

The role of electronic healthcare record databases in paediatric drug safety surveillance: a retrospective cohort study

Sandra de Bie,^{1,2} Preciosa M. Coloma,¹ Carmen Ferrajolo,^{1,3} Katia M. C. Verhamme,¹ Gianluca Trifirò,^{1,4} Martijn J. Schuemie,¹ Sabine M. J. M. Straus,^{1,2} Rosa Gini,⁵ Ron Herings,^{1,6} Giampiero Mazzaglia,⁷ Gino Picelli,⁸ Arianna Ghirardi,⁹ Lars Pedersen,¹⁰ Bruno H. C. Stricker,^{1,11,12} Johan van der Lei¹ & Miriam C. J. M. Sturkenboom¹ on behalf of the EU-ADR consortium

¹Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands, ²Dutch Medicines Evaluation Board, Utrecht, the Netherlands, ³Campania Regional Center of Pharmacovigilance and Pharmacoepidemiology, Department of Experimental Medicine, Pharmacology Section, Second University of Naples, Naples, Italy, ⁴Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Messina, Italy, ⁵Agenzia Regionale di Sanità della Toscana, Florence, Italy, ⁶PHARMO Institute, Utrecht, the Netherlands, ⁷Società Italiana di Medicina Generale, Florence, Italy, ⁸Pedianet-Società Servizi Telematici SRL, Padova, Italy, ⁹Department of Statistics, Università di Milano-Bicocca, Milan, Italy, ¹⁰Department of Clinical Epidemiology, Aarhus University Hospital Aarhus, Denmark, ¹¹Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands and ¹²Inspectorate of Health Care, The Hague, the Netherlands

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- Pre-approval clinical trials are rarely performed in children and adolescents and safety monitoring of drugs in this population relies heavily on the post-marketing phase, even more so than in adults.
- Population-based electronic healthcare records (EHRs) have become increasingly available and represent an additional resource for paediatric drug safety surveillance.

AIM

Electronic healthcare record (EHR)-based surveillance systems are increasingly being developed to support early detection of safety signals. It is unknown what the power of such a system is for surveillance among children and adolescents. In this paper we provide estimates of the number and classes of drugs, and incidence rates (IRs) of events, that can be monitored in children and adolescents (0–18 years).

METHODS

Data were obtained from seven population-based EHR databases in Denmark, Italy, and the Netherlands during the period 1996–2010. We estimated the number of drugs for which specific adverse events can be monitored as a function of actual drug use, minimally detectable relative risk (RR) and IRs for 10 events.

RESULTS

The population comprised 4 838 146 individuals (25 575 132 person years (PYs)), who were prescribed 2170 drugs (1 610 631 PYs drug-exposure). Half of the total drug-exposure in PYs was covered by only 18 drugs (0.8%). For a relatively frequent event like upper gastrointestinal bleeding there were 39 drugs for which

Correspondence

Dr Preciosa M. Coloma, MD, PhD, Department of Medical Informatics, Na-2611, Erasmus University Medical Center, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands.

Tel.: +31 10 704 4129

Fax: +31 10 704 4722

E-mail: p.coloma@erasmusmc.nl

Keywords

active drug safety surveillance, electronic healthcare records, EU-ADR, paediatric drug safety, pharmacovigilance, safety monitoring

Received

22 October 2014

Accepted

10 February 2015

Accepted Article Published Online

16 February 2015

WHAT THIS STUDY ADDS

- This study provides estimates of frequency range of adverse events and number and classes of drugs with enough exposure that can be monitored in the paediatric population using EHRs, based on actual exposure of almost 5 million children and adolescents in three countries.
- Use of additional sources of paediatric drug exposure information and global collaboration are necessary in order to optimise EHR data for paediatric safety surveillance.

an association with a $RR \geq 4$, if present, could be investigated. The corresponding number of drugs was eight for a rare event like anaphylactic shock.

CONCLUSION

Drug use in children is rare and shows little variation. The number of drugs with enough exposure to detect rare adverse events in children and adolescents within an EHR-based surveillance system such as EU-ADR is limited. Use of additional sources of paediatric drug exposure information and global collaboration are imperative in order to optimize EHR data for paediatric safety surveillance.

Introduction

Since very few pre-approval clinical trials are performed in children and adolescents, safety monitoring of drugs in this population relies heavily on the post-marketing phase, even more so than in adults. Spontaneously reported adverse drug reactions (ADRs) and post-marketing safety studies remain the most important sources for the identification of safety signals in both children and adults [1, 2]. There is a fair amount of experience with using spontaneous reporting systems (SRS) to study vaccine safety in children [3–8] but data are lacking as to the utility of such systems for routine safety surveillance of conventional medicines in children. Although SRS have proven their value for safety surveillance, there are well-recognized limitations and biases such as selective under-reporting, stimulated reporting and the lack of exposure data [9–11].

To complement SRS and other traditional monitoring systems, initiatives in the US and in Europe have previously set up population-based surveillance systems that make use of longitudinal healthcare data [12–14]. In Europe, the EU-ADR Project (Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge) was initiated in 2008 and is a collaboration of 18 public and private institutions. EU-ADR aimed to exploit information from various electronic healthcare record (EHR) and other biomedical databases in Europe to produce a computerized integrated system for the early detection of drug safety signals [12]. An earlier study conducted within EU-ADR showed that EHR databases are valuable data sources for drug safety surveillance in the general population but the statistical power might be low for infrequently used drugs or for rare outcomes [15]. While the study included paediatric data, no specific analyses were performed on the paediatric population.

In this current study, we provide estimates of the number and classes of drugs that have enough exposure to be monitored in children and adolescents for selected events investigated within EU-ADR. We further provide information on the frequency range of events that can be monitored based on the actual drug exposure within the study cohort and we give an approximation of how large a database network should be designed for the purpose of monitoring drug safety in children.

