MINING ELECTRONIC HEALTHCARE RECORD DATABASES

TO AUGMENT DRUG SAFETY SURVEILLANCE

Preciosa Miguel Coloma

ISBN: 978-94-6191-258-9
Cover design and layout by Catherine Coloma
Layout by Legatron Electronic Publishing, Rotterdam
Printing by Ipskampdrukkers, Enschede

The work in this thesis was conducted at the Department of Medical Informatics at the Erasmus MC, University Medical Center, Rotterdam, the Netherlands.

The partners of the EU-ADR Consortium and the contributors to the EU-ADR network of databases are gratefully acknowledged. The EU-ADR Project was funded by the European Commission under the Seventh Framework Programme.

Financial support for the publication of this thesis was provided by the Erasmus University Rotterdam and the Interdisciplinary Processing of Clinical Information (IPCI, formerly Integrated Primary Care Information) Project.

MINING ELECTRONIC HEALTHCARE RECORD DATABASES TO AUGMENT DRUG SAFETY SURVEILLANCE

AUTOMATISCH ANALYSEREN VAN DATABANKEN MET ELEKTRONISCHE GEZONDHEIDSZORGDOSSIERS TER VERSTERKING VAN GENEESMIDDELENBEWAKING

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag, 16 mei 2012 om 15.30 uur

door

Preciosa Miguel Coloma
geboren te Laoag (Filippijnen)

Promotiecommissie

Promotor:	Prof.dr. M.C.J.M. Sturkenboom
Overige leden:	Prof.dr. B.H.Ch. Stricker Prof.dr. E.M.E.H Lesaffre Prof.dr. H.G.M. Leufkens
Copromotoren:	Dr. G. Trifirò Dr. M.J. Schuemie

For my Dad

CONTENTS

Chapter 1

General Introduction: About this thesis

It is well known that placebos – and even no treatment – can be associated with adverse events.

INTRODUCTION

It is perhaps a fundamental truth in medicine that there is no intervention – be it a drug, a medical device or a procedure – that is without risks. Even with the most rigorous efforts in drug approval and regulation, there is not a drug out there that is 100% safe under all conditions. Randomized controlled trials (RCTs) are considered to be the most stringent approach to determining cause-and-effect relationship between an intervention and an outcome, but such trials are rarely designed or powered to detect uncommon or unexpected adverse events.[1-4] Once drugs are marketed, they are used in a more diverse group of people, often for much longer periods, and sometimes with a wider range of therapeutic indications. While monitoring the risks associated with drug use has come a long way since the thalidomide disaster in the 1960s and the institution of spontaneous reporting systems (SRS), it has become evident that adverse effects of drugs may be detected - and acted upon – too late, when millions of persons have already been exposed.[5-6]

There has been a growing clamor for improving the current passive-reactive paradigm of drug safety surveillance. Prominent issues in the last few years have emphasized the importance of a life-cycle approach to drug safety monitoring and the need to explore new methods to improve surveillance of drugs post-marketing.[7-8] It has been posited that electronic healthcare record (EHR) databases represent an important resource for proactive surveillance and can augment existing pharmacovigilance systems. Various public-private initiatives worldwide have been launched, and great investments have been made, to explore the secondary use of EHR for this purpose.[9-10] In this thesis, we draw on the experience of the EU-ADR network (**Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge**, http://www.euadr-project.org/), a federation of eight EHR databases in four countries in Europe, to demonstrate the feasibility of combining diverse and differently structured data and pave the way for large-scale drug safety monitoring. We describe the opportunities and challenges that come with heterogeneity in database structure, with differences in language and coding of both drugs and diseases, and with the diversity in the organization of European healthcare systems.

ABOUT THIS THESIS

In **Chapter 2**: *The Big Picture*, we provide an overview of ongoing initiatives exploring data from EHR for automatic signal detection vis-à-vis established SRS. We describe the role SRS has played in regulatory decision-making with respect to safety issues and further evaluate the potential added value of EHR-based signal detection systems to the current practice of drug safety surveillance.

In **Chapter 3**: *Breaking New Grounds,* we set the stage for large-scale monitoring of drug safety signals using multiple EHR databases. We start off with the identification of events that are considered important to monitor from a pharmacovigilance and public health perspective **(3.1)**. We describe the process of harmonizing event data extraction from different databases, the aim of

which is to come up with a common definition of events that is both clinically sound and agreeable to all stakeholders **(3.2)**. We then present a proof-of-concept study that describes the framework, distributed data processing, and preliminary results of combining data from eight European EHR databases for drug safety signal detection. This is the first study of its kind that explores the methodological, cultural, ethical, governance and political issues of sharing healthcare data across international borders **(3.3)**.

In **Chapter 4**: *Facing up to the Challenge*, we put into perspective the expectations about what these healthcare data-based surveillance systems can do. We demonstrate how much leverage EHR databases can provide for monitoring the safety of medicines. We provide estimates of the number and types of drugs that can be monitored in these systems as a function of actual drug use, minimal detectable relative risk, and incidence rates of outcomes of interest **(4.1)**. We further perform data simulation in an attempt to see if, and how, expansion of database size would make a difference in the capabilities of the system. Because the prevalence and nature of many diseases are not the same in children as in adults and because of age-related variation in drug pharmacology and drug utilization,[11-13] we performed a separate study to determine the capabilities of the system for pediatric drug safety signal detection **(4.2)**. Although several drugs have been withdrawn post-marketing by regulatory authorities after scientific evaluation of harms and benefits, there is no definitive list of confirmed signals (i.e., list of all known adverse reactions and which drugs can cause them). As there is no true gold standard, prospective evaluation of signal detection methods remains a challenge. We propose a surrogate reference standard of drug-adverse event associations based on existing published scientific literature, drug product labels, and expert opinion and developed for the primary purpose of evaluating performance of methods for signal detection using EHR **(4.2)**. We then evaluate the relative performance of 10 different statistical methods for detecting drug-adverse event associations in EHR data **(4.3)**.

While the motivation for merging disparate data sources primarily comes from the need to investigate drug safety in larger populations, it is a well-acknowledged limitation that EHR databases are only able to describe exposures and outcomes of interest to the extent that they are documented within the database systems.[14-15] We evaluate the positive predictive value of various coding-based algorithms in the identification of upper gastrointestinal bleeding (UGIB) and acute myocardial infarction (AMI) in different EHR databases by case validation and review of hospitalization charts and general practitioner (GP) records **(4.4, 4.5)**. Responding to a request of the Dutch Medicines Agency, which is currently facing drug safety problems with progressive multifocal leukoencephalopathy (PML) following the use of some biological agents, we estimated the incidence of PML in the general population using EHR data from three countries in Europe (**4.6**).

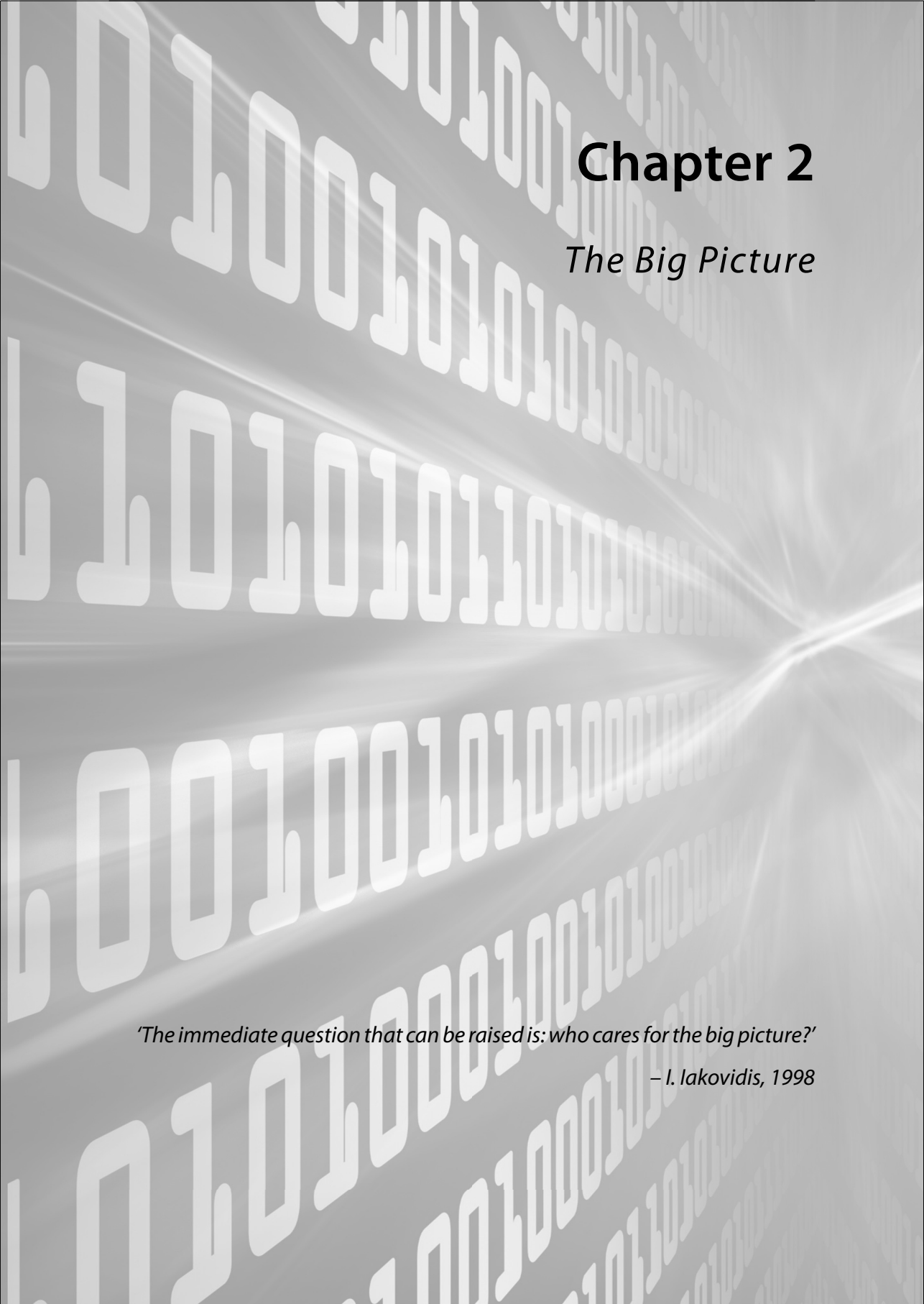
Like any signal detection system, there is a need to establish 'rules' how to trigger an alert, when to consider a signal likely enough to be real to warrant follow-up, and when a signal needs to be elevated to represent a potential safety risk. In **Chapter 5**: *Separating the Big Fish from the Small Fry*, we describe the process of evaluation and triage of potential causal drug-adverse event associations derived from signal detection using EHR databases. Taking the event acute myocardial

infarction (AMI) as an example, we propose a strategy for combining evidence from different data sources in the EU-ADR system to prioritize signals that may represent genuine risk and, hence, necessitate further investigation and formal pharmacoepidemiologic studies.

Finally, a summary discussion and future perspectives are presented in **Chapter 6**.

REFERENCES

1. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312(7040): 1215-1218.
2. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006;74:635–641.
3. Stricker BH, Psaty BM. Detection, verification, and quantification of adverse drug reactions. *BMJ* 2004 Jul 3; 329(7456):44-47.
4. Ray W. Population-based studies of adverse drug effects. *New Engl J Med.* 2003;349(1592-1594).
5. Avorn J. Evaluating drug effects in the post-Vioxx world: there must be a better way. *Circulation* 2006 May 9; 113(18):2173-2176.
6. Topol E. Failing the public health--rofecoxib, Merck, and the FDA. *N Engl J Med* 2004 Oct 21;351(17):1707-1709.
7. Institute of Medicine. The Future of Drug Safety - Promoting and Protecting the Health of thePublic. 2006. http://www.iom.edu/Reports/2006/The-Future-of-Drug-Safety-Promoting-and-Protecting-the-Health-of-the-Public.aspx. Last accessed 20 October 2011.
8. Krumholz HM, Ross JS, Presler AH, Egilman DS. What have we learnt from Vioxx? *BMJ* 2007 Jan 20; 334(7585):120-123.
9. 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 Aug 13;361(7):645-647.
10. Stang PE, Ryan P, 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(9):600-606.
11. Atuah KN, Hughes D, Pirmohamed M. Clinical pharmacology: special safety considerations in drug development and pharmacovigilance. *Drug Saf* 2004; 27(8): 535-554.
12. Klassen TP, Hartling L, Craig JC, Offringa M. Children are not just small adults: the urgent need for high-quality trial evidence in children. *PLoS Med.* Aug 12 2008;5(8):e172.
13. Napoleone E. Children and ADRs (Adverse Drug Reactions). *Ital J Pediatr.* 2010;36:4.
14. 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.
15. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects – advantages and disadvantages. *Nat Clin Pract Rheum* 2007;3(12):725-732.



Chapter 2

The Big Picture

'The immediate question that can be raised is: who cares for the big picture?'

– I. Iakovidis, 1998

A review of post-marketing drug safety surveillance systems: where automatic signal detection using Electronic Healthcare Records (EHR) fits into the big picture

Preciosa M Coloma, MD, MSc,[1] Gianluca Trifirò, MD, PhD,[1,2]
Vaishali Patadia,[1,3] and Miriam Sturkenboom, PharmD, PhD[1,4]

[1]Department of Medical Informatics, Erasmus Medical Center, Rotterdam, the Netherlands;
[2]Department of Clinical and Experimental Medicine and Pharmacology,
Section of Pharmacology, University of Messina, Messina, Italy;
[3]Astellas Pharma, Deerfield, Illinois, USA;
[4]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands

Submitted

ABSTRACT

The safety profile of a drug evolves over its lifetime on the market; there are bound to be changes in the circumstances of a drug's clinical use which may give rise to previously unobserved adverse effects, hence necessitating post-marketing surveillance. Post-marketing surveillance has traditionally been carried out by systematic manual review of spontaneous reports of adverse drug reactions (ADRs). Vast improvements in computing capabilities have provided opportunities to automate signal detection and several worldwide initiatives are exploring new approaches to earlier signal detection, primarily through mining of routinely-collected data from electronic healthcare records (EHR). This paper provides an overview of ongoing initiatives exploring data from EHR for automatic signal detection vis-à-vis established spontaneous reporting systems (SRS). We describe the role SRS has played in regulatory decision-making with respect to safety issues and evaluate the potential added value of EHR-based signal detection systems to the current practice of drug surveillance. Safety signal detection is both an iterative and dynamic process. It is in the best interest of public health to integrate and understand evidence from all possibly relevant information sources on drug safety. Proper evaluation and communication of potential signals identified remains an imperative and should accompany any signal detection activity.

INTRODUCTION

A drug's efficacy and safety must be demonstrated in a series of clinical trials conducted prior to approval. Phase III studies, consisting of randomized controlled trials (RCTs), are considered to be the most rigorous approach to determining cause-and-effect relationship between an intervention and an outcome. The controlled nature of such trials, however, calls for a limited number of patients who may not always be representative of the population of all potential users of the drug and a relatively short observation period, making it difficult to detect adverse drug reactions (ADRs) that are rare or with a long latency.[1-4] Hence, to protect public health, it is imperative to continue monitoring and evaluating the safety of a drug once it is on the market. The safety profile of a drug evolves over its lifetime on the market; after years, or even decades, of experience there are bound to be changes in the circumstances of a drug's clinical use (in the population for whom it is recommended, including off-label use, concomitant use with other drugs, and dosing regimen changes) which may give rise to previously unobserved adverse effects. Even over-the-counter products that have been available for a long time such as phenylpropanolamine and the nonsteroidal anti-inflammatory drugs (NSAIDs) have been found to be associated with adverse effects necessitating labeling changes several years after drug approval, or even market withdrawal.[5-8]

Post-marketing drug safety surveillance has traditionally been carried out by systematic manual review of reports of suspected ADRs sent by healthcare professionals, consumers, and pharmaceutical manufacturers and registered in national pharmacovigilance database systems. Qualitative review of all reports has become progressively more difficult and impractical because of the exponential increase in the number of cases over the years as well as the continuous influx of new drugs. In addition, vast improvements in computing capabilities in the last few decades have provided an opportunity to automate signal detection. For this reason, quantitative and automatic methods have been developed to supplement qualitative clinical evaluation, with quantitative signal detection being performed mostly, although not exclusively, on databases of spontaneous ADR reports.[9-13] Systems employing active ascertainment of adverse events related to specific drugs of interest have likewise been used for signal detection; these include the Prescription Event Monitoring systems in the United Kingdom and its counterpart in New Zealand.[14-15] More recently, high-profile safety issues such as those involving rofecoxib and rosiglitazone have stimulated initiatives in North America and Europe to explore new approaches to facilitate earlier signal detection, primarily through mining of routinely-collected, longitudinal data from electronic healthcare records (EHR), including medical records and claims for healthcare services.[16-17]

What constitutes a 'signal?'

The concept of a signal, from a drug surveillance point of view, has evolved from its definition by the World Health Organization (WHO) in 2002[18] to a more synthesized and comprehensive definition proposed by Hauben and Aronson:[19] (1) It is based on information from one or more sources (including observations and experiments), suggesting an association (either adverse or

beneficial) between a drug or intervention and an event or set of related events (e.g., a syndrome); (2) It represents an association that is new and important, or a new aspect of a known association, and has not been previously investigated and refuted; and (3) It demands investigation, being judged to be of sufficient likelihood to justify verificatory and, when necessary, remedial actions. It is thus evident that a signal in pharmacovigilance may – and will – arise from various data sources.

In this paper we provide an overview of ongoing initiatives exploring data from EHR for automatic signal detection vis-à-vis established spontaneous reporting systems (SRS). We describe the role SRS has played in regulatory decision-making with respect to safety issues. We further evaluate the potential added value of EHR-based signal detection systems to the current practice of drug safety surveillance.

Traditional Approach: Spontaneous Reporting

In the aftermath of the thalidomide tragedy in the late 1960s, the United States (US) Food and Drug Administration (FDA), the WHO, and the United Kingdom's Medicines and Healthcare products Regulatory Agency (MHRA) independently set up voluntary reporting systems geared towards the collection, and subsequent analysis, of post-marketing safety information. Establishment of other country-wide spontaneous reporting databases soon followed. More than 70 countries, including a number of developing countries, have their own SRS, which attempt to ensure that signals of possible ADRs are detected as soon as possible after licensing. Some of the largest SRS data repositories are briefly described below.

FDA – Adverse Event Reporting System (AERS) database

FDA's Adverse Event Reporting System (AERS) was established in 1969 to support the FDA's post-marketing safety surveillance program and currently contains over 4 million reports.[20] Manufacturers of prescription drugs are obliged to submit adverse event reports to the FDA while healthcare professionals and the public voluntarily report serious reactions and other problems regarding drugs and medical devices via the MedWatch program.[21] All data from MedWatch are entered into the AERS database. Reports submitted by manufacturers consist of post-marketing 15-day Alert reports (containing reports of adverse drug experiences that are both serious and unexpected, whether foreign or domestic, as soon as possible but in no case later than 15 calendar days of initial receipt of the information), follow-up of these 15-day Alert reports, and Periodic reports for new molecular entities (NME) within the first three years following FDA approval.[22] Although the database contains reports from worldwide sources, the majority (66%) come from the US. From 2000 to 2010, AERS received an average of 300 000 reports per year. An updated quarterly list of potential signals of serious risks or new safety information identified from AERS is available from the FDA website.[23] To preserve patient confidentiality, the publicly available AERS data do not include narratives which could be identifiable. However, links to other post-marketing information, including FDA communication to patients and healthcare providers, are accessible.

FDA Vaccine Adverse Event Reporting System (VAERS)

Similar to AERS, the Vaccine Adverse Event Reporting System (VAERS) is a nationwide post-marketing safety surveillance program, this time collecting information about adverse events occurring after the administration of vaccines licensed for use in the US. VAERS is jointly managed by the FDA and the Centers for Disease Control and Prevention (CDC) and reports come from healthcare providers, manufacturers, and the public.[24] VAERS receives around 30 000 reports annually, with 13% classified as serious (i.e., associated with hospitalization, disability, life-threatening illness, or death). Since its inception in 1990, VAERS has received over 200 000 reports, most of which (85-90%) describe mild adverse events such as fever and local reactions at injection site, but important signals about possible vaccination-related adverse events have also been identified.[25]

VigiBase

VigiBase is the data repository of the World Health Organization (WHO) Programme for International Drug Monitoring and is managed by the Uppsala Monitoring Centre (UMC) in Uppsala, Sweden. VigiBase is a relational database with tools for import, maintenance, and analysis of summary clinical reports on individual case safety reports (ICSRs) received from over 100 WHO member countries since 1968. It currently contains about 5.5 million reports.[26] A majority of the ICSRs are received from Europe and the US. In recent years some countries with smaller populations, such as New Zealand, Ireland, Switzerland, and the Nordic/Scandinavian countries, have taken a more prominent role in reporting.[27] ICSRs in VigiBase come from both regulatory and voluntary sources, depending on the national pharmacovigilance system. Some national centers (NCs) accept reports only from medical practitioners while others accept reports from other healthcare professionals. Some NCs include reports from pharmaceutical companies, while others do not. The UMC does not take part in any clinical review of case reports before they are entered in VigiBase; some NCs make an assessment of the likelihood that a drug caused the suspected reaction while others do not document such assessments on individual reports. Contributing NCs receive all regular data output from UMC, have access to the web-based tool VigiSearch to access the data, and can request database searches made by UMC. A separate tool – VigiFlow – is also available for NCs to manage their ICSRs and simplifies the process of sending ICSRs to VigiBase. Other stakeholders, including regulatory bodies apart from the reporting NC, the pharmaceutical industry, and academicians can request database searches made by UMC - this service is available to anyone with a legitimate interest in pharmacovigilance data.[27] An accompanying 'Caveat' document available on the UMC website sets out the guidelines for interpretation of the data obtained.[28] Although the data in VigiBase are more heterogeneous than in national pharmacovigilance databases because of (among other things) different healthcare practices and regulatory protocols, VigiBase presents opportunities to identify and analyze differences among countries or regions.

EudraVigilance

EudraVigilance, created by the European Medicines Agency (EMA) in 2001, is a centralized database containing reports of suspected ADRs to drugs licensed across the European Union (EU). Reports come from regulatory agencies of EU member states (national competent authorities, NCA) and from pharmaceutical companies (marketing authorization holders, MAH).[29] The system has been extended to allow commercial and non-commercial sponsors to report suspected unexpected serious adverse reactions (SUSARs) occurring during clinical trials. With the recent amendments to pharmacovigilance legislation, all ADRs (not just serious ones) are to be reported and NCAs and MAHs are directed to accept ICSRs sent to them by patients, caregivers, and consumers as well as healthcare practitioners. A list of drugs that are subject to additional monitoring, as well as other safety issues, shall be made publicly available via an EU safety portal.[30] The EMA has developed an Access Policy to provide stakeholders such as regulatory authorities, healthcare professionals, patients and consumers, as well as the pharmaceutical industry and research organizations with certain levels of access to suspected ADRs reported to EudraVigilance. This proactive access to the EudraVigilance data is envisioned to support signal detection and safety evaluation activities, inform healthcare professionals and the general public, and allow for the use of ADR data for research purposes.[31]

Signal detection in SRS

SRS employ data mining methods, primarily disproportionality analyses, to flag potential safety signals. The disproportionality analyses (which can be frequentist or Bayesian in approach) are based on statistical algorithms that detect drug-adverse event combinations occurring at higher than expected frequencies.[32-33] Techniques such as proportional reporting ratios (PRR, used in EudraVigilance) compare the proportion of events reported for a particular drug within a database with the background proportion for that same event for all drugs in the database.[34] Another method is the Reporting Odds Ratio (ROR), which is a reformulation of the PRR as an odds ratio.[35] The Multi-Item Gamma Poisson Shrinker (MGPS, used in the AERS)[9, 36] and the Bayesian Confidence Propagation Neural Network (BCPNN, used in Vigibase)[37] also examine disproportionality of reports for a specific drug compared to all other exposures, but draw on Bayesian models to shrink estimates of risk. Such methodologies have likewise been employed to assess time trends and drug-drug interactions.[10]

SRS gather real-life data on marketed drugs and, when review of individual case reports or case-series analysis is possible, may permit the identification of potential safety concerns. Examples of signals that have been generated or reinforced through SRS include hemolytic anemia associated with Temafloxacin, ventricular arrhythmias with Terfenadine and Cisapride, and cardiac valvulopathy with fenfluramine.[38-41] In addition, such reports have been useful in defining the nature of some ADRs. An understanding of factors involved in flucloxacillin-induced hepatitis, such as delayed time to onset, predominant cholestatic pattern and delayed recovery, were brought to light by ADR reports.[42-43] The delayed onset and typically cholestatic pattern of amoxicillin/clavulanic acid-induced hepatitis has likewise been recognized through such reports.[44-45] Higher

than expected reports of intussusception following administration of the RotaShield rotavirus vaccine were initially identified in VAERS in 1999.[46-47] The vaccine was voluntarily removed from the market by the manufacturer following the finding of an increased risk in epidemiologic studies.[48-49] The potential risk for development of Guillain-Barre syndrome (GBS) after administration of a meningococcal conjugate vaccine was first observed in VAERS.[50-51]

Despite their proven usefulness, there are several limitations for the use of SRS, primarily because SRS are mostly voluntary and studies have shown that only about 10% of serious adverse events are reported.[52] Underreporting can lead to protracted delays between marketing and discovery – and subsequent regulatory action – of an ADR. Close to 7 million patients were exposed to fenfluramine before the association with valvular heart disease led to its withdrawal from the market.[53] More than 80 million people worldwide (nearly 107 million prescriptions dispensed in the US alone) have been exposed to rofecoxib before it was voluntarily withdrawn by the manufacturer.[54-55] Case reports in SRS may not always be consistent or complete with respect to medical history or comorbidities and data quality varies by region, country, and reporting individual (i.e., consumer vs. healthcare professional) (see Table 1). SRS databases generally do not have exposure information and are therefore deficient in providing a true incidence rate of an event.[56-57] Furthermore, the phenomenon of masking has been shown to potentially cause signals of disproportionate reporting to be missed.[58] Because of the serious limitations in the use of SRS, some of the signals identified may be false positives, but before the signal is discarded important efforts need to be made, including performing expensive and time-consuming signal verification and hypothesis confirmation studies.

Active ascertainment: Prescription event monitoring

While SRS rely on passive data collection, there are few surveillance systems that undertake proactive safety monitoring on selected drugs in the early post-marketing period. Drugs of interest are usually those with potential safety issues identified during development, new chemical entities, or drugs belonging to a class previously known to cause problems. In the UK, reports of suspected ADRs are actively solicited through the Prescription Event Monitoring system, which surveys physicians regarding any adverse experiences among the first 10 000 patients prescribed a particular drug.[59] The Intensive Medicines Monitoring Programme in New Zealand likewise actively solicits reports of suspected ADRs, establishing cohorts of approximately 10 000 patients monitored for a mean of almost five years.[60] Prescription follow-up information provides a denominator for calculating event rates and the numerator originates from intensified reporting. Reporting rates are, hence, much higher than with traditional voluntary reporting although the degree of underreporting varies between systems and products and is difficult to ascertain. Some examples of adverse events identified by prescription event monitoring include cough with captopril and agranulocytosis with mianserin.[61-62]

Table 1 | Description of main SRS data sources.

