



# Risk Factor Considerations in Statistical Signal Detection: Using Subgroup Disproportionality to Uncover Risk Groups for Adverse Drug Reactions in VigiBase

Lovisa Sandberg<sup>1</sup> · Henric Taavola<sup>1</sup> · Yasunori Aoki<sup>1,2</sup> · Rebecca Chandler<sup>1</sup> · G. Niklas Norén<sup>1</sup>

Published online: 20 June 2020  
© The Author(s) 2020

## Abstract

**Introduction** In the treatment of the individual patient, a vision is to achieve the best possible balance between benefit and harm. Such tailored therapy relies upon the identification and characterisation of risk factors for adverse drug reactions. Information relevant to risk factor considerations can be captured in adverse event reports and could be utilised in statistical signal detection.

**Objective** The aim of this study was to explore whether statistical screening of a broad range of risk factors within a global database of adverse event reports could uncover signals of risk groups for adverse drug reactions.

**Methods** Subgroup disproportionality analysis was applied to 15.4 million reports entered in VigiBase, the World Health Organization (WHO) global database of individual case safety reports, up to August 2017. Disproportionality analyses for drug–adverse event pairs were performed (1) in the full database and (2) across a range of subgroups defined by the following covariates: patient age, sex, body mass index, pregnancy, underlying condition, reporting country, and geographical region. Drug–adverse event pairs disproportionately over-reported in such subgroups, but not in the full database, and with a substantial difference between the two observed-to-expected ratios, were highlighted as statistical signals. These were further prioritised, through filtering and sorting, for clinical assessment, whereafter clinically relevant signals were communicated to the pharmacovigilance community and the public.

**Results** Assessments were performed for 354 prioritised statistical signals, resulting in seven communicated signals describing previously unrecognised potential risk groups related to age (elderly), sex (male and female), body mass index (underweight and obese), and geographical region (Asia), all except one for already established adverse drug reactions. Important aspects considered in the assessments included an evaluation of the disproportionate over-reporting in the subgroup by reviewing alternative explanations and reporting patterns for similar drugs/adverse events/subgroups, and a search for plausible mechanisms to support the risk hypothesis.

**Conclusions** This study reveals that it is possible to uncover signals of risk groups for adverse drug reactions through incorporation of broad risk factor screening into statistical signal detection in a global database of adverse event reports. Our findings suggest the potential to use such statistical methodologies for risk characterisation in subpopulations of concern.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s40264-020-00957-w>) contains supplementary material, which is available to authorized users.

---

✉ Lovisa Sandberg  
lovisa.sandberg@who-umc.org

<sup>1</sup> Uppsala Monitoring Centre, Uppsala, Sweden

<sup>2</sup> National Institute of Informatics, Tokyo, Japan

## 1 Introduction

Within the scientific community, there is growing attention to precision medicine, most often considered in the context of using genetic diversity to predict which patients are more likely to receive benefit from certain medicines. However, it is also true that individual risk factors can be used to predict which patients are more likely to experience adverse drug reactions (ADRs). With the identification and characterisation of risk factors for ADRs, steps can be taken to minimise harms from medications.

### Key Points

The identification of specific patient subpopulations at increased risk for adverse drug reactions can help minimise harms from medicines.

This study identified clinically relevant signals of at-risk groups through the incorporation of risk factor considerations into statistical signal detection in adverse event reports.

The findings from this study may be a step on the way towards increased precision in pharmacovigilance and may inspire future research in the area.

Risk factors for the development of ADRs have been documented in a range of populations and from a variety of healthcare settings [1, 2]. Patient-level factors increase the risk of ADRs largely through an effect on either the pharmacokinetics or pharmacodynamics of drugs. Changes in physiology resulting from decreased renal or liver function, pregnancy, or increased/decreased body mass can affect the absorption, distribution, metabolism, and excretion of medications, while genetic polymorphisms or hormonal controls on genetic expression can affect both metabolic enzymes involved in pharmacokinetic pathways or the receptors/targets of pharmacodynamic pathways [3–6]. Risk factors for ADRs with an immunological pathophysiology, such as human leukocyte antigen (HLA)-associated ADRs, can also be genetically determined. Many of these patient-level risk factors, such as age, sex, and concomitant illness, can be captured in adverse event reports.

Historically, the main emphasis of signal detection in pharmacovigilance has been on the early identification of previously unknown causal associations between medicines and ADRs. Signal detection in large databases of spontaneous reports of adverse events is performed in a statistical screening process that uses a pair-wise analysis to detect disproportionality between the number of observed versus expected reports of a single drug and a single adverse event. Review of patient-level data for recognition of potential risk factors for the ADR occurs in general only after detection of a drug–ADR signal, during signal assessment.

The incorporation of risk factor considerations earlier in the signal management process has the potential to improve the efficiency and precision of action on post-marketing safety signals. The aim of this study was to explore whether statistical screening of a broad range of risk factors within a global database of adverse event reports could uncover signals of risk groups for ADRs.

## 2 Methods

### 2.1 Dataset

The source of data was VigiBase, the World Health Organization (WHO) global database of individual case safety reports, which includes reports of adverse events shared by national pharmacovigilance centres in the member countries of the WHO Programme for International Drug Monitoring [7]. The reports in VigiBase come from a variety of sources, including physicians, pharmacists, and consumers, are mostly spontaneously submitted, or, less commonly, come from studies, and—importantly—indicate varying degrees of probability that the adverse event is drug related.

The dataset used for this study consisted of all reports entered in VigiBase up to 28 August 2017. Suspected duplicate reports were identified with the *vigiMatch* algorithm [8] and excluded from the dataset, resulting in 15.4 million reports for analysis.

