ORIGINAL RESEARCH ARTICLE

Pediatric Drug Safety Signal Detection: A New Drug–Event Reference Set for Performance Testing of Data-Mining Methods and Systems

**Osemeke U. Osokogu · Federica Fregonese · Carmen Ferrajolo · Katia Verhamme · Sandra de Bie ·
Geert 't Jong · Mariana Catapano · Daniel Weibel · Florentia Kaguelidou · Wichor M. Bramer ·
Yingfen Hsia · Ian C. K. Wong · Madlen Gazarian · Jan Bonhoeffer · Miriam Sturkenboom**

Abstract

Background   Better evidence regarding drug safety in the pediatric population might be generated from existing data sources such as spontaneous reporting systems and electronic healthcare records. The Global Research in Paediatrics (GRiP)–Network of Excellence aims to develop pediatric-specific methods that can be applied to these data sources. A reference set of positive and negative drug–event associations is required.

On behalf of the Global Research in Paediatrics (GRiP)–Network of Excellence.

O. U. Osokogu (✉) · C. Ferrajolo · K. Verhamme · S. de Bie ·
G. Jong · D. Weibel · F. Kaguelidou · M. Sturkenboom
Department of Medical Informatics, Erasmus University
Medical Center, PO Box 2040, 3000 CA Rotterdam,
The Netherlands
e-mail: o.osokogu@erasmusmc.nl

F. Fregonese
Azienda Ospedaliera di Padova, University of Padua,
Padua, Italy

F. Fregonese
CRCHUM, Université de Montreal, Montreal, QC, Canada

C. Ferrajolo
Pharmacology Division, Department of Experimental Medicine,
Campania Regional Centre of Pharmacovigilance and
Pharmacoepidemiology, Second University of Naples,
Naples, Italy

G. Jong
Department of Pediatrics and Child Health, University of
Manitoba, Winnipeg, MB, Canada

Key Points

A pediatric-specific reference set of positive and negative drug–event associations was created.

The reference set may be utilized in evaluating various data-mining methods, and databases.

It is important to determine locally, when the positive associations became known, as this may impact methods' and database performance.

Objective   The aim of this study was to develop a pediatric-specific reference set of positive and negative drug–event associations.

M. Catapano
Gruppo Italiano per gli Studi di Farmacoeconomia (GISF),
Pavia, Italy

W. M. Bramer
Medical Library, Erasmus University Medical Center,
Rotterdam, The Netherlands

Y. Hsia · I. C. K. Wong
Department of Practice and Policy, Centre of Paediatric
Pharmacy Research, UCL School of Pharmacy, London, UK

Y. Hsia
Paediatric Infectious Diseases Research Group,
St. Georges University of London, London, UK

I. C. K. Wong
Department of Pharmacology and Pharmacy, Centre for Safe
Medication Practice and Research, University of Hong Kong,
Hong Kong, China

Methods Considering user patterns and expert opinion, 16 drugs that are used in individuals aged 0–18 years were selected and evaluated against 16 events, regarded as important safety outcomes. A cross-table of unique drug–event pairs was created. Each pair was classified as potential positive or negative control based on information from the drug's Summary of Product Characteristics and Micromedex. If both information sources consistently listed the event as an adverse event, the combination was reviewed as potential positive control. If both did not, the combination was evaluated as potential negative control. Further evaluation was based on published literature.

Results Selected drugs include ibuprofen, flucloxacillin, domperidone, methylphenidate, montelukast, quinine, and cyproterone/ethinylestradiol. Selected events include bullous eruption, aplastic anemia, ventricular arrhythmia, sudden death, acute kidney injury, psychosis, and seizure. Altogether, 256 unique combinations were reviewed, yielding 37 positive (17 with evidence from the pediatric population and 20 with evidence from adults only) and 90 negative control pairs, with the remainder being unclassifiable.

Conclusion We propose a drug–event reference set that can be used to compare different signal detection methods in the pediatric population.

1 Introduction

In the last 50 years, drug safety monitoring has developed rapidly in terms of increasing interest, broadening capacity, innovation of methods and availability of data [1–3]. This evolution has focused more on the adult population than the pediatric age group (individuals aged 0–18 years). However, drug safety monitoring in pediatrics is of particular importance because we continue to observe that many drugs are prescribed unlicensed and there is lack of adequate information on safety issues affecting this age group. This is of particular concern as the impact of adverse events during growth and maturation may be more serious and longer term compared with adults [4–8].

Globally, specific regulations are being implemented to generate better evidence on safety and efficacy in the

M. Gazarian
School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia

J. Bonhoeffer
Department of Paediatric Infectious Diseases and Vaccines, University Children's Hospital, Basel, Switzerland

J. Bonhoeffer
Brighton Collaboration Foundation, Basel, Switzerland

pediatric population, but mostly by clinical trials [9, 10]. Although useful for efficacy, such trials are usually too small and with too short a follow-up to yield adequate information on rare adverse drug reactions (ADR) and long-term safety [11]. Therefore, other and preferably existing data sources should be utilized to provide information on the safety of drugs in pediatrics [12]. Existing sources with lots of data comprise spontaneous reporting system (SRS) and electronic healthcare record (EHR) databases.

Although analysis of spontaneous reports is currently the most commonly used method for identifying safety signals, specific approaches to surveillance of the pediatric population are limited. The Council for International Organizations of Medical Sciences (CIOMS) Working Group VIII recently advocated for an increased pediatric focus in signal detection [13]. CIOMS also suggested methods to control for confounding in vaccines safety assessment, an issue specific to the pediatric population, and de Bie et al. [14] proposed further refinement of these methods.

Safety signal detection using SRS databases may be complemented by mining longitudinal data in EHRs, as described by the European Adverse Drug Reaction (EU-ADR) project—'Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge' and the 'Observational Medical Outcomes Partnership' (OMOP) project [15–17]. Although newly developed methods, i.e. Longitudinal Gamma Poisson Shrinker (LGPS), show promising results in pediatric data [18], more extensive and systematic testing is needed.