Database	Current number of reports available†	Average number of reports received per year	Catchment period	Source of reports	Participating Countries
FDA AERS	>4 million	300 000	1969-present	Patients/consumers, Healthcare professionals, Pharmaceutical companies	Mainly US (≈66%), other countries
FDA VAERS	200 000	30 000	1990-present	Patients/consumers, Healthcare professionals, Pharmaceutical companies	US
WHO-Vigibase	~5.5 million	200 000	1968-present	National pharmacovigilance centers (which may receive reports directly from Patients/consumers, Healthcare professionals, or Pharmaceutical companies)	106 official member countries and 34 associate members
Eudravigilance	Unknown	Unknown	2001-present	National competent authorities and Marketing authorization holders (soon to include direct reports from Patients/consumers and Healthcare professionals)	EU Member States

† As of October 2011

Complementary Approach: Mining Safety Signals from EHR

From the foregoing discussion it can be surmised that the greatest limitation in the current approach to safety surveillance is that most hitherto existing systems are passive and reactive. The imperative to shift the paradigm towards a more proactive approach has resulted in the exploration of existing and accessible data resources, whether or not the data are collected for the primary purpose of drug safety monitoring.[63-64] These potential resources include electronic medical records with detailed clinical information such as patients' symptoms, physical examination findings, diagnostic test results, and prescribed medications or other interventions. Automated electronic recording of filled prescriptions, laboratory and ancillary tests, as well as hospitalizations, are increasingly collected routinely for the payment and administration of health services. These EHR databases – medical records databases and administrative/claims databases – have been employed to characterize healthcare utilization patterns, monitor patient outcomes, and carry out formal pharmacoepidemiologic studies.[65-68] With regards to drug safety surveillance, such databases have been commonly used to confirm or refute potential signals detected initially by SRS, including vaccine-related signals.[69] EHR databases reflect practical clinical data culled from real-world settings. Being routine by-products of the healthcare delivery system, the use of these databases offers the advantage of efficiency in terms of time necessary to conduct a study, manpower, as well as financial costs.

There have been several efforts in recent years to evaluate the usefulness of EHR databases for drug safety signal detection. The WHO UMC adapted the BCPNN to the UK primary care database IMS Disease Analyzer MediPlus to show how longitudinal data may facilitate early signal detection.[70] Another study assessed the feasibility of using the MGPS algorithm to Medicare claims data in order to evaluate adverse outcomes associated with cyclooxygenase-2 inhibitors (coxibs).[71] A team of Danish investigators performed temporal data mining on EHR databases to evaluate adverse events potentially related to the measles mumps rubella (MMR) vaccine.[72] Employing traditional epidemiologic methods (nested case control analysis and self-controlled case series), the Meningococcal Vaccine Study demonstrated that a distributed network of administrative claims databases may facilitate large-scale surveillance of vaccine-related GBS.[73]

Further experience with safety monitoring using healthcare data fostered development of novel signal detection methods that can be employed specifically within the context of EHR. The maximized sequential probability ratio testing (maxSPRT), a signal detection method that supports continuous or time-period analysis of data as they are collected, was developed as part of the real-time surveillance system that has been used, among others, for evaluating meningococcal conjugated vaccine (MCV) vaccination among members of a US healthcare maintenance organization (HMO) network.[74] Two new methods – Longitudinal GPS (LGPS) and Longitudinal Evaluation of Observational Profiles of Adverse Events Related to Drugs (LEOPARD), have been evaluated using both simulated data and actual data from the Dutch Integrated Primary Care Information (IPCI) database. LGPS is a modification of the GPS that uses person-time rather than case counts for estimation of the expected number of events, while LEOPARD is a method designed to automatically discard false drug-event associations caused by protopathic bias or misclassification of the date of adverse event by comparing rates of prescription starts in a fixed window prior to and after the occurrence of an event.[75] Temporal pattern discovery is another method that looks into the chronology of drug prescription and occurrence of an adverse event and has been evaluated in the IMS Disease Analyzer containing observational healthcare data from the UK.[76]

Within the last five years international collaborations have been forged to venture beyond using EHR databases for signal validation to developing EHR data-based drug safety signal detection systems. Some of these collaborations are briefly described below and their major features summarized in Table 2.

The SENTINEL Network

The SENTINEL Initiative was established in 2008 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.[16, 77] Two pilot initiatives have been launched to help develop the eventual SENTINEL system: the Mini-Sentinel and the Federal Partners' Collaboration. Mini-Sentinel, launched at the end of 2009, will enable FDA to query privately-held electronic healthcare data representing approximately 60 million patients. Initial data sources include administrative

claims with pharmacy dispensing data, but data from outpatient and inpatient electronic health records and registries will be added later.[78] The Federal Partners' Collaboration, which includes the Centers for Medicare & Medicaid Services (CMS), the Veterans Health Administration at the Department of Veterans Affairs (VA), and the Department of Defense (DoD), will enable FDA to query federally-held electronic healthcare data, including administrative claims. The Mini-Sentinel pilot will focus on drugs, vaccines, other biologics, and medical devices regulated by the FDA. The vaccine safety activities together constitute the Post-Licensure Rapid Immunization Safety Measurement (PRISM) Program. There are 20 health outcomes of interest (HOI) being evaluated, including two outcomes which pertain specifically to medical devices (i.e., removal of implanted orthopedic device and surgical revision of implantable orthopedic device).[78]

OMOP

The Observational Medical Outcomes Partnership (OMOP) is a public-private partnership among the US FDA, academia, data owners, and the pharmaceutical industry and administered by the Foundation for the National Institutes of Health. It was initiated to identify the needs of an active drug safety surveillance system and to develop the necessary technology and methods to refine the secondary use of observational data for maximizing the benefit and minimizing the risk of pharmaceuticals.[79] OMOP is initially investigating 10 HOIs, which is a subset of all conditions considered important due to their historical associations with drug toxicities, their medical significance, and/or public health implications.[80] In 2009, OMOP organized a methods competition to facilitate development and evaluation of novel approaches for identifying drug safety issues in EHR.[81]

EU-ADR

The EU-ADR Project (Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge), launched in 2008, is funded by the European Commission (EC) under its Seventh Framework Programme.[82] EU-ADR is a collaboration of 18 public and private institutions representing academic research, general practice, health services administration, and the pharmaceutical industry. It currently has access to eight population-based EHR databases in four European countries (Denmark, Italy, the Netherlands, and the United Kingdom), all geared towards producing a computerized integrated system for the early detection of drug safety signals. The system is based on aggregated demographic, clinical, and prescription data pooled via a distributed network approach through generation of common input data followed by local aggregation using custom-built software, Jerboa©.[17] The distributed nature of the network allows data holders to maintain control over their protected data with only aggregated data being shared with the rest of the network. The safety signals are further substantiated by semantic mining of literature and computational analysis of pharmacological and biological information on drugs, molecular targets, and pathways.

Table 2 | International initiatives using EHR databases for drug safety signal detection.

	Data Sources	Catchment area	Source Population (available lives)	Adverse Events Currently Being Evaluated		Drugs Being Investigated
EU-ADR	Medical records	Denmark Italy	30 million	Hematologic	– Hemolytic anemia – **Aplastic anemia/pancytopenia** – Neutropenia – Thombocytopenia	All drugs in the database network
(Started 2008)	(primary care/ general practitioner) Administrative Claims	The Netherlands United Kingdom		Cutaneous/ Multi-systemic	– Maculo-papular erythematous eruptions – **Bullous eruptions** (Stevens Johnson Syndrome, Lyell's Syndrome) – Anaphylactic Shock	
				Liver and Gastrointestinal	– **Acute liver injury** – **Acute pancreatitis** – Upper gastrointestinal bleeding	
				Cardiac and vascular	– Acute myocardial infarction – **QT prolongation** – Cardiac valve fibrosis – Venous thrombosis	
				Musculoskeletal	– Rhabdomyolysis – **Hip Fracture**	
				Neuropsychiatric	– Convulsions – Peripheral neuropathy – Extrapyramidal disorders – Confusional state – **Mood changes (depression, mania)** – Amnesias – **Suicidal behavior/attempt** – Progressive multifocal leukoencephalopathy	
				Renal	– Acute renal failure	

	Data Sources	Catchment area	Source Population (available lives)	Adverse Events Currently Being Evaluated		Drugs Being Investigated
MINI-SENTINEL (Started 2009)	Medical records Administrative Claims	United States	60 million	Hematologic	– ABO incompatibility	Drugs, Biologics, and Devices regulated by the FDA
				Cutaneous/ Multi-systemic	– **Erythema multiforme** – Hypersensitivity reactions – **Anaphylaxis**	
				Liver and gastrointestinal	– **Pancreatitis**	
				Cardiac and vascular	– **Cardiac arrhythmias** – Atrial fibrillation – Congestive Heart Failure – Venous Thromboembolism	
				Neuropsychiatric	– Epilepsy – Stroke/Transient Ischemic Attack – **Depression** – **Suicide**	
				Respiratory	– Respiratory Failure – Pulmonary Fibrosis	
				Neoplastic	– Lymphomas	
				Infectious	– Transfusion Sepsis – Transfusion/Graft Infections	
				Other	– Orthpedic Device Removal – Implantable Device Revision	

	Data Sources	Catchment area	Source Population (available lives)	Adverse Events Currently Being Evaluated		Drugs Being Investigated
OMOP (Started 2009)	Medical records Administrative Claims	United States	325 million	Hematologic	– **Aplastic Anemia** – Bleeding	ACE inhibitors Amphotericin B Antibiotics Anti-epileptics Benzodiazepines Beta Blockers Tricyclic antidepressants Typical antipsychotics Warfarin
				Cutaneous/ Multi-systemic	– Angioedema	
				Liver and Gastrointestinal	– **Acute Liver Injury** – Gastrointestinal Ulcer Hospitalization	
				Cardiac and vascular	– **Myocardial Infarction** – Mortality after Myocardial Infarction	
				Musculoskeletal	– **Hip fracture**	
				Renal	– **Renal Failure**	
				Other	– Hospitalization – All other outcomes recorded in the databases	

NOTE: Events common among the three initiatives are indicated in bold.

EU-ADR takes an event-based approach to signal detection (i.e., all drugs are evaluated for their association with a set of specific events) using as guide a ranked list of 23 adverse events judged as important in pharmacovigilance based on predefined criteria (see Table 2).[83] Three additional events are being looked into (progressive multifocal leukoencephalopathy, acute pancreatitis, and hip fracture) subsequent to a request made by regulatory authorities and after consultation with other stakeholders. The rationale behind pursuing the event-based approach is to avoid unconstrained data mining, which is likely to raise excessive numbers of false positive signals. While the aim in the long-run is for the system to be able to detect a much broader range of events, this set of 'high-priority' events was deemed a good starting point.

Other Initiatives

The Canadian government has also established the Drug Safety and Effectiveness Network (DSEN) to increase the available evidence on drug safety and effectiveness by leveraging existing public resources such as the National Prescription Drug Utilization System.[84] Other recently launched initiatives partly focusing on improving methods for safety signal detection include Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT)[85] and Global Research Initiative in Paediatrics (GRIP).[86]

Sources of data on safety signals triggering withdrawal of drugs from the market

Table 3 gives a summary of the drugs that have been withdrawn from the market for safety reasons in the US and in the EU within the last 10 years. The year when the drug was initially marketed and the corresponding year when the drug was withdrawn are provided, as well as the reason for the withdrawal. It is not the objective of this paper to provide a comprehensive list of drug withdrawals, but rather to examine the nature and characteristics of these withdrawals, and particularly the type of data that provided the basis for these withdrawals.

Of the 24 safety-based withdrawals in the US or the EU, nine (38%) were for adverse cardiovascular events, while seven (29%) were for gastrointestinal – primarily hepatic – adverse events. Drugs acting on the gastrointestinal system comprised 29% (7 out of 24) of all drugs withdrawn; drugs acting on the neurologic and musculoskeletal systems each comprised 17% (4 drugs each). There were three immunomodulating agents (12%), two hematologic agents (8%), and there was one drug each of the therapeutic categories anti-infective, cardiovascular, dermatologic, and diagnostic radiopharmaceutical. 11 out of the 24 drugs (46%) were withdrawn from both the US and EU markets. There are two drugs (trovafloxacin and rosiglitazone) that have been removed from the EU market, but remain available in the US with restrictions or black box warnings.[87-88] Likewise, there are two other drugs (natalizumab and pergolide) that have been withdrawn from the US, but are still marketed in the EU with labeling changes and additional risk management.[89-90] In Figure 1 we further describe the characteristics of these safety-based withdrawals in terms of background frequency, latency or temporality, type of ADR,[91] and source of information used as basis for the withdrawal.

It is apparent from Figure 1 that spontaneous reports make substantial contribution to the decision to take regulatory action, case reports (both published and unpublished) being the primary source of information in 11 of the 24 withdrawals (46%). In two instances (8%), clinical trials were the sole source of the safety information, but for the rest of the withdrawals a combination of case reports and/or clinical trials and/or observational studies contributed to the regulatory action. Although spontaneous reports do not guarantee proof of causality, the scenario of multiple reports originating from different sources raises the degree of suspicion which, by itself, has, in some occasions, been considered sufficient basis for regulatory decisions.[92] Indeed, some signals generated by SRS have been deemed of sufficient strength that further epidemiologic investigation was not sought before taking regulatory action. Noteworthy examples include the FDA's action on L-tryptophan-related eosinophilia-myalgia syndrome and temafloxacin-induced hemolytic anemia.[93] Furthermore, evidence from spontaneous case reports has been vital for the removal from marketing or distribution of 23 drugs in the US from 1978 to 2000.[40]

While drug safety data from RCTs have been mostly utilized pre-approval, RCTs may still provide additional safety data after a drug has already been approved and marketed when there are RCTs investigating new indications for use as well as adverse events post-licensure. A well-known example is the non-steroidal anti-inflammatory drug rofecoxib, whose voluntary withdrawal by its manufacturer was brought about primarily by findings of increased cardiovascular risk in the APPROVe (Adenomatous Polyp PRevention On Vioxx) study, aimed at evaluating the efficacy of rofecoxib for the prophylaxis of colorectal polyps.[94] Data from RCTs, both pre-marketing and post-marketing, can also be pooled in meta-analyses to give a more comprehensive safety profile of the drug and to elucidate ADRs previously undetected in pre-marketing trials.[95]

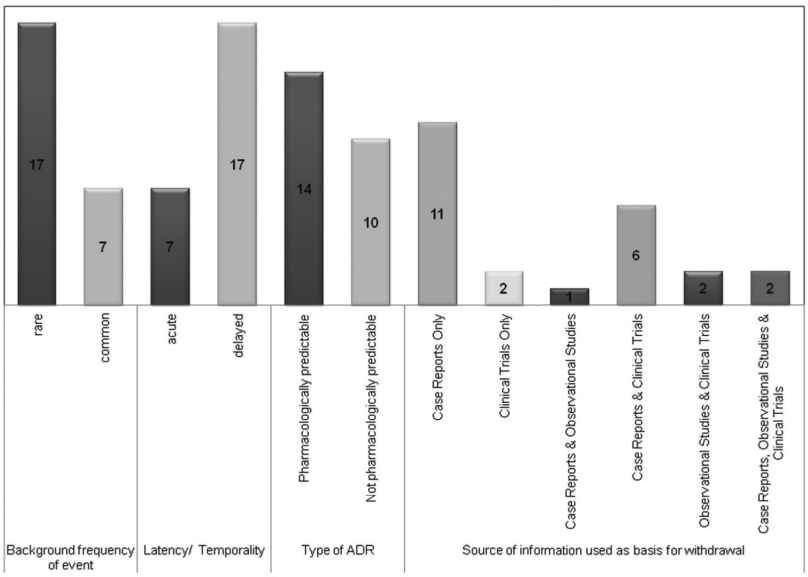


Figure 1 | Characteristics of Drug Safety Withdrawals in the last 10 years.

Clearly, regulatory decision-making is a complex process and is based on more data than what are readily available from published medical literature.[96] From a regulatory perspective, would-be consequences might not allow delaying decisions until all the information are available, especially if this is the kind of information that only a definitive clinical study can provide. At times, the decision to intervene before knowledge is complete becomes imperative in order to avoid potentially harmful consequences. At the same time, the balance of the benefit-to-risk ratio still remains an important factor in the decision to withdraw a drug from the market. While it is often safety concerns about the use of a drug that draw attention, the availability of viable alternative treatments and the impact the withdrawal of such drug would have on patients are equally important issues to consider.[97]

Table 3 | Drugs Withdrawn from the Market for Safety Reasons in the last years in the US/EU.

Drug	Year Initially Marketed (US/EU)	Year Withdrawn (US/EU)	Reason for withdrawal
Cisapride (Propulsid)	1993/1988	2000/2000 (UK) (EU – restricted indications only)	Fatal arrhythmia
Troglitazone (Rezulin)	1997/ 1997 (not centrally authorized)	2000/ 1997(UK)	Liver toxicity
Alosetron (Lotronex)	2000/ not marketed in EU	2000;reintroduced 2002 on a restricted basis	Ischemic colitis, severe constipation
Trovafloxacin (Trovan, Turvel)	1998/1998	Still available in the US but with restrictions/2001	Liver toxicity
Cerivastatin (Baycol)	1997/2001	2001/ 2001(UK), 2002 (EU)	Muscle damage leading to kidney failure
Rapacuronium (Raplon)	1999/ not marketed in EU	2001	Severe bronchospasm
Etretinate (Tegison)	1986/1983	2002/?	Birth defects
Levomethadyl (Orlaam)	1993/1997	2003/2001	Fatal arrhythmia
Rofecoxib (Vioxx)	1999/1999	2004/2004	Cardiovascular events (including myocardial infarction and stroke)
Valdecoxib (Bextra)	2001/2003	2005/2005	Serious skin reactions (TENS, SJS, EM)
Natalizumab (Tysabri)	2004/2006	2005/ still marketed, with additional risk management	Progressive multifocal leukoencephalopathy
Technetium fanolesomab (NeutroSpec)	2004/not marketed in EU	2005	Cardiopulmonary failure (respiratory distress, sudden hypotension)
Pemoline (Cylert)	1975/1960s	2005/ 1997(UK)	Liver failure
Ximelagatran (Exanta)	2004 – Refused by FDA/2003 (France; not centrally authorized)	2006	Liver toxicity

Drug	Year Initially Marketed (US/EU)	Year Withdrawn (US/EU)	Reason for withdrawal
Pergolide (Permax)	1988/1991	2007/still marketed with labeling changes	Cardiac valve damage
Tegaserod (Zelnorm)	2002/2005 – authorization refused	2007	Cardiovascular events (including myocardial infarction and stroke)
Lumiracoxib (Prexige)	2003 and 2007 – Refused by FDA/2005	2007	Liver toxicity, cardiovascular events
Aprotinin (Trasylol)	1993/1974	2008/2007	Renal and cardiac complications, death
Efalizumab (Raptiva)	2003/2004	2009/2009	Progressive multifocal leukoencephalopathy
Sibutramine (Meridia)	1997/1999	2010/2010	CV events (including myocardial infarction and stroke)
Gemtuzumab ozogamicin (Mylotarg)	2000/2007 – authorization refused	2010	Lack of efficacy, increased risk of death (due to liver toxicity/ veno-occlusive disease)
Propoxyphene (Darvon, Darvocet)	1957/1960s	2010/ 2009 (2005 –UK, Sweden) (US)	Cardiac arrhythmia
Rimonabant (Acomplia, Zimulti)	Not marketed in the US/2006	2009	Psychiatric problems, including depression and suicide
Rosiglitazone (Avandia)	1999/2000	Still marketed, but with black box warning/2010 (suspended)	Cardiovascular events, including congestive heart failure, myocardial infarction, and stroke

Note: The bases for withdrawal, as indicated in the respective regulatory agencies' press releases, were obtained from the websites of the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA), as well as other drug references available from the public domain.
Abbreviations: TENS - Toxic Epidermal Necrolysis; SJS - Stevens Johnson Syndrome; EM - Erythema Multiforme; CV - cardiovascular

How automatic signal detection using EHR data fits into the big picture

Despite the well-described limitations of SRS, spontaneous reports have unquestionably played a major role in highlighting potential safety signals post-marketing and have provided supporting evidence for various regulatory actions. Important gaps still remain, however, and one of the principal objectives of exploring EHR databases for safety surveillance is to be able to fill some of these gaps. EHR-based signal detection systems are not subject to many of the limitations of SRS, including the lack of 'denominator,' underreporting, variable and incomplete reporting, or lack of information regarding overall health status prior to and after a suspected ADR. Potential risk associated with drug use should be measured both in terms of risk to the individual and the population frequency, which requires knowledge of the level and duration of exposure. Time-

stamped, population-based healthcare data enables calculation of background incidence rates and allows comparison of event rates before and after drug exposure – information which spontaneous reports are not able to provide. The longitudinal nature of the data may allow identification of adverse events that have a long delay between exposure and clinical manifestations (e.g., cardiac valvulopathy or cancer), especially in databases with long patient follow-up and low turnover. While most spontaneous reports usually involve newly marketed drugs, EHR data may be able to highlight new risks associated with old drugs (as a consequence of new indications of use or new generation of users), as well as adverse events that have high background incidence rates (such as acute myocardial infarction) and events which are less likely to be suspected as drug-induced, thus less likely to be reported. Furthermore, data from EHR provide greater detail regarding patient demographics, drug use, and utilization of healthcare services which permit evaluation of the benefit-risk profile of drugs, hence putting safety issues in a broader perspective and fostering sound regulatory decisions.

While EHR databases may provide a wealth of drug use information and degree of detail that SRS databases are not able to provide and offer clear advantages over SRS, there remain caveats in the interpretation of signals derived from mining EHR data. Since these data are not primarily intended for recording drug-related adverse events, potential associations are inferred outside the actual patient-physician encounter that leads to suspicion of an ADR – something that is inherent in SRS. Data mining methods that filter out alternative explanations for these associations (by controlling for bias and confounding) attempt to simulate the causality assessment performed by reporting physicians. The literature is replete with discussions on the merits and challenges of the secondary use of EHR, including how the type of database influences the structure and content of the data.[67-68] Data in medical records databases are recorded in the course of clinical care and hence take a healthcare practitioner's view of what is going on with a patient. On the other hand, claims databases document information as by-product of fiscal transactions, and therefore provide an auditor's view of healthcare data and can be biased by differences (real or perceived) in reimbursement. There are limitations with respect to validity of coding systems and potential misclassification of both drug exposure and outcome and, like any other observational data, there can be inadequate control for bias and confounding.[98-99] EHR data derived from health maintenance organizations (HMOs) or social security systems could also be affected by a lack of incentive to record sufficient data to allow proper case classification. Drug use patterns derived from EHR are influenced by changes in clinical practice, including changes brought about by preferential prescribing and disease management guidelines, and may lead to underestimation of risks. It has been shown that even with large multi-country databases, the capability for signal detection may be low for drugs that are infrequently used and for very rare outcomes – situations wherein other surveillance systems, such as SRS and prescription event monitoring, may provide better benefits.[100] Furthermore, before an EHR database is used for signal detection purposes, the decision-makers should already anticipate the question of what happens if and when a signal is detected and whether the same database can be used for hypothesis confirmation studies related to

the signal identified. Clarifying beforehand the options for further use of the data in such an event becomes imperative.

CONCLUSION

Initiatives exploring EHR-based signal detection systems are intended to augment, not replace, existing drug safety surveillance systems. Signal detection – whether using EHR databases or otherwise – is, by definition, exploratory. Every signal demands further investigation and the goal of any surveillance system should be to make judicious use of available healthcare data to highlight potential safety problems earlier. Identification and elucidation of drug safety signals is both an iterative and dynamic process. It is in the best interest of public health to integrate – and understand – evidence from all possibly relevant information sources on drug safety.

ACKNOWLEDGMENT

The authors would like to thank Dr. Hector Izurieta of the Center for Biologics Evaluation and Research, US Food and Drug Administration and Dr Jim Slattery of the European Medicines Agency for reviewing the manuscript.

REFERENCES

1. Zarin DA, Young JL, West JC. Challenges to evidence-based medicine: a comparison of patients and treatments in randomized controlled trials with patients and treatments in a practice research network. *Soc Psychiatry Psychiatr Epidemiol.* Jan 2005; 40(1): 27-35.
2. Heiat A, Gross CP, Krumholz HM. Representation of the elderly, women, and minorities in heart failure clinical trials. *Arch Intern Med.* Aug 12-26 2002;162(15):1682-1688.
3. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; 312 (7040):1215-1218.
4. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006; 74: 635–641.
5. Cantu C, Arauz A, Murillo-Bonilla LM, López M, Barinagarrementeria F. Stroke associated with sympathomimetics contained in over-the-counter cough and cold drugs. *Stroke* 2003 Jul; 34(7): 1667-1672.
6. FDA issues public health warning on phenylpropanolamine. http://www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/ucm150763.htm. Accessed 11 November 2011.
7. FDA Requires Additional Labeling for Over-the-Counter Pain Relievers and Fever Reducers to Help Consumers Use Products Safely. http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2009/ucm149573.htm. Accessed 11 November 2011.
8. McGettigan P, Henry D. Cardiovascular risk with non-steroidal anti-inflammatory drugs: systematic review of population-based controlled observational studies. *PLoS Med* 2011 Sep; 8(9).
9. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA Spontaneous Reporting System. *Am Stat* 1999; 53:177-202.
10. Almenoff JS, DuMouchel W, Kindman LA, et al. Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting. *Pharmacoepi Drug Saf* 2003;12:517–521.
11. Hauben M, Zhou X. Quantitative Methods in pharmacovigilance: focus on signal detection. *Drug Saf* 2003; 26(3):159-186.
12. Bousquet C, Henegar C, Louet A, et al. Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach. *Int J Med Inform* 2005; 74:563-571.
13. Bate A, Edwards I, eds. Data Mining Techniques in Pharmacovigilance. in: Hartzema AG, Tilson HH, Chan KA, eds. *Pharmacoepidemiology and Therapeutic Risk Management*. Cincinnati: Harvey Whitney; 2008.
14. Heeley E, Wilton LV, Shakir SA. Automated signal generation in prescription-event monitoring. *Drug Saf* 2002;25(6):423-432.
15. Coulter DM. Signal generation in the New Zealand Intensive Medicines Monitoring Programme: a combined clinical and statistical approach. *Drug Saf* 2002;25(6):433-439.
16. 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 Aug 13;361(7):645-647.
17. Coloma PM, Schuemie M, Trifiró G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom MCJM, 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 Jan;20(1):1-11.
18. World Health Organization. Safety of medicines. A guide to detecting and reporting adverse drug reactions 2002; http://whqlibdoc.who.int/hq/2002/WHO_EDM_QSM_2002.2.pdf. Accessed 10 July 2011.
19. Hauben M, Aronson J. Defining 'Signal' and its Subtypes in Pharmacovigilance Based on a Systematic Review of Previous Definitions. *Drug Saf* 2009;32(2):99-110.
20. US Food and Drug Administration. Adverse Event Reporting System (AERS). http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm. Accessed 11 November 2011.
21. US Food and Drug Administration. MedWatch: The FDA Safety Information and Adverse Event Reporting Program. http://www.fda.gov/Safety/MedWatch/default.htm. Accessed 11 November 2011.

22. US Food and Drug Administration. Code of Federal Regulations. Sec. 314.80 Postmarketing reporting of adverse drug experiences. http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=314.80. Accessed 11 November 2011.

23. US Food and Drug Administration. US FDA Reports Received and Reports Entered into AERS by Year. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm070434.htm. Accessed 15 October 2011.

24. Chen RT, Rastogi SC, Mullen JR, Hayes SW, Cochi SL, Donlon JA, Wassilak SG. The Vaccine Adverse Event Reporting System (VAERS). *Vaccine* 1994 May;12(6):542-550.

25. US Food and Drug Administration. About the VAERS Program. http://vaers.hhs.gov/index/about/index Accessed 20 October 2011.

26. WHO Uppsala Monitoring Centre. Introduction to the WHO Programme for International Drug Monitoring. http://www.who-umc.org/. Accessed Last accessed 10 October 2011.

27. Lindquist M. VigiBase, the WHO Global ICSR Database System: Basic Facts. *Drug Inf J* 2008;42:409–419.

28. WHO Uppsala Monitoring Centre. WHO Collaborating Center for International Drug Monitoring Caveat Document. http://www.who-umc.org/graphics/25300.pdf. Accessed 20 October 2011.

29. European Medicines Agency. Eudravigilance http://eudravigilance.ema.europa.eu/human/index.asp. Accessed 15 October 2011.

30. Borg JJ, Aislaitner G, Pirozynski M, Mifsud S. Strengthening and rationalizing pharmacovigilance in the EU: where is Europe heading to? A review of the new EU legislation on pharmacovigilance. *Drug Saf* 2011 Mar 1; 34(3):187-197.