Methods

Data sources and setting

We performed a retrospective cohort study using data from the EU-ADR network of databases, of which detailed descriptions have been previously published [12, 16]. In summary, EU-ADR comprises data from eight EHR databases in four European countries. For this study we used paediatric data from seven of the databases from three European countries. Health Search/CSD Longitudinal Patient Database (HSD, Italy), Integrated Primary Care Information (IPCI, the Netherlands) and Pedianet (Italy) are population-based general practice databases, in which clinical information and medication prescriptions are recorded. Aarhus University Hospital Database (Aarhus, Denmark), PHARMO Network (Netherlands) and the regional Italian databases of Lombardy (Lombardy) and Tuscany (Agenzia Regionale di Sanità - ARS) are all comprehensive record-linkage systems in which drug dispensation data of regional/national catchment area are linked to a registry of hospital discharge diagnoses and other registries. The majority of healthcare services, including pharmaceutical services, are provided for, or subsidized by, the state in Italy and Denmark and covered by obligatory health insurance in the Netherlands.

In all of these countries general practitioners or family paediatricians function as gatekeepers of the healthcare system. All databases are population-based and represent all age groups, with the exception of Pédianet (which includes from birth to 14 years) and HSD (which includes individuals from 15 years onwards).

Study population

The study population included children and adolescents aged 0 to 18 years within the databases mentioned above. The study period ran from January 1 1996 to December 31 2010. Follow-up started after a run-in period of 365 days. This run-in period was required to determine the first occurrence of an event. The run-in period was omitted for children younger than 1 year at the start of observation. These children started to contribute follow-up person time from the date of birth or the date of registration, whichever came first.

Data from the different databases were pooled using a distributed network approach, in which data holders maintain control over their original data and only aggregated data are shared with the rest of the network. This was done through generation of all data into a common format followed by local aggregation using custom-built software, Jerboa[®] [12]. Jerboa was developed at Erasmus MC in Rotterdam, the Netherlands specifically for the EU-ADR project and is not commercially available.

Drug exposure

Drug exposure was categorized using the World Health Organization's (WHO) Anatomical Therapeutic Chemical (ATC) classification system [17] and measured in terms of person-years (PYs). We further analyzed drug use by anatomical main groups (ATC first level) and by chemical substances (ATC fifth level).

To study the exposure distribution, drugs were subsequently grouped according to the total amount of exposure in PYs as follows: <10 PYs, >10–≤50 PYs, >50–≤100 PYs, >100–≤500 PYs, >500–≤1000 PYs, >1000–≤5000 PYs, >5000–≤10 000 PYs and >10 000 PYs. The number of drugs that accounted for 50% and 90% of the total drug exposure in the study population were also calculated.

Events

The identification of the events of interest and the process of mapping and harmonization of event coding terminologies across the various databases within EU-ADR have been described in more detail in other publications [12, 16, 18, 19]. In summary, 10 pre-selected events, considered to be most serious and most relevant (generally within the context of safety monitoring in adults), were identified in the databases using an iterative process starting with event definitions based on clinical criteria established from the literature. The following are the

events of interest (all of these events occur at an annual IR of >1/100 000 PYs in children): (1) acute liver injury, (2) acute renal failure, (3) anaphylactic shock, (4) bullous eruptions, (5) cardiac valve fibrosis, (6) hip fractures, (7) neutropenia, (8) acute pancreatitis, (9) pancytopenia and (10) upper gastrointestinal bleeding. Events were extracted using diagnosis codes and free text as well as laboratory findings, when available.

Statistical analysis

Cohort-based approach

Given the pooled empirically derived IRs of each of the above-mentioned events, we calculated the total amount of PYs of exposure that would be required to detect an association between a particular drug and a particular event over varying magnitudes of relative risk (RR) of 2 ('weak' association), 4 ('moderate' association) and 6 ('strong' association) using a one-sided significance level $\alpha = 0.05$ and power of 80% ($\beta = 0.2$) [15]. We subsequently determined the number of drugs for which there would be sufficient data for safety monitoring, expressing this as the number of unique chemical substances (ATC 5th level). We calculated the proportion of the PYs of exposure for drugs with enough exposure (to detect the RR of interest) for these drugs, compared with the total exposure for all drugs. We then calculated the range of IRs of events that can be monitored to detect 'weak' ($RR \geq 2$), 'moderate' ($RR \geq 4$) or 'strong' ($RR \geq 6$) associations based on the actual drug exposure within the cohort. These results were stratified within categories of drug exposure (as specified under 'drug exposure') and age.

Based on the actual exposure and theoretical incidences of adverse events (1/100000 PYs, 10/100 000 PYs, 50/100 000 PYs, 100/100 000 PYs and 500/100 000 PYs) we also calculated for how many drugs within specific drug groups (ATC first level) would there be enough exposure to detect associations with varying magnitudes.

Stratification by age

Results were stratified according to four age categories: 0–<2 years, 2–≤5 years, 6–≤11 years and 12–<18 years [20].

Results

The paediatric population (aged 0–18 years) in this study comprised 4 838 146 individuals contributing 25 575 132 PYs of follow-up (1996–2010). Among those children contributing these PYs, 12.8% were aged 0–<2 years, 22.2% aged 2–≤5 years, 32.7% aged 6–≤11 years, while 32.3% were adolescents aged 12–<18 years.