The Global Research in Paediatrics (GRiP)–Network of excellence (http://www.grip-network.org/) was set up with the general objective of facilitating the development and safe use of medicines in the pediatric population, with a specific objective being to apply innovative approaches and standardized methodologies, as well as better utilization of existing healthcare and spontaneous reporting databases. GRiP aims to tailor existing signal detection methods to pediatric safety data. Comparison of the performance of existing methods within and across SRS and EHR databases is the first step in defining suitable methods to be implemented. For this purpose, creation of a reference set comprising pediatric-specific drug–event pairs serving as positive and negative control, is required to calculate baseline performance statistics. Coloma et al. [19] recently described the methodology for creating a reference set used to test methods in the EU-ADR project. Similarly, Ryan et al. [20] established a reference set for testing methods in the OMOP project. However, both were not specific to the pediatric population and comprise many drugs infrequently prescribed within this age group, and events that rarely (or never) affect them.

In this study we describe how we created a proposed reference set for comparing the performance of different methods in detecting drug safety signals in the pediatric population. This may be used for spontaneous reporting, as well as electronic healthcare record databases.

2 Methods

The first step in creating the reference set was to select a list of eligible drugs to be utilized. Based on four criteria, four (primary) lists of drugs were created: we compiled drugs that are frequently prescribed in pediatrics (including off-label use), on an outpatient basis in high-income countries (as per papers and reports of use) [21, 22]; to allow for inpatient databases to be assessed, we included drugs that are administered to hospitalized persons aged 0–18 years (or administered by specialists) [22]; to allow for databases from low- and middle-income countries (LMICs) to be assessed, we included drugs that are used in such countries [as per the World Health Organization (WHO) List of Essential Medicines for children] [23]; and to allow for testing signal detection performance by different age groups, we included drugs that are used in specific pediatric age groups (for example, adolescents) [22].

To obtain a final drug list, a stepwise procedure was implemented. First, if two or more drugs [fifth-level chemical substances, WHO Anatomical Therapeutic Chemical (ATC) Classification System] belonged to the same class ('WHO-ATC, fourth level'), and were listed in an equal number of primary lists (>1), we preferentially selected only the drug that had the oldest initial marketing authorization worldwide. This was done to have the most evidence available. For example, doxycycline (WHO-ATC code J01AA02) would be selected instead of minocycline (WHO-ATC code J01AA08) because although they both belong to the same class—'WHO-ATC, fourth level' (tetracyclines)—doxycycline was first marketed in 1967 [24], and minocycline in 1972 [25]. Second, we preferentially selected drugs that appeared in most of the lists, for example a drug appearing on three of four primary lists would be retained instead of another drug appearing on only two lists. The final list comprised more than 30 drugs, which was beyond our capacity and resources and was reduced to 16 for pragmatic reasons.

Events were chosen (independent of the drugs) with the aim of generating a set which may be used for methods development on spontaneous reporting as well as EHRs. Both rare and common events were included to allow for investigation of effect modification. Starting with common adverse events observed in pediatrics, as reported by Star et al. [26], we selected only events that were deemed

serious (as per the WHO definition [27]) and specific (to avoid misclassification). For example *aplastic anemia* was selected rather than *anemia* as the former connotes a more serious and specific medical condition. Some events (i.e. psychosis and seizure) were included by consensus in the research team because they were considered relevant for the pediatric population from a pharmacovigilance and public health point of view. Fifteen drugs and events were considered as the minimum required for generating enough positive and negative associations. Finally, the total number of drugs and events was set at 16 for pragmatic reasons.

Four researchers (MS, IW, JB, and GJ) with a range of expertise spanning pediatrics, pharmacology, and pharmacoepidemiology determined the final list of selected drugs and events. MS and IW are pharmacists/pharmacoepidemiologists, JB is a pediatrician, and GJ is a pediatrician/clinical pharmacologist/pharmacoepidemiologist.

All events of interest were defined using standard resources (i.e. medical textbooks, uptodate.com, and scientific societies such as the CIOMS) to increase the likelihood of comprehensive literature searches. The final reference set was generated by cross-tabulating the final lists of drugs and events, which led to a matrix of 256 unique drug–event pairs. In order to classify each unique drug–event pair as a 'positive', 'negative', or 'unclassifiable' association, evidence was reviewed in two sequential steps.

2.1 Review of Summary of Product Characteristics (SPC) and Micromedex

First, two researchers (OO and CF) with expertise in general medicine, pharmacy, and pharmacoepidemiology reviewed the SPC of each drug to ascertain that a specific event (for example, aplastic anemia) was listed as a possible adverse event under the appropriate section(s)—'Undesirable effects' (section 4.8) and/or 'Special warnings and precautions for use' (section 4.4) from the electronic Medicines Compendium (eMC) [28]. DailyMed (the 'Contraindications, Warnings, Precautions' and/or 'Adverse Reactions' sections) was consulted only if a drug was not listed in the eMC [29]. The eMC contains more than 9,000 up-to-date, freely accessible documents containing information about medicines licensed for use in the UK. Prior to publishing, these documents are usually checked and approved by either the UK Medicines and Healthcare products Regulatory Agency (MHRA) or the European Medicines Agency (EMA). DailyMed, published by the National Library of Medicine (NLM) in the US, contains up-to-date information about drugs licensed for use in the US. Both eMC and DailyMed are freely accessible online.