31. European Medicines Agency. EudraVigilance access policy for medicines for human use. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2011/07/WC500108538.pdf. Accessed 24 November 2011.

32. Szarfman A, Tonning JM, Doraiswamy PM. Pharmacovigilance in the 21st century: new systematic tools for an old problem. *Pharmacotherapy* 2004;24(9):1099-1104.

33. Hauben M, Madigan D, Gerrits CM, Walsh L, Van Puijenbroek EP. The role of data mining in pharmacovigilance. *Expert Opin Drug Saf* 2005;4(5):929-948.

34. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoeidemiol Drug Saf* 2001;10(6):483-486.

35. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol Drug Saf* 2004 Aug;13(8):519-523.

36. Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. *Drug Saf* 2002;25(6):381-392.

37. Bate A, Lindquist M, Edwards IR, Orre R. A data mining approach for signal detection and analysis. *Drug Saf* 2002;25(6):393-397.

38. Blum MD, Graham DJ, McCloskey CA. Temafloxacin syndrome: review of 95 cases. *Clin Infect Dis* 1994 Jun; 18(6):946-950.

39. Darpö B. Detection and reporting of drug-induced proarrhythmias: room for improvement. *Europace* 2007 Sep; 9(Suppl 4):iv23-36.

40. Wysowski DK, Swartz L SL. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. *Arch Intern Med.* 2005 Jun 27;165(12):1363-1369.

41. US Food and Drug Administration. FDA Announces Withdrawal Fenfluramine and Dexfenfluramine http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm179871.htm. Accessed 18 October 2011.

42. Fairley CK, McNeil JJ, Desmond P, Smallwood R, Young H, Forbes A, Purcell P, Boyd I. Risk factors for development of flucloxacillin associated jaundice. *BMJ* 1993;306(6872):233-235.

43. Desmond PV. Flucloxacillin hepatitis - an Australian epidemic. *Aust NZ J Med* 1995;25(3):195-196.

44. Thomson JA, Fairley CK, Ugoni AM, Forbes AB, Purcell PM, Desmond PV, Smallwood RA, McNeil JJ. Risk factors for the development of amoxycillin-clavulanic acid associated jaundice. *Med J Aust* 1995;162(12): 638-640.

45. Thomson JA, Fairley CK, McNeil JJ, Purcell P. Augmentin-associated jaundice. *Med J Aust* 1994;160(11):733-734.
46. Centers for Disease Control and Prevention. Intussusception among recipients of rotavirus vaccine--United States, 1998-1999. *MMWR Morb Mortal Wkly Rep* 1999 Jul 16;48(27):577-581.
47. Centers for Disease Control and Prevention. Withdrawal of rotavirus vaccine recommendation. *MMWR Morb Mortal Wkly Rep* 1999 Nov 5;48(43):1007.
48. Niu MT, ErwinDE, 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 Sep 14;19(32): 4627-4634.
49. Murphy TV, Gargiullo PM, Massoudi MS, Nelson DB, Jumaan AO, Okoro CA, Zanardi LR, Setia S, Fair E, LeBaron CW, Wharton M, Livengood JR, Rotavirus Intussusception Investigation Team. Intussusception among infants given an oral rotavirus vaccine. *N Engl J Med* 2001 Feb 22;344(8):564-572.
50. Centers for Disease Control and Prevention. Update: Guillain-Barré syndrome among recipients of Menactra meningococcal conjugate vaccine--United States, June 2005-September 2006. *MMWR Morb Mortal Wkly Rep* 2006 Oct 20;55(41):1120-1124.
51. US Food and Drug Administration. FDA and CDC Update Information on Menactra Meningococcal Vaccine and Guillain Barre Syndrome. http://www.fda.gov/BiologicsBloodVaccines/SafetyAvailability/ucm154937.htm. Accessed 20 October 2011.
52. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions. A systematic review. *Drug Saf* 2006; 29(5): 385-396.
53. Friedman MA, Woodcock J, Lumpkin MM, Shuren JE, Hass AE, Thompson LJ. The safety of newly approved medicines: do recent market removals mean there is a problem? *JAMA* 1999;281(18):1728-1734
54. Merck Pulls Arthritis Drug Vioxx from Market. http://www.npr.org/templates/story/story.php?storyId=4054991. Accessed 11 November 2011.
55. Krumholz HM, Ross J, Presler AH, Egilman DS. What have we learnt from Vioxx? *BMJ* 2007 Jan 20;334(7585): 120-123.
56. Goldman SA. Limitations and strengths of spontaneous reports data. *Clin Ther* 1998;20 (Suppl C):C40-44.
57. Trontell A. How the US Food and Drug Administration defines and detects adverse drug events. *Current Therapeutic Research* 2001;62:641-649.
58. Wang HW, Hochberg AM, Pearson RK, Hauben M. An experimental investigation of masking in the US FDA adverse event reporting system database. *Drug Saf* 2010 Dec 1;33(12):1117-1133.
59. Rawson NS, Pearce GL, Inman WH. Prescription-event monitoring: methodology and recent progress. *J Clin Epidemiol* 1990;43(5):509-522.
60. Coulter DM, Edwards I. The New Zealand Intensive Medicines Monitoring Programme. *Pharmacoepidemiol Drug Saf* 1998;7(2):79-90.
61. Coulter DM, Edwards I. Cough associated with captopril and enalapril. *BMJ* 1987;294:1521-1523.
62. Coulter DM, Edwards I. Mianserin and agranulocytosis. *Lancet* 1990;336:785-787.
63. Institute of Medicine. The Future of Drug Safety: Promoting and Protecting the Health of the Public. 2006; http://www.iom.edu/Reports/2006/The-Future-of-Drug-Safety-Promoting-and-Protecting-the-Health-of-the-Public.aspx. Accessed 20 October 2011.
64. Psaty BM, Burke SP. Protecting the health of the public--Institute of Medicine recommendations on drug safety. *N Engl J Med* 2006 Oct 26;355(17):1753-1755.
65. García Rodríguez LA, Pérez Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. *Br J Clin Pharmacol* 1998 May;45(5):419-425.
66. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects—advantages and disadvantages. *Nat Clin Pract Rheum* 2007;3(12):725-732.
67. Hennessy S. Use of health care databases in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol* 2006; 98(3): 311-313.
68. 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.

69. Kramarz P, France E, Destefano F, Black SB, Shinefield H, Ward JI, Chang EJ, Chen RT, Shatin D, Hill J, Lieu T, Ogren JM. Population-based study of rotavirus vaccination and intussusception. *Pediatr Infect Dis J* 2001 Apr;20(4):410-416.

70. Bate A, Edwards I, Edwards J, Swahn E, Noren GN, Lindquist M. Knowledge finding in IMS disease analyser Mediplus UK database – effective data mining in longitudinal patient safety data. . Paper presented at: ISoP Annual Conference: Pharmacovigilance - Current and Future Challenges2004; Dublin, Ireland.

71. Curtis J, Cheng H, Delzell E, Fram D, Kilgore M, Saag K, Yun H, DuMouchel W. Adaptation of Bayesian Data Mining Algorithms to Longitudinal Claims Data: Coxib Safety as an Example. *Med Care* 2008;46(9):969-975.

72. Svanström H, Callréus T, Hviid A. Temporal data mining for adverse events following immunization in nationwide Danish healthcare databases. *Drug Saf* 2010 Nov 1;33(11):1015-1025.

73. Velentgas P, BohnRL, Brown JS, Chan KA, Gladowski P, Holick CN, Kramer JM, Nakasato C, Spettell CM, Walker AM, Zhang F, Platt R. A distributed research network model for post-marketing safety studies: the Meningococcal Vaccine Study. *Pharmacoepidemiol Drug Saf* 2008 Dec;17(12):1226-1234.

74. Lieu TA, Kulldorff M, Davis RL, Lewis EM, Weintraub E, Yih K, Yin R, Brown JS, Platt R; for the Vaccine Safety Datalink Rapid Cycle Analysis Team. Real-time vaccine safety surveillance for the early detection of adverse events. *Med Care* 2007 Oct; 45(10 Supl 2):S89-95.

75. Schuemie MJ. Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD. *Pharmacoepidemiol Drug Saf* 2011 Mar; 20(3):292-9. doi: 10.1002/pds.2051. Epub 2010 Oct 13.

76. Norén GN. Hopstadius J, Bate A, et al. Temporal pattern discovery in longitudinal electronic patient records. *Data Min Knowl Disc* 2010(20):361–387.

77. US Food and Drug Administration. The FDA Sentinel Initiative. http://www.fda.gov/Safety/FDA SentinelInitiative. Accessed 12 July 2011.

78. FDA Mini-Sentinel. http://mini-sentinel.org/. Accessed 15 February 2011.

79. Stang PE, Ryan P, 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(9):600-606.

80. Observational Medical Outcomes Partnership. http://omop.fnih.org/HOI. Accessed 11 November 2011.

81. Observational Medical Outcomes Partnership. OMOP Cup. http://omop.fnih.org/omopcup. Accessed 10 October 2011.

82. EU-ADR Consortium. Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical records and Biomedical Knowledge - The EU-ADR Project http://www.euadr-project.org. Accessed 12 July 2011.

83. Trifirò G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salamé 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, 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):1176-1184.

84. Canadian Institutes of Health Research. About the Drug Safety and Effectiveness Network. http://www.cihr-irsc.gc.ca/e/39389.html. Accessed 01 March 2011.

85. Innovative Medicines Initiative. PROTECT Project. http://www.imi-protect.eu/about.html. Accessed 01 March 2011.

86. Global Resarch Initiative in Paediatrics (GRIP). http://www.euroneostat.org/paginas/publicas/euroneo/euroNeoStat/documents/GRIP-AEMPS_1.pdf. Accessed 11 November 2011.

87. US National Library of Medicine. Trovafloxacin drug label. http://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?id=442&CFID=66575927&CFTOKEN=a17d753a0754a3fb-24987D34-D80B-CD9C-39F668DB8C41A045&jsessionid=ca30e46b22b0112063a4. Accessed 13 July 2011.

88. US National Library of Medicine. Rosiglitazone drug label. http://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?id=38243. Accessed 13 July 2011.

89. European Medicines Agency. About Tysabri (natalizumab). http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000603/human_med_001119.jsp&murl=menus/medicines/medicines.jsp&mid=WC0b01ac058001d124. Accessed 13 July 2011.

90. UK Medicines and Healthcare products Regulatory Agency. Dopamine agonists for Parkinson's disease http://www.mhra.gov.uk/Safetyinformation/Generalsafetyinformationandadvice/Product-specificinformationandadvice/ Product-specificinformationandadvice-A-F/DopamineagonistsforParkinson146sdisease/index.htm. Accessed 13 July 2011.

91. Davies D, ed *Textbook of adverse drug reactions.* 3rd ed ed. Oxford: Oxford University Press; 1985. Rawlins MD TJ, ed. Pathogenesis of adverse drug reactions.

92. Auriche M, Loupi E. Does proof of causality ever exist in pharmacovigilance? *Drug Saf* 1993;9(3):230-235.

93. USFDA Center for Drug Evaluation and Research. The Clinical Impact of Adverse Event Reporting. http://www.fda.gov/downloads/Safety/MedWatch/UCM168505.pdf Accessed 10 April 2011.

94. Bresalier RS, Sandler RS, Quan H, Bolognese JA, Oxenius B, Horgan K, Lines C, Riddell R, Morton D, Lanas A, Konstam MA, Baron JA, Adenomatous Polyp Prevention on Vioxx (APPROVe) Trial Investigators. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med* 2005 Mar 17;352(11):1092-1102.

95. CIOMS. *Management of Safety Information from Clinical Trials - Report of CIOMS Working Group VI*. Geneva: CIOMS; 2005.

96. Edwards IR. What are the Real Lessons from Vioxx®? *Drug Saf* 2005;28(8):651-658.

97. Fung M, Thornton A, Mybeck K, et al. Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets – 1960 to 1999. *Drug Inf J* 2001;35:293-317.

98. Schneeweiss S. Understanding secondary databases: a commentary on "Sources of bias for health state characteristics in secondary databases". *J Clin Epidemiol* 2007 Jul;60(7):648-650.

99. Norén GN, Edwards I. Opportunities and challenges of adverse drug reaction surveillance in electronic patient records. *PharmacoVigilance Review* 2010 Feb;4(1):17-20.

100. Coloma PM, Trifirò G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Picelli G, Corrao G, Pedersen L, van der Lei J, Sturkenboom MCJM, on behalf of the EU-ADR consortium. Electronic healthcare databases for active drug safety surveillance: Is there enough leverage? *Pharmacoepidemiol Drug Saf* 2012. Feb 8. doi: 10.1002/pds.3197. [Epub ahead of print]

Chapter 3

Breaking New Grounds

'Discovery in an observational science such as pharmacovigilance depends on the capacity to recognize and investigate unexpected clinical events that are manifest once a new drug is in use. The detection of such unanticipated effects hinges on what Pasteur called "the prepared mind"." – A. Trontell, 2004

3.1

Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor?

Gianluca Trifirò, MD, PhD[1-3], Antoine Pariente, MD, PhD[4-6], Preciosa M. Coloma, MD, MSc,[1] Jan A. Kors, PhD[1], Giovanni Polimeni, PharmD, PhD[3], Ghada Miremont-Salamé, MD[4-6], Maria Antonietta Catania, MD[2,3], Francesco Salvo, MD, PhD[3,4], Anaelle David, MD[4-6], Nicholas Moore MD, PhD[4-6], Achille Patrizio Caputi, MD[2,3], Miriam Sturkenboom, PharmD, PhD[1], Mariam Molokhia, MD, PhD[7], Julia Hippisley-Cox, MD[8], Carlos Diaz Acedo[9], Johan van der Lei, MD, PhD[1] Annie Fourrier-Reglat, PharmD, PhD[4-6], on behalf of the EU-ADR group

[1]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands; [2]IRCCS Centro Neurolesi 'Bonino-Pulejo', Messina, Italy; [3]Department of Clinical and Experimental Medicine and Pharmacology, Pharmacology Unit, University of Messina, Messina, Italy; [4]Inserm U 657, Pharmacology Department, Bordeaux, France; [5]CHU Bordeaux, France; [6]Department of Pharmacology, University of Bordeaux, France; [7]Department of Epidemiology and Population Heath, London School of Hygiene & Tropical Medicine, London, UK; [8]Division of Primary care, School of Community Health Sciences, University of Nottingham, UK; [9]Fundacio' IMIM – European Projects Management Office, Barcelona, Spain

Pharmacoepidemiol and Drug Safety 2009 Dec; 18(12): 1176-84.

ABSTRACT

Purpose: Data mining on electronic health records (EHRs) has emerged as a promising complementary method for post-marketing drug safety surveillance. The EU-ADR project, funded by the European Commission, is developing techniques that allow mining of EHRs for adverse drug events across different countries in Europe. Since mining on all possible events was considered to unduly increase the number of spurious signals, we wanted to create a ranked list of high-priority events.

Methods: Scientific literature, medical textbooks, and websites of regulatory agencies were reviewed to create a preliminary list of events that are deemed important in pharmacovigilance. Two teams of pharmacovigilance experts independently rated each event on five criteria: (1) 'trigger for drug withdrawal'; (2) 'trigger for black box warning'; (3) 'leading to emergency department visit or hospital admission'; (4) 'probability of event to be drug-related'; and (5) 'likelihood of death'. In case of disagreement, a consensus score was obtained. Ordinal scales between 0 and 3 were used for rating the criteria, and an overall score was computed to rank the events.

Results: An initial list comprising 23 adverse events was identified. After rating all the events and calculation of overall scores, a ranked list was established. The top-ranking events were: bullous eruptions, acute renal failure, anaphylactic shock, acute myocardial infarction, and rhabdomyolysis.

Conclusions: A ranked list of 23 adverse drug events judged as important in pharmacovigilance was created to permit focused data mining. The list will need to be updated periodically as knowledge on drug safety evolves and new issues in drug safety arise.

INTRODUCTION

In pharmacovigilance, a signal is defined by the World Health Organization as information on a possible causal relationship between an adverse event and a drug, which is unknown or incompletely documented.[1,2] Spontaneous reporting systems (SRS) for adverse drug reactions (ADRs) have been the cornerstone of signal detection in pharmacovigilance for the last four decades.[3,4] In recent years, however, there has been a clamor for improved drug safety monitoring as a result of several high-impact safety issues.[5,6] It has become evident that some adverse effects of drugs may be detected too late, when millions of persons have already been exposed.

The need for earlier detection and for a more proactive approach to drug safety surveillance is receiving much attention both in Europe and in North America.[7–9] The increasing availability of electronic health records (EHRs) presents opportunities to investigate a wide spectrum of adverse drug effects and to detect signals closer to real time.[10,11] Compared to clinical trial data, population-based EHR databases contain data from clinical practice about larger populations and longer follow-up periods.[12–15] Additionally, in contrast to spontaneous reporting systems, EHR databases do not suffer from underreporting and reporting biases, thus potentially facilitating a more timely identification of signals.[11,16] It took five years for rofecoxib to be withdrawn from the market.[17] Using actual data on the penetration of rofecoxib in the market, it has been calculated 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.[18,19] A number of data mining techniques have been specifically developed for automatic detection of drug safety signals.[3,20–24]

It is within this context that the EU-ADR project (http://www.euadr-project.org) was funded by the European Commission with the aim to design, develop, and validate a computerized integrative system that exploits data from EHRs and biomedical databases for the early detection of ADRs. Beyond the current state-of-the-art, EU-ADR will lead to the federation of different population-based EHR databases, creating a resource of unprecedented size for drug safety monitoring in Europe (over 30 million patients from eight different databases). The initial stage of signal generation will be followed by signal substantiation and validation as described in Figure 1.

When using data mining to detect signals in EHR databases, a decision needs to be made about the type of approach, which can be drug-based or event-based. In a drug-based approach, a set of specific drugs are monitored for their association with all possible events.[25] In an event-based approach, a set of specific events are examined for their association with all possible drugs. The EU-ADR project is currently following the event-based approach. The definition of drugs across databases in various countries can more easily be harmonized compared to the definition of events. In addition, the event-based approach allows us to focuses on those events that are considered important from a public-health perspective irrespective of the drug causing the event.

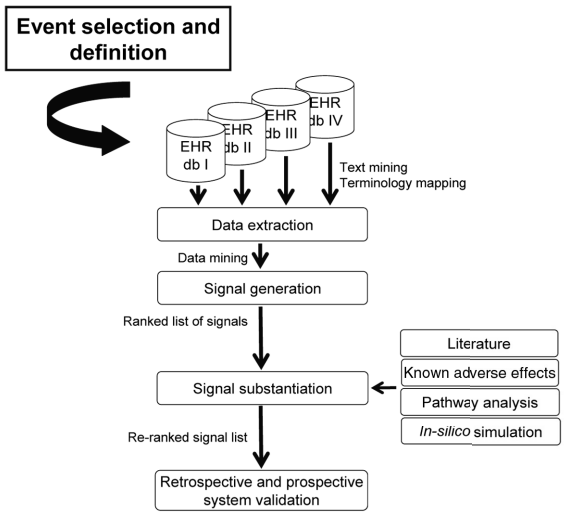


Figure 1 | Framework of the EU-ADR project.

One of the challenges in the event-based approach to signal detection through mining of EHR databases is the identification of events that are most important in pharmacovigilance and thus warrant priority for monitoring. Indeed, unconstrained data mining is likely to raise excessive numbers of spurious signals. Generating a long list of potential drug-event associations, especially in an automated fashion, may lead to so-called 'phantom ships',[26,27] which can have significant negative implications for public health, as demonstrated in the case of the atypical antipsychotic sertindole.[28] There is, however, no list available in the literature to date describing which are the primary events of interest for intensive monitoring in pharmacovigilance when applying data mining techniques. The aim of the current study was therefore to create such a list of events ranked by importance. This list of events shall be the starting point for signal detection in the EU-ADR project.

METHODS

For the priority event selection and ranking, a four-step procedure was outlined by two teams of four pharmacovigilance experts each from two institutions (Department of Pharmacology and Regional Pharmacovigilance Centre at the Universite' Victor-Segalen in Bordeaux, France, and the IRCCS Centro Neurolesi 'Bonino Pulejo', Messina, Italy).

Step 1: Identification of important events

A list of important adverse drug events was compiled, considering different system/organ classes. These events were identified from pharmacovigilance reference books and publications reviewing reasons for drug withdrawal (see Appendix), and from information in the websites of drug regulatory agencies (Food and Drug Administration, European Medicines Agency, French Agency for the Safety of Health Products, and Italian Drug Agency).

Step 2: Creation of a criteria set for event rating

In order to rank the events according to public health importance, five criteria were considered (Table 1): (1) frequency of the event as trigger for drug withdrawal; (2) frequency of the event as trigger for black box warning; (3) leading to emergency department visit or hospital admission; (4) probability of event to be drug-related; and (5) likelihood of death. For each criterion, ordinal scales were defined ranging from 0 (not relevant) to 3 (highly relevant).

Table 1 | Description and score definition of five criteria for the rating of Events.

Criterion	Description
Trigger for drug withdrawal	*Question*: How many times did the event lead to drug withdrawal?
	Score: 0= never; 1= 1-3; 2= 4-5; 3= >5.
	Note: The event causing the drug withdrawal should specifically be indicated (e.g., "cardiotoxicity" is not the same as "acute myocardial infarction").
Trigger for black box warning	*Question*: How many times was the event a trigger for a black box warning?
	Score: 0= never; 1= 1-5; 2= 6-10; 3= >10.
	Note: The event causing the black box warning should specifically be indicated.
Leading to emergency department visit or hospital admission	*Question*: How many published studies reported the event as an adverse drug reaction leading to an emergency department visit or to hospitalization?
	Score: 0= none; 1= 1-3; 2= 4-5; 3= >5.
	Note: Papers confined to emergency department visits and hospitalizations in Europe.
Probability of event to be drug-related	*Question*: What is the probability that the event is drug-related?
	Score: 0= zero probability; 1= <20%; 2= 20%-50%; 3= >50%.
	Note: In case of lack of information or equivocal findings, expert opinion is used to assign the score.
Likelihood of death	*Question*: What is the likelihood that a certain event leads to death?
	Score: 0= unlikely; 1= <10%; 2= 10%-30%; 3= >30%.
	Note: This evaluation is independent of the etiology of the event and potential association with the drug.

Step 3: Score assignment

The two pharmacovigilance teams independently assigned scores for each of the events. Scores were based on a comprehensive review of the scientific literature over the past 10 years (see Appendix) and the evaluation of other sources of information (e.g., websites of regulatory agencies,

medical textbooks). If literature review of the past 10 years was deemed insufficient, literature review was extended to the last 20 years. When review of the literature and other sources provided insufficient data, scoring of the criteria was based on expert opinion. In case of disagreement about a score, consensus was obtained after discussion among all members of the two teams.

Step 4: Ranking of events

For each event, an overall score was computed by summing up the scores for the five criteria. Based on the overall scores, a ranked event list was made. The top-ranked events were considered as having the highest priority for drug safety monitoring.

Table 2 | List of important events in pharmacovigilance, grouped according to system/organ involved.

System/organ	Event
Hematologic	– Hemolytic anemia
	– Aplastic anemia/pancytopenia
	– Neutropenia
	– Thrombocytopenia
Cutaneous	– Maculopapular erythematous eruptions
	– Bullous eruptions (Stevens Johnson Syndrome, Lyell's Syndrome)
Liver and gastrointestinal	– Acute liver injury
	– Acute pancreatitis
	– Upper gastrointestinal bleeding
Cardiac and vascular	– Acute myocardial infarction
	– QT prolongation
	– Cardiac valve fibrosis
	– Venous thrombosis
Neurologic/Musculoskeletal	– Convulsions
	– Peripheral neuropathy
	– Extrapyramidal disorders
	– Rhabdomyolysis
Psychiatric	– Confusional state
	– Mood changes: depression and mania
	– Amnesias
	– Suicidal behavior/attempt
Renal	– Acute renal failure
Multi-systemic	– Anaphylactic shock

RESULTS

Twenty-three events were identified in the first step of the event selection process (Table 2). These were classified by organ/body system into: hematologic;, cutaneous; liver and gastrointestinal; cardiac and vascular; neurologic/musculoskeletal; psychiatric; renal; and multi-systemic. Table 3 shows the scores on the five criteria for the 23 events. On the basis of the overall scores, the top-ranked events were: cutaneous bullous eruptions; acute renal failure; anaphylactic shock; acute myocardial infarction; and rhabdomyolysis. As an example, cutaneous bullous eruption (Stevens-Johnson syndrome or Lyell's syndrome) emerged as one of the most important events, garnering a score of 15 points: (1) at least five drugs have been withdrawn from the market due to this adverse reaction, viz., valdecoxib, chlormezanone, sulfamethoxypyridazine, sulfadimethoxine, and isoxicam, and multiingredients preparations containing phenobarbital[29–32] (criterion 'trigger for drug withdrawal', 3 points); (2) a black box warning for risk of cutaneous bullous reactions has been imposed on more than 10 drugs ('trigger for black box warning', 3 points); (3) more than five papers have reported this adverse event as being responsible for emergency department visit or hospitalization[33] ('leading to emergency department visit or hospital admission', 3 points); (4) at least 70% of cutaneous bullous reactions have been attributed to drug exposure[34,35] ('probability of event to be drug-related', 3 points); (5) more than 30% of Stevens-Johnson/Lyell's syndrome cases are fatal, mainly because of progression to sepsis or pulmonary involvement ('likelihood of death', 3 points).[36]

DISCUSSION

This is the first study aimed at systematically creating a prioritized list of events to monitor when applying data mining techniques on large EHR databases for safety signal detection. This list comprises 23 events with high attributable risks of drugs based on seven system/organ classes that are most commonly involved in ADRs.[37-41] Most of the events identified in this study constitute potentially clinically serious outcomes resulting in hospitalization, life-threatening situations, or death. While many adverse drug events do not require hospital admission or emergency department visit, the prioritized events do strain healthcare resources and thus have stronger implications for public health. Most studies report that ADRs cause 3-6% of all hospital admissions and are responsible for about 5-10% of inpatient costs[38,42-47] Six of the 23 events identified in this study were associated with prolonged hospitalization in the United Kingdom, based on hospitalization statistics from 1996 to 2000 (hemolytic anemia, liver injury, extrapyramidal effects, renal failure, rhabdomyolysis, anaphylaxis, and aplastic anemia).[48] While there are lists of most frequently reported adverse drug events available from various regulatory agencies and repositories of SRS, frequency of reporting does not necessarily translate to clinical significance. In a review of adverse drug event surveillance and drug withdrawals in the US, the top 20 events reported from 1969 to 2002 were also very different from the events cited as reasons for removing drugs from the

market, which were far more serious.[32] This is partly due to the fact that some ADRs are more easily recognizable than others because of the known pharmacology of the drug (so-called type A ADRs, which account for over 80% of all reactions), patients' previous experience or, probably most important, physician awareness.[47] On the other hand, more serious events that are unpredictable (idiosyncratic or type B ADRs) are less easily identified and may not be reported as often, but are frequently responsible for removal of drugs from the market.[49-51] In a French study that looked into drug withdrawal for pharmacovigilance reasons, type B reactions were responsible for the withdrawal of 11 out of 21 drugs.[52] Of the 23 events in our priority list, 18 have been implicated as reason for removal of drugs from the market (e.g., cutaneous bullous eruption for valdecoxib; liver injury for pemoline, benoxaprofen, and troglitazone; cardiac valve fibrosis for pergolide, fenfluramine, and dexfenfluramine; rhabdomyolysis for cerivastatin).[31-32,46,49] The adverse events identified in this study can be further characterized in terms of pathogenesis (i.e., predictable or unpredictable from the drug's pharmacology, role of immune or allergic mechanisms), clinical course (acute, delayed, chronic) and susceptibility of special populations (e.g., children or elderly) to experience the event. From the 23 events, four (17.4%) are known to be idiosyncratic, or immunologically mediated (thrombocytopenia, cutaneous bullous eruptions, anaphylactic shock, and maculopapular erythematous eruption). While the precise hazard function of adverse drug events is most of the time unknown, some events are acute by definition (e.g., acute myocardial infarction, acute renal failure, acute liver injury, and acute pancreatitis), while cardiac valve fibrosis and peripheral neuropathy are considered chronic and delayed, respectively. Some events, such as convulsions, upper gastrointestinal bleeding and acute renal failure, may be more critical to monitor in the extremes of age.