Analyses were performed on the active ingredient level of drugs, coded according to WHODrug Global [9], and on the Preferred Term level of adverse events, coded according to MedDRA®, the Medical Dictionary for Regulatory Activities terminology (version 20.0). Only reports where the drug was characterised as suspected or interacting were counted in the statistical screening.

### 2.2 Scope

A selection of risk factors for ADRs defined the set of covariates to be screened in the statistical signal detection. The risk factors were selected based on clinical relevance and technical feasibility and included patient age, sex, body mass index (BMI), pregnancy, underlying condition, reporting country, and geographical region. Within each covariate, one or several subgroups were defined, representing the potential risk groups to be explored. The covariates and their corresponding subgroups with justifications are presented in Table 1.

For patient age, the subgroups were 0–27 days, 28 days–23 months, 2–11 years, 12–17 years, 18–44 years, 45–64 years, 65–74 years, and  $\geq 75$  years, and for patient sex, female and male formed the two subgroups. For BMI, only underweight adults and obese adults were chosen as subgroups; the underweight subgroup comprised adults with  $BMI < 18.5$ , and the obese subgroup comprised adults with  $BMI \geq 30$ , as defined by the reference cut-off values described by WHO [10].

The pregnancy subgroup aimed to identify harms primarily for the pregnant woman, not the foetus; however, there is no single structured field in VigiBase indicating the

**Table 1** Covariates and the corresponding subgroups with justifications

Covariate	Subgroups	Justification for consideration as a risk factor
Age	0–27 days, 28 days–23 months, 2–11 years, 12–17 years, 18–44 years, 45–64 years, 65–74 years, ≥ 75 years	Variations in drug usage (e.g. dosage, route of administration), maturity/efficiency of renal clearance and hepatic metabolism, volume of distribution, comorbidities and polypharmacy especially in the elderly
Sex	Male, female	Pharmacokinetics affected due to e.g. body composition and variations in hepatic metabolism and renal clearance
Body mass index	Underweight adult, obese adult	Pharmacokinetics, e.g. drug distribution, affected due to body composition
Pregnancy	Pregnant	Pharmacokinetics affected due to e.g. changed activity of hepatic metabolising enzymes and increased renal blood flow
Underlying condition	Reported indications (MedDRA HLGs)	The patient's underlying medical condition/comorbidities may affect the pharmacokinetics or pharmacodynamics, e.g. chronic kidney disease may affect renal clearance, and polypharmacy may cause drug–drug interactions
Country	Individual reporting countries	Variations in drug usage and genetic polymorphisms of metabolising enzymes in different populations
Geographical region	Africa, Asia, Europe, Latin America and the Caribbean, Northern America, Oceania	Variations in drug usage and genetic polymorphisms of metabolising enzymes in different populations

HLG: High Level Group Terms

pregnancy status of a patient. To capture reports of likely pregnant patients, a simple algorithm was constructed to infer the pregnancy status from available information in Vigibase, such as reported terms, indications, medical history, tests and procedures, and the case narrative. For details, see Table S1 in the Electronic Supplementary Material.

Subgroups within the underlying condition covariate were defined using the reported drug indications grouped by MedDRA High Level Group Terms. The reported indications were used to infer the comorbidities of the patients, and thus the indications for all drugs on the reports were considered, not only the indication for the drug of interest.

Subgrouping by country or geographical region was primarily intended to be a surrogate for pharmaco-ethnic vulnerability. For the country covariate, each individual country holding reports in Vigibase represented a subgroup. The countries were grouped into regions to form subgroups within the geographical region covariate.

### 2.3 Statistical Signal Detection

Subgroup disproportionality analysis, as described by Hopstadius and Norén [11], was employed as the basis for the statistical signal detection. As in regular disproportionality analysis, observed-to-expected (OE) ratios contrast the number of reports on drug  $x$  with adverse event  $y$  to an expected value based on the total number of reports on drug  $x$  and the overall relative frequency of adverse event  $y$ . Here, the

computation is restricted to subgroup  $z$  (e.g. reports for a specific patient age group):

$$OE_{xy}^z = \frac{O_{xy}^z}{E_{xy}^z} = \frac{n_{xyz}}{n_{xz} \cdot n_{yz}/n_z}$$

Statistical shrinkage of the OE ratio towards 1 (the value implying no association) provides protection against spurious associations [12]. The base 2 logarithm of such a shrunk OE ratio is referred to as the Information Component (IC), and this is the disproportionality measure used throughout this paper:

$$IC_{xy}^z = \log_2 \frac{O_{xy}^z + 1/2}{E_{xy}^z + 1/2}$$

If the IC value exceeds zero, it means that the corresponding OE ratio exceeds 1, i.e. there are more reports than would be expected.

Bayesian credibility intervals indicating a range of IC values compatible with the data can be calculated, with pre-specified coverage probabilities. The lower limit of a 95% credibility interval for the IC is denoted  $IC_{0.25}$  and is commonly used in regular disproportionality analysis. For subgroup analyses, broader coverage credibility intervals are used to control the rate of spurious associations, because of the multiple comparisons [11].

In this study, two requirements were set up to identify statistical signals that may suggest risk groups for an ADR: (1) that the drug and adverse event be disproportionately

over-reported in subgroup  $z$ , but not in the database as a whole; and (2) that there be a substantial difference between the OE ratio for subgroup  $z$  and that for the rest of the database.

The first requirement seeks to highlight associations that would not be detected in regular disproportionality analysis. It identifies drug–adverse event–subgroup associations for which the 99.9% credibility interval of the IC exceeds zero in subgroup  $z$  while the 95% interval of the IC for the entire database does not. Or, alternatively, that the 99% interval of the IC exceeds zero simultaneously in two subgroups for the same covariate (for example, two different age groups) while the 95% interval of the IC for the entire database does not. The latter allows patterns that exist in multiple subgroups to be detected with a less strict credibility interval for each individual subgroup [11].