Second, two researchers (OO and CF) reviewed Micromedex to check if the event was listed under the
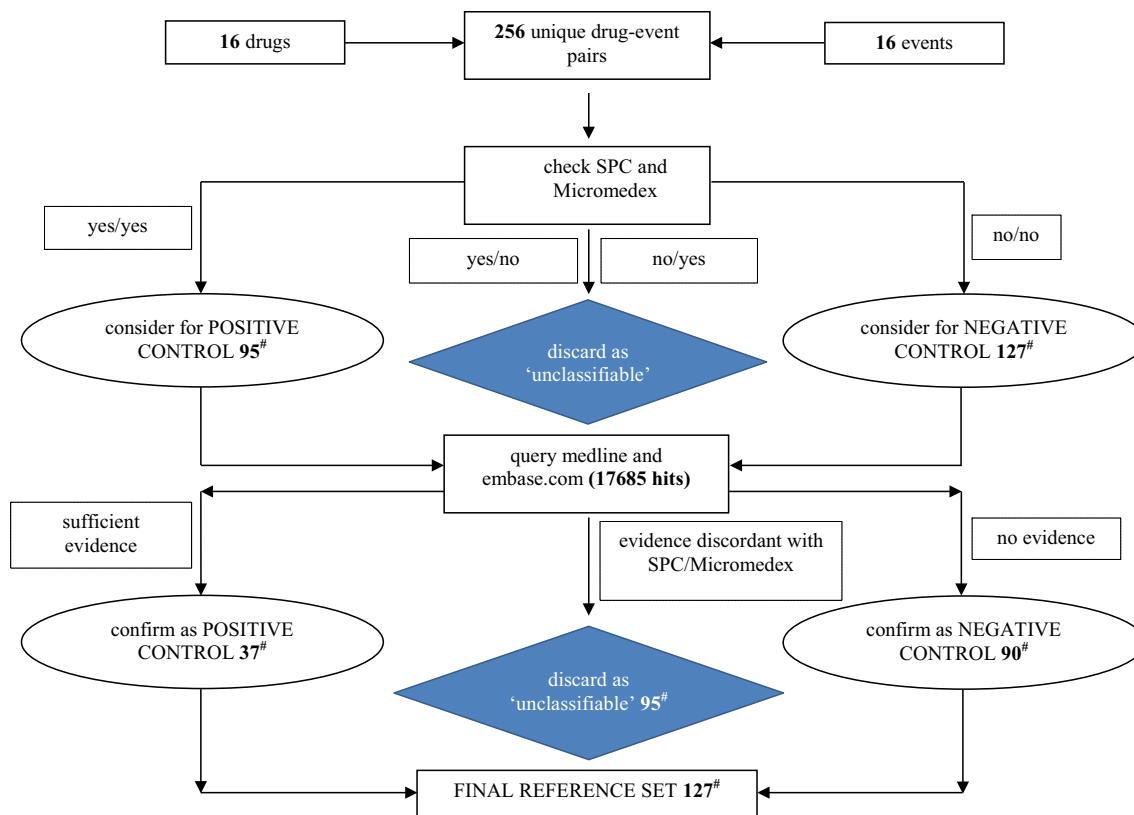
Fig. 1 Procedure adopted for the construction of the reference set (adapted from Coloma et al. [19]). *SPC* Summary of Product Characteristics, [#] drug–event pairs

section 'Adverse Reactions' within the Drugdex component. Micromedex is an online drug information system that contains referenced information from various sources needed for clinical decision making, including adverse effects of drugs (http://www.micromedex.com/).

After reviewing the SPC and Micromedex, drug–event pairs were classified as (1) 'potential positive control' (event was mentioned in both the SPC and Micromedex); (2) 'potential negative control' (event was mentioned in neither the SPC nor Micromedex); or (3) unclassifiable (discordant information between the SPC and Micromedex). 'Potential positive control' and 'potential negative control' pairs were retained and the relationship of each drug–event pair was further evaluated using published literature (Fig. 1).

2.2 Review of Published Literature

For each drug–event pair that was classified as a 'potential negative control', a systematic literature search was conducted in EMBASE.COM and MEDLINE (via OvidSP). The sensitive search algorithm applied to both title and abstract comprised controlled vocabulary plus free text for each of two concepts: 'event of interest' and drug.

For each 'potential positive control', the search algorithm was more specific (to avoid large numbers of papers) than for the potential negative controls, and included only controlled vocabulary for the drug name. However, the event was searched by using both controlled vocabulary and free text. In addition, controlled vocabulary was included for the concept 'general adverse drug reaction'; this was done to increase the probability of retrieving only those articles where adverse event and drug co-occurred in the context of drug safety [19].

For potential negative and positive control pairs we only considered articles published in English. Publications could be biological and/or epidemiological studies. Epidemiological studies could be case reports, observational studies (i.e. cohort, case-control), reviews, meta-analysis, and clinical trials. As an example, the search strings for the negative control *sudden death–cyproterone/ethinylestradiol*, and positive control *sudden death–clarithromycin* are presented in Appendix 1 of the electronic supplementary material (ESM).

One of five researchers (OO, CF, FF, MC, and YH) reviewed retrieved publications pertaining to a unique drug–event pair. All five researchers have received medical, biological, and/or pharmacology training. Based on

data extracted from relevant publications, unique drug–event pairs were classified according to the criteria outlined in Table 1. For example, a pair was assigned level I evidence if there was evidence from at least one randomized controlled trial or meta-analysis, while 'positive control, grade 1' (PC1) meant that in addition there was 'proven biological mechanism for causal association'. Level V evidence—(not mentioned in the SPC/Micromedex) AND (published evidence against causal association; OR no published evidence supporting causal association)—qualified a specific drug–event pair as a negative control, while 'negative control, grade 1' (NC1) meant that in addition there was 'proven biological mechanism against causal association'. 'Proven biological mechanism' meant that there was at least one publication providing relevant biological evidence regarding a unique drug–event pair. Two researchers (MS and FK; a pediatrician, clinical pharmacologist, and pharmacoepidemiologist) reviewed all associations that were classified as positive or negative control.

Whereas confirmation of negative control pairs required lack of association for either adults or the pediatric age group, positive control pairs were specifically assessed for availability of evidence pertaining to persons aged 0–18 years. However, such evidence was not mandatory for classification as positive control due to the acknowledged lack of pediatric-specific studies [30]. Those with lack of evidence in pediatrics are listed separately.

To further illustrate the process of reviewing the published literature, 126 unique references were retrieved following database search for articles supporting the potential positive control *sudden death–clarithromycin*. Of these, 103 articles were excluded following title/abstract screening, while 13 articles were excluded following full-text screening. Full-text copies of six articles could not be obtained. Finally, four articles—one clinical trial, two case-control studies, and one case report—presented sufficient evidence to support the association.