CONCLUSION

Prioritization of adverse events judged as important in pharmacovigilance was undertaken to optimize data mining and to contain the number of spurious signals that data mining on EHR databases may generate. The first five of the 23 events in this list will be the initial focus of signal detection in the EU-ADR project. Although the prioritization of adverse events for drug safety monitoring that was done in this study was based on thorough evaluation of evidence from various sources of information by pharmacovigilance experts, the list of events and their ranking is by no means definitive and is intended to be dynamic. As our knowledge on drug safety evolves and new issues in drug safety arise, the list will need to be updated.

Table 3. Ranked list of events.

	Trigger for drug withdrawal	Trigger for black box warning	Requiring ED visit or hospital admission	Probability of event to be drug-related	Likelihood of death	Total score
Bullous eruptions	3	3	3	3	3	15
Acute renal failure	3	3	3	2	3	14
Acute myocardial infarction	2	3	3	2	3	13
Anaphylactic shock	2	3	3	2	3	13
Rhabdomyolysis	1	3	3	2	3	12
Aplastic anemia/pancytopenia	2	3	3	2	2	12
Neutropenia	2	3	3	2	2	12
Cardiac valve fibrosis	3	1	2	2	3	11
Acute liver injury	3	3	2	1	1	10
Extrapyramidal disorders	0	1	3	3	2	9
QT prolongation	3	2	2	1	1	9
Suicidal behavior/attempt	1	3	1	1	3	9
Confusional state	0	3	3	2	1	9
Thrombocypenia	1	3	1	2	1	8
Upper gastrointestinal bleeding	0	3	2	2	1	8
Convulsions	1	2	2	1	2	8
Peripheral neuropathy	1	1	1	3	1	7
Maculopapular erythematous eruptions	1	1	1	3	1	7
Venous thrombosis	1	2	2	1	1	7
Mood changes: depression and mania	0	3	2	1	1	7
Amnesias	1	3	2	1	0	7
Hemolytic anemia	1	1	1	1	2	6
Acute pancreatitis	0	1	2	1	2	6

REFERENCES

1. Stahl M, Edwards IR, Bowring G, Kiuru A, Lindquist M. Assessing the impact of drug safety signals from the WHO database presented in 'SIGNAL': results from a questionnaire of National Pharmacovigilance Centres. *Drug Saf* 2003;26:721–727.
2. The Uppsala Monitoring Centre. Practical Pharmacovigilance. Signal Definition. http://www.who-umc.org/DynPage.aspx?id=22676#signaldef. (accessed on 5 February 2009).
3. Rodriguez EM, Staffa JA, Graham DJ. The role of databases in drug postmarketing surveillance. *Pharmacoepidemiol Drug Saf* 2001;10:407-410.
4. Meyboom RH, Lindquist M, Egberts AC, Edwards IR. Signal selection and follow-up in pharmacovigilance. *Drug Saf* 2002;25:459–465.
5. Avorn J. Evaluating drug effects in the post-Vioxx world: there must be a better way. *Circulation* 2006;113: 2173–2176.
6. Wadman M. Experts call for active surveillance of drug safety. *Nature* 2007;446:358–359.
7. Edwards IR. The accelerating need for pharmacovigilance. J R Coll Physicians Lond 2000;34:48–51.
8. EMEA. European Risk Management Strategy: 2008–2009 work programme adopted. London, 2 January 2008. http://www.hma.eu/uploads/media/ERMS_2008-2009_work_programme_adopted.pdf (accessed on 3 February 2009).
9. FDA. The Future of Drug Safety – Promoting and Protecting the Health of the Public. January 2007. http://www.fda.gov/oc/reports/iom013007.pdf (accessed on 3 February 2009).
10. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005; 58: 323–337.
11. Brown JS, KulldorffM, Chan KA, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. *Pharmacoepidemiol Drug Saf* 2007; 16: 1275–1284.
12. Vandenbroucke JP, Psaty BM. Benefits and risks of drug treatments: how to combine the best evidence on benefits with the best data about adverse effects. *JAMA* 2008;300:2417–2419.
13. Vandenbroucke JP. What is the best evidence for determining harms of medical treatment? *CMAJ* 2006;174: 645–646.
14. Kaufman DW, Shapiro S. Epidemiologic assessment of drug-induced disease. *Lancet* 2000;356:1339–1343.
15. Vray M, Hamelin B, Jaillon P, Round Table No. 2, Giens XX. The respective roles of controlled clinical trials and cohort monitoring studies in the pre- and postmarketing assessment of drugs. *Therapie* 2005;60:339–349.
16. Brown JS, Kulldorf M, Petronis KR, et al. Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. *Pharmacoepidemiol Drug Saf* 2009;18:226–234.
17. Juni P, Nartey L, Reichenbach S, Sterchi R, Dieppe PA, Egger M. Risk of cardiovascular events and rofecoxib: cumulative meta-analysis. *Lancet* 2004;364:2021–2029.
18. McClellan M. Drug safety reform at the FDA – pendulum swing or systematic improvement? *New Engl J Med* 2007;356:1700–1702.
19. Platt R. The future of drug safety: challenges for FDA. Presented at the Institute of Medicine Forum. Washington DC, 12 March 2007.
20. Edwards IR. Adverse drug reactions: finding the needle in the haystack. *Br Med J* 1997;315:500.
21. Bate A, Lindquist M, Orre R, Edwards IR, Meyboom RH. Data-mining analyses of pharmacovigilance signals in relation to relevant comparison drugs. *Eur J Clin Pharmacol* 2002;58:483–490.
22. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. *Drug Saf* 2003; 26:159–186.
23. Wilson AM, Thabane L, Holbrook A. Application of data mining techniques in pharmacovigilance. *Br J Clin Pharmacol* 2004;57:127–134.
24. Hauben M, Madigan D, Gerrits CM, et al. The role of data mining in pharmacovigilance. *Expert Opin Drug Saf* 2005;4:929–948.

25. i3 Aperio1 signal detection and assessment tool. http://www.i3global.com/Solutions/PharmaInformatics/i3Aperio/ (accessed on 3 February 2009).
26. Hauben M, Reich L, Van Puijenbroek EP, Gerrits CM, Patadia VK. Data mining in pharmacovigilance: lessons from phantom ships. *Eur J Clin Pharmacol* 2006;62:967–970.
27. Smyth P. Data mining: data analysis on a grand scale? *Stat Methods Med Res* 2000;9:309–327.
28. Moore N, Hall G, Sturkenboom M, Mann R, Lagnaoui R, Bégaud B. Biases affecting the proportional reporting ratio (PRR) in spontaneous reports pharmacovigilance databases: the example of sertindole. *Pharmacoepidemiol Drug Saf* 2003;12:271–281.
29. Platt R, Madre L, Reynolds R, Tilson H. Active drug safety surveillance: a tool to improve public health. *Pharmacoepidemiol Drug Saf* 2008;17:1175–1182.
30. European Medicines Agency. Valdecoxib withdrawal. http://www.emea. europa.eu/pdfs/human/press/pus/20480204en.pdf (accessed on 20 November 2008).
31. Lexchin J. Drug withdrawals from the Canadian market for safety reasons, 1963–2004. *CMAJ* 2005;172: 765–767.
32. Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969–2002: the importance of reporting suspected reactions. *Arch Intern Med* 2005;165:1363–1369.
33. 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;128:35–44.
34. Valeyrie-Allanore L, Poulalhon N, Fagot JP, et al. Stevens–Johnson syndrome and toxic epidermal necrolysis induced by amifostine during head and neck radiotherapy. *Radiother Oncol* 2008;87:300–303.
35. Roujeau JC, Stern RS. Severe adverse cutaneous reactions to drugs. *N Engl J Med* 1994; 331: 1272–1285.
36. Valeyrie-Allanore L, Sassolas B, Roujeau JC. Drug-induced skin, nail and hair disorders. *Drug Saf* 2007; 30: 1011–1030.
37. Thomsen LA, Winterstein AG, Søndergaard B, Haugbølle LS, Melander A. Systematic review of the incidence and characteristics of preventable adverse drug events in ambulatory care. *Ann Pharmacother* 2007;41: 1411–1426.
38. Trifiro` G, Calogero G, Ippolito FM, et al. Adverse drug events in emergency department population: a prospective Italian study. *Pharmacoepidemiol Drug Saf* 2005;14:333–340.
39. Lagnaoui R, Moore N, Fach J, Longy-Boursier M, Bégaud B. Adverse drug reactions in a department of systemic diseases-oriented internal medicine: prevalence, incidence, direct costs and avoidability. *Eur J Clin Pharmacol* 2000;55:181–186.
40. Budnitz DS, Pollock DA,Weidenbach KN, Mendelsohn AB, Schroeder TJ, Annest JL. National surveillance of emergency department visits for outpatient adverse drug events. *JAMA* 2006;296:1858–1866.
41. Gandhi TK, Weingart SN, Borus J, et al. Adverse drug events in ambulatory care. *N Engl J Med* 2003;348: 1556–1564.
42. Pirmohamed M, James S, Meakin S, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;329:15–19.
43. Bonn D. Adverse drug reactions remain a major cause of death. *Lancet* 1998;351:1183.
44. Chyka PA. How many deaths occur annually from adverse drug reactions in the United States? *Am J Med* 2000; 109: 122–1230.
45. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalised patients: a meta-analysis of prospective studies. *JAMA* 1998;279:1200–1205.
46. Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. *Ann Pharmacother* 2008;42:1017–1025.
47. Davies EC, Green CF, Mottram DR, Pirmohamed M. Adverse drug reactions in hospitals: a narrative review. *Curr Drug Saf* 2007;2:79–87.
48. Barrow P, Waller P, Wise L. Comparison of hospital episodes with 'drug-induced' disorders and spontaneously reported adverse drug reactions. *Br J Clin Pharmacol* 2005;61:233–237.

49. FDA Center for Drug Evaluation and Research 2005 Report to the Nation – Improving Public Health through Human Drugs. http://www.fda.gov/CDER/reports/rtn/2005/rtn2005.PDF (accessed on 3 February 2009).

50. Pirmohamed M, Breckenridge AM, Kitteringham NR, Park BK. Adverse drug reactions. *Br Med J* 1998;316: 1295–1298.

51. Lee A, Rawlins MD. Adverse drug reactions. In Clinical Pharmacy and Therapeutics (3rd edn), Walker R, Edwards C (eds). Churchill Livingstone: London, 2002;33–46.

52. Olivier P, Montastruc JL. The nature of the scientific evidence leading to drug withdrawals for pharmacovigilance reasons in France. Pharmacoepidemiol *Drug Saf* 2006;15:808–812.

APPENDIX

The following pharmacovigilance reference books and articles were used for the score assignment.

General

1. Black box warnings (from FDA, by product): http://www.formularyproductions.com/master/showpage.php?dir¼blackbox&whichpage¼9 (Accessed 20 September, 2008).
2. Budnitz DS, Shehab N, Kegler SR, et al. Medication use leading to emergency department visits for adverse drug events in older adults. *Ann Intern Med* 2007;147:755-65.
3. Easton KL, Chapman CB and Brien JA. Frequency and characteristics of hospital admissions associated with drug-related problems in paediatrics. *Br J Clin Pharmacol* 2004;575:611-5.
4. Franciotti D. Hospital admissions in Italy for liver and gastrointestinal diseases in the period 1999-2002: an epidemiological study. *Eur J Gastroenterol Hepatol* 2007;19:763-7.
5. Harrison's Principles of Internal Medicine: http://www.accessmedicine.com/resourceTOC.aspx?resourceID ¼4(Accessed 20 September, 2008).
6. Le J, Nguyen T, Law AV, et al. Adverse drug reactions among children over a 10-year period. *Pediatrics*. 2006; 118:555-62.
7. Olivier P and Montastruc JL. The nature of the scientific evidence leading to drug withdrawals for pharmacovigilance reasons in France. *Pharmacoepidemiol Drug Saf* 2006;15:808-12.
8. Onder G, Pedone C, Landi F, et al. Adverse drug reactions as cause of hospital admissions: results from the Italian Group of Pharmacoepidemiology in the Elderly (GIFA). *J Am Geriatr Soc* 2002;50:1962-8.
9. Patel H, Bell D, Molokhia M, et al. Trends in hospital admissions for adverse drug reactions in England: analysis of national hospital episode statistics 1998-2005. *BMC Clin Pharmacol* 2007;7:9.
10. Pirmohamed M, James S, Meakin S, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;329:15-9.
11. Pouyanne P, Haramburu F, Imbs JL, et al. Admissions to hospital caused by adverse drug reactions: cross sectional incidence study. French Pharmacovigilance Centres. *BMJ* 2000; 20:1036.
12. Recall and safety alerts FDA: http://www.fda.gov/oc/po/firmrecalls/archive_2003.html (Accessed 20 September, 2008).
13. Schneeweiss S, Gottler M, Hasford J, et al. First results from an intensified monitoring system to estimate drug related hospital admissions. *Br J Clin Pharmacol* 2001;52:196-200.
14. Schneeweiss S, Hasford J, Gottler M, et al. Admissions caused by adverse drug events to internal medicine and emergency departments in hospitals: a longitudinal population-based study. *Eur J Clin Pharmacol* 2002; 58:285–91.
15. van der Hooft CS, Dieleman JP, Siemes C, et al. Adverse drug reaction-related hospitalisations: a population-based cohort study. *Pharmacoepidemiol Drug Saf* 2008;17:365-71.
16. van der Hooft CS, Sturkenboom MC, van Grootheest K, et al. Adverse drug reaction-related hospitalisations: a nationwide study in The Netherlands. *Drug Saf* 2006;29:161-8.
17. Wagner AK, Chan KA, Dashevsky I, et al. FDA drug prescribing warnings: is the black box half empty or half full? *Pharmacoepidemiol Drug Saf* 2006;15:369-86.
18. Waller P, Shaw M, Ho D, et al. Hospital admissions for 'drug-induced' disorders in England: a study using the Hospital Episodes Statistics (HES) database. *Br J Clin Pharmacol* 2005;59: 213-9.
19. Wysowski DK and Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. *Arch Intern Med* 2005;165:1363-9.

Hematologic

1. Andersohn F, Konzen C, Garbe E. Systematic review: agranulocytosis induced by nonchemotherapy drugs. *Ann Intern Med* 2007;146:657-65.
2. Brodsky RA and Jones RJ. Aplastic Anemia. *Lancet* 2005;365:1647-56.
3. George JN, Raskob GE, Shah SR, et al. Drug-induced thrombocytopenia: a systematic review of published case reports. *Ann Intern Med* 1998;129:886-90.
4. Ibañez L, Vidal X, Ballarin E, et al. Population-based drug-induced agranulocytosis. *Arch Intern Med* 2005; 165:869-74.
5. Johnson ST, Fueger JT, Gottschall JL. One center's experience: the serology and drugs associated with drug-induced immune hemolytic anemia–a new paradigm. *Transfusion* 2007;47:697-702.
6. Lechner K. Management of adult immune thrombocytopenia. *Rev Clin Exp Hematol* 2001;5:222-35.
7. Montane E, Ibañez L, Vidal X, et al. Catalan Group for Study of Agranulocytosis and Aplastic Anemia. Epidemiology of aplastic anemia: a prospective multicenter study. *Haematologica* 2008;93:518-23.
8. Van den Bemt PM, Meyboom RH, Egberts AC. Drug-induced immune thrombocytopenia. *Drug Saf* 2004; 27:1243-52.
9. Young NS, Calado TR, Scheinberg CP. Current concepts in the athophysiology and treatment of aplastic anemia. *Blood* 2006;108:2509-19.
10. Young NS. Acquired aplastic anemia. *Ann Intern Med* 2002; 136:534-6.

Cutaneous

1. Scheinfeld N. Imatinib mesylate and dermatology Part 2: a review of the cutaneous side effects of imatinib mesylate. *Dermatol Venereol* 2004;131:571-3.
2. Valeyrie-Allanore L, Sassolas B, Roujeau JC. Drug-induced skin, nail and hair disorders. *Drug Saf* 2007;30: 1011-30.
3. Yawalkar N. Drug-induced exanthems. *Toxicology* 2005;209:131-4.
4. Olivier P and Montastruc JL. The nature of the evidence leading to drug withdrawals for pharmacovigilance reasons in France. *Pharmacoepidemiol Drug Saf* 2006; 15:808-12.
5. United Nations. Consolidated list of products – whose consumption and/or sale have been banned, withdrawn, severely restricted or not approved by governments, Eight Issue – Pharmaceuticals. 2003; 581 pages.

Liver and gastrointestinal

1. Carey EJ, Vargas HE, Douglas DD, et al. Inpatient Admissions for Drug-induced Liver Injury: Results from a Single Center. *Dig Dis Sci* 2008;53:1977-82.
2. Gerstner T, Bűsing D, Bell N, et al. Valproic acid-induced pancreatitis: 16 new cases and a review of the literature. *J Gastroenterol* 2007;42:39-48.
3. Sgro C, Clinard F, Ouazir K, et al. Incidence of drug-induced hepatic injuries: a French population-based study. *Hepatology* 2002;36:451-5.

Cardiac and vascular

1. Antonini A and Poewe W. Fibrotic heart-valve reactions to dopamine-agonist treatment in Parkinson's disease. *Lancet Neurol* 2007;6:826–29.
2. Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. *J Clin Pharmacol* 2006;46:498-507.

3. Deitcher SR and Gomes MP. The risk of venous thromboembolic disease associated with adjuvant hormone therapy for breast carcinoma: a systematic review. *Cancer* 2004;101:439-49.
4. El Menyar AA. Drug-induced myocardial infarction secondary to coronary artery spasm in teenagers and young adults. J Postgrad Med 2006;52:51-6.
5. Roden DM. Drug-induced prolongation of the QT interval. *N Engl J Med* 2004;350:1013-22.
6. Ruzicka E, Lınkova´H, Penicka M, et al. Low incidence of restrictive valvulopathy in patients with Parkinson's disease on moderate dose of pergolide. *J Neurol* 2007;254:1575-8.
7. Vandenbroucke JP, Rosing J, Bloemenkamp KW, et al. Oral contraceptives and the risk of venous thrombosis. *N Engl J Med* 2001;344:1527-35.
8. Zanettini R, Antonini A, Gatto G, et al. Valvular Heart Disease and the Use of Dopamine Agonists for Parkinson's disease. *N Engl J Med* 2007;356:39-46.
9. Shah RR. Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labeling of a new chemical entity. *Fundam Clin Pharmacol* 2002;16:147-56.

Neurologic/musculoskeletal

1. Arne-Bes MC. Neuropathies me´dicamenteuses: mise a` jour re´cente des donne´es (1996 a` 2003). *Rev Med Liege* 2004;59:118-23.
2. Esper CD and Factor SA. Failure of Recognition of Drug-Induced Parkinsonism in the Elderly. Movement Disorders 2008;23:401-4.
3. Gentile S. Extrapyramidal Adverse Events Associated With Atypical Antipsychotic Treatment of Bipolar Disorder. *J Clin Psychopharmacol* 2007;27:35-45.
4. Guis S. Drug-induced and toxic myopathies. *Best Pract Res Clin Rheumatol* 2003;17:877–907.
5. Haddad PM and Dursunx SM. Neurological complications of psychiatric drugs: clinical features and management. *Hum Psychopharmacol Clin Exp* 2008;23:15-26.
6. Lee KC, Finley PR, Alldredge BK. Risk of seizures associated with psychotropic medications: emphasis on new drugs and new findings. *Expert Opin Drug Saf* 2003;2:233-47.
7. Rollins N, Winick N, Bash R, et al. Acute methotrexate neurotoxicity: findings on diffusion-weighted imaging and correlation with clinical outcome. *Am J Neuroradiol* 2004;25:1688-95.
8. Ronaldson KJ O'Shea JM, Boyd IW. Risk factors for rhabdomyolysis with simvastatin and atorvastatin. *Drug Saf* 2006;29:1061-7.
9. Sieb JP and Gillessen T. Iatrogenic and toxic myopathies. *Muscle Nerve* 2003;27:142-56.
10. SilvaM, Matthews ML, Jarvis C, et al. Meta-Analysis of Drug-Induced Adverse Events Associated with Intensive-Dose Statin Therapy. *Clinical Ther* 2007;29:253-60.
11. Thundiyil JG, Kerney TE, Olson KR. Evolving Epidemiology of Drug-Induced Seizures Reported to a Poison Control Center System. *J Med Toxicol* 2007;3:15-9.

Psychiatric

1. Adamis D, Treloar A, Darwiche FZ, et al. Associations of delirium with in-hospital and in 6-months mortality in elderly medical inpatients. *Age Ageing* 2007;36:644-9.
2. Kakuma R, du Fort GG, Arsenault L, et al. Delirium in older emergency department patients discharged home: effect on survival. *J Am Geriatr Soc* 2003;51:443-50.
3. Müller-Oerlinghausen B, Bergho¨fer A, Ahrens B. The antisuicidal and mortality-reducing effect of lithium prophylaxis: consequences for guidelines in clinical psychiatry. *Can J Psychiatry* 2003;48:433-9.
4. Sachs GS. Unmet clinical needs in bipolar disorder. *J Clin Psychopharmacol* 2003;23:S2-8.
5. Tondo L and Baldessarini RJ. Reduced suicide risk during lithium maintenance treatment. *J Clin Psychiatry* 2000;61:97-104.

Renal

1. Evenepoel P. Acute toxic renal failure. *Best Pract Res Clin Anaesthesiol* 2004;18:37-52.
2. Guo X and Nzerue C. How to prevent, recognize, and treat drug-induced nephrotoxicity. *Cleve Clin J Med*. 2002;69:289-97.
3. Pannu N and Nadim MK. An overview of drug-induced acute kidney injury. *Crit Care Med* 2008;36:S216-23.

Multi-systemic

1. Brown SG, Mullins RJ, Gold MS. Anaphylaxis: diagnosis and management. *Med J Aust* 2006;185:283-9.
2. Wang DY, Forslund C, Persson U, et al. Drug-attributed anaphylaxis. *Pharmacoepidemiol Drug Saf* 1998;7: 269-74.

3.2

Harmonization process for the identification of medical events in eight European healthcare databases: experience from the EU-ADR project

Paul Avillach, MD, PhD*[1,2], Preciosa M Coloma, MD, MSc*[3], Rosa Gini, PhD[4], Martijn Schuemie,PhD[3], Fleur Mougin,PhD[1], Jean-Charles Dufour, MD, PhD[2], Giampiero Mazzaglia, MD, PhD [5], Carlo Giaquinto, MD[6], Carla Fornari,MD, PhD[7], Ron Herings, PhD, [8] Julia Hippisley-Cox, MD[9], Mariam Molokhia, MD, PhD[10], Lars Pedersen, PhD[11], Annie Fourrier-Réglat, PharmD, PhD[12], Marius Fieschi, MD, PhD[2], Miriam Sturkenboom, PharmD, PhD[3,13], Johan van der Lei, MD, PhD[3], Antoine Pariente, MD, PhD[1,2], Gianluca Trifirò, MD, PhD[3,14], on behalf of the EU-ADR consortium
*equal contribution

[1]LESIM, ISPED, University Bordeaux Segalen, Bordeaux, France; [2]LERTIM, EA 3283, Faculté de Médecine, Université Aix Marseille 2, Marseille, France; [3]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands; [4]Regional Health Agency of ARS, Florence, Italy; [5]Health Search – Italian College of General Practitioners, Italy; [6]Pedianet – Societa' Servizi Telematici SRL, Italy; [7]Centre on Public Health – University of Milano-Bicocca, Italy; [8]PHARMO Coöperation UA, the Netherlands; [9]University of Nottingham, United Kingdom; [10]Department of Primary Care & Public Health Sciences, Kings College London, United Kingdom; [11]Aarhus University Hospital, Århus Sygehus, Denmark; [12]INSERM U 657, University Bordeaux Segalen, Bordeaux, France; [13]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands; [14]Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Italy

Submitted

ABSTRACT

Background and Objective. Data from electronic healthcare records (EHR) can be used to monitor drug safety, but in order to compare and pool data from different EHR databases, extraction of potential adverse events must be harmonized. In this paper we describe the procedure used for harmonizing the extraction from eight European EHR databases of five events of interest deemed to be important in pharmacovigilance: acute myocardial infarction (AMI); acute renal failure (ARF); anaphylactic shock (AS); bullous eruption (BE), and rhabdomyolysis (RHABD).

Design and Methodology. The databases comprise general practitioners' records and claims for hospitalization and other healthcare services. Clinical information is collected using four different disease terminologies and free text in two languages. The Unified Medical Language System was used to identify concepts and corresponding codes in each terminology. A common database model was used to share and pool data and verify the semantic basis of the event extraction queries. Feedback from the database holders was obtained at various stages to refine the extraction queries.

Measurements. Age-adjusted incidence rates were calculated to facilitate harmonization of data extraction across the databases; this was an iterative process.

Results. The study population comprised overall 19 647 445 individuals with 59 929 690 person-years (PYs) of follow-up. Age-adjusted incidence rates for the five events of interest across databases were as follows: (1) AMI: 60-148/100 000 PYs; (2) ARF: 3-49/100 000 PYs; (3) AS: 2-12/100 000 PYs; (4) BE: 2-17/100 000 PYs; and (5) RHABD: 0.1-8/100 000 PYs.

Conclusion. The iterative harmonization process enabled a more homogeneous identification of events across differently structured databases using different coding-based algorithms. This workflow can facilitate transparent and reproducible event extractions and understanding of differences between databases.

INTRODUCTION

Spontaneous reporting of adverse drug reactions (ADRs) is currently the main source of data to monitor drug safety after drug licensing. It relies on healthcare professionals' ability and willingness to identify and report any suspected ADR to a centralized (nationwide or international) pharmacovigilance system.[1] However, the underreporting of ADRs remains an important limitation; it is estimated that only 1 to 10% of ADRs are reported through this channel[2-3] and it is likely to be subject to recording and ascertainment biases. It is increasingly being recognized that drug safety surveillance can benefit from the wide availability of healthcare databases to complement spontaneous reporting systems (SRS) and overcome some of their shortcomings.[4-11]

In 2007, the US Congress directed the Food and Drug Administration (FDA) to create a new post-marketing surveillance system, the Sentinel System.[12] The desired goal is to use by 2012 electronic health data from 100 million subjects for the prospective and systematic safety monitoring of marketed medical products in real-life setting.[13] The Observational Medical Outcomes Partnership (OMOP) was created in the US to design a common framework with the aim of setting up a system for drug surveillance through data mining of electronic health records.[14] In Europe, several projects, such as the Pharmacoepidemiological Research on Outcomes of Therapeutics (PROTECT), have been recently funded to link healthcare databases throughout Europe under the umbrella organization of the Innovative Medicines Initiative (IMI).[15] Other European Union (EU)-funded projects in which multiple healthcare databases are combined together to address specific safety issues include: NSAIDs-related gastrointestinal and cardiovascular risks (SOS, http://www.sos-nsaids-project.org/), the arrhythmogenic risk of drugs (ARITMO – http://www.aritmo-project.org), and the safety of vaccines (VAESCO – http://www.vaesco.net).

In February 2008, the EU-funded project 'Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge' (EU-ADR) was launched. The aim of this project was to design, develop and validate a computerized system to process data from eight electronic healthcare record (EHR) databases and several biomedical databases for drug safety signal detection.[16-17] In this project an event-based approach was adopted where events of special interest in pharmacovigilance were evaluated for their association with all available drugs in the EHR databases. Each of the eight databases participating in the project has unique characteristics, depending on its primary objective and local function (i.e., administrative/claims, medical records) and contains medical information coded according to different languages and terminologies. For these reasons, queries for data extraction concerning potential adverse events have to be created based on local expertise. Due to structural, syntactic, and semantic heterogeneities of the databases participating in the EU-ADR project, it is not possible to construct a single query for data extraction that could be used as such in all databases. In this context of large-scale drug safety monitoring using EHRs, the event data extraction from different databases requires a harmonization, i.e., a process geared towards reaching a common definition and identification of events which is both clinically sound and agreeable to all stakeholders. In this paper we describe the harmonization process for the data extraction concerning five events deemed to be important in pharmacovigilance from eight different databases participating in the EU-ADR project.