The second requirement seeks to highlight associations with pronounced contrast between the subgroup and the rest of the database. It identifies drug–adverse event–subgroup associations for which the OE ratio in subgroup  $z$  is at least twice as large as the OE ratio in the rest of the database, when adjusted for the corresponding covariate (e.g. age, if the subgroup of interest is children). The adjusted OE ratio for the rest of the database is computed as a weighted average of the OE ratios in the other subgroups. It simplifies to an OE ratio where the observed and expected counts are summed across subgroups [12]:

$$\frac{O'}{E'} = \frac{\sum_{j \neq z} O_{xy}^j}{\sum_{j \neq z} E_{xy}^j}.$$

Technically, a new OE ratio is constructed by dividing the subgroup-specific OE ratio and the adjusted OE ratio for the rest of the database [12]:

$$\frac{O^z/E^z}{O'/E'} = \frac{O^z}{E^z \cdot O'/E'}.$$

The corresponding IC value ( $IC_{\Delta}$ ) is then computed and is required to exceed 1, corresponding to a ratio of at least 2 between the subgroup-specific OE ratio and the adjusted OE ratio for the rest of the database:

$$IC_{\Delta} = \log_2 \frac{O^z + 1/2}{E^z \cdot O'/E' + 1/2}.$$

## 2.4 Triages for Signal Assessment

In the prioritisation for clinical review, triages were applied to filter and sort the statistical signals.

### 2.4.1 Filters

In a first phase of the study, the statistical signals fulfilling any of the below triage criteria were excluded:

- *No reports entered in VigiBase since 2012*  
To focus on current concerns. Related to reports on the drug–adverse event pair within the subgroup of interest.
- *Non-specific drug*  
To avoid non-actionable associations where the reported drug names were too general to support an assessment, such as “vaccines” or “analgesics”.
- *Drug with at least 80% of the reports within the subgroup*  
To avoid drugs that were almost exclusively reported within the subgroup of interest, thus difficult to argue for a risk group.
- *Adverse event with at least 80% of the reports within the subgroup*  
To avoid adverse events that were almost exclusively reported within the subgroup of interest, thus difficult to argue for a risk group.

In a second phase of the study, the scope was narrowed to focus on signals for serious events within certain covariates. In addition to the criteria listed above, the following exclusion criteria were applied:

- *Adverse event not classified as serious*  
To focus on the potentially more clinically significant events. Serious events were defined as either being an Important Medical Event as per the European Medicines Agency (EMA) [13] or at least 75% of the E2B reports in the case series being described as serious [14].
- *Selected covariates*  
To focus on the covariates more likely to qualify for in-depth assessment (based on interim results from the first phase of the study<sup>1</sup>). Patient age (restricted to the age groups within 0–17 years and  $\geq 65$  years), sex, country, and geographical region were kept, while BMI, pregnancy, and underlying condition were excluded.

### 2.4.2 Sorting

The remaining statistical signals were further prioritised using *vigiRank*, a predictive model for emerging safety signals that accounts for multiple aspects of strength of evidence in the case series [15]. To achieve the maximal *vigiRank* score, there needs to be disproportionate

<sup>1</sup> The preliminary assessments for some of the signals for BMI eventually selected for in-depth assessment (see Table 2) had not been finalized at the time of the interim evaluation.

over-reporting, reports from at least seven countries, four or more reports in the preceding 3 years, at least five reports with narratives, and at least five reports with high completeness of information. Signals with identical vigiRank scores were secondarily sorted through stratified random sampling across subgroups to favour the variety of subgroups explored, mitigating the risk of imbalanced representation of signals for certain subgroups. Equal weight was given to all subgroups within a covariate regardless of their total number of statistical signals.

### 2.5 Signal Assessment

Top prioritised statistical signals were clinically reviewed in a two-step process to decide whether a signal should be communicated.

#### 2.5.1 Preliminary Signal Assessment

Preliminary assessments were performed according to the review decision tree presented in Fig. 1. Three main aspects were considered: (1) the actionability of the signal (e.g. Is the drug/adverse event specific enough? Is the drug on the market?), (2) the previous awareness of the risk group (i.e. Is the risk group adequately described in the product label?), and (3) supportive evidence for a risk group from the case series (e.g. Are there any obvious alternative reasons for the disproportionate reporting in the subgroup? Are the cases informative enough to conduct an in-depth assessment?). The signals were then classified as “closed” or as “signal for in-depth assessment”.

To identify previous awareness of a risk group, the Summary of Product Characteristics in the electronic Medicines Compendium (eMC) [16] and the US Food and Drug Administration (FDA) drug labels in DailyMed [17] were consulted. To further support the assessments, a collection of points to consider was compiled for the purpose of this study (see Table S2 in the Electronic Supplementary Material).

The preliminary assessments were conducted by pharmacovigilance researchers at Uppsala Monitoring Centre (UMC), working in interdisciplinary teams of medical doctors, pharmacists, and data scientists. To ensure that signals from all covariates were explored in a balanced way, the different covariates were divided between the teams and reviewed separately. To reduce the risk of systematic bias, the covariates were regularly redistributed to different teams.