3 Results

As presented in Table 2, 16 drugs (unique WHO-ATC codes, fifth-level chemical substance) were selected for the reference set, comprising eight anti-infectives: flucloxacillin, clarithromycin, doxycycline, lopinavir (which is always administered in fixed-dose combination with ritonavir), isoniazid, praziquantel, mebendazole, and quinine. The remaining were respiratory drugs (fluticasone, administered as an inhalant, and montelukast), gastrointestinal drugs (loperamide and domperidone), antipyretic/analgesic (ibuprofen), a drug for attention-deficit hyperactivity disorder (methylphenidate), anti-acne

Table 1 Evaluation and grading of unique drug–event pairs based on SPC/Micromedex and literature evidence

Classification	Level of evidence	Description	Biological mechanism	Description	Grade
Positive control (PC)	I	(Included in SPC/Micromedex) AND (Evidence from at least one randomized controlled trial or meta-analysis)	Proven for causal association	Evidence from at least one publication explaining the mechanistic pathway	PC1
			Plausible for causal association	No published evidence	PC2
	II	(Included in SPC/Micromedex) AND [Evidence from at least one observational study (i.e. cohort, case-control, case-crossover, self-controlled case series) OR review of spontaneous reports OR systematic review OR (at least three published case reports from different sources and concerning different patients with causality evaluation of definite or probable)]	Proven for causal association	Evidence from at least one publication explaining the mechanistic pathway	PC1
			Plausible for causal association	No published evidence	PC2
Indeterminate	III	(Included in SPC/Micromedex) AND (Evidence from less than three published case reports and no further substantiation in the literature)			
	IV	Included in SPC/Micromedex BUT no published case reports or studies			
Negative control (NC)	V	(Not mentioned in SPC/Micromedex) AND (published evidence against causal association OR no published evidence supporting causal association)	Proven against causal association	Evidence from at least one publication explaining the mechanistic pathway	NC1
			Plausible against causal association	No published evidence	NC2

SPC Summary of product characteristics

Table 2 Classification of each drug–event pair as positive control (green: PC1 or PC2) or negative control (red: NC2)

		Bullous eruption	Aplastic anemia	Agranulocytosis	Thrombocytopenia	Psychosis	Suicide	Vent. arrhythmia	Sudden death	QT prolongation	Venous thromboembolism	Anaphylaxis	Seizure	Acute kidney injury	Acute liver injury	Sepsis	SIDS
Selected Drugs	flucloxacillin						R			R			R		G		
	clarithromycin	G			G	G	R	G	G	G	R				G		
	doxycycline	G			G		R		R	R							
	lopinavir						R								G		
	isoniazid	G				G		R	R	R			G			R	
	praziquantel		R	R	R	R			R	R	R					R	
	mebendazole			G								R		R	G		
	quinine	G	R	R	G			G	R	G	R			G	G		
	fluticasone		R	R						R				R		R	
	montelukast		R			G	G	R	R		R					R	
	isotretinoin					G	G									R	
	loperamide		R	R			R	R	R	R						R	
	domperidone	R					R	G	G		R	R				R	
	methylphenidate					G						R				R	
	ibuprofen	G			G							G		G	G	R	
	cyproterone /eth.est.						R	R	R	R	R	R		R		R	

Abbreviations: Vent. - ventricular; SIDS - Sudden Infant Death Syndrome; eth.est.- ethinylestradiol

(isotretinoin), and a hormonal oral contraceptive (cyproterone/ethinylestradiol).

We selected 16 events for the reference set—bullous eruption [comprising fixed drug eruption (FDE), erythema multiforme (EM), Stevens–Johnson syndrome (SJS), and toxic epidermal necrolysis (TEN)], aplastic anemia, agranulocytosis, thrombocytopenia, psychosis, suicide, ventricular arrhythmia, sudden death, QT prolongation, venous thromboembolism, anaphylaxis, seizure, acute kidney injury (AKI), acute liver injury (ALI), sepsis, and sudden infant death syndrome (SIDS) (Table 2). Medical definitions for all events and their proposed (unvalidated) Medical Dictionary for Regulatory Activities (MedDRA) codes are presented in Appendix 2 of the ESM.

From the total number of combinations (256), we discontinued assessment of 34 unclassifiable drug–event pairs since we found discrepant information between the SPC and Micromedex. For the remaining 222 pairs, the literature search generated 17,685 hits. Based on review of these hits, 127 pairs were confirmed as positive control (37 pairs) or negative control (90 pairs) (Tables 2, 3); for 95 'unclassifiable' pairs there was discrepant information between the published literature on one hand and the SPC and Micromedex on the other hand.

In confirming the 37 positive controls, evidence was used from 171 relevant publications, comprising 14 biological studies, 10 clinical trials, 23 observational studies, 34 reviews, and 90 case reports/series. The association between quinine and thrombocytopenia had the highest number of supporting publications, i.e. 20 (of 171); eight publications pertained to biological evidence, while 12 reported on epidemiologic evidence. Table 4 shows how the positive controls (*quinine–thrombocytopenia* and *clarithromycin– sudden death*) were established. For complete evaluation of all positive controls, see Appendix 3 of the ESM.

As presented in Table 3, we generated 37 positive controls; of these, level I evidence was available for only 8 (22 %), and 13 (35 %) were supported by both biological and epidemiological evidence. Only four associations (clarithromycin–thrombocytopenia, montelukast–psychosis, montelukast–suicide AND methylphenidate–psychosis) were supported by evidence generated exclusively from the pediatric age group, while 13 associations were supported by evidence from both adults and the pediatric population. Overall, 17 (46 %) of all positive associations were based on evidence from the pediatric population. Twenty associations were supported by evidence from only adults.

