

ORIGINAL REPORT

Electronic healthcare databases for active drug safety surveillance: is there enough leverage?

Preciosa M. Coloma^{1*}, Gianluca Trifirò^{1,2}, Martijn J. Schuemie¹, Rosa Gini³, Ron Herings⁴, Julia Hippisley-Cox⁵, Giampiero Mazzaglia⁶, Gino Picelli⁷, Giovanni Corrao⁸, Lars Pedersen⁹, Johan van der Lei¹, Miriam Sturkenboom^{1,10} and on behalf of the EU-ADR consortium

¹Department of Medical Informatics, Erasmus Medical Center, Rotterdam, The Netherlands

²Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Messina, Italy

³Agenzia Regionale di Sanità della Toscana, Florence, Italy

⁴PHARMO Institute, Utrecht, The Netherlands

⁵Division of Primary Care, School of Community Health Sciences, University of Nottingham, UK

⁶Società Italiana di Medicina Generale, Florence, Italy

⁷Pedianet-Società Servizi Telematici SRL, Padova, Italy

⁸Department of Statistics, Università di Milano-Bicocca, Milan, Italy

⁹Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

¹⁰Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands

ABSTRACT

Purpose To provide estimates of the number and types of drugs that can be monitored for safety surveillance using electronic healthcare databases.

Methods Using data from eight European databases (administrative claims, medical records) and in the context of a cohort study, we determined the amount of drug exposure required for signal detection across varying magnitudes of relative risk (RR). We provide estimates of the number and types of drugs that can be monitored as a function of actual use, minimal detectable RR, and empirically derived incidence rates for the following adverse events: (i) acute myocardial infarction; (ii) acute renal failure; (iii) anaphylactic shock; (iv) bullous eruptions; (v) rhabdomyolysis; and (vi) upper gastrointestinal bleeding. We performed data simulation to see how expansion of database size would influence the capabilities of such system.

Results Data from 19 647 452 individuals (59 594 132 person-years follow-up) who used 2289 drugs in the EU-ADR network show that for a frequent event such as acute myocardial infarction, there are 531 drugs (23% of total) for which an association with RR = 2, if present, can be investigated. For a rare event such as rhabdomyolysis, there are 19 drugs (1%) for which an association of same magnitude can be investigated.

Conclusion Active surveillance using healthcare data-based networks for signal detection is feasible, although the leverage to do so may be low for infrequently used drugs and for rare outcomes. Extending database network size to include data from heterogeneous populations and increasing follow-up time are warranted to maximize leverage of these surveillance systems. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS—active drug safety surveillance; drug safety monitoring; signal detection; electronic healthcare records; electronic healthcare databases; EU-ADR

Received 4 July 2011; Revised 22 November 2011; Accepted 28 November 2011

INTRODUCTION

Safety data of newly marketed drugs are limited because of small and selective groups of individuals included in clinical trials. Characterization of the full safety profile of drugs thus relies on the careful observation and systematic monitoring of their effects

in ‘real-world’ practice. Over the past four decades, individual case safety reports — also known as spontaneous reports — have been the cornerstone of drug safety surveillance post-marketing. Spontaneous reporting systems elicit reports from both health professionals and patients, but they derive their key strength from the reporting physician’s astute clinical observation and insight in assessing the likelihood of an adverse event being drug induced. Although pharmaceutical companies are obliged to submit reports of suspected

*Correspondence to: P. M. Coloma, Ee 2116, Department of Medical Informatics, Erasmus Medical Center, PO Box 2040, 3000CA Rotterdam, The Netherlands. E-mail: p.coloma@erasmusmc.nl

severe and serious reactions to regulatory authorities the mostly voluntary, and hence passive, nature of spontaneous reporting systems lends itself to various limitations. These limitations include significant under-reporting, inability to provide incidence rates (IRs) due to lack of information on user population and patterns of drug use, double-counting of reports pertaining to the same incident, and bias in reporting of adverse events related to drugs that get media attention.^{1–3}

Prompted by the safety issues surrounding rofecoxib (Vioxx), there has been a shift towards utilizing linked electronic healthcare databases for active drug safety surveillance.^{4–7} The Sentinel Initiative was established after the US Food and Drug Administration (FDA) Amendments Act mandated the creation of a new post-marketing surveillance system that, by 2012, will be using electronic health data from 100 million people to prospectively monitor the safety of marketed medical products.^{8,9} The Observational Medical Outcomes Partnership (OMOP), a public–private partnership among the FDA, academia, data owners, and the pharmaceutical industry, and administered by the Foundation for the National Institutes of Health, has been tasked with the development and evaluation of methods and data infrastructure to address the needs of an active surveillance system.^{10,11} Similar initiatives are ongoing in Canada¹² and in Europe. The EU-ADR project (Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge, <http://www.euadr-project.org>) is a collaboration of 18 public and private institutions representing academic research, general practice, healthcare services administration, and the pharmaceutical industry. The project, initiated in 2008, is funded by the European Commission.¹³

Data mining on healthcare databases generates statistical associations between drug exposure and adverse outcomes and can augment the current passive–reactive system and facilitate earlier detection of potential safety issues. Although such an approach for safety surveillance is promising, the question arises as to what this type of surveillance can add to existing systems and whether these database platforms have enough power to adequately detect safety signals. In this paper, we estimate for how many and for which types of drugs safety signals might be detected as a function of actual drug use, minimal detectable relative risk (RR), and IRs of events as observed in the EU-ADR database network.