A total of 2170 drugs were used (i.e. prescribed or dispensed) during the study period with an overall exposure time of 1 610 631 PYs. An overview of drug exposure, at

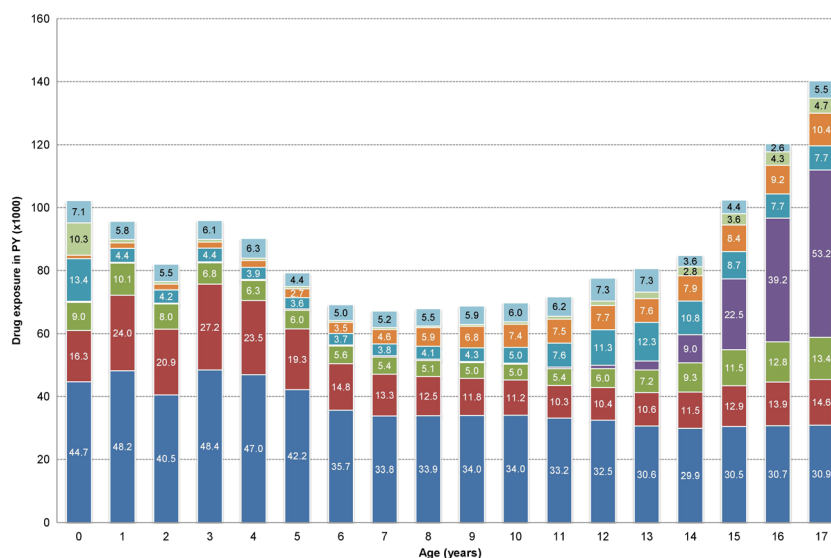


Figure 1

Drug exposure in person-years by age. Note: Drug exposure is aggregated on the first ATC level (anatomical main group). ‘Other’ represents all other drug groups with a total exposure of <5000 PYs. ■ respiratory, ■ anti-infectives, ■ dermatological, ■ genitourinary, ■ alimentary, ■ neurologic, ■ blood, ■ other (<5 000 PYs)

the anatomical level of the ATC classification across different age categories, is shown in Figure 1. Across the entire paediatric population the drug classes with the highest exposure were respiratory drugs (661 000 PYs), anti-infective drugs (279 000 PYs), dermatological drugs (138 000 PYs), genitourinary drugs (132 000 PYs) and alimentary drugs (121 000 PYs). The remaining drug classes had a total exposure of <100 000 PYs (data not shown). Respiratory drugs and anti-infective drugs in particular accounted for the majority of exposure up to 12 years of age, while genitourinary drugs (mainly oral

contraceptives) comprised the bulk of drug exposure from age 14 years onwards.

The number of drugs that have enough exposure to detect, if present, weak (RR ≥2), moderate (RR ≥4) or strong (RR ≥6) associations for the 10 events are presented in Table 1. Since the numbers are low, these results were not further stratified by age. As expected, the stronger the association to be investigated, the higher the number of drugs that can be investigated. Also, more drugs can be studied for frequently occurring events. For a relatively frequent event such as upper

Table 1

Amount of required drug exposure to identify potentially drug-induced adverse events

Event type	IR per 100 000 PY	Weak association (RR ≥2)			Moderate association (RR ≥4)			Strong association (RR ≥6)		
		Required exposure (PY)	Drugs n	% of Exp	Required exposure (PY)	Drugs n	% of Exp	Required exposure (PY)	Drugs n	% of Exp
Hip fracture	15.3	52 501	6	29.5	8039	42	67.8	3589	81	80.4
Upper GI bleeding	14.4	55 725	5	26.2	8532	39	66.3	3810	79	79.9
Neutropenia	8.1	99 259	2	13.0	15 198	25	56.9	6786	48	70.5
Acute liver injury	4.0	202 733	0	0	31 041	9	37.3	13 860	26	57.8
Pancytopenia	3.7	215 469	0	0	32 991	9	37.3	14 730	25	56.9
Bullous eruption	3.6	224 394	0	0	34 358	9	37.3	15 341	24	56.0
Anaphylactic shock	3.2	248 526	0	0	38 053	8	35.0	16 990	20	52.1
Cardiac valve fibrosis	2.9	275 840	0	0	42 235	8	35.0	18 858	15	46.6
Acute renal failure	1.6	517 050	0	0	79 168	3	17.9	35 348	9	37.3
Acute pancreatitis	1.6	519 664	0	0	79 568	3	17.9	35 527	9	37.3

Drugs (n): Number of drugs at fifth ATC, chemical substance level that have enough PY of exposure to detect a potential signal (total 2170). % of Exp: Proportion of PYs of exposure of the drugs with enough exposure compared with the total PYs of exposure for all drugs. IR, incidence rate; PY, person years; RR, relative risk; upper GI bleeding, upper gastrointestinal bleeding

gastrointestinal bleeding (UGIB) ($IR=14.4/100\ 000$ PYs), five drugs had the minimal exposure of 55 725 PYs required to detect a weak association ($RR \geq 2$). These five drugs comprised 26.2% of the total drug exposure in PYs. Thirty-nine drugs (representing two-thirds of total drug exposure) had the minimal required exposure of 8532 PYs to detect a moderate association ($RR \geq 4$) with UGIB while 79 drugs (79.9% of total exposure) had the minimal exposure of 3810 PYs required to assess a strong association ($RR \geq 6$). On the other hand, for acute pancreatitis, a rare event among children and adolescents ($IR = 1.6/100\ 000$ PYs), none of the drugs had enough exposure to detect a weak association ($RR \geq 2$), three drugs (17.9% of exposure) had enough exposure to detect a moderate association ($RR \geq 4$) and only nine drugs (37.3% of exposure) had enough exposure to detect a strong association ($RR \geq 6$).

The number of drugs, stratified at the anatomical level of the ATC classification, with enough exposure to investigate weak, moderate or strong associations is given in **Table 2**. Respiratory drugs and anti-infectives were among those drugs having enough exposure to monitor associations of $RR \geq 2$, $RR \geq 4$ and $RR \geq 6$ for events having (theoretical) IRs 10/100 000 PYs and higher. Anti-neoplastic, antiparasitic and cardiovascular drugs, because they are rarely used in the paediatric population (and/or rarely documented in the databases), did not have enough exposure to monitor an association with $RR \geq 2$ for any of the events with (theoretical) incidences ranging from 1 to 500/100 000 PYs.