#### 2.5.2 In-Depth Signal Assessment

Experienced pharmacovigilance assessors further evaluated the signals selected for in-depth assessment to decide if there was enough supporting evidence to communicate a

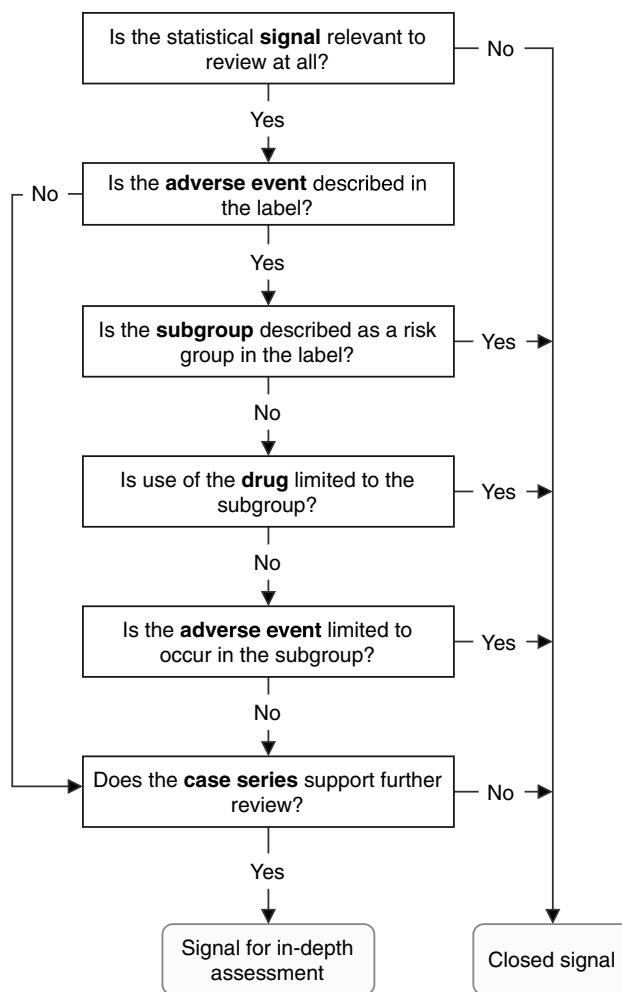


Fig. 1 Review decision tree with the main aspects of the preliminary signal assessment

signal. The individual case reports were reviewed in detail to investigate if these were supportive of a risk group or if the disproportionate reporting in the subgroup was more likely due to other reasons, and the literature was consulted for any further supporting evidence.

### 2.6 Signal Communication

Signals which were considered to potentially describe previously unrecognised risk groups for ADRs were communicated in the form of a report to the national pharmacovigilance centres participating in the WHO Programme for International Drug Monitoring and to the public through the WHO Pharmaceuticals Newsletter.

### 3 Results

Preliminary assessments were performed for 354 statistical signals, representing the highest prioritised signals for each covariate according to the filtering and sorting approaches.<sup>2</sup> Of these, 19 (5.4%) were selected for in-depth assessment, resulting in seven (2.0%) signals for communication: hepatitis with ceftriaxone in patients 75 years and older [18], myoclonus with levofloxacin in patients 75 years and older [19], anaphylactic shock with omalizumab in females [20], deep vein thrombosis and pulmonary embolism with aflibercept in males [21], gynaecomastia with esomeprazole in obese adults [22], hypoglycaemia with selegiline in underweight adults [23], and palpitations with glibenclamide in the Asian population [24].

#### 3.1 Characteristics of Assessed Signals

Breaking down the results, 293 statistical signals were preliminarily assessed in the first phase of the study and 61 in the second phase. In the first phase, 14 signals (4.8%) qualified for in-depth assessment, resulting in five signals (1.7%) for communication [20–24]. In the second phase, five signals (8.2%) were assessed in-depth, resulting in two signals (3.3%) for communication [18, 19]. Tables 2 and 3 present the number of assessments and signals for communication by covariate.

The preliminarily assessed signals were relatively evenly distributed across covariates and subgroups. However, for some specific subgroups, there were no assessed signals: countries and underlying conditions had too many subgroups to consider each one, and for subgroups such as Northern America and Oceania, the vigiRank score was limited by the low number of countries in these regions. Patient sex and BMI yielded the highest number of signals selected for in-depth assessment, while pregnancy and underlying condition yielded none. The communicated signals related to patient age (elderly patients), sex (female and male), BMI (underweight and obese adults), and geographical region (Asia).

Three of the five communicated signals identified in the first phase [20, 21, 23] and both signals identified in the second phase [18, 19] described serious events. The adverse event terms most commonly assessed in the two phases are presented in Tables 4 and 5.

Six of the communicated signals represented previously unrecognised risk groups for already known ADRs, whereas

**Table 2** Signals by covariate in the first phase of the study

Covariate Subgroup	Preliminary assessed signals	Signals assessed in-depth (%)	Communicated signals (%)
<b>Sex</b>	<b>32</b>	<b>5 (16)</b>	<b>2 (6.3)</b>
Female	16	4 (25)	1 (6.3)
Male	16	1 (6.3)	1 (6.3)
<b>Body mass index</b>	<b>41</b>	<b>5 (12)</b>	<b>2 (4.9)</b>
Underweight adult	16	3 (19)	1 (6.3)
Obese adult	25	2 (8.0)	1 (4.0)
<b>Geographical region</b>	<b>50</b>	<b>2 (4.0)</b>	<b>1 (2.0)</b>
Asia	18	2 (11)	1 (5.6)
<b>Country</b>	<b>35</b>	<b>1 (2.9)</b>	<b>0 (0.0)</b>
Korea, Republic of	2	1 (50)	0 (0.0)
<b>Age</b>	<b>43</b>	<b>1 (2.3)</b>	<b>0 (0.0)</b>
75 years and older	5	1 (20)	0 (0.0)
<b>Pregnancy</b>	<b>32</b>	<b>0 (0.0)</b>	<b>0 (0.0)</b>
<b>Underlying condition</b>	<b>60</b>	<b>0 (0.0)</b>	<b>0 (0.0)</b>
<b>Total</b>	<b>293</b>	<b>14 (4.8)</b>	<b>5 (1.7)</b>