METHODS

Data sources

The EU-ADR platform currently comprises data from eight electronic healthcare databases in four European

countries.¹⁴ Health Search/CSD LPD (HSD, Italy), Integrated Primary Care Information (IPCI, The Netherlands), Pédianet (Italy), and QRESEARCH (UK) are population-based general practice databases, in which clinical information and medication prescriptions are recorded. Aarhus University Hospital Database (Aarhus, Denmark), PHARMO Network (The Netherlands), and the regional Italian databases of Lombardy and Tuscany are all comprehensive record-linkage systems, in which drug dispensing 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, Denmark, and the UK and covered by obligatory health insurance in the Netherlands. In all of these countries, general practitioners function as ‘gatekeepers’ of the healthcare system. Data from the different databases were pooled using a distributed network approach, in which data holders maintain control over their protected data, and only aggregated data are shared with the rest of the network. This was carried out through generation of the data into a common format followed by local aggregation using custom-built software, Jerboa[®].¹⁴

Drug exposure

We categorized drug use using the World Health Organization’s (WHO) Anatomical Therapeutic Chemical (ATC) classification system¹⁵ and measured exposure in terms of person-years (PYs). We further analyzed drug use by anatomic groups (ATC first level), by therapeutic category (ATC second level), and by distinct chemical entities (ATC fifth level).

Events

We have chosen in EU-ADR an event-based approach to active drug safety surveillance, focusing only on the most serious and relevant events. These events were chosen from a ranked list of 23 adverse events judged as important in pharmacovigilance on the basis of the following criteria: (i) ‘trigger for drug withdrawal’; (ii) ‘trigger for black box warning’; (iii) ‘leading to emergency department visit or hospital admission’; (iv) ‘probability of event to be drug related’; and (v) ‘likelihood of death’.¹⁶ For this paper, we considered the following top-ranked six events because they vary in occurrence from frequent to rare and thereby represent a broad range of events: (i) acute myocardial infarction; (ii) upper gastrointestinal bleeding; (iii) acute renal failure; (iv) anaphylactic shock; (v) bullous eruptions; and (vi) rhabdomyolysis.

We identified the events of interest in the databases using an iterative process that included definition of events based on clinical criteria established from literature, using diagnosis codes and free text as well as laboratory findings when available. Databases in EU-ADR use one of four nomenclature systems to code the events.¹⁴ To extract the same events across databases, these different terminologies were mapped using the Unified Medical Language System¹, a biomedical terminology integration system handling more than 150 terminologies.¹⁷ The mapping ensured that all events were described using a common language. We inspected differences in event ascertainment by comparing data queries and age-specific and standardized IR of the events (direct standardization was carried out using the WHO World Standard Population).¹⁸ The processes of terminology mapping, harmonization, and benchmarking of event extractions from the various databases have been described in more detail in other publications.^{13,16,19}

Required amount of drug exposure to detect safety signals

Given the pooled population-based IR of six events that were estimated directly within EU-ADR, 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 RR of 2, 4, and 6 using one-sided significance level $\alpha = 0.05$ and a power of 80% ($\beta = 0.2$). To estimate what the required amount of exposure would be for specific strengths of signals to be detected, we used the formula as described in the Appendix.^{20–22} We subsequently determined the number and types of drugs for which there would be sufficient data for safety monitoring. We likewise investigated how correction for multiple comparisons would influence the number of drugs that can be monitored by applying the Bonferroni correction.²³

Simulation: how much power would be gained with an increase in the size of the database network?

Following on the premise that an increase in the size of the database network would translate to an increase in the power to detect safety signals, we performed a simulation to determine how the percentage of drugs that can be monitored would change if more data become available. We calculated the maximum percentage of drugs that could be investigated for 'strong' (i.e., RR=6), 'moderate' (i.e., RR=4), and 'weak' (i.e., RR=2) associations if the EU-ADR platform

were expanded to 10 times its current size, assuming the same patterns of use.

RESULTS

The current population of the EU-ADR network considered for this analysis comprised 19 647 452 individuals with 59 594 132 PYs of follow-up (from 1996 to 2008). A total of 2289 drugs (i.e., distinct chemical entities, ATC fifth level) were used across the network within the study period. An overview of exposure to all drugs, at the anatomical level of the ATC classification and across different age categories, is illustrated in Figure 1. Cardiovascular medications accounted for the largest exposure time (14 110 725 PYs, 38% of exposure time to all drugs), followed by drugs acting on the alimentary tract/metabolism (4 104 330 PYs, 11%), and nervous systems (4 070 445 PYs, 11%). More than half of the exposure time to all drugs came from the elderly population (65 years old and above): 19 564 584 PYs (53%). Children made up 20% of the population but only contributed 3% of total drug exposure time (1 278 639 PYs).

Required amount of exposure for signal detection

The amount of drug exposure necessary to enable detection of 'weak' (RR=2), 'moderate' (RR=4), or 'strong' (RR=6) associations for each of the six events of interest is given in Table 1. As expected, a greater amount of exposure is required to detect an association for events with very low background IR (Figure 2). In the case of rhabdomyolysis (IR=2.5/100 000 PYs), 323 897 PYs of exposure to a particular drug would be necessary to

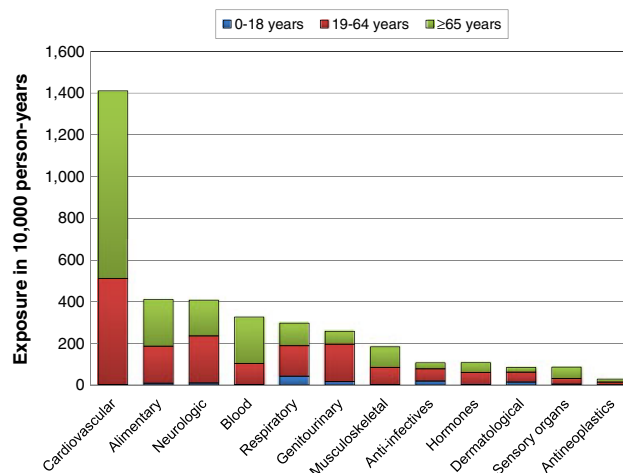


Figure 1. Overview of exposure to all drugs (grouped according to ATC first level) by age categories