METHODS

Data sources

The eight databases involved in the EU-ADR project contain information from the healthcare records of almost 20 million European citizens (Table 1). Health Search/CSD Patient (HSD, Italy), Integrated Primary Care Information (IPCI, the Netherlands), Pedianet (Italy), and QRESEARCH (United Kingdom) are general practice (GP) databases, where both clinical information and drug prescriptions are recorded. The Aarhus University Hospital Database (Aarhus, Denmark), the 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 a well-defined population are linked to a registry of hospital discharge diagnoses and various other registries. The databases are heterogeneous in both structure and content.[17-18] For this analysis (which was done in the beginning of the project), databases contributed data from the period 1996 to 2007. Four disease terminologies are used to code the clinical events in the eight databases: the International statistical Classification of Diseases and related health problems – 9th and 10th revisions (ICD9-CM[19] and ICD10[20]); the International Classification of Primary Care (ICPC);[21] and the READ CODE (RCD) classification.[22] Two GP databases also describe events in clinical notes using free text in either Dutch (IPCI) or Italian (HSD).

Linking disparate databases using a distributed network

A distributed database network approach was chosen in EU-ADR,[17] allowing database holders to maintain local control of their data, while reaching the goal of sharing data in a standardized manner. Without this control, database holders might be reluctant to participate in a large network, primarily because of concerns regarding privacy and data confidentiality.[23] The decentralized data storage avoids (or reduces) many of the security, proprietary, legal, and privacy concerns of data owners both at the institution and country levels. Moreover, this approach allows local database experts to keep the data within their protected environment, hence they may easily and more effectively trouble-shoot unexpected findings or data irregularities and inconsistencies.[24] Local experts are, naturally, in the best position to understand the context within which their own data are recorded.

Different distributed data approaches can be used when combining databases. In the OMOP (http://omop.fnih.org/) and Mini-Sentinel (http://mini-sentinel.org/data_activities/) projects, original data tables are transferred to a standardized table, which can then be queried for extraction of relevant information and analyzed using standardized scripts. Databases within the EU differ considerably in content, coding, as well as origin (see Table 2 for a description of the content that can be utilized for event extractions in the different databases).

Table 1 | Characteristics of healthcare databases participating in the EU-ADR project.

Characteristics	**Pedianet*** (Italy)	**Health Search/ CSD*** (Italy)	**Lombardy Regional** (Italy)	**Tuscany Regional** (Italy)	**IPCI** (Netherlands)	**PHARMO** (Netherlands)	**QRESEARCH** (UK)	**Aarhus** (Denmark)
	PEDIANET	**HSD**	**UNIMIB**	**ARS**	**IPCI**	**PHARMO**	**QRESEARCH**	**AARHUS**
Current source population	129 742 children	771 907	9 924 758	3 585 560	479 585	1 280 805	1 515 116 (based on a 20% sample)	1 959 972
Type of database	General Practice pediatric database	General Practice database	Data warehouse record linkage system with: 1. Registry inhabitants 2. Regional Drug dispensation records 3. Hospital claims database	Data warehouse record linkage system with: 1. Registry inhabitants 2. Regional Drug dispensation records 3. Hospital claims database 4. Death registry	General Practice database	Data warehouse record linkage system with: 1. Registry inhabitants 2. Regional Drug dispensation records 3. Hospital claims database 4. Lab values	General Practice clinical database	Data warehouse record linkage system with: 1. Registry inhabitants 2. Regional Drug dispensation records 3. Hospital claims database 4. Lab values 5. Death registry
Symptoms (Yes/No)	Yes, as free text/codes	Yes, as free text/codes	No	No	Yes, as free text/codes	Yes for some	Yes, as codes	No
Outpatient primary care diagnoses	Yes, as free text /codes	Yes Free text /codes	No	No	Yes, as free text /codes	No	Yes	No

Characteristics	**Pedianet*** (Italy)	**Health Search/ CSD*** (Italy)	**Lombardy Regional** (Italy)	**Tuscany Regional** (Italy)	**IPCI** (Netherlands)	**PHARMO** (Netherlands)	**QRESEARCH** (UK)	**Aarhus** (Denmark)
	PEDIANET	**HSD**	**UNIMIB**	**ARS**	**IPCI**	**PHARMO**	**QRESEARCH**	**AARHUS**
Outpatient specialist care diagnoses	Yes, as free text/codes	Yes	No	No	Yes	No	Yes	No
Hospital discharge diagnoses	Yes, as free text /codes	Yes, as free text /codes	Yes	Yes	Yes, as free text /codes	Yes	Yes	Yes
Diagnosis coding scheme	ICD-9CM	ICD-9CM	ICD-9CM	ICD-9CM	ICPC	ICD-9CM	RCDv2 and v3	ICD-10
Language of free text	Italian	Italian	*No free text*	*No free text*	Dutch	*No free text*	*No free text*	*No free text*
Diagnostic procedures	Yes	Yes	Yes	Yes	No	Yes for in-hospital interventions	Yes	Yes, in-hospital only
Laboratory tests	Yes	Yes	No	No	Yes	Yes (for a subset of patients)	Yes	Yes, in-hospital only

*In Italy children are cared for until 14 years of age by family pediatrician and in the subsequent years by general practitioners. **ICPC**: International Classification of Primary Care; **RCD**: READ CODE Classification; **ICD9-CM** and **ICD-10**: International Classification of Diseases – 9th revision Clinical Modification and 10th revision, respectively.

Table 2 | Database content and attributes used for event extraction.

Table	Fields	AARHUS	ARS	LOMBARDY	HSD	IPCI	PEDIANET	PHARMO	QRESEARCH
HOSP	**main diagnosis**	ICD10	ICD9-CM	ICD9-CM				ICD9CM	
	secondary diagnosis	ICD10	ICD9-CM	ICD9-CM				ICD9CM	
	Procedures	ICD10	ICD9-CM	ICD9-CM				ICD9CM	
DEATH	**cause of death**	ICD10	ICD9-CM						
GP	**Diagnosis**				ICD9-CM and free text (in Italian)	ICPC and free text (in Dutch)	ICD9-CM and free text (in Italian)		READ
	Specialist/ Hospital				free text	free text	free text		
	Lab				free text/ referrals	referrals	free text/referrals		READ
	Death				Text /ICD-9	Text /ICPC	Text /ICD-9CM		READ
LAB	**classification**	NPU				WCIA		WCIA	READ
	Result	NPU				numbers		numbers	numbers

HOSP: discharge summary from hospitalizations recorded by hospitals; Main diagnosis: principal cause that led to the hospitalization episodes; Secondary diagnoses: 5 or more fields that contain diagnoses that refer either to preexisting diseases or to complications that arose during the hospital stay
DEATH: registry of causes of death; Cause of death: principal cause of death
GP: information recorded by GPs during their clinical practice: Symptoms and/or physical examination and/or diagnosis; Specialist visit prescriptions and diagnoses; Laboratory results
LAB: information obtained from laboratories Classification of laboratory analysis
NPU : Nomenclature, Properties and Units.

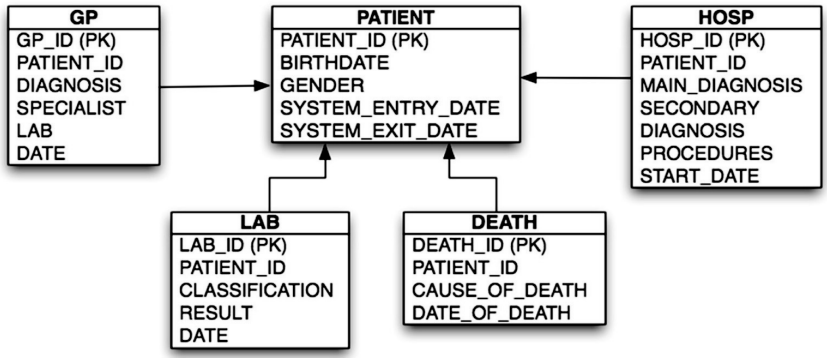


Figure 1 | Common database model.

HOSP: discharge summary from hospitalizations recorded by administrative/claims databases
DEATH: registry of causes of death
GP: information recorded by GPs during their clinical practice:
LAB: information obtained from laboratory test results

In EU-ADR simple common data files were utilized as basis for standardized analyses: files containing information on (1) patient demographics and follow-up time; (2) events; and (3) drug exposures. The common event files required combination of information from different files within each database (e.g., mortality, diagnosis, GP visits, procedures, laboratory test values). The three common data files are uploaded in a purpose-built Java-based software called Jerboa©. The software processes patient-level data to create aggregated, de-identified statistics which are then sent in an encrypted format to a central repository for evaluation and further analyses.[17]

Harmonization process for the event data extraction

The complete stepwise process adopted for the definition and harmonization of the queries for event data extraction is outlined in Figure 2. A task force was created within the EU-ADR project to spearhead the process of database query harmonization. This task force was composed of researchers with medical background and expertise in pharmacovigilance and medical informatics in addition to the database holders. Guidelines for the query harmonization process have been developed with the purpose of making this process as precise and transparent as possible.

Events. Within the EU-ADR project, 23 events of interest were identified as 'priority' events according to a ranking criteria based on relevance from a drug safety perspective.[16] For this paper, we describe the harmonization process for data extraction pertaining to the top-five ranked events: (1) acute myocardial infarction (AMI); (2) acute renal failure (ARF); (3) anaphylactic shock (AS); (4) bullous eruptions (BE); and (5) rhabdomyolysis (RHABD). For each event, a clinical definition was first provided in a structured event definition form (EDF) using medical textbooks and published guidelines of diagnostic criteria from scientific societies concerning the events of

interest. This definition was subsequently validated by medical specialists. The EDF for the event bullous eruptions (BE) is shown as an example in Appendix 1.

Mapping of terminologies. To overcome the differences in terminologies, we built a shared semantic foundation for the events by selecting disease concepts in the Unified Medical Language System® (UMLS, version 2008AA),[25-26] based on the medical definitions reported in the EDF. Since the four different terminologies encountered in the databases are part of the UMLS, the concepts could easily be projected into codes for the four disease coding terminologies used in the EU-ADR project. Examples of BE concepts are C0235818: Bullous eruption; C0014742: Erythema Multiforme; C0038325: Stevens-Johnson Syndrome (see Table 3 for the whole list). This part of the workflow has been previously described for the event upper gastrointestinal bleeding (UGIB).[27]

Revision of the concepts and related terms. The relevance of each medical concept, as well as the corresponding codes for each event of interest, were discussed via the EU-ADR consortium's web-based forum and during conference calls led by a team of experts and database owners. As a final result of this activity, database owners were provided with a list of concepts and corresponding codes and terms for each terminology, which were then used in the query for the event data extraction. The list of codes and terms provided was non-restrictive; database owners were free to employ additional codes or terms which they deemed pertinent in identifying a given event within their own database as long as these codes or terms were in line with the selected concepts. A few additional concepts with a lower specificity were identified; these concepts and related codes were used only in those databases where it was possible to refine the event search by taking into account these codes in combination with additional information (e.g., laboratory test results, free text specifications). For this reason, those concepts were marked as "only with refinement use" concepts. The list of concepts with the corresponding codes is presented in Table 3. An average number of five concepts were selected per event, ranging from three for ARF to nine for AMI.

Translation of concepts and coding algorithms into search queries. The common database model (Figure 2, Table 2), the concepts projection into different terminologies (Table 3) and the algorithms for the definition of the event of interest (Table 4) were used to build the queries in each database. The translation into the common algorithm language was done in the following manner: "search for concepts (Table 3) in the fields listed in Table 5 that belong to the database AND refine with the refinement strategy if applicable."

Query analysis. The database owners were asked to construct initial queries for the identification of events of interest, using the recommended codes and terms. All queries from each database for each event were analyzed in content (i.e., which codes or terms were used) and structure (i.e., which record(s) and corresponding field(s) from the database). Then each code, term and/or laboratory test result was mapped into the common database model. These queries were subsequently compared across the different databases. The query analysis was aimed at assessing the consistency across different databases with respect to the use of similar information (i.e., codes, free text, laboratory test results, query refinements) and to the search within the same type of data (e.g., primary and/or secondary hospital discharge diagnoses in claims databases).

If major differences in the query analysis were identified, a consensus was reached among the respective database holders to adopt similar strategies for the event detection.

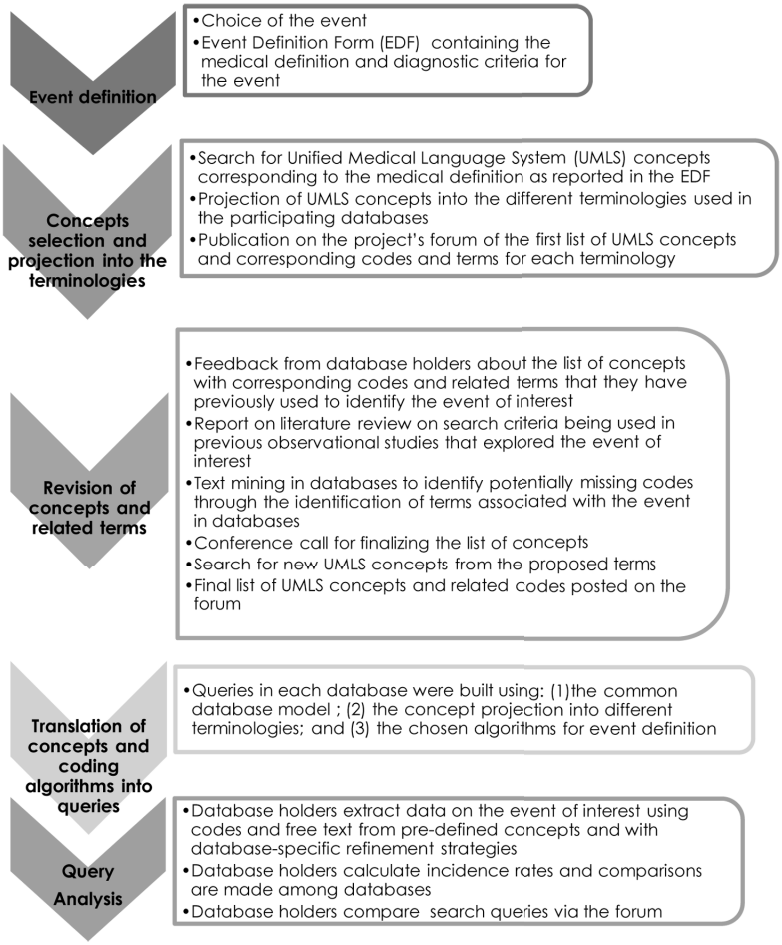


Figure 2 | Workflow of the harmonization of event queries across the different databases.

Evaluation of event data extraction. For each event, an internal evaluation within the database network was first performed through the evaluation of standardized and age-specific incidence rates (IRs) as estimated in different databases. The differences in IRs across databases can either be due to different strategies being used to extract the event or to differences in the source of information (e.g., GPs' electronic medical records vs. hospitalization records) and the underlying population (i.e., Italian vs. Dutch vs. Danish vs. British population; only pediatric population in Pedianet). To account for differences in age, IRs were standardized using the WHO World Standard Population as a reference.[28]

Table 3 | UMLS concepts projection into the four terminologies for the events: Acute myocardial infarction (AMI), Acute renal failure (ARF), Anaphylactic shock (AS), Bullous eruptions (BE) and Rhabdomyolysis (RHABD).

Event	UMLS Concept	Preferred term	ICD9CM	ICD10	RCD	ICPC
AMI	C0155626	Acute myocardial infarction	410.x	I21.x	G30z.,XE0Uh	K75, K75002
	C0428953	ECG: myocardial infarction			323.., 323Z.	
	C0232320	ECG: antero-septal infarct			3233	
	C0428956	ECG:posterior/inferior infarct			3234	
	C0428955	ECG: subendocardial infarct			3235	
	C0232325	ECG: lateral infarction			3236	
	C0428953	ECG: myocardial infarction			323.., 323Z.	
	C0340324	Silent myocardial infarction			X200a	
	C0340283 *Only for refinement use**	Other acute and subacute ischemic heart disease NOS	411		G31..,G31yz	
ARF	C0022660	Kidney Failure, Acute	584	N17		U99005
	C0022672	Kidney Tubular Necrosis, Acute	584.5	N17.0	K040.	U05001
	C0003460 *Only for refinement use**	Anuria	788.5	R34	1AC0., R0851, R34	
AS	C0002792	Anaphylactic shock		T78.2	SN50	A12004, A92005
	C0375697	Other anaphylactic shock	995			
	C0685898	Anaphylactic shock due to adverse food reaction	995.6	T78.0	X70vm, X70w1	
	C0161840	Anaphylactic shock due to serum	999.4	T80.5	SP34, X70vl	
	C0274304	Anaphylactic shock, due to adverse effect of correct medicinal substance properly administered		T88.6	SN501	

Event	UMLS Concept	Preferred term	ICD9CM	ICD10	RCD	ICPC
BE	C0235818	Bullous eruption			XM05i	
	C0014742	Erythema Multiforme	695.1	L51	M151z,XE1B0	S99007
	C0038325	Stevens-Johnson Syndrome	695.1, 695.14, 695.15		M1517, X50CE	A12005, S99032
	C0014518	Toxic Epidermal Necrolysis	695.1, 695.14, 695.15	L51.2	M1518	
	C0085932	Skin Diseases, Bullous	694	L10-L14.9	M14.., M14z., My1.,	
RHABD	C0035410	Rhabdomyolysis	728.88	M62.8	X70AI	
	C1135344	Acute necrotizing myopathy	359.81			
	C1401301	ischemia; muscle				
	C0027080	Myoglobinuria	791.3	R82.1	X709S, R113	

*NOTE: These concepts and related codes were used only in some databases wherein it was possible to increase their specificity through a refinement of the query for the event data extraction, by including additional information (i.e., laboratory test results, free-text specifications).

Table 4 | Summary of coding algorithms for event data extraction for AMI, ARF, AS, BE and RHABD.

Event	Chosen fields	Conditions for extractions from LAB or GP-lab	Refinement of concepts	Other refinement
AMI	GP-diag HOSP-Main DEATH		Other acute and subacute ischemic heart disease NOS concept refined by HSD through free text	Case 1
ARF	GP-diag; HOSP-Main DEATH		Anuria concept refined by AARHUS, HSD, IPCI and PHARMO through lab values: elevated serum creatinine within 30 days of date recorded diagnosis	Case 1
AS	GP-diag HOSP-Main			Case 1
BE	GP-diag; HOSP-Main			Case 2
RHABD	GP-diag GP-lab HOSP-Main HOSP-Sec LAB	serum CK > 10xnormal value within 30 days of date of recorded diagnosis		Case 2

Definitions of fields are in legend of Table 2
Case 1: ARS and AARHUS discarded duplicates coming from both DEATH and HOSP. HSD and IPCI discarded through free text algorithms records referring to chronic events, familial cases, and prevention. PEDIANET did manual validation.
Case 2: HSD and IPCI discarded through free text algorithms records referring to chronic events, familial cases, and prevention. PEDIANET performed manual validation of all cases.
AMI: Acute myocardial infarction
ARF: Acute renal failure
AS: Anaphylactic shock
BE: Cutaneous bullous eruption
RHABD: Rhabdomyolysis

Database holders ran several queries for each event: a basic query with a search in the primary hospital discharge diagnosis field of an administrative database record or the diagnosis field in a GP database record. Thereafter, queries employing other sources of data were added to the initial query, as applicable: 1) secondary diagnoses in hospitalization data; 2) death registry; and 3) laboratory results or corresponding UMLS concepts with refinement. The impact of the additional data source on the incidence rates was then analyzed. Not all possible additional queries were relevant for each event (e.g., no laboratory results were available to identify the event AS). The final agreement as to which query each database would adopt for each event data extraction was reached on the grounds of the following criteria: databases having a similar structure search in the same information sources but databases having broader sources of information can exploit them as much as possible, provided that the resulting IR is consistent with what is described in the literature and with the IRs obtained from the other databases. After internal comparisons among the EU-ADR databases, a literature review was conducted for each event and the observed IRs were compared to the IRs estimated in previous epidemiologic studies. For the literature review, the specified diagnosis was entered as a keyword in PubMed and articles which estimated IRs

were retrieved.[29] We gave the priority to recent articles, describing age-specific IRs and whenever possible, concerning the same underlying population as in the EU-ADR databases.

Table 5 | Comparison of age-standardized incidence rates (per 100 000 PY) for AMI, ARF, AS based on basic queries and the effect of additional information.

		IR for basic query		Additional data: IR and increment	
Event	**DB**	HOSP-main	GP	add DEATH	add concept with refinement
AMI	AARHUS	101.4		126.5 (+25%)	
	ARS	77.8		90.2 (+15%)	
	HSD		58.7		59.1 (+0.5%)
	IPCI		148.4		
	PHARMO	93.4			
	LOMBARDY	82.5			
ARF	AARHUS	6.3		7.1 (+19%)	17.9 (+150%)
	ARS	12.1		12.6 (+4%)	
	HSD		3.2		3.3 (+3%)
	IPCI		48.9		
	PHARMO	2.3			3.6 (+52%)
	LOMBARDY	15.4			
		IR for basic Query		**Additional data: IR and increment**	
Event	**DB**	HOSP-main	GP	add HOSP-sec	add DEATH
AS	AARHUS	5.7		6.4 (+12%)	6.4 (+0%)
	ARS	12.0		12.7 (+6%)	12.8 (+0%)
	HSD		5.2		
	IPCI		7.9		
	PHARMO	1.9		2.4 (+26%)	
	LOMBARDY	2.2		2.8 (+27%)	

RESULTS

The study population of the EU-ADR network for this analysis comprised overall 19 647 445 individuals with 59 929690 person-years (PYs) of follow-up. After discussions with all database holders, query recommendations were proposed to facilitate harmonization of the process of event data extraction. As an example in Appendix 2, the query analysis for AS (finalized after consensus discussion) is shown. The following are the age-adjusted IRs obtained using the harmonized queries for the five events of interest across the databases: (1) AMI: 60-148/100 000 PYs; (2) ARF: 3-49/100 000 PYs; (3) AS: 2-12/100 000 PYs; (4) BE: 2-17/100 000 PYs; and (5) RHABD: 0.1-8/100 000 PYs. The age-specific incidence rates per database for all events are illustrated in Appendix 3.

Review and harmonization of the event extraction processes changed the age-adjusted IRs across the databases to various extents. This is shown for three of the events of interest in Table 5. For AMI, a considerable increase in the incidence rates (25% and 15% in Aarhus and ARS, respectively) resulted from the inclusion of records extracted from death registries. For ARF, this same inclusion also resulted in an increase in IR (19% and 4% in Aarhus and ARS, respectively), although introduction of a concept with refinement had a higher impact on the IR for PHARMO (*'anuria'*- 52% increase). For AS, the information from death registries gave no additional contribution to the extraction of cases in Aarhus or to ARS, while search of specific codes within secondary discharge diagnoses had additional value for ARS (6% increase) and for Lombardy (27% increase).

DISCUSSION

This study shows how event extractions may differ across databases and how different choices impact on the estimated incidence of the event. We described a workflow that has been successful for combining data across databases of various origins and constructs in the context of drug safety signal detection. The UMLS was used as the common terminological system to map events across different terminologies. The knowledge described in the various terminologies, which are included in the UMLS, was inadequate to define all the clinical aspects of an event for our purposes, and so expert knowledge and experience from the database holders were necessary to build a more comprehensive definition of the event. Use of EHR databases requires an understanding of how the healthcare data are generated from the initial patient encounter all the way to completion of the database entry.[30]

Inspection of the queries for the event data extraction revealed two main types of differences. The first is due to deviation from proposed concepts, resulting mainly from a database holder finding additional value in using other concepts that were found to be useful based on their own previous studies. Databases with free text information refined the queries according to *ad hoc* search algorithms, while databases with information on laboratory test results used the numeric value associated with a diagnostic concept. As an example, 'anuria' was considered only if a value of serum creatinine was greater than 4 mg/dL and if the examination was done within 30 days of the date of diagnosis. Using the concept 'anuria' alone, without any laboratory results, would have decreased the specificity of the query. The second type of difference is due to modifications in the source of information considered in the query. Some databases searched only for codes in the 'primary hospital discharge diagnosis,' field while others searched for concepts in both 'primary hospital discharge diagnosis' and 'secondary hospital discharge diagnosis.' Moreover, when using laboratory test findings, the threshold values employed by the databases were not always compatible.

The small difference in the IR for AMI observed between the two Italian regional claims databases is probably explained by the fact that Lombardy could not contribute AMI-related

deaths occurring outside of the hospital, while the lower IR observed in the Italian GP database is probably due to the non-routine recording of hospital deaths (including those attributed to AMI) by the GP. In line with this hypothesis, similar estimates were observed in another GP database, QRESEARCH from the UK (67.4/100 000 PYs). Despite using a search strategy similar to the one used by ARS, in the Danish claims database (Aarhus), a higher IR (126.5/100 000 PYs) was observed. This disparity may be due to inherent differences between the two underlying populations (Italian and Danish), as a consequence of the so-called south-north trend in cardiovascular diseases, largely attributable to the Mediterranean diet.[31-34] The same trend could probably explain the difference between Lombardy and PHARMO (93.4/100 000 PYs), which used the same query but have different populations. The highest IR for AMI was observed in the Dutch GP database IPCI. This is probably an overestimation since the IR observed in the Dutch administrative database PHARMO is much lower. This overestimation is most likely due to an extensive use of free text in the search strategy of IPCI. Case validation by manual review of medical records is currently ongoing to assess and quantify the effect of event misclassification and how this would translate in the context of drug safety signal detection.

The pattern of estimates for RHABD was noted to be quite different from the other events. This particular event is not captured by the administrative databases that are unable to link to a data source having laboratory results (i.e., the two Italian regional databases, ARS and Lombardy). It is also important to note that rhabdomyolysis is part of a spectrum of conditions (myopathy → rhabdomyolysis → renal failure) and is more of a clinical manifestation than an actual diagnosis. Thus, if patients are admitted to a hospital for rhabdomyolysis due to whatever cause, by the time they are discharged, the etiology of the rhabdomyolysis has already been found and this is what is recorded as the discharge diagnosis (e.g., trauma, burns, sepsis, poisoning). Finally, if the rhabdomyolysis progresses to acute renal failure, then the case is likely to be registered as ARF.

Pedianet is an Italian nationwide database which contains medical information concerning children until 14 years of age, as recorded by family pediatrician. Hence, the IRs obtained from this database cannot be directly compared with those of the other databases, although the data can be explored to further study specific differences within the pediatric populations of each database.

Comparison with incidence rates in published literature

In the Spanish EPIC cohort study[35] involving over 33 000 individuals with ~300 000 PYs of follow-up, age-standardized incidence rates of AMI were found to be in the range of 302-330/100 000 PYs in men and 60-114/100 000 PYs in women. In a large US community-based population, the age and sex-adjusted incidence of AMI was found to be 208/100 000 PYs.[36] Those figures are higher than those we obtained in EU-ADR (59.1 to 148.4 per 100 000 PYs) most likely because these studies employed broader search queries and other sources of information aside from coded hospital diagnoses, death registries, and free text. In the Spanish study, ICD9CM diagnostic codes 410-414 and ICD10 I20-I25 codes (all codes for ischemic heart disease), as well as ICD9 procedure codes for stent placement and bypass operation (36.0 and 36.1, respectively), were used. Cases were also ascertained by means of self-reported questionnaires, population-based specific

AMI registries, and autopsy data. The US study included only individuals who were 30 years of age or older. There are conflicting views as to whether coronary heart disease, or associated cardiovascular risk factors, is more prevalent in North America or in Europe.[37-40] When defining the extraction strategy in EU-ADR, specificity of the query took precedence over sensitivity so as to avoid having too many false positive drug safety signals. For example, in identifying the event AMI, we did not use the concept 'myocardial infarction,' but rather '*acute* myocardial infarction.' In other publications, the approach is to have a much broader definition of an event, comprising the whole spectrum of the acute coronary syndrome, which includes unstable angina.[41-43]

There are very few population-based studies on the incidence of ARF; most are focused on special populations such as the elderly, patients in intensive care units (ICUs), or those requiring renal replacement therapy.[44-45] It has been suggested that ARF is nearly as common as myocardial infarction.[46] Indeed, a population-based study in Scotland obtained an IR of 181/100 000 PYs.[47] This study employed only laboratory values in the identification of cases (ARF defined as having baseline serum creatinine below the threshold (150 µmol/L in men or 130 µmol/L in women) that subsequently increased by a factor of 1.5 or more, or the glomerular filtration rate (GFR) was reduced by at least 25%). The lower estimates we obtained in EU-ADR (3.3-48.9/100,000 PYs) are probably due to the fact that laboratory data were not uniformly available in all the databases to supplement the coded diagnoses.