Shows all covariates as well as individual subgroups with at least one signal subjected to in-depth signal assessment

**Table 3** Signals by covariate in the second phase of the study

Covariate Subgroup	Preliminary assessed signals	Signals assessed in-depth (%)	Communicated signals (%)
<b>Age</b>	<b>15</b>	<b>2 (13)</b>	<b>2 (13)</b>
75 years and older	7	2 (29)	2 (29)
<b>Geographical region</b>	<b>19</b>	<b>2 (11)</b>	<b>0 (0.0)</b>
Europe	18	2 (11)	0 (0.0)
<b>Sex</b>	<b>11</b>	<b>1 (9.1)</b>	<b>0 (0.0)</b>
Female	7	1 (14)	0 (0.0)
<b>Country</b>	<b>16</b>	<b>0 (0.0)</b>	<b>0 (0.0)</b>
<b>Total</b>	<b>61</b>	<b>5 (8.2)</b>	<b>2 (3.3)</b>

Shows all covariates as well as individual subgroups with at least one signal subjected to in-depth signal assessment

one [19] suggested a risk group for a previously non-established ADR.

#### 3.2 Rationales of Signal Assessment Outcomes

##### 3.2.1 Closed Signals

Signals were closed if the subgroup of interest was already characterised as a risk group in the product label or if the adverse event was known to occur very commonly with the drug. Signals were also closed if there were no explanations

<sup>2</sup> Preliminary assessments were performed for an additional 53 statistical signals which are not presented in the results as these were identified further down the order of priority and had been selected to be assessed together with higher prioritized signals, for example, due to similarities in the type of event or closely connected subgroups.

**Table 4** Top assessed adverse events in the first phase of the study (total number of assessed signals = 293)

Adverse event (MedDRA Preferred Term)	Preliminary assessed signals (%)
Headache	19 (6.5)
Pyrexia	11 (3.8)
Drug ineffective	10 (3.4)
Pruritus	10 (3.4)
Urticaria	7 (2.4)
Dizziness	6 (2.0)
Dyspnoea	6 (2.0)
Malaise	6 (2.0)
Nausea	6 (2.0)
Palpitations	6 (2.0)

**Table 5** Top assessed adverse events in the second phase of the study (total number of assessed signals = 61)

Adverse event (MedDRA Preferred Term)	Preliminary assessed signals (%)
Overdose	3 (4.9)
Anaemia	2 (3.3)
Anaphylactic reaction	2 (3.3)
Angioedema	2 (3.3)
Death	2 (3.3)
Loss of consciousness	2 (3.3)
Tachycardia	2 (3.3)

found for the disproportionate reporting in the subgroup (e.g. lack of plausible mechanisms), if the disproportionate over-reporting was not unique to the specific subgroup (e.g. too broad subgroup), or if there was no evidence of a similar reporting pattern for the subgroup for a clinically related adverse event. Finally, if review of the subgroup case series identified alternative explanations for the disproportionate reporting, or if the case series did not contain enough information for an assessment, the signal was closed. Individual examples of reasons for closing signals are presented in Table 6.

### 3.2.2 Communicated Signals

The rationales for the seven communicated signals were multi-layered with some common themes. In all signals, the assessor brought forward the unique disproportionate over-reporting of the drug–adverse event pair in the subgroup of interest in contrast to the full data. All assessors also proposed plausible mechanisms for why the subgroup of patients may be at increased risk for the adverse event. Even though a causal relationship between the drug and the adverse event was previously known for most signals,

standard causality assessments were performed for the case series subgroup. Some assessors also characterised the case series outside the subgroup to exclude any obvious alternative explanations for the imbalanced reporting patterns. Table 7 lists the communicated signals and suggested mechanisms for risk group susceptibility.

## 4 Discussion

This study reveals that it is possible to uncover signals of risk groups for ADRs through incorporation of broad risk factor screening into statistical signal detection in a global database of adverse event reports. Using subgroup disproportionality analysis, we identified seven clinically relevant signals related to patient age, sex, BMI, and geographical region for communication to the pharmacovigilance community and the public. The proportion of signals for communication identified in this study (2%) is in line with previous experience of statistical signal detection in VigiBase for both pairwise drug–adverse event associations [25] and specific subgroups [26].

The risk factor signal detection approach evaluated here is based on subgroup disproportionality analysis. Stratified analyses, commonly used in epidemiology to adjust for possible confounding factors, also divide data into subgroups, but then compute a single measure, for example as a weighted average across the subgroups; by design they will fail to appreciate any variation in risk between the strata. A previous study evaluated the performance of subgroup and stratified disproportionality analyses within spontaneous report databases of differing sizes and characteristics, using a reference set of established ADRs [27]. Overall, subgroup analyses were found to perform better than stratified analyses for use in first pass signal detection; subgrouping by age, country, and geographical region showed the highest improvement in precision and sensitivity, while subgrouping by sex showed modest improvement (BMI, pregnancy, and underlying condition covariates were not investigated) [27]. The findings for patient sex and BMI in the current study might thus be especially worth noting; to our knowledge, screening for these risk factors is not currently applied in routine statistical signal detection.