Table 1. Amount of drug exposure that would be required to identify potential safety signals of varying strengths concerning six events of interest with varying incidence rates

Event	Incidence rates (<i>per 100 000 person-years</i>)	Relative risk of event to be detected								
		Weak association (relative risk = 2)			Moderate association (relative risk = 4)			Strong association (relative risk = 6)		
		Required exposure (person-years)	No. of drugs that can be investigated (<i>% of total</i>)		Required exposure (person-years)	No. of drugs that can be investigated (<i>% of total</i>)		Required exposure (person-years)	No. of drugs that can be investigated (<i>% of total</i>)	
			Actual *	Simulated [†]		Actual *	Simulated [†]		Actual *	Simulated [†]
Acute myocardial infarction	170.8	4706	531 (23)	1018 (44)	721	939 (41)	1435 (63)	322	1107 (48)	1607 (70)
Upper gastrointestinal bleeding	66.8	12 028	361 (16)	836 (37)	1842	755 (33)	1237 (54)	822	905 (40)	1403 (61)
Acute renal failure	26.4	30 397	221 (10)	638 (28)	4654	534 (23)	1021 (45)	2078	730 (32)	1201 (52)
Bullous eruptions	5.9	135 993	74 (3)	341 (15)	20 823	282 (12)	729 (32)	9297	404 (18)	885 (39)
Anaphylactic shock	5.7	141 939	72 (3)	330 (14)	21 733	272 (12)	723 (32)	9704	396 (17)	877 (38)
Rhabdomyolysis	2.5	323 897	19 (1)	214 (9)	49 593	170 (7)	522 (23)	22 143	270 (12)	717 (31)

*Number (count) of drugs refers to number of distinct ATC codes recorded with adequate exposure in the current EU-ADR database platform; total number of ATC codes recorded in EU-ADR = 2289. Some drugs have multiple ATC codes for different indications of use.

[†]Number (count) of drugs with adequate exposure if the EU-ADR database platform were to be expanded to 10 times its current size.

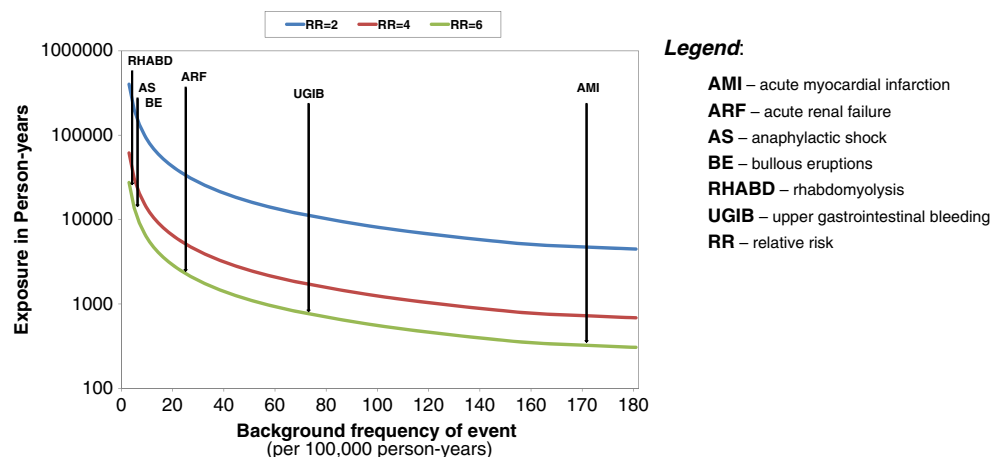


Figure 2. Relationship between background incidence rate and amount of drug exposure required to identify a potential safety signal

detect a 'weak' association, whereas 49 593 PYs and 22 143 PYs would be needed to detect 'moderate' and 'strong' associations, respectively. Exposure requirements for the similarly rare events anaphylactic shock (IR = 5.7/100 000 PYs) and bullous eruptions (IR = 5.9/100 000 PYs) are also given in Table 1. The amount of drug exposure data required to detect associations of certain risk would be correspondingly lower for relatively more frequent events. At least 4706 PYs of drug exposure would be required to identify a safety signal for acute

myocardial infarction (IR = 171/100 000 PYs), whereas at least 12 028 PYs would be required for upper gastrointestinal bleeding (IR = 66.8/100 000 PYs), both at RR of 2. For acute renal failure (IR = 26.4/100 000 PYs), the requisite exposure would be 30 397 PYs to detect a 'weak' association.

Table 1 also shows the number (count) of drugs that can be investigated to detect 'weak,' 'moderate,' and 'strong' associations for the six events, on the basis of the current size of the network. It shows that the

number of drugs for which signals can be detected with enough power increases if the minimal RR that we want to detect increases, which is expected from the power calculations. In acute myocardial infarction, for example, the current EU-ADR network would allow for detection of RR estimates of 2 and above for 531 drugs (23% of total). For 939 drugs, an association with RR of 4 can be detected with enough power and for 1107 drugs, an association with RR of 6 or more can be detected. For the least frequent event rhabdomyolysis, these numbers are 19, 170, and 270 drugs for 'weak,' 'moderate,' and 'strong' associations, respectively. Table 2 shows the exposure requirements and the corresponding number of drugs that can be investigated when multiple comparisons are taken into account using Bonferroni correction (total number of potential signals = 4990). For acute myocardial infarction, the number of drugs for which RR estimates of 2 and above would be detected decreases to 306 (13% of total). For 702 drugs, an association with RR of 4 can be detected with enough power; for 883 drugs an association with RR of 6 or more can be detected. For the least frequent event rhabdomyolysis, these numbers are 1, 63, and 151 drugs for 'weak,' 'moderate,' and 'strong' associations, respectively.