As presented in Appendix 4 of the ESM, we compared the reference set we created with the reference sets that were created within EU-ADR and OMOP. Of the 16 drugs that were selected for GRiP, four were also included in EU-ADR and/or OMOP: fluticasone, ibuprofen, isoniazid, and mebendazole. Ibuprofen was classified to be a positive control for AKI in each of the three reference sets, while the same drug was classified to be associated with ALI only

Table 3 Level of epidemiological and biological evidence; population in which association was found (adults, 'children'[a], or both) and grading of positive drug–event associations

Event	Positive associations					
	ATC code	Drug name	Level of epidemiological evidence[c]	Population (A/B/C)[d]	Biological evidence (Pr/Pl)	Grade[e]
Bullous eruption[b]	J01FA09	Clarithromycin	II	B	Pl	PC2
	J01CF05	Doxycycline	II	B	Pl	PC2
	J04AC01	Isoniazid	II	B	Pl	PC2
	P01BC01	Quinine	II	A	Pl	PC2
	M01AE01	Ibuprofen	II	B	Pl	PC2
Aplastic anemia	P01BC01	Quinine	II	A	Pr	PC1
Agranulocytosis	P02CA01	Mebendazole	II	A	Pl	PC2
	P01BC01	Quinine	II	A	Pr	PC1
Thrombocytopenia	J01FA09	Clarithromycin	II	C	Pl	PC2
	J01CF05	Doxycycline	I	A	Pl	PC2
	P01BC01	Quinine	II	A	Pr	PC1
	M01AE01	Ibuprofen	I	A	Pl	PC2
Psychosis	J01FA09	Clarithromycin	II	A	Pl	PC2
	J04AC01	Isoniazid	II	A	Pl	PC2
	R03DC03	Montelukast	II	C	Pl	PC2
	D10BA01	Isotretinoin	II	B	Pr	PC1
	N06BA04	Methylphenidate	I	C	Pr	PC1
Suicide	R03DC03	Montelukast	II	C	Pl	PC2
	D10BA01	Isotretinoin	II	B	Pr	PC1
Ventricular arrhythmia	J01FA09	Clarithromycin	II	A	Pl	PC2
	P01BC01	Quinine	II	A	Pl	PC2
	A03FA03	Domperidone	II	A	Pr	PC1
Sudden death	J01FA09	Clarithromycin	I	A	Pl	PC2
	A03FA03	Domperidone	II	A	Pr	PC1
QT prolongation	J01FA09	Clarithromycin	II	A	Pr	PC1
	P01BC01	Quinine	I	B	Pr	PC1
Anaphylaxis	M01AE01	Ibuprofen	II	B	Pr	PC1
Seizure	J04AC01	Isoniazid	II	B	Pr	PC1
Acute kidney injury	P01BC01	Quinine	II	A	Pl	PC2
	M01AE01	Ibuprofen	II	B	Pr	PC1
Acute liver injury	J01CF05	Flucloxacillin	II	A	Pl	PC2
	J01FA09	Clarithromycin	II	B	Pl	PC2
	J05AE06	Lopinavir	I	A	Pl	PC2
	J04AC01	Isoniazid	I	B	Pl	PC2
	P02CA01	Mebendazole	I	B	Pl	PC2
	P01BC01	Quinine	II	A	Pl	PC2
	M01AE01	Ibuprofen	II	A	Pl	PC2

ATC Anatomical Therapeutic Chemical, *Pr* proven biological evidence, *Pl* plausible biological evidence, *PC* positive control

[a] In this context 'Children' refers to individuals aged 0–18 years

[b] Epidemiological evidence levels I and II are defined in Table 1

[c] Population in which epidemiological evidence was found: A, adults; B, both adults and 'children'; C, 'children'

[d] As defined in Table 1

[e] Bullous eruption includes fixed drug eruption, erythema multiforme, Stevens–Johnson syndrome, and toxic epidermal necrolysis

Table 4 Examples of evaluation of a positive drug–event association: (1) quinine–thrombocytopenia and (2) clarithromycin–sudden death

ATC code	Drug name	Event type	Labeled as AE in SPC (yes/no)	Type/no. of supporting literature citations
P01BC01	Quinine	Thrombocytopenia	Yes *eMC (special warnings and precautions for use; undesirable effects) #Micromedex (summary): black-box warning; (contraindications/warnings → contraindications; precautions); (adverse effects → serious)	Total number of supporting citations = 20 Biological studies = 8 Review of biological studies = 4 Systematic review = 1 Case-series = 1 Case reports = 4 Review of spontaneous reports = 2
J01FA09	Clarithromycin	Sudden death	Yes *eMC (undesirable effects) #Micromedex (summary): (contraindications/warnings → precautions); (adverse effects → serious)	Total number of supporting citations = 4 Clinical trial = 1 Case-control = 2 Case report = 1

ATC Anatomical Therapeutic Chemical, *AE* adverse effect, *SPC* Summary of Product Characteristics, *eMC* electronic Medicines Compendium

* https://www.medicines.org.uk/emc/

http://micromedex.com/

within GRiP and EU-ADR. Isoniazid was classified as positive control for ALI, both in GRiP and OMOP. Neither OMOP nor EU-ADR labeled mebendazole with AKI, nor fluticasone with ALI.

4 Discussion

We describe a pediatric-focused reference set of drug–event associations that may be used for testing the performance of different signal detection methods and databases. To our knowledge, this is the first structured approach to creating a reference set that is specific to pediatric safety outcomes. This approach yielded 37 positive and 90 negative drug–event associations; 17 positive associations were supported by evidence in pediatric age group, and 20 were based on adult information only.

Projects such as OMOP and EU-ADR have also created reference sets but none was targeted to pediatrics; in addition, the construct of these reference sets was different [19, 31–33]. In the current project, drugs and events were selected independently, unlike EU-ADR and OMOP [19, 20]. In addition, the EU-ADR network restricted the list of drugs based on the amount of drug exposure that would be required to identify associations with selected adverse events at pre-specified relative risk (RR) values. This was done so that such drug–event associations could actually be identified if indeed they occurred within the network. Similar calculations were not done for the current project, although most of the selected drugs are frequently administered in the pediatric population (based on reported evidence in the literature). Furthermore, the reference set

resulting from the current project will be applied to SRS databases (in addition to EHRs) and therefore should preferably be unbiased to one or the other.

The GRiP reference set focused on diversity of drugs and events which may allow us to stratify by outpatient/inpatient care, and frequent and rare events. Sets with drugs for inpatient use may favour performance of data mining on SRS databases, while sets utilizing drugs prescribed for outpatient treatments may favour mining performance on EHR databases [34]. In order to have enough power for both, we focused on drugs with longer license status.