The screening of a broad variety of risk factors to explore different types of risk groups concurrently is a novel approach and a strength of this study, allowing for more open-ended signal detection with less pre-set hypothesis around a certain risk factor. Inclusion of single risk group considerations into statistical signal detection approaches have been described, especially for age. UMC has previously performed a signal detection exercise to identify signals from the paediatric subgroup of case reports contained within VigiBase [26]. Similarly, the EMA has implemented

**Table 6** Examples of reasons for closing signals, grouped by category

Category	Reason for closing signal
Too broad subgroup	The adverse event was disproportionately over-reported with the drug in several age groups, resulting in too broad an age group to assess as a risk group (e.g. 12–64 years) The adverse event was disproportionately over-reported with the drug for both subgroups (underweight and obese adults), hence, difficult to argue for a risk group The underlying condition was too non-specific, and no common pattern was identified among the reported indications (e.g. general system disorders)
Alternative explanations for subgroup disproportionality	The drug was used for different indications in males vs females; the overrepresentation of the adverse event in one of the subgroups was more likely due to the underlying disease (e.g. prostate cancer vs endometriosis for leuprorelin) The subgroup was, to a greater extent, concomitantly using another drug known to cause the adverse event The underlying condition was closely related to the adverse event, and it was difficult to argue for a potential risk group (e.g. acetylsalicylic acid and pruritus in allergic conditions) Most cases in the subgroup described non-compliance with prescription instructions, including disregard of drug–drug interactions and dosage recommendations More likely that the increased reporting among females was related to a large number of reports describing overdoses, suicidal attempts, and brand switching/reported ineffectiveness
Lack of biological plausibility	Unlikely biologic plausibility that the reaction would be influenced by weight No logical genetic or cultural explanation, and no pharmacological mechanism or metabolism found to differ for the geographical region of interest
Lack of subgroup disproportionality for related adverse events/drugs	Other related adverse events commonly reported with the drug, but none of them reproduced the disproportionate reporting in the geographical region of interest The adverse event appeared to be a symptom reported with hypersensitivity; however, a similar pattern for other terms related to hypersensitivity was not seen for females
Drug/adverse event exclusive to subgroup	The drug was exclusively used in pregnancy; hence, not relevant to talk about a risk group
Targeted monitoring in subgroup	The disproportionate reporting within the specific country was likely due to close monitoring of the drug
Misclassification of subgroup	The patients were not pregnant at the time of exposure
Very common adverse drug reaction	The adverse event was labelled as very common in general, hard to justify further highlighting of this event specifically for patients 75 years and older

approaches of using within-group disproportionality to identify signals specific to the paediatric and the geriatric subpopulations in EudraVigilance [28, 29]. The FDA has described its principles to monitor the at-risk groups of children, elderly, and pregnant women in a recent draft document of best practices for post-marketing safety surveillance within the FDA Adverse Event Reporting System (FAERS) [30].

However, the broad screening approach restricted the number of statistical signals assessed for each covariate in this study, limiting the ability to draw conclusions for each separate risk factor, especially in the cases where no signals for a covariate or subgroup were deemed to merit in-depth review. A narrower scope, for example limited to a single covariate, such as patient sex or even a specific subgroup such as females, might also allow assessors to more rapidly learn to assess the subgroup signals of interest. It may furthermore allow the statistical signal detection algorithms to be customised, with a possible improvement in performance.

Missing data on covariates is an impediment. This is particularly problematic for covariates such as pregnancy and BMI with high degrees of missing data. Improved algorithms

for identifying reports related to pregnant women may be needed in view of the many ways that this can be recorded on adverse event reports. The algorithm used in this study was designed to promote precision over sensitivity, resulting in true pregnancies being missed and thus being represented also in the data outside the subgroup. Apart from potentially missing relevant signals, another implication was that signals for drugs exclusively used in pregnancy passed the triages set up to avoid signals for which no risk group could be argued. Missing data on BMI could possibly be approached by using weight as a surrogate.

All except one of the communicated signals in this study related to risk groups for already established ADRs. By design, no highlighted drug–adverse event pair was disproportionately over-reported in the entire database. However, requiring absence of disproportionality in the full database should perhaps be reconsidered as it may lead to missed risk group signals when the adverse event is not unique to the risk group and is also disproportionately over-reported in the entire database.

Signal assessment of a drug–adverse event–subgroup association was found to be more challenging than routine



**Table 7** Communicated signals with suggested mechanisms for risk group susceptibility

Communicated signal	Mechanisms for risk group susceptibility suggested in signal
Hepatitis with ceftriaxone in patients 75 years and older [18]	Elderly patients have a longer elimination half-life of ceftriaxone than young adults, which may increase the risk of hepatotoxicity, as raised liver enzymes is a known adverse drug reaction
Myoclonus with levofloxacin in patients 75 years and older [19]	Elderly patients with risk factors for myoclonus such as renal impairment and CNS conditions may predispose to the event
Anaphylactic shock with omalizumab in females [20]	Females suggested to be more sensitive to anaphylactic reactions, possibly attributable to oestradiol
Deep vein thrombosis and pulmonary embolism with aflibercept in males [21]	Multiple lines of evidence suggest a gender influence on the amount of circulating VEGF, with females having higher serum levels than males. A hypothesis would be that females have a higher concentration of un-antagonised VEGF at the same doses of aflibercept than males (although a subgroup analysis in clinical development did not reveal any “clinical relevant” influence of gender on plasma concentrations of free aflibercept or aflibercept–VEGF complex)
Gynaecomastia with esomeprazole in obese adults [22]	Obesity is associated with increased oestrogen due to extragonadal conversion of androgen by tissue aromatase, and omeprazole known to be associated with gynaecomastia due to inhibition of CYP3A4, which is responsible for catabolism of oestradiol
Hypoglycaemia with selegiline in underweight adults [23]	Underweight is a risk factor of hypoglycaemia and may present an additive effect to the known ability of the drug to cause the reaction
Palpitations with glibenclamide in the Asian population [24]	Glibenclamide is metabolised by CYP2D6, of which specific polymorphisms (*2 and *3), common in Asian populations, lead to poor metabolisation. Palpitations may be an adrenergic response to hypoglycaemia secondary to poor metabolism of glibenclamide

CNS central nervous system, VEGF vascular endothelial growth factor

signal assessment of a pairwise drug–adverse event association. The disproportionate over-reporting in the subgroup was an important consideration and often involved an initial review to determine that it was not due simply to a statistical artefact. Often this required a review of the drug–adverse event pair in other subgroups, as well as exploration of subgroup over-reporting for other clinically relevant adverse events or drugs. Referring to Bradford Hill’s criteria for evidence of causal relationships [31], the strength of the association, an analogy with other drugs/adverse events/subgroups, and the identification of a plausible mechanism seemed to be the most prominent aspects weighed into the risk group signal assessments.