Table 3 shows the total exposure time and number of drugs (distinct ATC fifth level codes, chemical substance) within each main anatomical group in the network. For most of the anatomical main groups,

the percentage of drugs for which we would have 5000 PYs of exposure (which would allow for detection of a 'weak' association for a frequent event such as acute myocardial infarction) or 50 000 PYs of exposure (this amount would allow for detection of a 'moderate' association for a rare event such as rhabdomyolysis) is far below 50% of the total number of drugs in the network. It is also evident from Table 3 that only for the most frequently used classes of drugs is there power to detect safety signals within this range. Furthermore, that percentage is less than 15% for some important drug classes such as antineoplastic and anti-infective agents.

What size is needed?

Data simulation showed that if the EU-ADR platform were to be expanded to 10 times its current size, assuming the same patterns of use, the maximum percentage of drugs that can be investigated for acute myocardial infarction with adequate power is 44% (1018 drugs) at an RR of 2, 54% (1237 drugs) for upper gastrointestinal bleeding at RR of 4, and 9% (214 drugs) for rhabdomyolysis at RR of 2. For strong associations (RR of 6) with acute renal failure, 1201 drugs (52%) would be investigated, whereas less than 40% of all drugs in the database network would be investigated for both anaphylactic shock and bullous eruptions (Table 1). With Bonferroni correction, the maximum

Table 2. Amount of drug exposure that would be required to identify potential safety signals of varying strengths concerning six events of interest with varying incidence rates, taking into account multiple comparisons (Bonferroni correction)

Event	Incidence rates (<i>per 100 000 person-years</i>)	Relative risk of event to be detected								
		Weak association (relative risk = 2)			Moderate association (relative risk = 4)			Strong association (relative risk = 6)		
		Required exposure (person-years)	No. of drugs that can be investigated (<i>% of total</i>)		Required exposure (person-years)	No. of drugs that can be investigated (<i>% of total</i>)		Required exposure (person-years)	No. of drugs that can be investigated (<i>% of total</i>)	
			Actual*	Simulated [†]		Actual*	Simulated [†]		Actual*	Simulated [†]
Acute myocardial infarction	170.8	17 422	306 (13)	765 (33)	2301	702 (31)	1175 (51)	937	883 (39)	1387 (61)
Upper gastrointestinal bleeding	66.8	44 524	184 (8)	542 (24)	5882	487 (21)	972 (42)	2395	691 (30)	1164 (51)
Acute renal failure	26.4	112 520	89 (4)	372 (16)	14 864	323 (14)	795 (35)	6054	485 (21)	965 (42)
Bullous eruptions	5.9	503 410	9 (<i><1</i>)	169 (7)	66 502	146 (6)	466 (20)	27 083	236 (10)	662 (29)
Anaphylactic shock	5.7	525 420	9 (<i><1</i>)	164 (7)	69 410	140 (6)	464 (20)	28 267	231 (10)	652 (28)
Rhabdomyolysis	2.5	1 198 983	1 (<i><1</i>)	83 (4)	158 390	63 (3)	319 (14)	64 505	151 (7)	473 (21)

*Number (count) of drugs refers to number of distinct ATC codes recorded with adequate exposure in the current EU-ADR database platform; total number of ATC codes recorded in EU-ADR = 2289. Some drugs have multiple ATC codes for different indications of use.

†Number (count) of drugs with adequate exposure if the EU-ADR database platform were to be expanded to 10 times its current size.

Table 3. Total exposure time in person-years (PYs) and number of drugs (distinct ATC fifth level codes, chemical substance) within each main anatomical group in the current EU-ADR database network

Anatomical class (ATC I level)	Total exposure in PYs	Total number of drugs [†]	% of drugs with 5000 PYs exposure [‡]	% of drugs with 50 000 PYs exposure [§]
Cardiovascular (C)	14 110 725	271	39	23
Alimentary tract and metabolism (A)	4 104 330	307	21	6
Nervous (N)	4 070 445	320	28	6
Blood (B)	3 262 772	120	21	8
Respiratory (R)	2 968 902	172	24	12
Genitourinary (G)	2 579 467	147	31	9
Musculoskeletal (M)	1 833 924	108	25	12
Anti-infectives (J)	1 067 271	268	14	1
Sensory (S)	860 164	152	14	3
Dermatologicals (D)	845 916	180	18	4
Antineoplastics (L)	285 206	114	10	1
Antiparasitic (P)	86 411	39	8	0
Various (V)	9721	45	0	0
Total	36 085 254	2243	253	85

[†]Total number (count) of drugs within each class recorded with adequate exposure in EU-ADR to detect at least one of the events of interest.

[‡]Exposure of 5000 person-years corresponds to the exposure that would allow for detection of a 'weak' association (i.e., relative risk of 2) with a frequent event such as acute myocardial infarction.

[§]Exposure of 50 000 person-years corresponds to the exposure that would allow for detection of a 'moderate' association (i.e., relative risk of 4) with a rare event such as rhabdomyolysis.

percentage of drugs that can be investigated for acute myocardial infarction with adequate power becomes 33% (765 drugs) at an RR of 2, 42% (972 drugs) for upper gastrointestinal bleeding at RR of 4, and 4% (83 drugs) for rhabdomyolysis at RR of 2. For strong associations (RR of 6) with acute renal failure, 965 drugs (42%) would be investigated, whereas less than 30% of all drugs in the database network would be investigated for both anaphylactic shock and bullous eruptions (Table 2).