As illustrated in Figure 2, only a small proportion of drugs that are being used in the EU-ADR paediatric population of around 5 million subjects had a high person-time exposure. About half of the drugs had a total exposure of <10 PYs. This was most pronounced in the youngest age group. In the table accompanying Figure 2, the minimal detectable IRs for the exposure categories for each RR is given. For drugs with exposure of 10 PYs, events having IRs of 765/1000 PY and higher can be detected for $RR \geq 2$, 12/1000 PY and higher for $RR \geq 4$ and 5.2/1000 PY and higher for $RR \geq 6$. The proportion of the drugs with an exposure of more than 1000 PYs (necessary to detect IRs of up to 16/1000 PYs with a $RR \geq 2$) represented only about 9% of total exposure for the entire paediatric population.

Eighteen of the 2170 drugs (0.8%) made up 50% of the total drug exposure in PYs (0–<18 years), each of these 18 drugs representing exposure of at least 18 236 PY (see Table 3). For 0–<2 years, 2–≤5 years, 6–≤11 years and 12–<18 years there were eight (0.6% of total exposure within the age group), eight (0.5%), 14 (0.9%) and 20 (1.0%) drugs, respectively. Based on these exposures, events (other than those 10 already considered in this study) could be detected with $IR >44/100\ 000$ PYs ($RR \geq 2$), $IR >6.7/100\ 000$ PYs ($RR \geq 4$) and $IR >3.0/100\ 000$ PYs ($RR \geq 6$) (data not shown). Data from

90% of the total drug exposure for the entire population was represented by 158 drugs. The distribution per age category is as follows. For 0–<2 years, there were 67 drugs (representing 3.1% of total exposure within this group), for 2–≤5 years, 86 drugs (4.0% of exposure), for 6–≤11 years, 125 drugs (5.8% of exposure) and for 12–<18 years 165 drugs (7.6% of exposure). Each of these 158 drugs had drug exposures of ≥ 1334 PYs, allowing detection of events (other than those 10 already considered in this study) with $IR >603/100\ 000$ PYs ($RR \geq 2$), $IR >92/100\ 000$ PYs ($RR \geq 4$) and $IR >41/100\ 000$ PYs ($RR \geq 6$) (data not shown).

Discussion

The number of initiatives evaluating the use of EHR databases as a source for drug safety surveillance is growing [13, 14, 21, 22]. Some of these include data on children and adolescents. However, none of these initiatives has yet focused on the paediatric population, for whom safety data are actually lacking the most.

Despite the fact that almost 5 million children and adolescents within the EU-ADR database network are included in this study, the number of drugs that have enough exposure to study weak, moderate or strong associations with the events monitored within this network is limited. For a rare but serious event like anaphylactic shock there were no drugs with enough exposure to study a weak association ($RR \geq 2$) and only 20 drugs to study a strong association ($RR \geq 6$). These numbers are low compared with the total of 2170 drugs used. It is mainly for drugs that are known to be frequently used in children (i.e. anti-infectives, respiratory drugs and hormones) [23] that there was enough exposure to monitor a wide range of IRs for varying magnitudes of risks. An important group of drugs for which safety alerts concerning the use in children and adolescents have been issued in recent years are central nervous system drugs for the treatment of ADHD (attention deficit-hyperactivity disorder) [24]. Methylphenidate was the only neurological drug among 18 drugs that covered 50% of the total drug exposure in PYs. Thus, safety concerns regarding methylphenidate can be studied, although this may not be possible for other neurological drugs where the total exposure is too low for a potential risk to be detectable.

Our findings showed that in this paediatric population, the amount of drug exposure (person-time) is low and a limited number of drugs cover the majority of the exposure. Children made up about 20% of the entire EU-ADR network population and only contributed 3% of total drug exposure time (see also Supplementary Figure 1 for overview of distribution of drug exposure across age groups in the overall EU-ADR population, including adults). The 1.6 million PYs of exposure were distributed over 2170 individual drugs, compared with 2289 for the overall population (all ages) in the database network (95%). Of these,

Table 2

Number of drugs (fifth ATC level) by class with enough exposure to study the given incidences with the given relative risks (RRs)