Most of the published studies on the incidence of AS are based on emergency room visits or hospitalizations, which makes direct comparisons difficult.[48-51] A Swiss study showed that the incidence of life-threatening anaphylaxis ranged from 7.9-9.6/100 000 PYs, which is not far off from the estimates we obtained in EU-ADR (1.9-12.1/100 000 PYs).

Toxic Epidermal Necrolysis (TEN) and Stevens Johnson Syndrome (SJS), which collectively comprise the event BE, are rare conditions occurring worldwide and most often in adults, women more likely than men. Its incidence is estimated at 2-3 cases/million population/year in Europe but is up to three times higher in the HIV-infected population.[52] The incidence rates we obtained in EU-ADR range from 16-178 per million PYs and are not directly comparable because these rates reflect the incidence rates of both TEN and SJS as well as of other forms of bullous eruptions such as erythema multiforme. Diagnosis of TEN/SJS usually requires histologic confirmation,[53-54] information that is not uniformly available in the databases in EU-ADR.

As discussed earlier, estimating the incidence of RHABD is difficult because it is part of a spectrum of conditions that start with myopathy progressing to muscle necrosis and renal failure, and most studies in the literature investigate the endpoint renal failure or the entire spectrum.[55-56] Thus, use of diagnostic codes alone, without values of serum creatine phosphokinase (CPK), may underestimate the true incidence. At the same time, with case identification algorithms that employ CPK values there is a need to distinguish elevations due to cardiac ischemia or infarction (usually detected by the CPK-MB isoenzyme) with those due to skeletal muscle injury (detected by the CPK-MM isoenzyme which, although predominantly found in skeletal muscle, is also the primary CPK isoenzyme present in heart muscle). We only employed CPK-total in the search algorithms.

Coding changes in international disease classification have posed new challenges for the comparability of indicators for various diseases. For example, in the worldwide MONICA Project which studied fatal and nonfatal coronary events through population-based registers, different versions of the ICD were used in different countries.[37] This is in addition to the innovations made in the last decade with respect to technologies that have enabled diagnosis of disease at earlier stages (such as novel biomarkers).[57-59] Healthcare systems and physician practices that vary between countries may also play a role in how clinical outcomes are captured in databases.[60] All these factors need to be considered when analyzing trends in disease frequency, severity, prognosis and subsequent variations in medical practice. It is hoped that the harmonization process we have described will enable the identification of outcomes across differently structured databases using compatible definitions and facilitate better comparison of IRs among various data sources and countries.

Limitations

While the aim of our study was not to estimate the most accurate IR for each event, the harmonization process enabled estimation and evaluation of rates of events with a common definition across databases having a similar structure. This work is based on currently available data which does not capture sources of bias and residual differences, including the effects of immigration and ethnic variation. Case validation using manual review of hospitalization records and GP records remains an indispensable part of the process.

CONCLUSION

No single data source is likely to be sufficient to meet all expected drug safety analysis needs, hence it is valuable to assess the feasibility and utility of analyzing multiple data sources concurrently. We provided an external shared semantic basis in content and structure for the creation of queries adapted to the heterogeneous EHR databases within the EU-ADR network. The iterative harmonization process enabled a more homogeneous identification of events across differently structured databases using different coding schemes. This workflow can facilitate transparent and reproducible extraction of events using EHR databases as well as a better understanding of differences between databases.

APPENDICES

Appendix 1 | Event definition form for the Bullous Eruption event

1 Definition

Erythema multiforme, Stevens-Johnson syndrome and toxic epidermal necrolysis are conditions characterized by blisters (bullous reactions; bullae are circumscribed, elevated lesions, containing fluid, that are larger than 5 mm.). Erythema multiforme is not necessarily caused by drugs, while Stevens-Johnson syndrome and toxic epidermal necrolysis in general are adverse drug reactions.

In some countries, the term erythema exudativum or erythema exudativum multiforme is used as a synonym of erythema multiforme. The term Lyell's syndrome is considered a synonym of toxic epidermal necrolysis. *Erythema multiforme* is an acute disease characterized by symmetrically distributed papular lesions affecting mainly the extremities, often with mucosal erosions. The typical lesion is target-shaped: it is concentrically organized with three different-colored zones, often with a blister in the centre, and it is clearly demarcated from the surrounding skin.

Stevens-Johnson syndrome (formerly also called erythema multiforme of major type) shows widespread skin lesions, which may either be target-shaped or consist of erythematous macules with epidermal detachment, together with severe mucosal erosions. Erosions of the skin do not exceed 10 per cent of body surface area. The general symptoms are more marked than in erythema multiforme.

Toxic epidermal necrolysis is characterized by widespread erythematous areas with epithelial necrosis and epidermal detachment (> 10 per cent body surface area), leaving bare dermis. Initially there are often also small erythematous or purpuric lesions with or without blisters. Extensive mucosal erosion is frequent.

2 Inclusion criteria for research and identification in databases

2.1 Primary criteria (signs & symptoms)

Erythema multiforme, Stevens-Johnson syndrome

- ✓ skin lesions (target or iris lesions, 0.5-2 cm; the lesions predominate on the extensor surfaces of the extremities)
- ✓ erosive mucosal lesions may be observed in about one-third of the cases
- ✓ fever
- ✓ malaise
- ✓ myalgia
- ✓ arthralgia

Toxic epidermal necrolysis

- ✓ skin lesions
- ✓ high fever
- ✓ malaise
- ✓ painful skin

3 Suggested criteria to exclude false positive cases during the process from broad search to final case identification

3.1 Clinical criteria (signs & symptoms)

Erythema multiforme, Stevens-Johnson syndrome

- ✓ time of onset: the suggestive time sequence is 1-3 weeks after the introduction of a new drugs, but other timings are compatible; the course is favorable and self-limiting in 10-20 days
- ✓ target lesions are mostly atypical in a drug-induced reaction, varying in size, with an irregular outline and the intermediate edematous ring is often absent; more target lesions are typical, more they suggest an infectious etiology

Toxic epidermal necrolysis

- ✓ Prodromal period (2-3 days): fever, conjunctivitis, pharyngitis, pruritus and occasionally, difficulty in urination.
- ✓ Time of onset: 1-3 weeks
- ✓ In the beginning, an atypical erythema multiforme is present, spreading quickly and becoming confluent, whereupon the epidermis comes off in large sheets (as a burn); the patient is at risk of severe complications (infections)
- ✓ In severe cases other organ systems can be involved: hepatocellular damage, pneumonia, nephritis, myocardial damage

3.2. Complementary investigations

Nikolski's sign is positive

- – Leucopenia is frequent and eosinophilia unusual
- – Physical causes, autoimmune blistering diseases and infectious diseases (herpes, Mycoplasma pneumoniae infections) may have to be excluded; skin biopsy and clinical photographs are helpful.

References

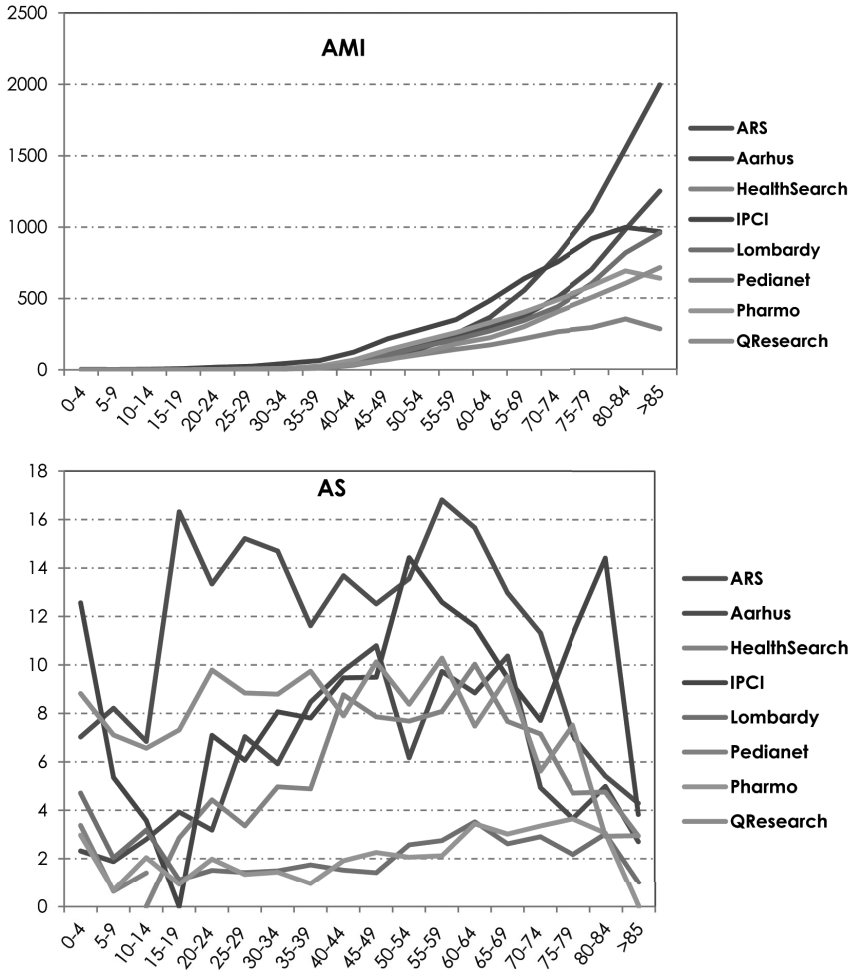
- CIOMS. Reporting Adverse Drug Reactions. Definition of terms and criteria for their use. WHO. Geneva.1999.
- Adverse Drug Reactions. A practical guide to diagnosis and management. Bènichou C (Editor). Wiley, 1994.
- Lee A. Adverse Drug Reactions. Pharmaceutical Press, 2001.
- Valeyrie-Allanore L et al. Drug-induced skin, nail and hair disorders. Drug Saf 2007; 30: 1011-30

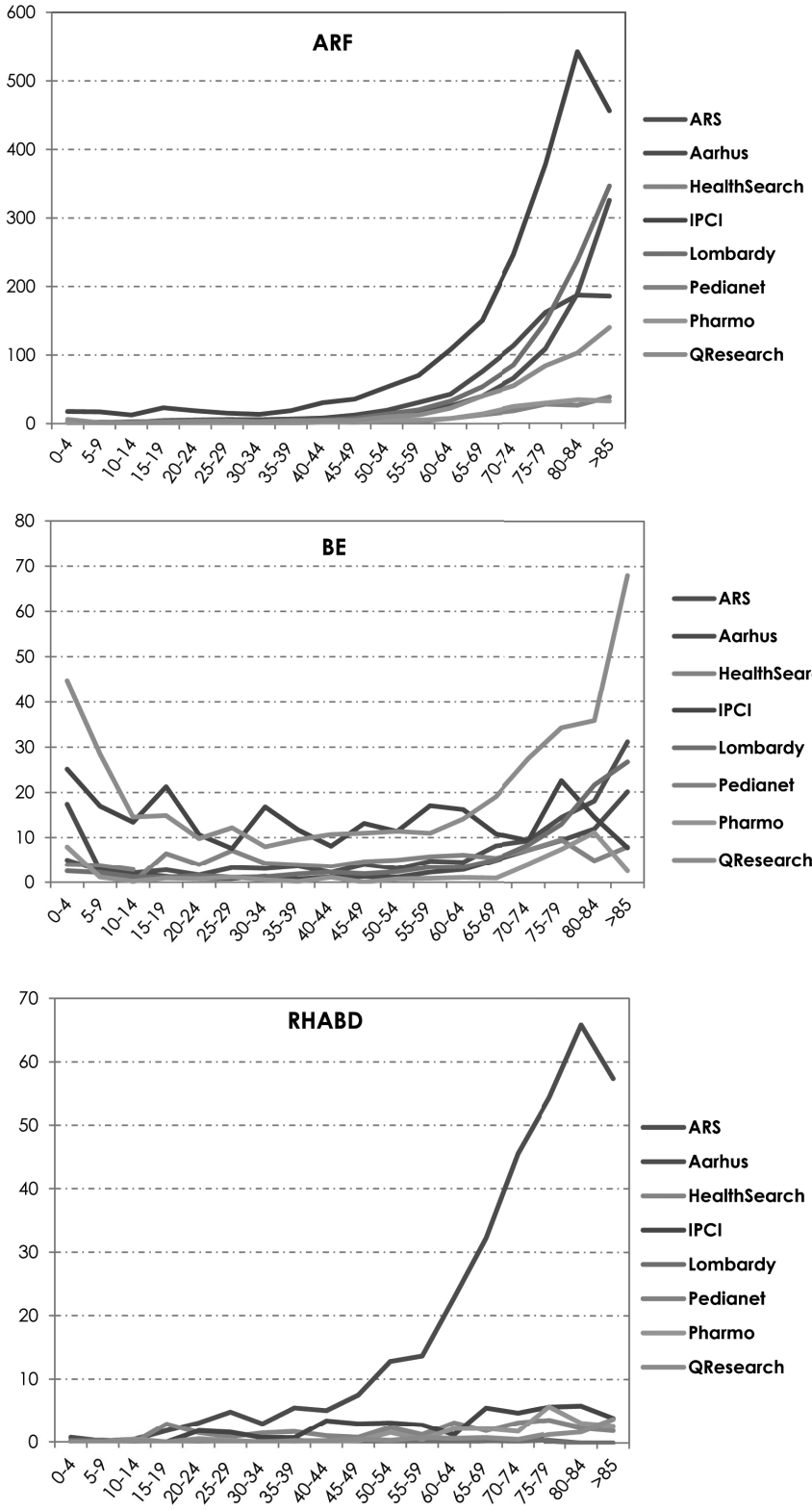
Appendix 2 | Query analysis table for the event Anaphylactic Shock before and after consensus

Table	Fields	AARHUS	ARS	HSD	IPCI	PEDIANET	PHARMO	QRESEARCH	LOMBARDY
HOSP	main diagnosis	T78.2, T78.0, T88.6, T80.5	995.0, 995.6, 999.4				995.0, 999.4, 995.6		995.0, 995.6, 999.4
	secondary diagnosis	~~T78.2, T78.0, T88.6~~	~~995.0, 995.6, 999.4~~						~~995.0, 995.6, 999.4~~
	procedures								
DEATH	cause of death	T78.2, T78.0, T88.6, T80.5	995.0, 995.6, 999.4						
GP	diagnosis			995.0 OR 995.6 OR 999.4 OR 'shock anafilattico' OR 'shock anafil' OR 'reaz anafil' OR 'reaz ipotens anafil' OR 'ipotens reaz anafil' OR 'shock allerg' OR 'reaz ipotens allerg' OR 'ipotens reaz allerg' OR 'ipotens bradicard' OR 'bradicard ipotens'	A12.1 OR anaphyl OR (shock AND anaf)	(shock AND allergia) OR (shock AND anaf) OR anafil		SN50, ~~SN501~~ SN500, SP34., SN502	
	specialist								
	lab								
LAB	classification								
	result								

Legend: After negotiation: In bold: additional code; ~~with a bar~~: deleted code

Appendix 3 | Age-specific incidence rates of the different events across the databases (per 100 000 person-years) ***(see pages 305 and 306 for colour figures)*****.**




ARF
600
500
400
300
200
100
0
0-4
5-9
10-14
15-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80-84
>85
ARS
Aarhus
HealthSearch
IPCI
Lombardy
Pedianet
Pharmo
QResearch
BE
80
70
60
50
40
30
20
10
0
ARS
Aarhus
HealthSearch
IPCI
Lombardy
Pedianet
Pharmo
QResearch
RHABD
70
60
50
40
30
20
10
0
ARS
Aarhus
HealthSearch
IPCI
Lombardy
Pedianet
Pharmo
QResearch

REFERENCES

1. Lindquist M. VigiBase, the WHO Global ICSR Database System : Basic Facts. 2008;42(5):11.
2. De Bruin ML, van Puijenbroek EP, Egberts AC, Hoes AW, Leufkens HG. Non-sedating antihistamine drugs and cardiac arrhythmias -- biased risk estimates from spontaneous reporting systems? *Br J Clin Pharmacol.* Apr 2002;53(4):370-374.
3. Begaud B, Martin K, Haramburu F, Moore N. Rates of spontaneous reporting of adverse drug reactions in France. *Jama.* Oct 2 2002;288(13):1588.
4. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *J Am Med Inform Assoc.* Mar-Apr 2003;10(2):115-128.
5. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc.* Jul-Aug 2005;12(4):448-457.
6. Murff HJ, Forster AJ, Peterson JF, Fiskio JM, Heiman HL, Bates DW. Electronically screening discharge summaries for adverse medical events. *J Am Med Inform Assoc.* Jul-Aug 2003;10(4):339-350.
7. Nadkarni PM. Drug safety surveillance using de-identified EMR and claims data: issues and challenges. *J Am Med Inform Assoc.* Nov 1 2010;17(6):671-674.
8. Reisinger SJ, Ryan PB, O'Hara DJ, et al. Development and evaluation of a common data model enabling active drug safety surveillance using disparate healthcare databases. *J Am Med Inform Assoc.* Nov 1 2010;17(6):652-662.
9. Bates DW, Gawande AA. Improving safety with information technology. *N Engl J Med.* Jun 19 2003;348(25):2526-2534.
10. Richesson RL, Fung KW, Krischer JP. Heterogeneous but "standard" coding systems for adverse events: Issues in achieving interoperability between apples and oranges. *Contemp Clin Trials.* Sep 2008;29(5):635-645.
11. Sturkenboom MCJM. *Other Databases in Europe for the Analytic Evaluation of Drug Effects*: John Wiley & Sons, Ltd; 2007.
12. FDA Sentinel Initiative. http://www.fda.gov/Safety/FDAsSentinelInitiative.
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.* Aug 13 2009;361(7):645-647.
14. 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.* Nov 2 2010;153(9):600-606.
15. Weeber M VR, Klein H, De Jong-Van Den Berg LT, Aronson AR, Molema G. Generating hypotheses by discovering implicit associations in the literature: a case report of a search for new potential therapeutic uses for thalidomide. *J Am Med Inform Assoc* 2003;10(3):252-259.
16. Trifirò G, Pariente A, Coloma PM, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiology and drug safety.* Dec 2009;18(12):1176-1184.
17. Coloma PM, Schuemie MJ, Trifirò G, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiology and drug safety.* Jan 2011;20(1):1-11.
18. Avillach P, Mougin 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. *Studies in health technology and informatics.* 2009;150:190-194.
19. ICD-9-CM. International Classification of Diseases, 9th edition, Clinical Modification: update. Official authorized addendum, effective October 1, 1986. *J Am Med Rec Assoc.* Sep 1986;57(9):suppl 1-32.
20. Pavillon G, Maguin P. [The 10th revision of the International Classification of Diseases]. *Rev Epidemiol Sante Publique.* 1993;41(3):253-255.
21. Lamberts H, Wood M, eds. *ICPC: International Classification of Primary Care*. Oxford: Oxford University Press; 1987.
22. O'Neil M, Payne C, Read J. Read Codes Version 3: a user led terminology. *Methods Inf Med.* Mar 1995;34(1-2):187-192.

23. Moore KM, Duddy A, Braun MM, Platt R, Brown JS. Potential population-based electronic data sources for rapid pandemic influenza vaccine adverse event detection: a survey of health plans. *Pharmacoepidemiology and drug safety.* Dec 2008;17(12):1137-1141.
24. US Food and Drug Administration. Adverse Event Reporting System (AERS). http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm. Accessed 11 November 2011.
25. Humphreys BL. The 1994 Unified Medical Language System knowledge sources. *Health Libr Rev.* Sep 1994;11(3):200-203.
26. Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med.* Aug 1993;32(4):281-291.
27. Avillach P, Joubert M, Thiessard F, et al. Design and evaluation of a semantic approach for the homogeneous identification of events in eight patient databases: a contribution to the European EU-ADR project. *Studies in health technology and informatics.* 2010;160:1085-1089.
28. Korman BD TK, Korman NJ. Progressive multifocal leukoencephalopathy, efalizumab, and immunosuppression: a cautionary tale for dermatologists. . *Arch Dermatol* 2009;145:937–942.
29. Black S, Eskola J, Siegrist CA, et al. Importance of background rates of disease in assessment of vaccine safety during mass immunisation with pandemic H1N1 influenza vaccines. *Lancet.* Dec 19 2009;374(9707):2115-2122.
30. Alecu I, Bousquet C, Jaulent MC. A case report: using SNOMED CT for grouping Adverse Drug Reactions Terms. *BMC Med Inform Decis Mak.* 2008;8 Suppl 1:S4.
31. 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.* May-Jun 2006;30(3):161-168.
32. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation.* Nov 27 2001;104(22):2746-2753.
33. Menotti A. Food patterns and health problems: health in southern Europe. *Ann Nutr Metab.* 1991;35 Suppl 1:69-77.
34. Menotti A, Lanti M, Puddu PE, Kromhout D. Coronary heart disease incidence in northern and southern European populations: a reanalysis of the seven countries study for a European coronary risk chart. *Heart.* Sep 2000;84(3):238-244.
35. Larranaga N, Moreno C, Basterretxea M, et al. Incidence of acute myocardial infarction in the Spanish epic cohort. *An Sist Sanit Navar.* Jan-Apr 2009;32(1):51-59.
36. Yeh RW, Sidney S, Chandra M, Sorel M, Selby JV, Go AS. Population trends in the incidence and outcomes of acute myocardial infarction. *N Engl J Med.* Jun 10 2010;362(23):2155-2165.
37. Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas AM, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation.* Jul 1994;90(1):583-612.
38. Ecological analysis of the association between mortality and major risk factors of cardiovascular disease. The World Health Organization MONICA Project. *Int J Epidemiol.* Jun 1994;23(3):505-516.
39. Higgins M. Patients, families and populations at high risk for coronary heart disease. *Eur Heart J.* Sep 2001;22(18):1682-1690.
40. Wolf-Maier K, Cooper RS, Banegas JR, et al. Hypertension prevalence and blood pressure levels in 6 European countries, Canada, and the United States. *JAMA.* May 14 2003;289(18):2363-2369.
41. Perugini E, Maggioni AP, Boccanelli A, Di Pasquale G. [Epidemiology of acute coronary syndromes in Italy]. *G Ital Cardiol (Rome).* Oct 2010;11(10):718-729.
42. Ruff CT, Braunwald E. The evolving epidemiology of acute coronary syndromes. *Nat Rev Cardiol.* Mar 2011;8(3):140-147.

43. Varas-Lorenzo C, Castellsague J, Stang MR, Tomas L, Aguado J, Perez-Gutthann S. Positive predictive value of ICD-9 codes 410 and 411 in the identification of cases of acute coronary syndromes in the Saskatchewan Hospital automated database. *Pharmacoepidemiol Drug Saf.* Aug 2008;17(8):842-852.

44. Uchino S. The epidemiology of acute renal failure in the world. *Curr Opin Crit Care.* Dec 2006;12(6):538-543.

45. Hoste EA, Schurgers M. Epidemiology of acute kidney injury: how big is the problem? *Crit Care Med.* Apr 2008;36(4 Suppl):S146-151.

46. Srisawat N, Kellum JA. Acute kidney injury: definition, epidemiology, and outcome. *Curr Opin Crit Care.* Dec 2011;17(6):548-555.

47. Ali T, Khan I, Simpson W, et al. Incidence and outcomes in acute kidney injury: a comprehensive population-based study. *J Am Soc Nephrol.* Apr 2007;18(4):1292-1298.

48. Moro Moro M, Tejedor Alonso MA, Esteban Hernandez J, Mugica Garcia MV, Rosado Ingelmo A, Vila Albelda C. Incidence of anaphylaxis and subtypes of anaphylaxis in a general hospital emergency department. *J Investig Allergol Clin Immunol.* 2011;21(2):142-149.

49. Harduar-Morano L, Simon MR, Watkins S, Blackmore C. A population-based epidemiologic study of emergency department visits for anaphylaxis in Florida. *J Allergy Clin Immunol.* Sep 2011;128(3):594-600 e591.

50. Koplin JJ, Martin PE, Allen KJ. An update on epidemiology of anaphylaxis in children and adults. *Curr Opin Allergy Clin Immunol.* Oct 2011;11(5):492-496.

51. Tang ML, Osborne N, Allen K. Epidemiology of anaphylaxis. *Curr Opin Allergy Clin Immunol.* Aug 2009;9(4):351-356.

52. Fritsch P. European Dermatology Forum: skin diseases in Europe. Skin diseases with a high public health impact: toxic epidermal necrolysis and Stevens-Johnson syndrome. *Eur J Dermatol.* Mar-Apr 2008;18(2):216-217.

53. Kaufman DW. Epidemiologic approaches to the study of toxic epidermal necrolysis. *J Invest Dermatol.* Jun 1994;102(6):31S-33S.

54. French LE. Toxic epidermal necrolysis and Stevens Johnson syndrome: our current understanding. *Allergol Int.* Mar 2006;55(1):9-16.

55. Bollaert PE, Frisoni A. Epidemiology, mechanisms and clinical features of rhabdomyolysis. *Minerva Anestesiol.* May 1999;65(5):245-249.

56. Bosch X, Poch E, Grau JM. Rhabdomyolysis and acute kidney injury. *N Engl J Med.* Jul 2 2009;361(1):62-72.

57. Madsen M, Gudnason V, Pajak A, et al. Population-based register of acute myocardial infarction: manual of operations. *Eur J Cardiovasc Prev Rehabil.* Dec 2007;14 Suppl 3:S3-22.

58. Di Pasquale G, Lombardi A, Casella G. The redefinition of acute myocardial infarction. *Ital Heart J.* Jun 2004;5 Suppl 6:9S-18S.

59. Kavsak PA, MacRae AR, Lustig V, et al. The impact of the ESC/ACC redefinition of myocardial infarction and new sensitive troponin assays on the frequency of acute myocardial infarction. *Am Heart J.* Jul 2006;152(1):118-125.

60. von dem Knesebeck O, Bonte M, Siegrist J, et al. Country differences in the diagnosis and management of coronary heart disease - a comparison between the US, the UK and Germany. *BMC Health Serv Res.* 2008;8:198.

3.3

Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project

Preciosa M. Coloma, MD, MSc[1], Martijn J. Schuemie, PhD[1], Gianluca Trifirò, MD, PhD[1], Rosa Gini, PhD[2], Ron Herings, PhD[3], Julia Hippisley-Cox, MD[4], Giampiero Mazzaglia, MD, PhD[5], Carlo Giaquinto,[6] Giovanni Corrao, PhD[7], Lars Pedersen, PhD[8], Johan van der Lei, MD, PhD[1], Miriam Sturkenboom, PharmD, PhD[1,9] on behalf of the EU-ADR consortium

[1]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands; [2]Agenzia Regionale di Sanita` della Toscana Florence, Italy; [3]PHARMO Institute, Utrecht, the Netherlands; [4]Division of Primary Care – School of Community Health Sciences, University of Nottingham, Nottingham UK; [5]Societa 'Italiana di Medicina Generale, Florence, Italy; [6]Pedianet-Società Servizi Telematici SRL, Padova, Italy; [7]Department of Statistics, Università di Milano-Bicocca, Milan, Italy; [8]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; [9]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

Pharmacoepidemiol Drug Saf 2011 Jan; 20(1):1-11.
doi: 10.1002/pds.2053. Epub 2010 Nov 8.