In selecting adverse events, we considered both frequent and rare events. Thus, the resulting reference set can be tested in a wide variety of databases with unique adverse event profiles, such as SRSs, and hospital-based and general practice healthcare databases. Previous reference sets focused mostly on rare and well-known drug-induced events which may favour SRSs [19]. Such events may be reported more often than common, multifactorial events because they are easier to identify as being caused by drugs. Given that the composition of the lists of drugs and adverse events to be tested may have an extensive impact on performance assessment [35], we tried to ensure that the criteria and data sources that were utilized to create the reference set were independent of the data on which they will eventually be tested.

We conducted extensive reviews to list evidence for both positive and negative controls. Fewer publications were retrieved for the potential positive control pairs (7,745 hits) compared with the potential negative control pairs (9,940 hits), possibly because the search algorithm for the former was more specific. However, this was considered

necessary to increase the probability of retrieving relevant publications (i.e. publications that reported on adverse event and drug in the context of drug safety), an approach similar to that adopted by the EU-ADR project [19].

To validate potential negative control pairs, terms that were related to the actual event term were considered. For example suicide–isoniazid was initially classified as potential negative control because suicide was not mentioned (in relation to isoniazid), both in the SPC (DailyMed) and Micromedex. However, a case report described the occurrence of suicide attempt following ingestion of isoniazid [36]; therefore, this association could not be confirmed as negative control. Whereas the negative drug–event associations required lack of association for adults or the pediatric population, the positive drug–event associations were specifically (or primarily) assessed for availability of evidence pertaining to the pediatric age group. However, due to the general lack of pediatric pharmacoepidemiological data, only four associations (clarithromycin–thrombocytopenia, montelukast–psychosis, montelukast–suicide AND methylphenidate–psychosis) were supported by evidence generated exclusively from this age group: a case-control study for clarithromycin–thrombocytopenia [37]; case reports (more than three) for montelukast–psychosis [38]; review of spontaneous reports for montelukast–suicide [39]; and clinical trials as well as case series for methylphenidate–psychosis [40]. The scarcity and quality of pediatric-specific data further highlight the difficulties in generating safety evidence in the pediatric population, thereby underlining the importance of developing a tool to define appropriate signal detection methods in this population. We recommend that the 20 positive associations that come from adult evidence only, be treated separately in the performance testing in pediatric data.

We chose to classify all pairs with inconsistent evidence as unclassifiable, to avoid misclassification. We searched for biological (in addition to epidemiological) evidence to further strengthen retrieved evidence for positive controls. However, we were able to find such evidence for only 13 of 37 positive associations: quinine–aplastic anemia [41]; quinine–agranulocytosis [42]; quinine–thrombocytopenia [43]; isotretinoin–psychosis [44, 45]; methylphenidate–psychosis [46, 47]; isotretinion–suicide [44, 45]; domperidone–ventricular arrhythmia [48]; domperidone–sudden death [49]; clarithromycin–QT prolongation [50]; quinine–QT prolongation [51, 52]; ibuprofen–anaphylaxis [53]; isoniazid–seizure [54]; and ibuprofen–AKI [55]. Of these, quinine–thrombocytopenia had the highest number of supporting publications, i.e. eight regarding biological evidence (in addition to 12 others pertaining to epidemiological evidence). This is possibly because quinine has been in use for a long time, both as over-the-counter (OTC) and prescription drug [56]; therefore, its safety

profile has been well investigated. Otherwise, the limited biological evidence for most of the other positive associations may reflect the current gap of knowledge and understanding of ADRs.

Comparing our reference set with others, we found little overlap in the choice of drugs, possibly because we aimed to be pediatric-specific in our selection while also including drugs used in specific subpopulations (i.e. adolescents) and context (LMICs). Of 16 drugs considered in GRiP, only four were also considered in EU-ADR and/or OMOP: isoniazid, ibuprofen, mebendazole, and fluticasone. Perhaps this, as well as differences in adverse event selection, explains the few similarities we found across the three reference sets. Nevertheless, ibuprofen was found to be associated with AKI in all sets.

There are several limitations in the creation and use of a reference set. Some potential positive associations that are well known (i.e. *domperidone–QT prolongation* and *cyproterone/ethinylestradiol–venous thromboembolism*, both of which have been well investigated) could not be validated. The search query we used to retrieve the publications may have been too specific. For other unconfirmed potential positive control pairs, events mentioned in the SPC and Micromedex may have been reported through means other than peer-reviewed literature (for example, US FDA reports).

Time is an important limiting aspect in building a reference set, both for the positive as well as negative controls. We labeled drug–event associations as negative if there was lack of evidence, which in itself is something that may rapidly change over time; checking of the absence of evidence should always be carried out prior to using the reference set. For the positive controls, it is important to know at which point in time the association was 'known' as this may lead to changes in reporting behaviour to spontaneous reporting databases and to changes in clinical care. Those changes may have an impact on the ability to detect associations (for example, in spontaneous reporting databases it may increase the association, whereas it may decrease in electronic healthcare databases) [57–59]. Time stamping of the 'known' associations would be important. However, this was impossible for this reference set since we chose drugs that are available for a long time and have been registered nationally. Inclusion of information in an SPC may vary from country to country. We recommend investigators who will use this set, to assess in their reality when associations were 'known' in order to evaluate the impact of that on performance.

In order to use the reference set, the events need to be translated into codes. This is an important step and may impact on the performance testing. In Appendix 2 of the ESM we have provided initial MedDRA codes as most of the events have Standardised MedDRA Queries (SMQs). These codes should be reviewed and the impact of choices

should be carefully evaluated; they may differ between spontaneous reporting databases and EHRs. Within the GRiP project, we aim to perform this work for MedDRA, International Classification of Diseases, Ninth Revision (ICD-9), ICD, Tenth Revision (ICD-10), READ and International Classification of Primary Care (ICPC), and a full code list with the impact of choice on performance will become available later.