First ATC tclass *	≤1/100 000 PYs			≤10/100 000 PYs			≤50/100 000 PYs			≤100/100 000 PYs			≤500/100 000 PYs		
	RR ≥2	RR ≥4	RR ≥6	RR ≥2	RR ≥4	RR ≥6	RR ≥2	RR ≥4	RR ≥6	RR ≥2	RR ≥4	RR ≥6	RR ≥2	RR ≥4	RR ≥6
Alimentary (n = 391)	N (%)	-	-	1 (0.3)	2 (0.5)	1 (0.3)	11 (2.8)	21 (5.4)	2 (0.5)	20 (5.1)	37 (9.5)	14 (3.6)	52 (13.3)	66 (16.9)	
Respiratory (n = 160)	N (%)	-	3 (1.9)	2 (1.3)	15 (9.4)	21 (13.1)	12 (7.5)	29 (18.1)	18 (11.3)	35 (21.9)	47 (29.4)	30 (18.8)	54 (33.8)	63 (39.4)	
Dermatological (n = 203)	N (%)	-	-	-	2 (1.0)	8 (3.9)	1 (0.5)	14 (6.9)	23 (11.3)	5 (2.5)	23 (11.3)	19 (9.4)	48 (23.6)	62 (30.5)	
Anti-infectives (n = 232)	N (%)	-	2 (0.9)	1 (0.4)	5 (2.2)	8 (3.4)	3 (1.3)	12 (5.2)	22 (9.5)	6 (2.6)	19 (8.2)	17 (7.3)	40 (17.2)	58 (25.0)	
Neurologic (n = 269)	N (%)	-	-	-	2 (0.7)	2 (0.7)	1 (0.4)	9 (3.3)	19 (7.1)	2 (0.7)	19 (7.1)	14 (5.2)	41 (15.2)	57 (21.2)	
Sensory organs (n = 169)	N (%)	-	-	-	-	1 (0.6)	-	5 (3.0)	12 (7.1)	-	12 (7.1)	11 (6.5)	34 (20.1)	42 (24.9)	
Cardiovascular (n = 192)	N (%)	-	-	-	-	-	-	5 (2.6)	5 (2.6)	-	3 (1.6)	9 (4.7)	18 (9.4)	33 (17.2)	
Genitourinary (n = 153)	N (%)	-	1 (0.7)	-	2 (1.3)	4 (2.6)	2 (1.3)	5 (3.3)	8 (5.2)	3 (2.0)	8 (5.2)	8 (5.2)	24 (15.7)	32 (20.9)	
Hormones (n = 72)	N (%)	-	-	-	2 (2.8)	8 (11.1)	-	10 (13.9)	14 (19.4)	4 (5.6)	14 (19.4)	18 (25.0)	24 (33.3)	30 (41.7)	
Musculoskeletal (n = 104)	N (%)	-	-	-	-	-	-	2 (1.9)	5 (4.8)	-	4 (3.8)	7 (6.7)	3 (2.9)	8 (7.7)	
Blood (n = 97)	N (%)	-	-	-	-	2 (2.1)	-	3 (3.1)	6 (6.2)	2 (2.1)	5 (5.2)	7 (7.2)	3 (3.1)	11 (11.3)	
Antiparasitic (n = 35)	N (%)	-	-	-	-	-	-	-	1 (2.9)	-	1 (2.9)	-	5 (14.3)	11 (31.4)	
Antineoplastics (n = 76)	N (%)	-	-	-	-	-	-	-	2 (2.6)	-	2 (2.6)	-	6 (7.9)	10 (13.2)	
Total (n = 2170)	N (%)	-	6 (0.3)	3 (0.1)	29 (1.3)	56 (2.6)	20 (0.9)	100 (4.6)	175 (8.1)	42 (1.9)	165 (7.6)	251 (11.6)	365 (16.8)	498 (22.9)	

* First ATC level 'Various' (n = 35) had no drugs with enough exposure for any of the incidences and is not presented.

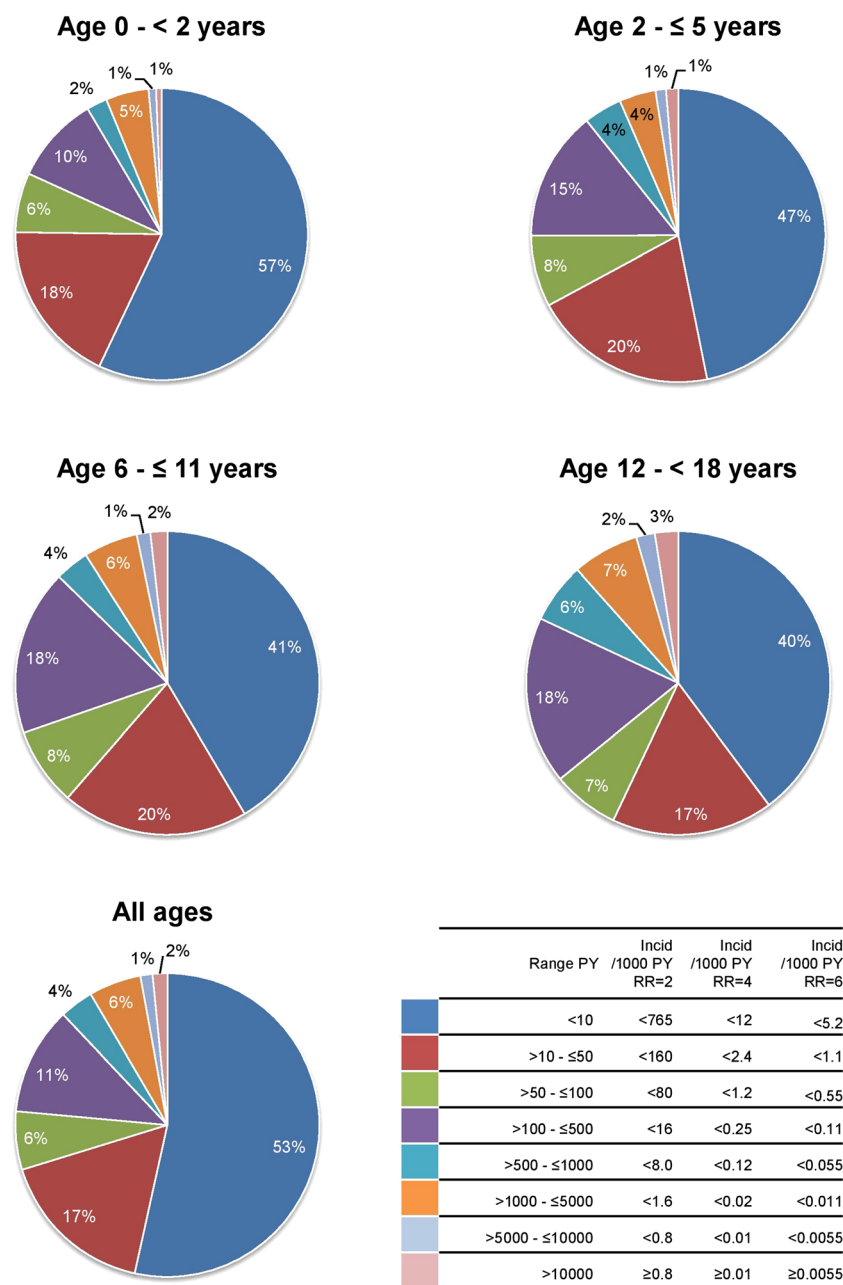


Figure 2

Distribution of exposure in PY by age groups (fifth ATC level, chemical subgroup). The range in PY is given with the corresponding incidence rates of events that can be monitored. PY person-years; incid incidence