The statistical signals identified for the underlying condition covariate were considered the most challenging to assess. The underlying conditions were inferred from all the reported drug indications, regardless of the timing of the drug use and regardless of which drug was highlighted in the signal, which caused uncertainty and complexity. Conversely, the identification of underlying conditions could possibly be further advanced beyond drug indications, utilising additional information on the report, for example, the patient’s medical history. The subgrouping by MedDRA High Level Group Terms could also be further explored as it sometimes created case series where the underlying conditions were clinically too diverse to consider the comorbidity

a common risk factor. With that said, we do not think there is evidence to dismiss this covariate as non-relevant, but rather that there is room for future improvements.

Since disproportionality analysis is a widely used methodology, familiar to many pharmacovigilance professionals, it should be possible to adapt the subgroup approach for other databases. Whereas this study used VigiBase, similar analyses may be possible in smaller datasets, depending on the scope. A small database may not have enough reports to support subgroup analysis, especially for rare events, drugs, or small subgroups. On the other hand, a potential advantage is that national databases may contain more complete information on covariates such as BMI.

While the results in this study are promising, the subgroup signal detection methodology can be further improved. Here, we have considered each subgroup in isolation, but methods to automatically identify coherent subgroups, such as adjacent age groups or neighbouring countries, in a data-driven manner may yield better results. It should also be emphasised that we only used one approach per risk factor for subgrouping and that exploration of other approaches may reveal more appropriate groupings. Another possible improvement would be to flag systematic differences between the subgroup of interest and other reports to support and enhance the assessments [32]. This might highlight possible alternative explanations to disproportionality unique

to a subgroup, such as a co-reported drug in that subgroup, known to cause the adverse event in question. Moreover, as this study was limited to a selection of risk factors and some of them, such as genetic susceptibility and renal function, were only indirectly approached, future research could involve further exploration of these and additional risk factors.

## 5 Conclusions

This novel approach, incorporating broad screening of several different risk factors of ADRs into statistical signal detection in a global collection of adverse event reports, is promising. Clinically relevant signals of risk groups were identified related to patient age, sex, BMI, and geographical region, while refined methods for pregnancy and underlying condition may be needed to reach their full potential. Our findings suggest the potential to use such statistical methodologies for risk characterisation in subpopulations of concern, thereby increasing the precision of pharmacovigilance.

**Acknowledgements** The authors are indebted to the national centres which make up the WHO Programme for International Drug Monitoring and provide reports to VigiBase. However, the opinions and conclusions of this study are not necessarily those of the various centres, nor of the WHO. The authors would also like to thank the assessors who performed the clinical reviews of the signals in this study.

## Compliance with Ethical Standards

**Funding** No funding to conduct the study was received.

**Conflict of interest** Lovisa Sandberg, Henric Taavola, Rebecca Chandler, Yasunori Aoki, and G. Niklas Norén have no conflicts of interest that are directly relevant to the content of this study. As of October 2018, Yasunori Aoki is a full-time employee of Astra Zeneca. However, all of his contributions to this work were made prior to the time of departure, as part of his employment at UMC.

**Data sharing** The datasets generated and analysed during the current study are not publicly available due to agreements between contributors of data to the database used (VigiBase) and the custodian of this database. National centres (mainly national drug regulatory authorities) constituting the WHO Programme for International Drug Monitoring (PIDM) contribute data to VigiBase, and UMC is the custodian in its capacity as a WHO collaborating centre for international drug monitoring.