ABSTRACT

Purpose. In this proof-of-concept paper we describe the framework, process, and preliminary results of combining data from European electronic healthcare record (EHR) databases for large-scale monitoring of drug safety.

Methods. Aggregated demographic, clinical, and prescription data from eight databases in four countries (Denmark, Italy, Netherlands, the UK) were pooled using a distributed network approach by generation of common input data followed by local aggregation through custom-built software, Jerboa©. Comparison of incidence rates of upper gastrointestinal bleeding (UGIB) and nonsteroidal anti-inflammatory drug (NSAID) utilization patterns were used to evaluate data harmonization and quality across databases. The known association of NSAIDs and UGIB was employed to demonstrate sensitivity of the system by comparing incidence rate ratios (IRRs) of UGIB during NSAID use to UGIB during all other person-time.

Results. The study population for this analysis comprised 19 647 445 individuals corresponding to 59 929 690 person-years of follow-up. 39 967 incident cases of UGIB were identified during the study period. Crude incidence rates varied between 38.8 and 109.5/100 000 person-years, depending on country and type of database, while age-standardized rates ranged from 25.1 to 65.4/100 000 person-years. NSAID use patterns were similar for databases within the same country but heterogeneous among different countries. A statistically significant age- and sex-adjusted association between use of any NSAID and increased risk for UGIB was confirmed in all databases, IRR from 2.0 (95%CI: 1.7-2.2) to 4.3 (95%CI: 4.1-4.5).

Conclusions. Combining data from EHR databases of different countries to identify drug-adverse event associations is feasible and can set the stage for changing and enlarging the scale for drug safety monitoring.

INTRODUCTION

Every year numerous drugs are introduced into the international healthcare market. In 2008 alone, a total of 66 medicinal products received a positive opinion from the European Medicines Agency (EMA)[1] while 99 drugs were approved by the US Food and Drug Administration (FDA).[2] These new drugs often represent important advances that improve the care or quality of life for many patients worldwide. While a drug's efficacy and safety must be demonstrated in randomized controlled clinical trials prior to approval, these trials are rarely designed or powered to detect uncommon or unexpected adverse events.[3-6] Safety surveillance of drugs in the post-marketing phase has traditionally been performed by analyses of spontaneous case reports for signal generation and of observational healthcare databases for signal evaluation.[7-9] The sudden worldwide withdrawal of rofecoxib (Vioxx) in 2004[10] stimulated an international review of how best to measure and monitor drug safety so that the balance of risks and benefits can be continually and rapidly evaluated.[11–14] As mandated by the FDA Amendments Act of 2007, the US is establishing the Sentinel System, a nationwide network of databases targeted to capture by the year 2012 data on more than 100 million subjects for prospective drug safety surveillance.[15] The European Commission (EC) has similarly issued initiatives to develop new methodologies for drug safety monitoring based on analysis of large databases. Under the auspices of the EC, the EU-ADR Project ("Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge") (http://www.euadr-project.org) was launched in 2008. EU-ADR aims to exploit information from various electronic healthcare record (EHR) databases in Europe to produce a computerized integrated system for the early detection of drug safety signals. Once detected, the safety signals will be substantiated by semantic mining of literature and computational analysis of pharmacological and biological information on drugs, molecular targets, and pathways (Figure 1).[16]

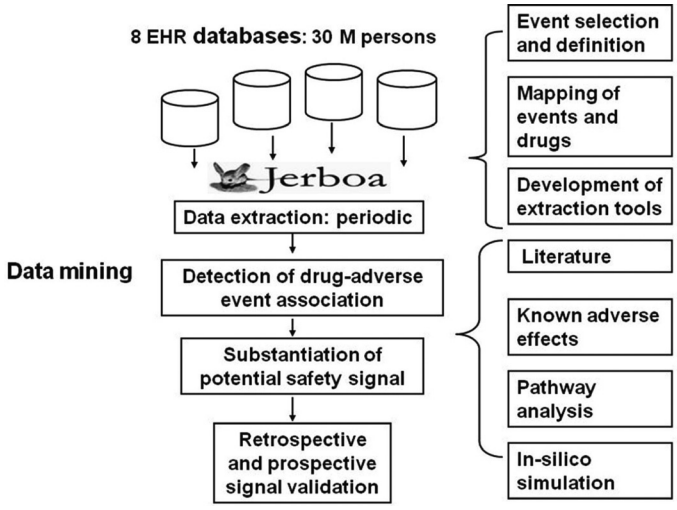


Figure 1 | EU-ADR Project Schema.

In this proof-of-concept paper, we endeavor to set the stage for large-scale drug safety monitoring by describing the framework, process, and preliminary results of combining data from a federation of eight EHR databases in four countries, a resource of unprecedented size for monitoring of drug safety in Europe. We further illustrate the challenges encountered in the amalgamation of data from diverse locations into a uniform repository and the obstacles that had to be overcome in terms of heterogeneity in database structure, language, coding of drugs and diseases, and diversity in the organization of European healthcare systems.

METHODS

Study population and data sources

The EU-ADR database platform currently comprises anonymous healthcare data from eight established European databases located in four countries. Health Search/CSD Patient (HSD, Italy),[17,18] Integrated Primary Care Information (IPCI, Netherlands),[19,20] Pedianet[21,22] (Italy), and QRESEARCH[23,24] (United Kingdom) are general practice (GP) databases, where both clinical information and drug prescriptions are recorded. The Aarhus University Hospital Database (Aarhus, Denmark),[25,26] the PHARMO Network[27,28] (Netherlands), and the regional Italian databases of Lombardy[29,30] and Tuscany[31,32] are all comprehensive record-linkage systems in which drug dispensing data of a well-defined population are linked to a registry of hospital discharge diagnoses and various other registries. Table 1 provides an overview of characteristics of each database. Most healthcare services, including pharmaceutical services, are provided for, or subsidized by, the state in Italy (ITA), Denmark (DK), and the United Kingdom (UK) and covered by obligatory health insurance in the Netherlands (NL) and turnover is low. In all of the countries with GP databases, GPs function as "gatekeepers" of the healthcare system.

Common data framework: distributed data processing

Founded on the basic governance principle that database owners should be involved in the elaboration of data as they best understand the context within which the data are recorded, we have chosen in the EU-ADR Project a distributed network approach that requires standardization of input files from the different databases (Figure 2). These common input files (patient, drug, and event files) are created locally and are subsequently managed by purpose-built software called Jerboa©, which has been tested against different scripts. Jerboa© uses flat text files as input and is written entirely in Java™ to ensure that it will run in a wide variety of computational environments. The software queries patient-level data in the different databases, which are later aggregated, de-identified and sent in encrypted format to a central repository for evaluation and further analysis. This repository is managed by the Department of Medical Informatics at Erasmus Medical Center in the Netherlands, the project's coordinating center.

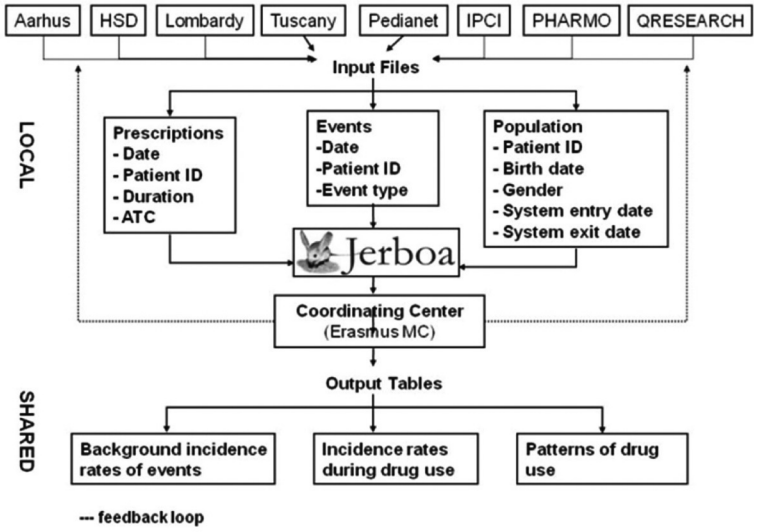


Figure 2 | Common data framework: distributed network.

Cohort definition and follow-up time

To harmonize follow-up definitions across databases, we defined three dates for each patient file: (1) the eligibility period begins on the date the patient is registered in the database; (2) the eligibility period ends on the date the patient transfers out of the system, with the last supply of data, or on the patient's death, whichever is earlier; and (3) the date of birth. From these dates and in combination with the drug prescription file, Jerboa© marks the beginning of follow-up, which is the start of eligibility date plus one year, or the date of birth (for subjects whose date of birth-start of eligibility is less than one year).

Drug exposure

Drug prescriptions and dispensings are locally coded using the national product codes, which differ among countries (see Table 1). Most countries, however, link these product codes to the Anatomical Therapeutic Chemical (ATC) classification system. The ATC code is used as the drug code in the EU-ADR input files. Each database owner estimates the duration covered by each prescription or dispensing according to the legend duration (if dosing regimen is available), or is otherwise based on the defined daily dose (DDD).[33]

Event definition

The EU-ADR Project started out by defining a selected number of events that are subsequently mapped to a common terminology system. This process has been described in more detail in separate publications.[16,34,35] Databases in EU-ADR use one of four nomenclature systems to describe

the events: the International Classification of Diseases (ICD9-CM and ICD10), the International Classification of Primary Care (ICPC), and the READ Code (RCD) classification. These different terminologies are mapped using the Unified Medical Language System1 (UMLS1). The UMLS is a biomedical terminology integration system handling more than 150 terminologies including the four used in the EU-ADR project.[36] Ascertainment of the event of interest from the databases follows an iterative process with seven stages: (1) event definition using clinical criteria established from literature; (2) identification of UMLS concepts corresponding to the event; (3) revision and validation of medical concepts by database owners and pharmacovigilance experts; (4) translation of the medical concepts into each database terminology; (5) extraction of data and computation of event rates; (6) comparison of query structure – to detect and harmonize eventually any major disagreement across databases; and (7) creation of event input files for Jerboa©.

Statistical analyses

Incidence rates. We calculated age- and sex-specific incidence rates of the event of interest within each database as a means to compare and benchmark the data extraction. We performed direct standardization using the WHO World Standard Population as reference to account for age differences when comparing the overall rates.[37] Jerboa© manages and aggregates the data over all patients locally, producing as output the amount of person-time of follow-up and the number of events per sex and age group. To calculate the incidence rates, we only considered the first recorded occurrence of the event of interest after a one year run-in period.

Associations. Drug prescription and/or dispensing data were used to estimate drug utilization and event rates during drug exposure. Overlapping treatment episodes with the same drug (same ATC code) are combined into a single episode of drug use that starts when the first prescription begins and stops when the last prescription ends. When a patient uses more than one drug at a time, the corresponding person-time is labeled accordingly. Using individual data on start date and end date of prescription or dispensing, Jerboa© determines and marks as unexposed those periods during which an individual is included in the study but is not using any drug. Events are then assigned to the episodes (drug use/non-use) in which they occurred.

The validity of the methodology previously described has to be evaluated in terms of the system's objective of drug safety monitoring. The system's ability to identify a drug-adverse event pair for which the association is established (i.e., true positive signal) will provide an indication of its sensitivity. We illustrate this methodology on actual data, first focusing on the known association between use of NSAIDs and increased risk of upper gastrointestinal bleeding (UGIB). For the initial estimation of the association between event and drug use, we summarized exposure to each drug in tables, stratified by age, sex, and calendar year. We performed the analyses at three different ATC levels: pharmacological subgroup (i.e., M01A); chemical subgroup (i.e., M01AB, M01AC, etc.); and chemical substance (i.e., M01AC01, etc.). Using the Mantel–Haenszel method, we calculated age- and sex-adjusted rate ratios, with all remaining person-time (i.e., non-use and use of all other drugs) as reference.

Table 1 | Characteristics of the databases in the EU-ADR network.

Characteristics	Pedianet (Italy)	HSD (Italy)	Lombardy Regional (Italy)	Tuscany Regional (Italy)	IPCI (Netherlands)	PHARMO (Netherlands)	QRESEARCH (UK)	Aarhus (Denmark)
Current source population	160 000 children	1 500 000	9 000 000	3 500 000	1 500 000	3 000 000	4 000 000	1 800 000
Years covered for this study	2003-2007	2003-2007	2003-2005	2003-2006	1996-2006	1998-2007	2000-2007	2001-2006
Type of database	General Practice pediatric database	General Practice database	Administrative	Administrative	General Practice database	Hybrid (administrative and medical record/ registries)	General Practice database	Administrative
Age range	0-14	From 15 onwards	All ages	All ages	All ages	All ages	All ages	All ages
% Males	52.2	47.2	48.8	48.1	49.6	45.8	49.6	49.9
Demographic information available								
Date of registration	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date of transferring out	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date of birth	MM-YY	MM-YY	DD-MM-YY	DD-MM-YY	MM-YY	DD-MM-YY	YY	MM-YY
Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity/Race	No	No	No	No	No	No	No	No
Drug information available								
Product coding	MINSAN	MINSAN	MINSAN	MINSAN	HPK	Z index	EMIS	VAerets
Active international principle coding system	ATC	ATC	ATC	ATC	ATC	ATC	BNF	ATC
Date of prescription/dispensing	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dosing regimen	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Quantity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indication of use	Yes	Yes	No	No	Yes	Yes for in-hospital	No	Yes

Characteristics	Pedianet (Italy)	HSD (Italy)	Lombardy Regional (Italy)	Tuscany Regional (Italy)	IPCI (Netherlands)	PHARMO (Netherlands)	QRESEARCH (UK)	Aarhus (Denmark)
Outcome information available								
Symptoms (Yes/No)	Yes, as free text/ codes	Yes, as free text/codes	No	No	Yes, as free text/ codes	Yes for some	Yes, as codes	No
Outpatient primary care diagnoses	Yes, as free text / codes	Yes Free text /codes	No	No	Yes, as free text / codes	No	Yes	No
Outpatient specialist care diagnoses	Yes, as free text/ codes	Yes	No	No	Yes	No	Yes	No
Hospital discharge diagnoses	Yes, as free text / codes	Yes, as free text /codes	Yes	Yes	Yes, as free text / codes	Yes	Yes	Yes
Diagnosis coding scheme	ICD-9CM	ICD-9CM	ICD-9CM	ICD-9CM	ICPC	ICD-9CM	RCD	ICD-10
Diagnostic procedures	Yes	Yes	Yes	Yes	No	Yes for in-hospital interventions	Yes	Yes, in-hospital only
Laboratory tests	Yes	Yes	No	No	Yes	Yes subset	Yes	Yes, in-hospital only

ICPC: International Classification of Primary Care
ICD9-CM: International Classification of Diseases – 9th revision Clinical Modification
RCD: READ CODE Classification
ICD-10: International Classification of Diseases – 10th revision
MINSAN: Italian Ministry of Health

RESULTS

Study population

The total study population comprised 19 647 445 individuals corresponding to 59 929 690 person-years (PYs) of follow-up. The databases contributed varying follow-up time to the study period which covered the years 1996–2007. Comparison of follow-up time across databases revealed similar age and sex distributions, except for Pedianet which, by design, only includes patients aged less than 14 years and HSD which covers only patients older than 14 years old (see Figure 3).

Incidence rates of upper gastrointestinal bleeding

We identified a total of 39 967 incident cases of UGIB in the study population. Analysis of crude overall incidence rates demonstrated heterogeneity with non-standardized rates of UGIB ranging from 38.8 (PHARMO) to 61.1 (IPCI) per 100 000 PYs in the Netherlands to 87.7 (Aarhus) and 109.5 (HSD) per 100 000 PYs in Denmark and Italy, respectively (Table 2). Incidence rates for the UK database QRESEARCH (84.3) and the regional Italian databases of Tuscany (70.7) and Lombardy (52.5) were somewhere in between. Direct age standardization attenuated the variation in overall incidence rates between databases. Table 2 also shows the incidence rates of UGIB separately for children (14 years old and below) and those older than 15 years. Across all databases, incidence rates increased with advancing age (Figure 4), the risk for UGIB in individuals 70 years and older was 5.2 times greater compared to those aged 50-59 years old (95% CI: 5.1- 5.4). There were also consistently higher incidence rates in males overall compared to females (incidence rate ratio (IRR) 1.24 (95% CI: 1.22-1.27)).

Table 2 | Incidence rates of UGIB (per 100 000 PYs).

Country	Database	No. of Events	Population-time	Overall incidence rate	*(age-standardized)*	Incidence rate in <15 years	*(age-standardized)*	Incidence rate in ≥15 years	*(age-standardized)*
ITA	Pedianet	65	407 150	16.0	*(14.5)*	16.0	*(14.5)*	–	
	HSD	3 639	3 324 089	109.5	*(65.4)*	–		109.7	*(71.6)*
	Lombardy	10 127	19 282 776	52.5	*(29.2)*	9.8	*(9.3)*	59.5	*(36.2)*
	Tuscany	7 024	9 940 626	70.7	*(31.6)*	13.3	*(13.0)*	78.5	*(38.2)*
NL	IPCI	1 139	1 862 932	61.1	*(44.3)*	7.6	*(7.5)*	73.2	*(57.4)*
	PHARMO	2 944	7 591 284	38.8	*(25.1)*	5.0	*(4.9)*	46.3	*(32.3)*
UK	QRESEARCH[a]	5 721	6 788 121	84.3	*(60.1)*	15.4	*(16.0)*	99.0	*(75.8)*
DK	Aarhus	9 308	10 611 047	87.7	*(54.6)*	3.7	*(3.8)*	108.0	*(72.6)*
Total		39,967	59 807 984	66.8		9.1		77.4	

[a]represent data from 30% of database population.

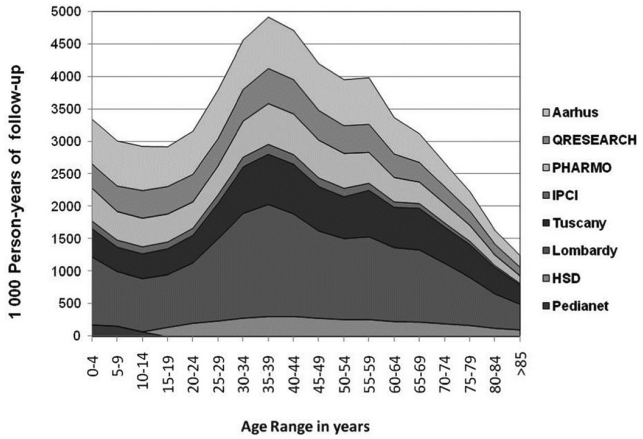


Figure 3 | Distribution of follow-up time per age group across databases.

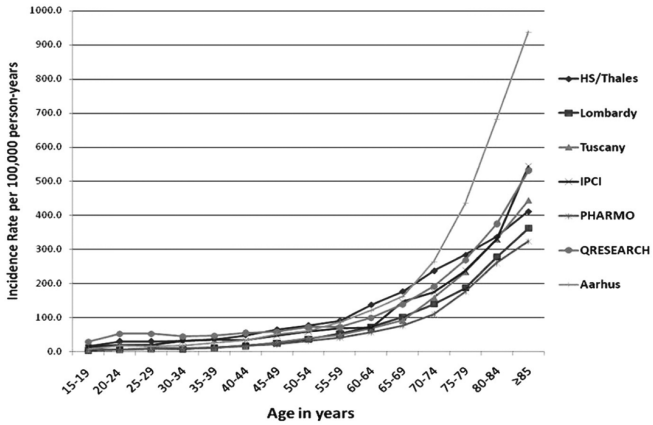


Figure 4 | Incidence rates of UG1B across databases, age 15 years and above ***(see page 306 for colour figure).***

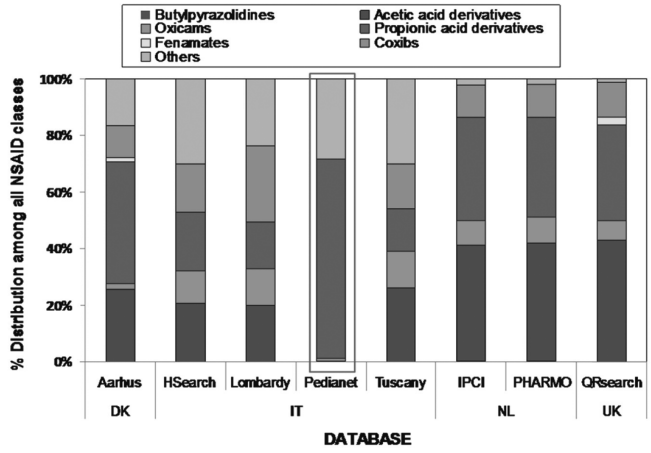


Figure 5 | Comparison of use of specific NSAID classes across databases ***(see page 306 for colour figure).***

Patterns of NSAID use

Patterns of use of NSAID classes varied among different countries but were similar among different databases in the same country (Figure 5). The Dutch databases IPCI and PHARMO showed identical utilization profiles, with the acetic acid derivatives (e.g., diclofenac) representing about 40% of all NSAID exposure time. The UK database QRESEARCH had a profile similar to NL, while the Danish database Aarhus showed more use of propionic acid derivatives (e.g., naproxen) compared to the acetic acid derivatives (47% vs. 30%). Except for the pediatric database Pedianet, the other three Italian databases had similar drug use patterns with relatively high use of the ATC M01AX class of NSAIDs (other NSAIDs). This group of drugs, which includes nimesulide, represented the lowest percentage of NSAID use in the other databases (except Aarhus). Heterogeneity in exposure among countries became more apparent when we explored the use of individual drugs (data not shown).

Association of NSAID use with upper gastrointestinal bleeding

Overall, we detected 4 934 incident cases of UGIB during NSAID use. Data from Pedianet were not taken into account in this analysis due to low number of events in children. The incidence rates were consistently around 3-4 per 1000 PYs of NSAID exposure, except for PHARMO which had lower incidence rates (1.9/1000 PYs) and Aarhus which had higher incidence rates (6.5/1000 PYs). The rates of UGIB were significantly increased during use of NSAIDs as compared to all other follow-up time in each of the individual databases (Table 3), with IRRs ranging from 2.0 (95% CI 1.7 to 2.2) to 4.3 (95% CI 4.1 to 4.5).

Table 3 | Incidence Rate Ratios (IRRs) of UGIB during NSAID use.

Country	Database	No. of Events	Exposure[a]	Incidence Rate[b]	Rate Ratio[c] *(95%CI)*
ITA	**HSD**	250	81 734	3.1	2.0 (1.7 to 2.2)
	Lombardy	991	314 852	3.1	2.9 (2.7 to 3.1)
	Tuscany	698	205 012	3.4	2.4 (2.3 to 2.6)
NL	**IPCI**	116	26 780	4.3	4.0 (3.3 to 4.9)
	PHARMO	342	177 698	1.9	2.8 (2.5 to 3.2)
UK	**QRESEARCH**	467	158 783	2.9	2.4 (2.2 to 2.6)
DK	**Aarhus**	2,070	316 348	6.5	4.3 (4.1 to 4.5)
TOTAL		4,934	1 281 207	3.9	

Legend:
a - in person-years (PYs)
b - per 1 000 PYs
c - adjusted for age and sex, non-NSAID use as comparator; p value <<<0.01

DISCUSSION

We have developed and tested a methodology that enables combining data from EHR databases of various countries and origins (medical records, administrative registries, record-linkage databases). Revisiting the known association of UGIB and nonsteroidal anti-inflammatory drug (NSAID) use, we have shown that data sharing can take place within this distributed networking system to provide consistent incidence rates and detect a known drug-adverse event association. This network system yields other interesting side products such as age- and sex-specific disease rates and drug utilization patterns across a wide variety of settings. Together this opens a new avenue for conducting observational research on a wider, more global perspective.

In combining the data of various databases it was crucial to take into account ethical issues regarding the processing of anonymized healthcare data. The databases involved in EU-ADR have considerable experience in using patient data for research purposes, and have well-developed safeguard mechanisms ensuring compliance with the European directives and national regulations, as well as database governance rules. Since no new data are collected, other than those made available by the participating databases, the cornerstone of ensuring proper ethical and legal conduct is set down in the rules and regulations that govern each database. Rather than imposing a one-size- fits-all approach and compel people to change their data, we leveraged on the diversity of the databases to use local expertise and to maximize extraction of relevant information, hence effectively dealing with all methodological, cultural, ethical, governance, and political issues of sharing data across borders. Databases retain ownership of their respective data, extraction being done locally, and only the aggregated, de-identified data are shared with the rest of the network.

A similar concept of distributed processing of healthcare data has been employed by other research collaborations, albeit with different research objectives. This model has previously been described in bioterrorism and syndromic surveillance as well as vaccine safety surveillance.[38-44] The ongoing Sentinel Initiative of the FDA is also adopting a distributed data architecture for combining healthcare databases to improve drug safety monitoring.[45-47] While the scale may be comparable, there are issues in combining data that are unique to Europe. Challenges stem from the fact that different countries have distinct natural languages, aside from having different drug and disease coding systems. The diversity of healthcare systems throughout Europe makes merging data from databases a more complex task that requires striking a balance between international cooperation and adequate protection of patient confidentiality.[48]

Incidence rates

Standardized incidence rates were generally higher than the crude incidence rates, the European EU-ADR network population being older than the reference WHO World Standard Population. Comparison of incidence rates was used for benchmarking of the data extraction process and the heterogeneity in incidence rates of UGIB reflects interesting peculiarities in the populations themselves, in the type of database, and in the respective healthcare systems. There may be inter-country differences as to causes of bleeding, risk factors, and variations in the extent

of diagnostic evaluation. Different sources of data are apt to capture UGIB of varying severity, the events identified from hospitalization records likely to be more severe than the events ascertained from general practitioner (GP) visits. This may explain the lower incidence rates observed in administrative databases compared to general practice databases in the same country (e.g., HSD vs. Tuscany or IPCI vs. PHARMO).

Although there is not yet a gold standard to validate our results, the observed incidence rates were in line with previously published literature. Most studies on the epidemiology of UGIB are hospital-based studies that cite incidence varying from 48/100 000 adults in the Netherlands to 103/100 000 adults in the UK.[49-52] Other reports estimate the incidence to be 95/ 100 000 adults (USA) and 172 (Scotland).[53,54] EUADR provides data specifically for children. To date, there are no published population-based studies that include children below 15 years of age, which underlines the importance of this type of information.

The increase in risk of UGIB with advancing age and with men compared to women, which we observed in all of the databases, are consistent with the literature.[55] Exact comparisons between studies are difficult because of discrepancies in event and drug exposure definitions, selection criteria of study populations, and methodologies. The advantage of EU-ADR is that the variations in event definition are documented in a standard way and the methodological differences in rate estimations have been eliminated which allows for easier comparison across countries and databases.

Sensitivity of the system

In general, the relative risk of UGIB during NSAID use is around four.[56] This is similar to what we obtained in this study even if the rate ratios were adjusted only for age and sex and were compared to all other person-time (i.e., exposed to other drugs or unexposed). The consistency of the degree of association between NSAID use and UGIB within the databases studied reiterates the huge burden of drug-related disease. Its concordance with published literature increases the face validity of the methodology and reinforces the potential of drug safety monitoring and risk quantification being performed on a very different scale, without limitations as to the size or to the number of databases. Key steps for success in this process were the chosen distributed approach that dealt with different governance issues, the definition of common data input files to deal with database heterogeneity, the availability of common customized software that allowed for local elaboration of data, and terminology mapping using the Unified Medical Language System. Improving the sensitivity of the database network system will require further development of scalable methods for better control of bias and confounding to allow further inferences regarding causality.

While the motivation for merging disparate data sources primarily comes from the need to investigate drug safety in larger populations, it is a well-acknowledged limitation that EHR databases are only able to describe exposures and outcomes of interest to the extent that they are documented within the database systems.[57-59] The databases capture information on outpatient (pharmacy-based) drug use. Most databases do not capture all vaccinations (e.g., those provided in

the childhood vaccination programs) but do register, for example influenza, vaccination. Specific drug groups of interest, such as biologicals, are captured if provided through routine dispensing system. The type of drugs that can be captured in the system is being addressed in a separate paper. Furthermore, the issue of accuracy and completeness of information concerning diagnoses and actual drug use is common to all types of databases and some degree of misclassification is unavoidable. Although review of every individual medical record in each database would incur prohibitive costs, some sensitivity analyses should be done to determine the validity of the extracted data.