only 18 represented 50% and 158 drugs represented 90% of the total drug exposure time. This knowledge places the number of drugs having enough exposure to detect weak, moderate or strong associations in another context. The 20 drugs that have enough exposure to study a strong association with anaphylactic shock (at $RR \geq 6$) represent 52.1% of the total drug exposure. As illustrated in the current study, moderate associations can be studied for half of the total drug exposure, for events having IRs of $\geq 10/100\ 000$ (with 29 drugs), while for events having IRs of $\geq 50/100\ 000$ weak associations can also be studied

(with 20 drugs). It should be noted that these results have not been corrected for multiple testing.

The number of drugs that can be investigated is also limited by the low IRs in the paediatric population of the 10 events of interest in EU-ADR. The low IRs were not surprising, since the events were chosen based on safety issues that were perceived to be relevant in the general adult population. An example of an event that is particularly relevant and frequently occurring in children, but was not considered a priority event in EU-ADR, is febrile seizures (incidence estimated at 14/1000

Table 3

Drugs that cover 50% of the total drug exposure in person years by age categories

Age 0–< 2 years *	Age 2–≤ 5 years*	Age 6–≤ 11 years*	Age 12–< 18 years*	Total*	
Beclomethasone [R03BA01] (13.1)	Beclomethasone [R03BA01] (12.5)	Salbutamol [R03AC02] (6.8)	Levonorgestrel and estrogen [G03AA07] (11.9)	Beclomethasone [R03BA01] (6.8)	1
Salbutamol [R03AC02] (10.5)	Salbutamol [R03AC02] (9.1)	Beclomethasone [R03BA01] (6.4)	Sodium fluoride [A01AA01] (3.9)	Salbutamol [R03AC02] (6.2)	2
Amoxicillin [J01CA04] (6.5)	Amoxicillin/Clavulanic acid [J01CR02] (7.9)	Amoxicillin/Clavulanic acid [J01CR02] (5.2)	Amoxicillin/Clavulanic acid [J01CR02] (3.4)	Amoxicillin/Clavulanic acid [J01CR02] (4.9)	3
Amoxicillin/Clavulanic acid [J01CR02] (4.9)	Amoxicillin [J01CA04] (5.0)	Fluasone [R03BA05] (4.8)	Salbutamol [R03AC02] (3.2)	Levonorgestrel and estrogen [G03AA07] (4.6)	4
Phytomenadione (vitamin K) [B02BA01] (4.4)	Fluticasone [R03BA05] (5.0)	Cetirizine [R06AE07] (4.1)	Cyproterone and estrogen [G03HB01] (3.2)	Amoxicillin [J01CA04] (3.6)	5
Fluticasone [R03BA05] (3.8)	Budesonide [R03BA02] (4.3)	Budesonide [R03BA02] (3.5)	Cetirizine [R06AE07] (2.5)	Fluticasone [R03BA05] (3.4)	6
Budesonide [R03BA02] (3.6)	Clarithromycin [J01FA09] (3.8)	Amoxicillin [J01CA04] (3.4)	Beclomethasone [R03BA01] (2.4)	Budesonide [R03BA02] (2.9)	7
Flunisolide [R03BA03] (3.6)	Flunisolide [R03BA03] (3.1)	Methylphenidate [N06BA04] (3.2)	Amoxicillin [J01CA04] (2.3)	Cetirizine [R06AE07] (2.6)	8
		Salmeterol and other drugs for obstructive airway diseases [R03AK06] (2.7)	Ferrous sulfate [B03AA07] (2.1)	Clarithromycin [J01FA09] (2.2)	9
		Clarithromycin [J01FA09] (2.7)	Methylphenidate [N06BA04] (1.8)	Sodium fluoride [A01AA01] (1.9)	10
		Desmopressin [H01BA02] (2.3)	Salmeterol and other drugs for obstructive airway diseases [R03AK06] (1.7)	Flunisolide [R03BA03] (1.7)	11
		Montelukast [R03DC03] (1.7)	Desloratadine [R06AX27] (1.6)	Methylphenidate [N06BA04] (1.6)	12
		Fluticasone (nasal) [R01AD08] (1.7)	Budesonide [R03BA02] (1.6)	Salmeterol and other drugs for obstructive airway diseases [R03AK06] (1.6)	13
		Terbutaline [R03AC03] (1.6)	Fluticasone [R03BA05] (1.6)	Terbutaline [R03AC03] (1.5)	14
			Levocetirizine [R06AE09] (1.4)	Cyproterone and estrogen [G03HB01] (1.2)	15
			Gestodene and estrogen [G03AA10] (1.4)	Fluticasone [R01AD08] (1.1)	16
			Clarithromycin [J01FA09] (1.3)	Montelukast [R03DC03] (1.1)	17
			Fluticasone (nasal) [R01AD08] (1.3)	Salbutamol and other drugs for obstructive airway diseases [R03AK04] (1.1)	18
			Terbutaline [R03AC03] (1.2)		
			Mometasone [R01AD09] (1.1)		

*drug [fifth ATC level] (% of total exposure in person years)

PYs) [25]. Considering this incidence, a total of 132 drugs within the database network would have enough exposure to detect a possible association with febrile seizures at $RR \geq 4$. What this essentially illustrates is that extrapolation of relevant safety outcomes from adults to children does not always work and that it is very important to choose age appropriate events and definitions when setting up EHR-based paediatric surveillance systems.