**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. Alhawassi TM, Krass I, Bajorek BV, Pont LG. A systematic review of the prevalence and risk factors for adverse drug reactions in the elderly in the acute care setting. *Clin Interv Aging*. 2014;9:2079–86. <https://doi.org/10.2147/CIA.S71178>.
2. Miguel A, Azevedo LF, Araújo M, Pereira AC. Frequency of adverse drug reactions in hospitalized patients: a systematic review and meta-analysis. *Pharmacoepidemiol Drug Saf*. 2012;21(11):1139–54. <https://doi.org/10.1002/pds.3309>.
3. Kearns GL, Abdel-Rahman SM, Alander SW, Blowey DL, Leeder JS, Kauffman RE. Developmental pharmacology—drug disposition, action, and therapy in infants and children. *Wood AJJ, ed. N Engl J Med*. 2003;349(12):1157–67. <https://doi.org/10.1056/nejmra035092>.
4. Feghali M, Venkataramanan R, Caritis S. Pharmacokinetics of drugs in pregnancy. *Semin Perinatol*. 2015;39(7):512–9. <https://doi.org/10.1053/j.semperi.2015.08.003>.
5. Anderson GD. Sex and racial differences in pharmacological response: where is the evidence? Pharmacogenetics, pharmacokinetics, and pharmacodynamics. *J Womens Health*. 2005;14(1):19–29. <https://doi.org/10.1089/jwh.2005.14.19>.
6. Zhou Y, Ingelman-Sundberg M, Lauschke V. Worldwide distribution of cytochrome P450 alleles: a meta-analysis of population-scale sequencing projects. *Clin Pharmacol Ther*. 2017;102(4):688–700. <https://doi.org/10.1002/cpt.690>.
7. Lindquist M. VigiBase, the WHO Global ICSR database system: basic facts. *Drug Inf J*. 2008;42(5):409–19. <https://doi.org/10.1177/009286150804200501>.
8. Norén GN, Orre R, Bate A, Edwards IR. Duplicate detection in adverse drug reaction surveillance. *Data Min Knowl Discov*. 2007;14(3):305–28. <https://doi.org/10.1007/s10618-006-0052-8>.
9. Lagerlund O, Strese S, Fladvad M, Lindquist M. WHODrug: a global, validated and updated dictionary for medicinal information. *Ther Innov Regul Sci*. February 2020. <https://doi.org/10.1007/s43441-020-00130-6>.
10. World Health Organization. Body mass index—BMI. <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>. Accessed 17 Apr 2017.
11. Hopstadius J, Norén GN. Robust discovery of local patterns: subsets and stratification in adverse drug reaction surveillance. In: *Proceedings of the 2nd ACM SIGHIT international health informatics symposium. IHI '12*. New York, NY, USA: ACM; 2012. pp. 265–74. <https://doi.org/10.1145/2110363.2110395>.
12. 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(1):57–69. <https://doi.org/10.1177/0962280211403604>.
13. European Medicines Agency. Inclusion/exclusion criteria for the “Important Medical Events” list. 2020. [https://www.ema.europa.eu/en/documents/other/eudravigilance-inclusion/exclusion-criteria-important-medical-events-list\\_en.pdf](https://www.ema.europa.eu/en/documents/other/eudravigilance-inclusion/exclusion-criteria-important-medical-events-list_en.pdf). Accessed 25 Mar 2020.
14. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. ICH harmonised tripartite guideline—Clinical safety data management: definitions and standards for expedited reporting E2A. 1994. [https://database.ich.org/sites/default/files/E2A\\_Guideline.pdf](https://database.ich.org/sites/default/files/E2A_Guideline.pdf). Accessed 25 Mar 2020.

15. Caster O, Juhlin K, Watson S, Norén GN. Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in *vigiRank*: retrospective evaluation against emerging safety signals. *Drug Saf.* 2014;37(8):617–28. <https://doi.org/10.1007/s40264-014-0204-5>.
16. Datapharm Communications Limited. *electronic Medicines Compendium (eMC)*. <https://www.medicines.org.uk/emc/>. Accessed 2 Oct 2017.
17. U.S. National Library of Medicine. *DailyMed*. <https://dailymed.nlm.nih.gov/dailymed/>. Accessed 2 Oct 2017.
18. Boyd I. Ceftriaxone and hepatitis in patients 75 years and older. *WHO Pharm Newsl.* 2018;(6):World Health Organization; 2018.
19. Zappacosta S, Savage R. Levofloxacin and myoclonus in the elderly over 75 years: susceptibilities and prescribing issues. *WHO Pharm Newsl.* 2019;(1):World Health Organization; 2019.
20. Venegoni M. Omalizumab and anaphylactic shock in females. *WHO Pharm Newsl.* 2019;(1):World Health Organization; 2019.
21. Chandler RE, Aoki Y, Sandberg L. Aflibercept and deep vein thrombosis/pulmonary embolism. *WHO Pharm Newsl.* 2018;(6):World Health Organization; 2018.
22. Wong A. Esomeprazole and gynaecomastia in obese adults. *WHO Pharm Newsl.* 2019;(2):World Health Organization; 2019.
23. Hill R. Selegiline and hypoglycaemia in underweight adults. *WHO Pharm Newsl.* 2019;(2):World Health Organization; 2019.
24. Herrera Comoglio R. Glibenclamide/glyburide and palpitations in the Asian population. *WHO Pharm Newsl.* 2019;(2):World Health Organization; 2019.
25. Caster O, Sandberg L, Bergvall T, Watson S, Norén GN. *vigiRank* for statistical signal detection in pharmacovigilance: first results from prospective real-world use. *Pharmacoepidemiol Drug Saf.* 2017;26(8):1006–10. <https://doi.org/10.1002/pds.4247>.
26. Star K, Sandberg L, Bergvall T, Choonara I, Caduff-Janosa P, Edwards IR. Paediatric safety signals identified in *VigiBase*: methods and results from Uppsala Monitoring Centre. *Pharmacoepidemiol Drug Saf.* 2019;28(5):680–9. <https://doi.org/10.1002/pds.4734>.
27. Seabroke S, Candore G, Juhlin K, Quarcoo N, Wisniewski A, Arani R, Painter J, Tregunno P, Norén GN, Slattery J. Performance of stratified and subgrouped disproportionality analyses in spontaneous databases. *Drug Saf.* 2016;39(4):355–64. <https://doi.org/10.1007/s40264-015-0388-3>.
28. Blake KV, Saint-Raymond A, Zaccaria C, Domergue F, Pelle B, Slattery J. Enhanced paediatric pharmacovigilance at the European Medicines Agency: a novel query applied to adverse drug reaction reports. *Paediatr Drugs.* 2016;18(1):55–63. <https://doi.org/10.1007/s40272-015-0154-0>.
29. European Medicines Agency. Screening for adverse reactions in *EudraVigilance*. 2016. [https://www.ema.europa.eu/en/documents/other/screening-adverse-reactions-eudravigilance\\_en.pdf](https://www.ema.europa.eu/en/documents/other/screening-adverse-reactions-eudravigilance_en.pdf). Accessed 10 Mar 2020.
30. Food and Drug Administration. Best practices in drug and biological product postmarket safety surveillance for FDA staff DRAFT. 2019. <https://www.fda.gov/media/130216/download>. Accessed 10 Mar 2020.
31. Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965;58(5):295–300.
32. 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(10):1256–65. <https://doi.org/10.1002/pds.4285>.