Adjusting level of significance in view of exploratory nature of signal detection

As a post hoc analysis, we re-calculated the exposure requirements with a more lenient significance level (i.e., $\alpha=0.10$). The person-time of drug exposure required to detect a signal was predictably lower for all the events, and the increase in number of drugs that may be investigated ranged from 1% to 3% across all magnitudes of RR (Figure 3). This was true for the current database network as well as for the simulated expanded platform (**Supplementary Table 1A**). With correction for multiple comparisons taken into account, the maximum gain in the number of drugs that may be investigated was 1% (**Supplementary Table 2A**).

DISCUSSION

In the new era of active drug safety surveillance using linked electronic healthcare databases, we need to

know how big a database network or system must be to have the power to detect signals of potential safety issues. Database size (variously measured as total population, total follow-up time, or total exposure time to drugs) is important in understanding its capability for meaningful signal detection. The overall size of a database is not the determinant of the statistical power to detect safety signals but rather the drug exposure data.²⁴ Hence, it is not sufficient to know the total number of individuals contributing to the database; we specifically want to determine whether drug exposure in the database is 'big enough' that an adverse effect of such magnitude as to be scientifically significant will also be statistically significant. Our findings show that (i) linking of healthcare databases for active drug safety surveillance is feasible in Europe, although the leverage to do so may be too low for very rare events and for drugs that are infrequently used, or captured, in the databases; and (ii) such a system has more power to detect signals with lower strength of association for the relatively frequent events such as acute myocardial infarction and upper gastrointestinal bleeding. The number (i.e., count) of drugs that can be investigated to detect associations with adverse events at varying magnitudes of RR gives a more quantitative picture of the capability of such a system for safety surveillance: as the strength of the association to be detected goes up, the number of drugs that can be investigated also increases.

We used the classifications 'weak,' 'moderate,' and 'strong' for associations having RR of 2, 4, and 6,

HEALTHCARE DATABASES AND POWER TO DETECT SAFETY SIGNALS

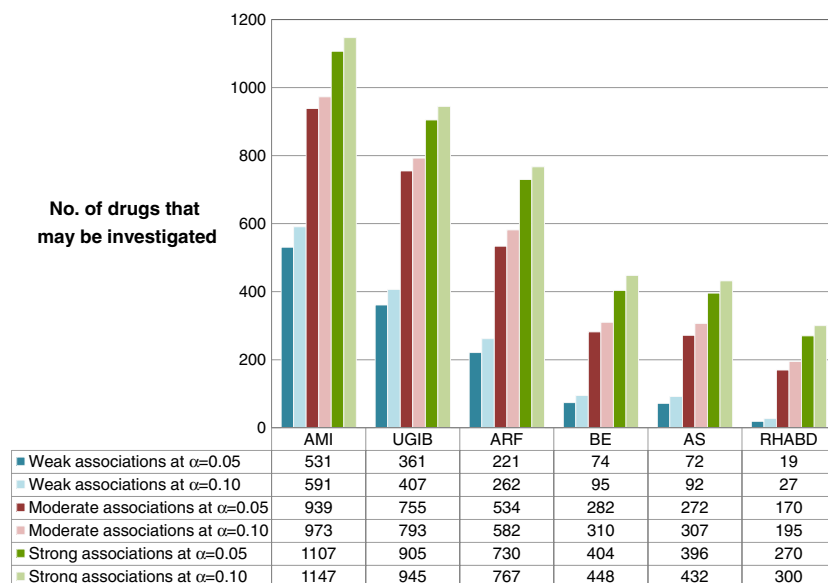
**Legend:****AMI** – acute myocardial infarction**BE** – bullous eruptions**ARF** – acute renal failure**RHABD** – rhabdomyolysis**AS** – anaphylactic shock**UGIB** – upper gastrointestinal bleeding

Figure 3. Change in the number of drugs that may be investigated for signal detection when the acceptable probability of error is increased from $\alpha=0.05$ to $\alpha=0.10$

respectively, which is arbitrary as we have no defined rules for this in epidemiology. To date, there is nothing in the literature that categorically defines which magnitude of risk would be considered strong, weak, or moderate, although, conventionally, epidemiologists consider an association with RR of less than 2.0 a weak association.²⁵

Although relatively frequent in the general population, acute myocardial infarction is an event that is weakly associated with drug use (RR for rofecoxib is 1.43, 95% confidence interval [CI]: 1.16 to 1.76).²⁶ Population-based healthcare data from EU-ADR, however, can still provide risk information for acute myocardial infarction on a considerable number of drugs. A larger amount of exposure is required to detect signals for events having very low IRs such as anaphylactic shock, bullous eruptions, and rhabdomyolysis. This may be less of a problem when such event also happens to be strongly associated with drug use, as is the case for anaphylactic shock and bullous eruptions (i.e., Stevens–Johnson syndrome and toxic epidermal necrolysis). The RR of anaphylactic shock with use of penicillins in hospitalized patients, for example, has been estimated to range from 7.5 (95%CI: 2.9 to 19.6) to 10.7 (95%CI: 3.5 to 32.5),²⁷ whereas the RR of bullous eruptions associated with

use of trimethoprim–sulfamethoxazole and other sulfonamide antibiotics has been estimated to range from 21 (95%CI: 2.3 to 172) to 172 (95%CI: 75 to 396).^{28,29} On the other hand, rhabdomyolysis (despite its known association with statins) is not an event that is usually attributed to use of medications, traumatic injuries, and recreational drug use being more likely etiologies.³⁰ Moreover, rhabdomyolysis itself belongs to a spectrum of conditions ranging from clinically benign myalgia to overt renal failure, and distinguishing among these conditions may not always be easy when mining healthcare databases. The estimated RR for rhabdomyolysis with use of statins is 1.4 (95%CI: 0.45 to 4.4).^{31,32}

It is also clear that use of electronic healthcare databases for safety surveillance will be most powerful for more frequently used classes of drugs. Very few or only very ‘strong’ signals may be detected for drugs that are seldom used or whose use is incompletely captured in the databases, either because they are utilized only in-hospital or in specialized institutions not covered by the databases, or because they are not reimbursed. It is worthwhile to note that although ‘weak’ adverse event associations (e.g., increased cardiovascular risk with pain medications) may not be significant on the individual level, they may have a large

public health impact if a huge proportion of the population is exposed. Non-steroidal anti-inflammatory drugs (NSAIDs), for example, are among the most widely used drugs.