An important question that remains unaddressed is whether the positive predictive value of mining EHR data

for safety surveillance will be higher than data mining in SRS [10]. Trifirò and colleagues compared potential signals derived from the EU-ADR network with signals derived from SRS [26]. SRS appeared to be more likely to detect potential associations for events with a low incidence in the general population and commonly regarded as drug-induced like Stevens Johnson syndrome and anaphylactic shock. The results we obtained in this study are in line with Trifirò *et al.*'s findings that EHR-based surveillance may complement traditional

SRS in the detection of adverse events that are frequent in the general population and are not usually regarded to be drug-related. For events with a low IR and a high probability to be drug-induced only a small number of drugs had enough exposure to detect potentially drug-induced events. For events with a high IR a larger number of drugs could be studied. This makes EU-ADR and other EHR-based systems an important complement to existing SRS. While ADRs are more likely to be detected (and reported) at the beginning of drug therapy (since at this time both the treating physician and the patients are most aware of potential adverse events), the longitudinal nature of the data collection in EHRs, may allow further observation, after long term use of drugs and, possibly, even for rare diseases.

Limitations and future directions

Our study illustrates that the capacity of EHR databases as a source for paediatric drug safety surveillance is limited not by the size of the population, but is mainly hampered by the fact that the majority of the drugs are prescribed very rarely or for a very short period of time. We emphasize that the results should be interpreted within the context of the data that gave rise to these results. Because the majority of the databases are primary care-based, specialist prescriptions (e.g. for antineoplastic drugs, immunologic agents and other biologicals) are only captured in the system if continued by the general practitioner or if provided through the routine dispensing system. In-hospital drug use is not captured. Furthermore, over-the-counter drugs are not documented in these databases. Expansion of the database network to include other populations and linkage to other data sources of drug exposure (e.g. hospitals and specialty clinics) would be necessary to capture the greatest possible number of drugs used in the population. The purpose of this is not only to increase the size of the studied population, but also to increase the variation in prescribing patterns.

Other relevant sources of paediatric data include the FDA Mini-Sentinel and the Observational Medical Outcomes Partnership, OMOP, (now combined under the Innovation in Medical Evidence Development and Surveillance (IMEDS) Programm) [14, 21]. The paediatric population within Mini-Sentinel comprised approximately 27 million children and adolescents up to 19 years (21.6% of total) [27] while OMOP comprised approximately 39.5 million children and adolescents up to 18 years. With the possibility to combine all these data sources together, the current study population will be enlarged by a factor of approximately 15. Assuming similar patterns of follow-up and patterns of exposure to drugs in all databases, this (hypothetical) population will have a total drug exposure of approximately 24 million PYs. Consequently, for an event like anaphylactic shock the number of drugs having enough exposure to study

a moderate association ($RR \geq 4$) will increase from 8 to 100 and for a more frequent event like UGIB 242 drugs, instead of 39, could be investigated to study a moderate association. Global collaboration will be necessary for further development of paediatric drug safety monitoring systems using EHRs and for external validation of newly identified safety signals.

Estimation of adequate statistical power for this study was based on cohort methodology. However, there are other more efficient study designs such as the case-only or self-controlled designs that are not as predisposed to the limitations of rare outcomes or low and intermittent exposures. Such alternative designs may give different results and need to be explored. There are ongoing efforts worldwide that aim to improve paediatric drug therapy (e.g. Global Research in Paediatrics, <http://www.grip-network.org>), also by developing new statistical methodologies for drug safety surveillance. Such new methodologies can further advance the use of EHR data for paediatric surveillance.

In conclusion, drug use in children is rare and only 18 out of the total 2170 prescribed or dispensed drugs (<1%) made up half of the total exposure to drugs in the paediatric population of almost 5 million within the EU-ADR network. The number of drugs with enough exposure within an EHR-based surveillance system such as EU-ADR to detect rare adverse events in children and adolescents is limited. Mining within EHR databases seems especially promising for events that have a high background incidence in the paediatric population and for drugs with a large amount of exposure. Additional sources of paediatric drug exposure information and global collaboration are imperative in order to optimize EHR data for paediatric safety surveillance.

Competing Interests

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare Katia Verhamme has been involved as project coordinator in analyses conducted by various pharmaceutical companies and received unconditional research grants from Pfizer, Yamanouchi and Boehringer-Ingelheim, none of which is related to the subject of this study. Miriam Sturkenboom leads a research group that is conducting research for pharmaceutical companies through non-conditional grants, none of which is related to this research. All other authors have no other relationships or activities that could appear to have influenced the submitted work.

This research has been funded by the European Commission Seventh Framework Programme (FP7/2007–2013) under grant no. 215847—the EU-ADR project.

