definitions and their effect on the performance of methods for observational outcome studies. Drug Saf. 2013;36(Suppl 1):S181–93.
43. Fusaroli M, Raschi E, Giunchi V, Menchetti M, Rimondini Giorgini R, De Ponti F, et al. Impulse control disorders by dopamine partial agonists: a pharmacovigilance-pharmacodynamic assessment through the FDA adverse event reporting system. Int J Neuropsychopharmacol. 2022;25:727–36.
44. Raschi E, Fusaroli M, Ardizzoni A, Poluzzi E, De Ponti F. Cyclin-dependent kinase 4/6 inhibitors and interstitial lung disease in the FDA adverse event reporting system: a pharmacovigilance assessment. Breast Cancer Res Treat. 2021;186:219–27.
45. Golder S, Smith K, O'Connor K, Gross R, Hennessy S, Gonzalez-Hernandez G. A comparative view of reported adverse effects of statins in social media, regulatory data, drug information databases and systematic reviews. Drug Saf. 2021;44:167–79.
46. Smith K, Golder S, Sarker A, Loke Y, O'Connor K, Gonzalez-Hernandez G. Methods to compare adverse events in twitter to FAERS, drug information databases, and systematic reviews: proof of concept with adalimumab. Drug Saf. 2018;41:1397–410.
47. Oster ME, Shay DK, Su JR, Gee J, Creech CB, Broder KR, et al. Myocarditis cases reported after mRNA-based COVID-19 vaccination in the US from December 2020 to August 2021. JAMA. 2022;327:331–40.
48. Raschi E, Poluzzi E, Godman B, Koci A, Moretti U, Kalaba M, et al. Torsadogenic risk of antipsychotics: combining adverse event reports with drug utilization data across Europe. PLoS ONE. 2013;8: e81208.
49. Cornet L, Khouri C, Roustit M, Guignabert C, Chaumais M-C, Humbert M, et al. Pulmonary arterial hypertension associated with protein kinase inhibitors: a pharmacovigilance-pharmacodynamic study. Eur Respir J. 2019;53:1802472.
50. Mahé J, de Campaigno EP, Chené A-L, Montastruc J-L, Despas F, Jolliet P. Pleural adverse drugs reactions and protein kinase inhibitors: identification of suspicious targets by disproportionality analysis from VigiBase: pleural adverse drugs reactions and protein kinase inhibitors. British Journal of Clinical Pharmacology [Internet]. 2018 [cited 2018 Sep 7]. Available from: https://doi.org/10.1111/bcp.13693
51. Nguyen TTH, Pariente A, Montastruc J-L, Lapeyre-Mestre M, Rousseau V, Rascol O, et al. An original pharmacoepidemiological-pharmacodynamic method: application to antipsychotic-induced movement disorders. Br J Clin Pharmacol. 2017;83:612–22.
52. Khouri C, Fusaroli M, Salvo F, Raschi E. Transparency and robustness of safety signals. BMJ. 2022;379:o2588.