How big should the network be?

Timely detection of safety signals requires a large population that is representative of the entire spectrum of medication users as well as an extensive observation period, particularly for events that are rare or have a long latency.^{25,33} Intuitively, it makes sense to expect that a single database may not have sufficient data to allow for adequate and timely signal detection; there is strength in numbers. Indeed, it has been pointed out that reliance on a single database could reduce statistical power and data diversity.²⁴ One of the presumed benefits of combining international healthcare databases for safety surveillance is the ability to assess exposures to a larger variety of drugs and to characterize use of drugs within a wider range of the population. In a previous publication, we showed that patterns of use varied among different countries but were similar among different databases in the same country.¹⁴ Extending the database network, in terms of both follow-up time and additional databases, is desirable and is becoming more feasible with the establishment of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance.³⁴ If the EU-ADR network would be expanded to 10 times its current size, assuming the same patterns of use, the maximum percentage of drugs that can be investigated for acute myocardial infarction is 44% at an RR of 2, 54% for upper gastrointestinal bleeding at an RR of 4 (RR of upper gastrointestinal bleeding during use of NSAIDs is around 4)³⁵, and 23% for rhabdomyolysis at an RR of 4. Even for events strongly associated with drug use, the maximum percentage would still be less than 50% of the total number of drugs in the network. Thus, it is likely not feasible to monitor the safety of *all* drugs in the databases, regardless of whether the associated adverse events are rare or frequent.

Statistical significance and the exploratory nature of signal detection

A 5% significance level is generally deemed desirable for the purpose of minimizing the probability of observing false positive associations. This value is fairly arbitrary, however. Although this significance level is customarily employed in most epidemiologic studies seeking to confirm drug–event associations, this requirement may not be appropriate in the context of

exploratory signal detection. As a post hoc analysis, we re-calculated the exposure requirements when the acceptable probability of error is lower (i.e., $\alpha = 0.10$). The exposure requirements were predictably reduced for all events, although the increase in the number of drugs that may be investigated was marginal, ranging from 1% to 3% across all magnitudes of RR. This was true for the current database network and for the simulated expanded database platform. With multiple testing already factored in the exposure calculations, there is a very small gain in number of drugs that may be investigated (at most 1% across all events), which indicates that maintaining $\alpha = 0.05$ may be suitable, even for signal detection.

How reliable are the data from the network?

All of the databases in EU-ADR have been widely used for pharmacoepidemiologic research, have well-developed safeguard mechanisms ensuring patient data protection, and have been validated for a variety of drug exposures and clinical outcomes.^{36–45} Because data from medical records databases are derived from physician office records, they capture information of a different nature from that of administrative databases, which are derived from claims for hospitalization and other healthcare services. In an earlier publication, we reported that IRs of upper gastrointestinal bleeding observed in EU-ADR were in line with previously published literature.¹⁴ The databases capture information on outpatient (pharmacy-based) drug use. Drugs utilized in-hospital or in specialized institutions are not registered by the databases, and specific drug groups of interest, such as biologicals (including cancer chemotherapeutic agents and vaccines), are captured only if provided through the routine dispensing system.

Implications for safety surveillance and clinical practice

The driving force behind exploring electronic healthcare databases for active surveillance is the earlier detection, and hence earlier management, of potential safety issues. However, new drugs that may slowly penetrate the market will require a large amount of patient data to comprise a significant user population within a reasonable period.^{46,47} Although it took five years for rofecoxib to be withdrawn from the market, it has been posited that if the medical records of 100 million patients would have been available for safety monitoring, the adverse cardiovascular effect would have been discovered in just three months, given the drug utilization patterns in the US.⁴⁸ On the basis of

utilization patterns in Europe, the EU-ADR network would have had enough exposure data to detect an association between rofecoxib and acute myocardial infarction within two years after being marketed. It is evident that if one aims to create a comprehensive platform that would allow for monitoring the largest number of drugs possible, the greatest benefit would be achieved by linking databases from different countries with diverse utilization patterns. A good way to start would be to foster collaboration with existing initiatives such as EU-ADR, FDA Sentinel, and OMOP.

Although we have raised the issue of statistical power as being important in safety signal detection, this is not to say that the usefulness of a database network is measured only by *how much data* it can provide. It can be argued that especially in the case of longitudinal EHR, there is even greater value in *what kind of data* such networks can provide in terms of estimating incidence of adverse events during drug use and across various population subgroups. These data are, by themselves, important for pharmacovigilance and public health regardless of whether there is sufficient statistical power to demonstrate a difference with non-use of a drug. Particularly for adverse events that are rare or inadequately described in other data sources, if it can be shown that the upper bound of the confidence interval of the incidence estimate is within the limits of what is considered acceptable from a public health perspective, this may be sufficient to downgrade, or exclude, a substantial safety problem.

CONCLUSION

In this paper, we have demonstrated that electronic healthcare data may allow for safety signal detection of a large number of drugs, although the leverage to do so may be low for very rare events and for drugs that are infrequently used, or captured, in the databases. Extending the size of the database network, in terms of both follow-up time and additional databases, may ameliorate this problem.

CONFLICT OF INTEREST

There is no specific product being studied. M.S. is running a research group that occasionally performs studies for pharmaceutical companies according to unconditional grants. The companies include AstraZeneca, Pfizer, Lilly, and Boehringer. M.S. has been consultant

to Lundbeck and Pfizer in the past. G.C. received educational grants from Abbott, GSK, GSK-Bio, Gilead, BMS, Tibotec, and SP-MSD. J.H.C. is director of QRESEARCH, which is a partnership between the University of Nottingham and EMIS (commercial supplier of 56% of GP clinical computer systems in the UK).