REFERENCES

- 1 Hartmann K, Doser AK, Kuhn M. Postmarketing safety information: how useful are spontaneous reports? *Pharmacoepidemiol Drug Saf* 1999; 8(Suppl 1): S65–71.
- 2 Harmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. *Eur J Clin Pharmacol* 2008; 64: 743–52.
- 3 Niu MT, Erwin DE, Braun MM. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. *Vaccine* 2001; 19: 4627–34.
- 4 Aagaard L, Hansen EW, Hansen EH. Adverse events following immunization in children: retrospective analysis of spontaneous reports over a decade. *Eur J Clin Pharmacol* 2011; 67: 283–8.
- 5 Letourneau M, Wells G, Walop W, Duclos P. Improving global monitoring of vaccine safety: a quantitative analysis of adverse event reports in the WHO Adverse Reactions Database. *Vaccine* 2008; 26: 1185–94.
- 6 Zeinoun Z, Seifert H, Verstraeten T. Quantitative signal detection for vaccines: effects of stratification, background and masking on GlaxoSmithKline's spontaneous reports database. *Hum Vaccin* 2009; 5: 599–607.
- 7 Woo EJ, Ball R, Burwen DR, Braun MM. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS). *Drug Saf* 2008; 31: 667–74.
- 8 Varricchio F, Iskander J, Destefano F, Ball R, Pless R, Braun MM, Chen RT. Understanding vaccine safety information from the Vaccine Adverse Event Reporting System. *Pediatr Infect Dis J* 2004; 23: 287–94.
- 9 Goldman SA. Limitations and strengths of spontaneous reports data. *Clin Ther* 1998; 20(Suppl C): C40–4.
- 10 Practical Aspects of Signal Detection in Pharmacovigilance. Report of CIOMS Working Group VIII. Geneva: CIOMS, 2010.
- 11 Moore N, Hall G, Sturkenboom M, Mann R, Lagnaoui R, Begaud B. Biases affecting the proportional reporting ratio (PPR) in spontaneous reports pharmacovigilance databases: the example of sertindole. *Pharmacoepidemiol Drug Saf* 2003; 12: 271–81.
- 12 Coloma PM, Schuemie MJ, Trifiro G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom M. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiol Drug Saf* 2011; 20: 1–11. Epub 2010/12/25.
- 13 Platt R, Wilson M, Chan KA, Benner JS, Marchibroda J, McClellan M. The new Sentinel Network—improving the evidence of medical-product safety. *N Engl J Med* 2009; 361: 645–7. Epub 2009/07/29.
- 14 Stang PE, Ryan PB, Racoosin JA, Overhage JM, Hartzema AG, Reich C, Welebob E, Scarnecchia T, Woodcock J. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med* 2010; 153: 600–6. Epub 2010/11/03.
- 15 Coloma PM, Trifiro G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Picelli G, Corrao G, Pedersen L, van der Lei J, Sturkenboom M, on behalf of the EUADRC. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? *Pharmacoepidemiol Drug Saf* 2012; 21: 611–21.
- 16 Trifiro G, Fourrier-Reglat A, Sturkenboom MC, Diaz Acedo C, Van Der Lei J. The EU-ADR project: preliminary results and perspective. *Stud Health Technol Inform* 2009; 148: 43–9. Epub 2009/09/12.
- 17 WHO Collaborating Centre for Drug Statistics Methodology. Guidelines for ATC classification and DDD assignment. [cited 02-02-2012]; Available from <http://www.whocc.no/>.
- 18 Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salame G, Catania MA, Salvo F, David A, Moore N, Caputi AP, Sturkenboom M, Molokhia M, Hippisley-Cox J, Acedo CD, van der Lei J, Fourrier-Reglat A. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf* 2009; 18: 1176–84. Epub 2009/09/17.
- 19 Avillach P, Mougou F, Joubert M, Thiessard F, Pariente A, Dufour JC, Trifiro G, Polimeni G, Catania MA, Giaquinto C, Mazzaglia G, Baio G, Herings R, Gini R, Hippisley-Cox J, Molokhia M, Pedersen L, Fourrier-Reglat A, Sturkenboom M, Fieschi M. A semantic approach for the homogeneous identification of events in eight patient databases: a contribution to the European eu-ADR project. *Stud Health Technol Inform* 2009; 150: 190–4. Epub 2009/09/12.
- 20 International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use. Clinical Investigation Of Medicinal Products In The Pediatric Population E11. 2000 [02-02-2012]; Available at http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E11/Step4/E11_Guideline.pdf.
- 21 Curtis LH, Weiner MG, Boudreau DM, Cooper WO, Daniel GW, Nair VP, Raebel MA, Beaulieu NU, Rosofsky R, Woodworth TS, Brown JS. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiol Drug Saf* 2012; 21(Suppl 1): 23–31. Epub 2012/01/25.
- 22 Johansson S, Wallander MA, de Abajo FJ, Garcia Rodriguez LA. Prospective drug safety monitoring using the UK primary-care General Practice Research Database: theoretical framework, feasibility analysis and extrapolation to future scenarios. *Drug Saf* 2010; 33: 223–32. Epub 2010/02/18.
- 23 Sturkenboom MC, Verhamme KM, Nicolosi A, Murray ML, Neubert A, Caudri D, Picelli G, Sen EF, Giaquinto C, Cantarutti L, Baiardi P, Felisi MG, Ceci A, Wong IC. Drug use in children: cohort study in three European countries. *BMJ* 2008; 337: a2245.
- 24 Clavenna A, Bonati M. Adverse drug reactions in childhood: a review of prospective studies and safety alerts. *Arch Dis Child* 2009; 94: 724–8.

- 25** Sillanpaa M, Camfield P, Camfield C, Haataja L, Aromaa M, Helenius H, Rautava P, Hauser WA. Incidence of febrile seizures in Finland: prospective population-based study. *Pediatr Neurol* 2008; 38: 391–4. Epub 2008/05/20.
- 26** Trifiro G, Patadia V, Schuemie MJ, Coloma PM, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Scotti L, Pedersen L, Avillach P, Sturkenboom MC, van der Lei J, Eu-Adr G. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Stud Health Technol Inform* 2011; 166: 25–30. Epub 2011/06/21.
- 27** Mini-Sentinel. Mini-Sentinel Distributed Database "At A Glance". December 2011; Available at http://www.mini-sentinel.org/about_us/MSDD_At-a-Glance.aspx (last accessed 20 October 2014).

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Supplementary Figure 1

Overview of drug exposure in person-years across all age groups in the EU-ADR network. Note: Drug exposure is aggregated on the first ATC level (anatomical main group). (Reproduced with permission from Coloma *et al.* [15].)