5 Conclusions

We have generated a pediatric-focused reference set that can be applied for testing performance of methods and databases for drug safety signal detection in the pediatric population. This reference set may be viewed as dynamic. The status of drug–event associations may change over time, particularly as more evidence derived specifically from the pediatric population becomes available in the future. Therefore, periodic review and checking against the local situation is advisable.

References

1. Scurti V, Romero M, Tognoni G. A plea for a more epidemiological and patient-oriented pharmacovigilance. Eur J Pharmacol. 2012;68(1):11–9.
2. Meyboom RHB. Pharmacovigilance in a changing world. Klinicka Farmakologie a Farmacie. 2011;25(3):102–11.
3. Van Grootheest K. The dawn of pharmacovigilance: an historical perspective. Int J Pharmaceut Med. 2003;17(5–6):195–200.
4. Vlahovic-Palcevski V, Mentzer D. Postmarketing surveillance. Pediatric clinical pharmacology. Handbook of experimental pharmacology, vol 205. Heidelberg: Springer; 2011. p. 339–51.
5. Autret-Leca E, Marchand MS, Cissoko H, Beau-Salinas F, Jonville-Bera AP. Pharmacovigilance in children (in French). Arch Pediatr. 2012;19(8):848–55.
6. Phan H, Leder M, Fishley M, Moeller M, Nahata M. Off-label and unlicensed medication use and associated adverse drug events in a pediatric emergency department. Pediatr Emerg Care. 2010;26(6):424–30.
7. Conroy S. Unlicensed and off-label drug use: issues and recommendations. Pediatr Drugs. 2002;4(6):353–9.
8. Knellwolf AL, Bauzon S, Alberighi ODC, Lutsar I, Bacsy E, Alfarez D, et al. Framework conditions facilitating paediatric clinical research. Ital J Pediatr. 2011;37:12.
9. United States of America. Pediatric Research Equity Act. 2003. Available at: www.fda.gov/downloads/drugs/development approvalprocess/developmentresources/ucm077853.pdf. Accessed 12 June 2014.
10. European Medicines Agency. Paediatric Regulation. 2006. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000068.jsp. Accessed 12 June 2014.
11. Pincus T. Limitations of randomized clinical trials to recognize possible advantages of combination therapies in rheumatic diseases. Semin Arthritis Rheum. 1993;23(2 Suppl):2–10.
12. Valkhoff VE, Schade R, 't Jong GW, Romio S, Schuemie MJ, Arfe A, et al. Population-based analysis of non-steroidal anti-inflammatory drug use among children in four European countries in the SOS project: what size of data platforms and which study designs do we need to assess safety issues? BMC Pediatr. 2013;13:192.
13. CIOMS. Practical aspects of signal detection in pharmacovigilance. Report of CIOMS Working Group VIII. Geneva: CIOMS; 2010.
14. de Bie S, Verhamme KMC, Straus SMJM, Stricker BHC, Sturkenboom MCJM. Vaccine-based subgroup analysis in VigiBase: effect on sensitivity in paediatric signal detection. Drug Saf. 2012;35(4):335–46.
15. Schuemie MJ, Coloma PM, Straatman H, Herings RM, Trifiro G, Matthews JN, et al. Using electronic healthcare records for drug safety signal detection: a comparative evaluation of statistical methods. Pharmacoepidemiol Drug Saf. 2012;21:342.
16. De Bie S, Coloma P, Ferrajolo C, Trifiro G, Verhamme K, Schuemie M, et al. The power of electronic healthcare databases for active drug safety surveillance in children and adolescents: an EUADR study. Pharmacoepidemiol Drug Saf. 2012;21:329–30.
17. Coloma PM, Trifiro G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? Pharmacoepidemiol Drug Saf. 2012;21(6):611–21.
18. Ferrajolo C, Coloma PM, Verhamme KMC, Schuemie MJ, de Bie S, Gini R, et al. Signal detection of potentially drug-induced acute liver injury in children using a multi-country healthcare database network. Drug Saf. 2014;37(2):99–108.
19. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, et al. A reference standard for evaluation of methods for drug safety signal detection using electronic healthcare record databases. Drug Saf. 2013;36(1):13–23.
20. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. Drug Saf. 2013;36(Suppl 1):S33–47.
21. Sturkenboom MCJM, Verhamme KMC, Nicolosi A, Murray ML, Neubert A, Caudri D, et al. Drug use in children: cohort study in three European countries. BMJ. 2008;337(7682):1338–41.

22. European Medicines Agency. Report on the survey of all paediatric uses of medicinal products in Europe. 2010. http://www.ema.europa.eu/ema/pages/includes/document/opendocument.jsp?webContentId=WC500101006. Accessed 1 Dec 2014.

23. World Health Organization. WHO model lists of essential medicines. 2013. Available at: http://www.who.int/medicines/publications/essentialmedicines/en/index.html. Accessed 1 Dec 2014.

24. Alexander J. Doxycycline medKit regulatory background. 2012. Available at: http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/anti-infectivedrugsadvisory committee/ucm299210.pdf. Accessed 16 June 2014.

25. Lawrenson RA, Seaman HE, Sundstrom A, Williams TJ, Farmer RDT. Liver damage associated with minocycline use in acne: a systematic review of the published literature and pharmacovigilance data. Drug Saf. 2000;23(4):333–49.

26. Star K, Noren GN, Nordin K, Edwards IR. Suspected adverse drug reactions reported for children worldwide: An exploratory study using vigibase. Drug Safety. 2011;34(5):415–28.

27. World Health Organization, Quality Assurance and Safety of Medicines Team. Safety of medicines. A guide to detecting and reporting adverse drug reactions: why health professionals need to take action. 2002. Available at: http://www.who.int/iris/handle/10665/67378. Accessed 16 June 2014.

28. Datapharm Communications Limited. Electronic medicines compendium. Available at: http://www.medicines.org.uk/emc/. Accessed 21 June 2014.