KEY POINTS

- There is growing interest in the use of electronic healthcare databases for active drug safety surveillance, but an important concern is whether these databases have the data and enough power to monitor all types of drugs and all types of events.
- Healthcare data from the European EU-ADR database network are used to estimate the number and types of drugs that can be monitored as a function of actual use, minimal detectable relative risk, and empirically derived IRs for six adverse events deemed important in pharmacovigilance. Findings show that the potential for healthcare data-based surveillance systems can be maximized in the detection of safety signals associated with relatively frequent events such as acute myocardial infarction. A greater amount of drug exposure would have to be present in the databases to enable detection of safety signals concerning events with low IRs such as rhabdomyolysis. Only for the most commonly used classes of drugs (e.g., cardiovascular, gastrointestinal, and neurologic medications) is there power to detect safety signals for a 'weak' association with a frequent event (such as acute myocardial infarction) and a 'moderate' association for a rare event (such as rhabdomyolysis).
- Active safety surveillance is feasible using healthcare data-based systems in Europe, although the leverage to do so may be low for drugs that are not frequently used. Extending the EU-ADR database network, both in terms of follow-up time and additional databases from different countries with diverse utilization patterns, may allow for monitoring the largest number of drugs possible.

ACKNOWLEDGEMENT

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

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Supplementary Table 1A. Amount of drug exposure that would be required to identify potential safety signals of varying strengths concerning six events of interest with varying incidence rates (at $\alpha = 0.10$).

Supplementary Table 2A. Amount of drug exposure that would be required to identify potential safety signals of varying strengths concerning six events of interest with varying incidence rates, taking into account multiple comparisons (at $\alpha = 0.10$).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