29. National Library of Medicine. DailyMed. Available at: http://dailymed.nlm.nih.gov/dailymed/about.cfm?CFID=74319392&CFTOKEN=69c4f4d2dc469ecc-E9581537-B0F2-4DB4-C2BF7E65FAC84A0E&jsessionid=84303863c51992297375697f41594567791b. Accessed 21 June 2014.

30. Choonara I. Unlicensed and off-label drug use in children: implications for safety. Expert Opin Drug Saf. 2004;3(2):81–3.

31. Hochberg AM, Hauben M, Pearson RK, O'Hara DJ, Reisinger SJ, Goldsmith DI, et al. An evaluation of three signal-detection algorithms using a highly inclusive reference event database. Drug Saf. 2009;32(6):509–25.

32. Hauben M, Reich L, Chung S. Postmarketing surveillance of potentially fatal reactions to oncology drugs: potential utility of two signal-detection algorithms. Eur J Clin Pharmacol. 2004;60(10):747–50.

33. Lindquist M, Stahl M, Bate A, Edwards IR, Meyboom RHB. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. Drug Saf. 2000;23(6):533–42.

34. Pearson JF, Brownstein CA, Brownstein JS. Potential for electronic health records and online social networking to redefine medical research. Clin Chem. 2011;57(2):196–204.

35. Zhou X-H, Obuchowski N, McClish D. Statistical methods in diagnostic medicine. New York: John Wiley & Sons; 2011.

36. Madan A, Jain NK, Sharma TN, Sharma DK, Madan R. Attempted suicide following treatment with isoniazid. Tubercle. 1989;70(2):147–9.

37. Bertuola F, Morando C, Menniti-Ippolito F, Da Cas R, Capuano A, Perilongo G, et al. Association between drug and vaccine use and acute immune thrombocytopenia in childhood: a case-control study in Italy. Drug Saf. 2010;33(1):65–72.

38. Wallerstedt SM, Brunlof G, Sundstrom A, Eriksson AL. Montelukast and psychiatric disorders in children. Pharmacoepidemiol Drug Saf. 2009;18(9):858–64.

39. Bygdell M, Brunlof G, Wallerstedt SM, Kindblom JM. Psychiatric adverse reactions reported during a 10-year period in the Swedish pediatric population. Pharmacoepidemiol Drug Saf. 2012;21(1):79–86.

40. Shibib S, Chalhoub N. Stimulant-induced psychosis. Child Adolesc Ment Health. 2009;14(1):20–3.

41. Maguire RB, Stroncek DF, Campbell AC. Recurrent pancytopenia, coagulopathy, and renal failure associated with multiple quinine-dependent antibodies. Ann Intern Med. 1993;119(3):215–7.

42. Sutherland R, Vincent PC, Raik E, Burgess K, Sutherland R, Vincent PC, et al. Quinine-induced agranulocytosis: toxic effect of quinine bisulphate on bone marrow cultures in vitro. BMJ. 1977;1(6061):605–7.

43. Peterson JA, Nelson TN, Kanack AJ, Aster RH. Fine specificity of drug-dependent antibodies reactive with a restricted domain of platelet GPIIIA. Blood. 2008;111(3):1234–9.

44. Bremner JD, Shearer K, McCaffery P. Retinoic acid and affective disorders: the evidence for an association. Neuropsychopharmacology. 2010;35:S115.

45. Bremner JD, McCaffery P. The neurobiology of retinoic acid in affective disorders. Prog Neuropsychopharmacol Biol Psychiatry. 2008;32(2):315–31.

46. Pliszka SR. Pharmacologic treatment of attention-deficit/hyperactivity disorder: efficacy, safety and mechanisms of action. Neuropsychol Rev. 2007;17(1):61–72.

47. Ross RG. Psychotic and manic-like symptoms during stimulant treatment of attention deficit hyperactivity disorder. Am J Psychiatry. 2006;163(7):1149–52.

48. Claassen S, Zunkler BJ. Comparison of the effects of metoclopramide and domperidone on HERG channels. Pharmacology. 2005;74(1):31–6.

49. Stork D, Timin EN, Berjukow S, Huber C, Hohaus A, Auer M, et al. State dependent dissociation of HERG channel inhibitors. Br J Pharmacol. 2007;151(8):1368–76.

50. Stanat SJC, Carlton CG, Crumb WJ Jr, Agrawal KC, Clarkson CW. Characterization of the inhibitory effects of erythromycin and clarithromycin on the HERG potassium channel. Mol Cell Biochem. 2003;254(1–2):1–7.

51. Sheldon R, Duff H, Koshman ML. Antiarrhythmic activity of quinine in humans. Circulation. 1995;92(10):2944–50.

52. White NJ, Looareesuwan S, Warrell DA. Quinine and quinidine: a comparison of EKG effects during the treatment of malaria. J Cardiovasc Pharmacol. 1983;5(2):173–5.

53. Nanau RM, Neuman MG. Ibuprofen-induced hypersensitivity syndrome. Transl Res. 2010;155(6):275–93.

54. Wood JD, Peesker SJ. A correlation between changes in GABA metabolism and isonicotinic acid hydrazide-induced seizures. Brain Res. 1972;45(2):489–98.

55. Carmichael J, Shankel SW. Effects of nonsteroidal anti-inflammatory drugs on prostaglandins and renal function. Am J Med. 1985;78(6 Pt 1):992–1000.

56. Brinker AD, Beitz J. Spontaneous reports of thrombocytopenia in association with quinine: clinical attributes and timing related to regulatory action. Am J Hematol. 2002;70(4):313–7.

57. Noren GN, Caster O, Juhlin K, Lindquist M. Zoo or savannah? Choice of training ground for evidence-based pharmacovigilance. Drug Saf. 2014;37(9):655–9.

58. Alvarez Y, Hidalgo A, Maignen F, Slattery J. Validation of statistical signal detection procedures in eudravigilance post-authorization data: a retrospective evaluation of the potential for earlier signalling. Drug Saf. 2010;33(6):475–87.

59. Harpaz R, Odgers D, Gaskin G, DuMouchel W, Winnenburg R, Bodenreider O, et al. A time-indexed reference standard of adverse drug reactions. Sci Data. 2014;1. doi:10.1038/sdata.2014.43.