REFERENCES

- Goldman SA. Limitations and strengths of spontaneous reports data. *Clin Ther* 1998; **20** Suppl C: C40–4.
- Stephenson WP, Hauben M. Data mining for signals in spontaneous reporting databases: proceed with caution. *Pharmacoepidemiol Drug Saf* 2007; **16**(4): 359–65.
- Molokhia M, Tanna S, Bell D. Improving reporting of adverse drug reactions: systematic review. *Clin Epidemiol* 2009; **1**: 75–92.
- The Future of Drug Safety: Promoting and Protecting the Health of the Public. National Academies Press, Washington, DC: 2007. <http://www.nap.edu>. (Accessed 20 March 2011).
- Krumholz HM, Ross JS, Presler AH, et al. What have we learnt from Vioxx? *BMJ* 2007; **334** (7585): 120–3.
- Edwards IR. What are the real lessons from Vioxx? *Drug Saf* 2005; **28**(8): 651–8.
- Härmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. *Eur J Clin Pharmacol* 2008; **64**(8): 743–52. Epub 2008 Jun 4.
- Platt R, Wilson M, Chan KA, et al. The new Sentinel Network—improving the evidence of medical-product safety. *N Engl J Med* 2009; **361**(7): 645–7. Epub 2009 Jul 27.
- The Sentinel Initiative — National Strategy for Monitoring Medical Product Safety. May 2008. <http://www.fda.gov/downloads/Safety/FDAasSentinelInitiative/UCM124701.pdf>. (Accessed 15 April 2011).
- Observational Medical Outcomes Partnership. <http://omop.fnih.org>. (Accessed 15 February 2011).
- Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med* 2010; **153**(9): 600–6.
- About the Drug Safety Effectiveness Network. <http://www.cihir-irsc.gc.ca/39389.html>. (Accessed 01 March 2011).
- Trifirò G, Fourrier-Reglat A, Sturkenboom MC, et al. The EU-ADR project: preliminary results and perspective. *Stud Health Technol Inform* 2009; **148**: 43–9.
- Coloma PM, Schuemie MJ, Trifirò G, et al., on behalf of the EU-ADR consortium. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiol Drug Saf* 2011; **20** (1): 1–11. doi: 10.1002/pds.2053. Epub 2010 Nov 8.
- WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2011. http://www.whooc.no/atc_ddd_index/. (Accessed 14 February 2011).
- Trifirò G, Pariente A, Coloma PM, et al., on behalf of the EU-ADR consortium. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf* 2009; **18**(12): 1176–84.
- Unified Medical Language System. http://www.nlm.nih.gov/research/umls/about_umls.html. (Accessed 28 February 2011).
- Age standardization of rates: a new WHO standard. <http://www.who.int/healthinfo/paper31.pdf>. (Accessed 20 February 2011).
- Avillach P, Mouglin F, Joubert M, et al. 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**: 190–194.
- Beaumont JJ, Breslow NE. Power considerations in epidemiologic studies of vinyl chloride workers. *Am J Epidemiol* 1981; **114**(5): 725–34.
- Breslow NE, Day NE. Volume II—The Design and Analysis of Cohort Studies. Statistical Methods in Cancer Research. Lyon: IARC Scientific Publications, No. 82. 1987. <http://www.iarc.fr/en/publications/pdfs-online/stat/>. (Accessed 20 March 2011).
- Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials*. 1990; **11**(2): 116–28.
- Altman D. *Practical Statistics for Medical Research*. Boca Raton: Chapman and Hall London, 1991. pp.210–212.
- Hammond IW, Gibbs TG, Seifert HA, et al. Database size and power to detect safety signals in pharmacovigilance. *Expert Opin Drug Saf* 2007; **6**(6): 713–721.
- Strom BL, Kimmel SE (eds.). *Textbook of Pharmacoepidemiology*. John Wiley & Sons: Chichester, 2006.
- Ross JS, Madigan D, Hill KP, et al. Pooled analysis of rofecoxib placebo-controlled clinical trial data: lessons for postmarket pharmaceutical safety surveillance. *Arch Intern Med* 2009; **169** (21): 1976–85.
- van der Klauw MM, Stricker BH, Herings RM, et al. A population based case-cohort study of drug-induced anaphylaxis. *Br J Clin Pharmacol* 1993; **35** (4): 400–8.
- Mockenhaupt M, Viboud C, Dunant A, et al. Stevens–Johnson syndrome and toxic epidermal necrolysis: assessment of medication risks with emphasis on recently marketed drugs. The EuroSCAR-study. *J Invest Dermatol* 2008; **128** (1): 35–44. Epub 2007 Sep 6.
- Roujeau JC, Kelly JP, Naldi L, et al. Medication use and the risk of Stevens–Johnson syndrome or toxic epidermal necrolysis. *N Engl J Med* 1995; **333** (24): 1600–7.
- Linares LA, Golomb BA, Jaojoco JA, et al. The modern spectrum of rhabdomyolysis: drug toxicity revealed by creatine kinase screening. *Curr Drug Saf* 2009; **4** (3): 181–7. Epub 2009 Sep 1.
- Thompson PD, Clarkson P, Karas RH. Statin-associated myopathy. *JAMA* 2003; **289** (13): 1681–90.
- Joy TR, Hegge RA. Narrative review: statin-related myopathy. *Ann Intern Med* 2009; **150** (12): 858–68.
- Meyboom RH, Egberts AC, Gribnau FW, et al. Pharmacovigilance in perspective. *Drug Saf* 1999; **21**(6): 429–47.
- European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. <http://www.encepp.eu>. (Accessed 01 May 2011).
- Hernandez Diaz S, Garcia Rodriguez LA. Association of nonsteroidal anti-inflammatory drugs and upper gastrointestinal bleeding/perforation. *Arch Intern Med* 2000; **160** (14): 2093–2099.
- Vlug AE, van der Lei J, Mosseveld BM, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf Med* 1999; **38**(4–5): 339–44.
- Filippi A, Vanuzzo D, Bignamini AA, et al. The database of Italian general practitioners allows a reliable determination of the prevalence of myocardial infarction. *Ital Heart J* 2005; **6**(4): 311–4.
- Filippi A, Bianchi C, Parazzini F, et al. A national survey on aspirin patterns of use and persistence in community outpatients in Italy. *Eur J Cardiovasc Prev Rehabil* 2011 Mar 1. [Epub ahead of print].
- Sturkenboom M, Nicolosi A, Cantarutti L, et al., NSAIDs Paediatric Research Group. Incidence of mucocutaneous reactions in children treated with niflumic acid, other nonsteroidal antiinflammatory drugs, or nonopioid analgesics. *Pediatrics* 2005; **116**(1): e26–33. Epub 2005 Jun 1.
- Hippisley-Cox J, Stables D, Pringle M. QRESEARCH: a new general practice database for research. *Inform Prim Care* 2004; **12**(1): 49–50.
- Ehrenstein V, Antonsen S, Pedersen L. Existing data sources for clinical epidemiology: Aarhus University Prescription Database. *Clin Epidemiol* 2010; **2**: 273–9.
- Corrao G, Zambon A, Conti V, et al. Menopause hormone replacement therapy and cancer risk: an Italian record linkage investigation. *Ann Oncol* 2007; **19**: 150–155. DOI: 10.1093/annonc/mdm404; Published online: 4 September 2007.
- Barchielli A, Balzi D, Pasqua A, Buiatti E. Incidence of acute myocardial infarction in Tuscany, 1997–2002: data from the Acute Myocardial Infarction Registry of Tuscany (Tosc-AMI). *Epidemiol Prev* 2006; **30**(3): 161–168.
- van den Boogaard CH, Breeckveldt-Postma NS, Borggreve SE, Goettsch WG, Herings RM. Persistent bisphosphonate use and the risk of osteoporotic fractures in clinical practice: a database analysis study. *Curr Med Res Opin* 2006; **22**(9): 1757–64.
- Corrao G, Romio SA, Zambon A, Merlino L, Bosi E, Scavini M. Multiple outcomes associated with the use of metformin and sulphonylureas in type 2 diabetes: a population-based cohort study in Italy. *Eur J Clin Pharmacol* 2011; **67**(3): 289–99. Epub 2010 Nov 19.
- Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005; **58** (4): 323–337.
- Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects—advantages and disadvantages. *Nat Clin Pract Rheumatol* 2007; **3**(12): 725–732.
- McClellan M. Drug safety reform at the FDA—pendulum swing or systematic improvement? *N Engl J Med* 2007; **356** (17): 1700–2. Epub 2007 Apr 13.

APPENDIX

Sample size calculations were carried out within the context of a cohort study using a hypothesis testing approach. The approximate power formula assumes that the observed number of adverse events resulting from exposure to the drug of interest follows a Poisson distribution with mean RE , where E is the expected number of events based on background population rates and R is the relative risk.^{20–22} It follows that for large values of RE , the square root of the observed events is approximately normally distributed with mean $\sqrt{RE}$ and variance $1/4$. Now if a random quantity

has a normal distribution with mean μ and variance σ^2 , the power $(1 - \beta)$ of a test to reject the null hypothesis $\mu = \mu_0$, given that the alternative is true, at the α level of significance is given by

$$Z_{1-\beta} = Z_\alpha - \{(\mu - \mu_0)(1/\sigma)\}$$

Substitution of $\mu = \sqrt{RE}$, $\mu_0 = \sqrt{E}$, and $\sigma = 1/2$ into this expression yields the following:

$$Z_{1-\beta} = Z_\alpha - 2(\sqrt{R} - 1)(\sqrt{E}).$$