CONCLUSION

In this proof-of-concept paper, we have demonstrated the feasibility of combining diverse and differently structured data in an effective way to detect comparative risks of potential adverse drug events and pave the way for large-scale drug safety monitoring. The common data framework described takes advantage of multiple, routinely collected, aggregated healthcare data while minimizing sharing of confidential patient-level information. Although still a work-in-progress, this system can facilitate the planning of permanent, global-scale, EHR-based structures for the early detection of drug safety signals.

REFERENCES

1. http://www.emea.europa.eu/pdfs/general/direct/emeaar/AnnualReport2008.pdf [Accessed 21 October 2009].
2. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm?fuseaction=Reports. ReportsMenu [Accessed 01 October 2009].
3. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996;312: 1215–1218.
4. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006;174:635–641. DOI: 10.1503/cmaj.050873.
5. Stricker BHCh, Psaty BM. Detection, verification, and quantification of adverse drug reactions. *BMJ* 2004; 329:44–47.
6. Ray WA. Population-based studies of adverse drug effects. *New Engl J Med* 2003;349:1592–1594.
7. Curri CJ, Macdonald TM. Use of routine healthcare data in safe and cost-effective drug use. *Drug Saf* 2000; 22(2):97–102.
8. Rodriguez EM, Staffa JA, Graham DJ. The role of databases in drug postmarketing surveillance. *Pharmacoepidemiol Drug Saf* 2001;10:407–410. DOI: 10.1002/pds.615.
9. Harmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. *Eur J Clin Pharmacol* 2008;64:743–752. DOI:10.1007/s00228-008-0475-9.
10. Singh D. Merck withdraws arthritis drug worldwide. *BMJ* 2004;329(7470):816. DOI:10.1136/bmj. 329.7470.816-a.
11. Avorn J. Evaluating drug effects in the post-Vioxx world: there must be a better way. *Circulation* 2006; 113: 2173–2176. DOI: 10.1161/CIRCULATIONAHA. 106.625749.
12. Wadman M. Experts call for active surveillance of drug safety. *Nature* 2007;446(7134):358–359. DOI: 10.1038/446358b; Published online: 21 March 2007.
13. EMEA. European Risk Management Strategy: 2008–2009 work programme adopted. London, 2 January, 2008. http://www.hma.eu/uploads/media/ERMS_2008-2009_work_programme_adopted.pdf [Accessed 03 February 2009].
14. FDA. The Future of Drug Safety-Promoting and Protecting the Health of the Public, January 2007. http://www.fda.gov/oc/reports/iom013007.pdf [Accessed 03 February 2009].
15. http://www.fda.gov/oc/initiatives/advance/reports/report0508.pdf [Accessed 23 February 2009].
16. Trifirò G, Pariente A, Coloma PM, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom M, 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): 1176–1184. DOI: 10.1002/pds.1836; Published online: 15 September 2009.
17. Sacchetti E, Trifirò G, Caputi A, et al. Risk of stroke with typical and atypical anti-psychotics: a retrospective cohort study including unexposed subjects. *J Psychopharmacol* 2008;22(1):39–46. DOI: 10.1177/0269881107080792.
18. Mazzaglia G, Mantovani LG, Sturkenboom MC, et al. Patterns of persistence with antihypertensive medications in newly diagnosed hypertensive patients in Italy: a retrospective cohort study in primary care. *J Hypertens* 2005;23(11):2093–2100.
19. Vlug AE, van der Lei J, Mosseveld BM, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf Med* 1999;38:339–344.
20. Verhamme K, Mosis G, Dieleman J, Stricker B, Sturkenboom M. Spironolactone and risk of upper gastrointestinal events: population based case-control study. *BMJ* 2006;333(7563):330. DOI: 10.1136/ bmj.38883.479549.2F; Published online: 13 July 2006.
21. Sturkenboom MCJM, Nicolosi A, Cantarutti L, et al. for the NSAIDs Paediatric Research Group. Incidence of mucocutaneous reactions in children treated with niflumic acid, other non steroidal anti-inflammatory drugs and analgesics. *Pediatrics* 2005; 116(1): e26–e33. DOI: 10.1542/peds.2004-0040; Published online: June 1 2005.
22. Giaquinto C, Van Damme P, Huet F, et al. Clinical consequences of rotavirus acute gastroenteritis in Europe, 2004–2005: The REVEAL Study. *J Infect Dis* 2007;195,S26–S35. DOI: 10.1086/516717.

23. Hippisley-Cox J, Coupland CAC, Logan R. Risk of adverse gastro-intestinal outcomes in patients taking cyclooxygenase-2 inhibitors or conventional nonsteroidal anti-inflammatory drugs: population based nested case-control analysis. *BMJ* 2005;331:1310–1316. DOI: 10.1136/bmj.331.7528.1310.
24. Hippisley-Cox J, Stables D, Pringle M. QRESEARCH: a new general practice database for research. *Inform Prim Care* 2004;12(1):49–50.
25. Christensen S, Riis A, Nørgaard M, Thomsen W, Sørensen HT. Introduction of newer selective cyclooxygenase-2 inhibitors and rates of hospitalization with bleeding and perforated peptic ulcer. *Aliment Pharmacol Ther* 2007; 25:907–912. DOI: 10.1111/j.1365-2036.2007.03274.x.
26. Dalton SO, Johansen C, Poulsen AH, et al. Cancer risk among users of neuroleptic medication: a population-based cohort study. *Br J Cancer* 2006;95:934–939. DOI: 10.1038/sj.bjc.6603259; Published online: 22 August 2006.
27. Penning-van Beest FJ, Termorshuizen F, Goettsch WG, Klungel OH, Kastelein JJ, Herings RM. Adherence to evidence-based statin guidelines reduces the risk of hospitalizations for acute myocardial infarction by 40%: a cohort study. *Eur Heart J* 2007;28(2):154–159. DOI: 10.1093/eurheartj/ehl391; Published online: 7 December 2006.
28. Goettsch WG, de Jong RB, Kramarz P, Herings RM. Developments of the incidence of osteoporosis in The Netherlands: a PHARMO study. *Pharmacoepidemiol Drug Saf* 2007;16(2):166–172. DOI: 10.1002/pds.1245.
29. Corrao G, Zambon A, Conti V, et al. Menopause hormone replacement therapy and cancer risk: an Italian record linkage investigation. *Ann Onco*l 2007;19:150– 155. DOI: 10.1093/annonc/mdm404; Published online: 4 September 2007.
30. Corrao G, Botteri E, Bertu` L, Zambon A, Favilli S. Exploring the effect of transient exposure on the risk of acute events by means of time-window designs: an application to fluoroquinolone antibacterials and arrhythmia. *Pharmacoepidemiol Drug Saf* 2006;15:31–37. DOI: 10.1002/pds.1125.
31. 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.
32. MonamiM, Balzi D, Lamanna C, et al. Are sulphonylureas all the same? A cohort study on cardiovascular and cancer-related mortality. *Diabetes Metab Res Rev* 2007;23(6):479–484. DOI: 10.1002/dmrr.736; Published Online: 23 March 2007.
33. http://www.whocc.no/atcddd/ [Accessed 14 February 2009].
34. 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;148:43–49. DOI: 10.3233/978-1-60750-043-8-43.
35. Avillach P, Mougin 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. DOI: 10.3233/978-1-60750-044-5-190.
36. http://www.nlm.nih.gov/research/umls/about_umls.html [Accessed 14 February 2009].
37. http://www.who.int/healthinfo/paper31.pdf [Accessed 06 January 2009].
38. Platt R, Davis R, Finkelstein J, et al. Multicenter epidemiologic and health services research on therapeutics in the HMO Research Network Center for Education and Research on Therapeutics. *Pharmacoepidemiol Drug Saf* 2001; 10: 373–377. DOI: 10.1002/pds.607; Published Online: 14 December 2001.
39. Lazarus R, Yih K, Platt R. Distributed data processing for public health surveillance. *BMC Public Health* 2006; 6:235. DOI: 10.1186/1471- 2458-6-235.
40. Lieu TA, KulldorffM, Davis RL, et al. Real-time vaccine safety surveillance for the early detection of adverse events. *Med Care* 2007;45(10 Supl 2):S89–S95. DOI: 10.1097/MLR.0b013e3180616c0a.
41. Brown JS, Kulldorff M, Chan KA, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. *Pharmacoepidemiol Drug Saf* 2007; 16: 1275–1284. DOI: 10.1002/ pds.1509; Published Online: 22 October 2007.
42. McMurry AJ, Gilbert CA, Reis BY, et al. A self-scaling, distributed information architecture for public health, research, and clinical care. *J Am Med Inform Assoc* 2007;14(4):527–533. DOI: 10.1197/jamia.M2371.
43. Velentgas P, Bohn RL, Brown JS, et al. A distributed research network model for post-marketing safety studies: the meningococcal vaccine study. *Pharmacoepidemiol Drug Saf* 2008;17:1226–1234. DOI: 10.1002/pds.1675; Published Online: 27 October 2008.

44. Brown JS, Kulldorff M, Petronis KR, et al. Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. *Pharmacoepidemiol Drug Saf* 2009; 18(3): 226–234. DOI: 10.1002/pds.1706; Published Online: 15 January 2009.
45. http://www.fda.gov/oc/initiatives/criticalpath/behrman_interview.html [Accessed 23 February 2009].
46. http://omop.fnih.org/?q=node/34 [Accessed 23 February 2009].
47. Brown J, Lane K, Moore K, Platt R. Defining and Evaluating Possible Database Models to implement the FDA Sentinel Initiative. http://www.regulations.gov/ fdmspublic/component/main?main=Document Detail &d=FDA-2009-N-0192-0005. [Accessed 12 June 2009].
48. Rynning E. Public trust and privacy in shared electronic health records. *Eur J Health Law* 2007;14(2):105–112.
49. van Leerdam ME, Vreeburg EM, Rauws EA, et al. Acute upper GI bleeding: did anything change? Time trend analysis of incidence and outcome of acute upper GI bleeding between 1993/1994 and 2000. *Am J Gastroenterol* 2003;98(7):1494–1499. DOI: 10.1111/j.1572-0241.2003.07517.x.
50. Rockall TA, Logan RF, Devlin HB, Northfield TC. Incidence of and mortality from acute upper gastrointestinal haemorrhage in the United Kingdom. Steering committee and members of the National Audit of Acute Upper Gastrointestinal Haemorrhage. *BMJ* 1995;311(6999):222–226.
51. Laporte JR, Vidal X, Carne X. Non-steroidal anti-inflammatory drugs in elderly people. Gastrointestinal bleeding is common. BMJ 1995;311:391–392.
52. Loperfido S. Incidence of gastrointestinal bleeding in Italy is similar to that in Britain. *BMJ* 1995;311(7009): 874.
53. Longstreth GF. Epidemiology of hospitalization for acute upper gastrointestinal hemorrhage: a population-based study. *Am J Gastroenterol* 1995; 90(2): 206–210.
54. Blatchford O, Davidson LA, Murray WR, Blatchford M, Pell J. Acute upper gastrointestinal haemorrhage in west of Scotland: case ascertainment study. *BMJ* 1997;315:510–514.
55. van Leerdam ME. Epidemiology of acute upper gastrointestinal bleeding. *Best Pract Res Clin Gastroenterol* 2008;22(2):209–224. DOI: 10.1016/j.bpg.2007. 10.011.
56. Hernandez Diaz S, Garcia Rodriguez LA. Association of nonsteroidal anti-inflammatory drugs and upper gastrointestinal bleeding/perforation. *Arch Intern Med* 2000;160:2093–2099.
57. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005;58:323–337. DOI: 10.1016/j.jclinepi.2004.10.012.
58. Strom BL, Kimmel SE (eds.). Textbook of Pharmacoepidemiology. John Wiley & Sons: Chichester, 2006;pp. 167–170.
59. 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. DOI: 10.1038/ncprheum0652.

Chapter 4

Facing up to the Challenge

'No signal generation technique ensures identification of all adverse effects; in reality, signal generation works like a fish net: certain fish escape.'

– W.K. Amery, 1999

4.1

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

Preciosa M Coloma, MD, MSc[1], Gianluca Trifirò, MD, PhD[1,2], Martijn J Schuemie, PhD[1], Rosa Gini, PhD[3], Ron Herings, PhD[4], Julia Hippisley-Cox, MD[5], Giampiero Mazzaglia, MD, PhD[6], Gino Picelli, PhD[7], Giovanni Corrao, PhD[8], Lars Pedersen, PhD[9], Johan van der Lei, MD, PhD[1], Miriam Sturkenboom, PharmD, PhD[1,10], on behalf of the EU-ADR consortium

[1]Department of Medical Informatics, Erasmus Medical Center, Rotterdam, the Netherlands; [2]Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Messina, Italy; [3]Agenzia Regionale di Sanità della Toscana, Florence, Italy; [4]PHARMO Institute, Utrecht, the Netherlands; [5]Division of Primary Care, School of Community Health Sciences, University of Nottingham, UK; [6]Società Italiana di Medicina Generale, Florence, Italy; [7]Pedianet-Società Servizi Telematici SRL, Padova, Italy; [8]Department of Statistics, Università di Milano-Bicocca, Milan, Italy; [9]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; [10]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands

Pharmacoepidemiology and Drug Safety
Article first published online: 8 Feb 2012. DOI: 10.1002/pds.3197

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.

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 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 underreporting, 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 Canada[12] 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), Pedianet (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 system[15] 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]

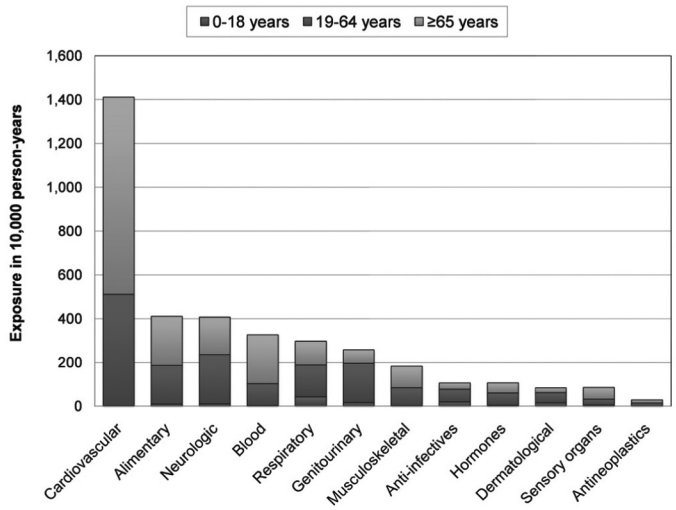


Figure 1 | Overview of exposure to all drugs (grouped according to ATC first level) by age categories ***(see page 307 for colour figure).***

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).

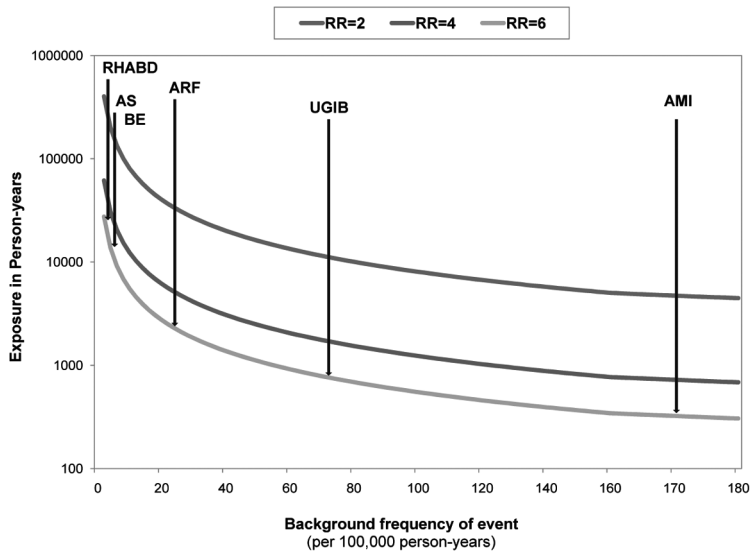


Figure 2 | Relationship between background incidence rate and amount of drug exposure required to identify a potential safety signal ***(see page 307 for colour figure).***

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 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.

Table 1 | Amount of d rug 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 *(per 100 000 person-years)*	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 *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*	
			Actual*	Simulated†		Actual*	Simulated†		Actual*	Simulated†
Acute Myocardial Infarction	170.8	4 706	531 *(23)*	1 018 *(44)*	721	939 *(41)*	1 435 *(63)*	322	1 107 *(48)*	1 607 *(70)*
Upper Gastrointestinal Bleeding	66.8	12 028	361 *(16)*	836 *(37)*	1 842	755 *(33)*	1 237 *(54)*	822	905 *(40)*	1 403 *(61)*
Acute Renal Failure	26.4	30 397	221 *(10)*	638 *(28)*	4 654	534 *(23)*	1 021 *(45)*	2 078	730 *(32)*	1 201 *(52)*
Bullous Eruptions	5.9	135 993	74 *(3)*	341 *(15)*	20 823	282 *(12)*	729 *(32)*	9 297	404 *(18)*	885 *(39)*
Anaphylactic Shock	5.7	141 939	72 *(3)*	330 *(14)*	21 733	272 *(12)*	723 *(32)*	9 704	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 = 2 289. 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 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 *(per 100 000 person-years)*	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 *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*	
			Actual*	Simulated†		Actual*	Simulated†		Actual*	Simulated†
Acute Myocardial Infarction	170.8	17 422	306 *(13)*	765 *(33)*	2 301	702 *(31)*	1 175 *(51)*	937	883 *(39)*	1 387 *(61)*
Upper Gastrointestinal Bleeding	66.8	44 524	184 *(8)*	542 *(24)*	5 882	487 *(21)*	972 *(42)*	2 395	691 *(30)*	1 164 *(51)*
Acute Renal Failure	26.4	112 520	89 *(4)*	372 *(16)*	14 864	323 *(14)*	795 *(35)*	6 054	485 *(21)*	965 *(42)*
Bullous Eruptions	5.9	503 410	9 *(<1)*	169 *(7)*	66 502	146 *(6)*	466 *(20)*	27 083	236 *(10)*	662 *(29)*
Anaphylactic Shock	5.7	525 420	9 *(<1)*	164 *(7)*	69 410	140 *(6)*	464 *(20)*	28 267	231 *(10)*	652 *(28)*
Rhabdomyolysis	2.5	1 198 983	1 *(<1)*	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 = 2 289. 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.

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 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).

Table 3 | Total exposure time in person-years (PYs) and number of drugs (distinct ATC 5th 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 5 000 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)	9 721	45	0	0
TOTAL	36 085 254	2 243	253	85

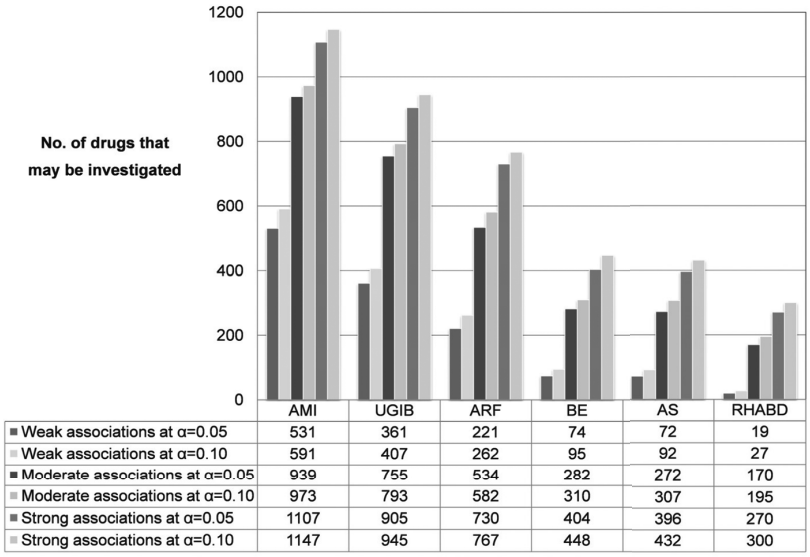
† Total number (count) of drugs within each class recorded with adequate exposure in EU-ADR to detect at least one of the events of interest.

‡ Exposure of 5 000 person-years corresponds to the exposure which 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 which would allow for detection of a 'moderate' association (i.e., relative risk of 4) with a rare event such as rhabdomyolysis.

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., α= 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).



	AMI	UGIB	ARF	BE	AS	RHABD
Weak associations at α=0.05	531	361	221	74	72	19
Weak associations at α=0.10	591	407	262	95	92	27
Moderate associations at α=0.05	939	755	534	282	272	170
Moderate associations at α=0.10	973	793	582	310	307	195
Strong associations at α=0.05	1107	905	730	404	396	270
Strong associations at α=0.10	1147	945	767	448	432	300

Figure 3 | Change in the number of drugs that may be investigated for signal detection when the acceptable probability of error is increased from α= 0.05 to α= 0.10 ***(see page 308 for colour figure).***

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, 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 EUADR 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.

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 ¼. 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:

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

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 α= 0.10).

Event	Incidence Rates *(per 100 000 person-years)*	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 *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*	
			Actual *	Simulated†		Actual *	Simulated†		Actual *	Simulated†
Acute Myocardial Infarction	170.8	3 577	591 *(26)*	1 075 *(47)*	572	973 *(43)*	1 463 *(64)*	262	1 147 *(50)*	1 641 *(72)*
Upper Gastrointestinal Bleeding	66.8	9 143	407 *(18)*	885 *(39)*	1 462	793 *(35)*	1 285 *(56)*	669	945 *(41)*	1 445 *(63)*
Acute Renal Failure	26.4	23 105	262 *(11)*	700 *(31)*	3 694	582 *(25)*	1 067 *(47)*	1 691	767 *(34)*	1 260 *(55)*
Bullous Eruptions	5.9	103 373	95 *(4)*	379 *(17)*	16 525	310 *(14)*	774 *(34)*	7 564	448 *(20)*	923 *(40)*
Anaphylactic Shock	5.7	107 892	92 *(4)*	375 *(16)*	17 247	307 *(13)*	764 *(33)*	7 895	432 *(19)*	918 *(40)*
Rhabdomyolysis	2.5	246 204	27 *(1)*	253 *(11)*	39 357	195 *(8)*	570 *(25)*	18 015	300 *(13)*	756 *(33)*

* 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= 2 289. 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

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 α=0.10).

Event	Incidence Rates *(per 100 000 person-years)*	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 *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*		Required Exposure (Person-Years)	No. of drugs that can be investigated *(% of total)*	
			Actual *	Simulated†		Actual *	Simulated†		Actual *	Simulated†
Acute Myocardial Infarction	170.8	16 431	310 *(14)*	775 *(34)*	2 181	721 *(32)*	1 179 *(52)*	891	888 *(39)*	1 390 *(61)*
Upper Gastrointestinal Bleeding	66.8	41 991	190 *(8)*	556 *(24)*	5 575	496 *(22)*	978 *(43)*	2 278	706 *(31)*	1 176 *(51)*
Acute Renal Failure	26.4	106 119	92 *(4)*	376 *(16)*	14 088	329 *(14)*	799 *(35)*	5 756	490 *(21)*	975 *(42)*
Bullous Eruptions	5.9	474 772	10 *(<1)*	178 *(8)*	63 029	152 *(7)*	476 *(21)*	25 752	244 *(11)*	678 *(30)*
Anaphylactic Shock	5.7	495 530	9 *(<1)*	170 *(7)*	65 784	147 *(6)*	470 *(21)*	26 878	238 *(10)*	661 *(29)*
Rhabdomyolysis	2.5	1 130 774	1 *(<1)*	89 *(4)*	150 117	67 *(3)*	321 *(14)*	61 334	154 *(7)*	483 (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= 2 289. 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

REFERENCES

1. Goldman SA. Limitations and strengths of spontaneous reports data. *Clin Ther* 1998;20 Suppl C: C40-4.
2. Stephenson WP, Hauben M. Data mining for signals in spontaneous reporting databases: proceed with caution. *Pharmacoepidemiol Drug Saf* 2007;16(4):359–65.
3. Molokhia M, Tanna S, Bell D. Improving reporting of adverse drug reactions: systematic review. *Clin Epidemiol* 2009;1:75–92.
4. 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).
5. Krumholz HM, Ross JS, Presler AH, et al. What have we learnt from Vioxx? *BMJ* 2007;334(7585):120–3.
6. Edwards IR. What are the real lessons from Vioxx? *Drug Saf* 2005;28(8):651–8.
7. Härmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. *Eur J Clin Pharmacol* 2008;64(8):743–52. Epub2008 Jun 4.
8. 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. Epub2009 Jul 27.
9. The Sentinel Initiative — National Strategy for Monitoring Medical Product Safety. May 2008. http://www.fda.gov/downloads/Safety/FDAsSentinelInitiative/UCM124701.pdf. (Accessed 15 April 2011).
10. Observational Medical Outcomes Partnership. http://omop.fnih.org. (Accessed 15 February 2011).11. 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.
11. About the Drug Safety Effectiveness Network. http://www.cihr-irsc.gc.ca/e/39389.html. (Accessed 01 March 2011).
12. 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.
13. Coloma PM, Schuemie MJ, Trifiró G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom M, 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.
14. WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2011. http://www.whocc.no/atc_ddd_index/. (Accessed 14 February 2011).
15. 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.
16. Unified Medical Language System. http://www.nlm.nih.gov/research/umls/about_umls.html. (Accessed 28 February 2011).
17. Age standardization of rates: a new WHO standard. http://www.who.int/healthinfo/paper31.pdf. (Accessed 20 February 2011).
18. Avillach P, Mougin 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.
19. Beaumont JJ, Breslow NE. Power considerations in epidemiologic studies of vinyl chloride workers. *Am J Epidemiol* 1981;114(5):725–34.
20. 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).
21. Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials* 1990;11(2):116–28.
22. Altman D. Practical Statistics for Medical Research. Boca Raton: Chapman and Hall London, 1991. pp.210-212.
23. 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.

24. Strom BL, Kimmel SE (eds.). Textbook of Pharmacoepidemiology. John Wiley & Sons: Chichester, 2006.
25. 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.
26. 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.
27. 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.
28. 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.
29. 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.
30. Thompson PD, Clarkson P, Karas RH. Statin-associated myopathy. *JAMA* 2003;289(13):1681–90.
31. Joy TR, Hegele RA. Narrative review: statin-related myopathy. *Ann Intern Med* 2009;150(12):858–68.
32. Meyboom RH, Egberts AC, Gribnau FW, et al. Pharmacovigilance in perspective. *Drug Saf* 1999;21(6):429–47.
33. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. http://www.encepp.eu. (Accessed 01 May 2011).
34. 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.
35. 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.
36. 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.
37. 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].
38. Sturkenboom M, Nicolosi A, Cantarutti L, et al., NSAIDs Paediatric Research Group. Incidence of mucocutaneous reactions in children treated with niflumic acid, other nonsteroidal anti-inflammatory drugs, or nonopioid analgesics. *Pediatrics* 2005;116(1):e26-33. Epub 2005 Jun 1.
39. Hippisley-Cox J, Stables D, Pringle M. QRESEARCH: a new general practice database for research. *Inform Prim Care* 2004;12(1):49–50.
40. Ehrenstein V, Antonsen S, Pedersen L. Existing data sources for clinical epidemiology: Aarhus University Prescription Database. *Clin Epidemiol* 2010;2:273–9.
41. 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.
42. 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; 0(3):161–168.
43. van den Boogaard CH, Breekveldt-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.
44. 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.
45. 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.
46. 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.
47. 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.

4.2

The power of electronic healthcare databases for active drug safety surveillance in children and adolescents

Sandra de Bie, MD, MSc[1,2], Preciosa M Coloma, MD, MSc[1], Carmen Ferrajolo, PharmD, MSc[1,3] Katia MC Verhamme, MD, PhD[1], Gianluca Trifirò, MD, PhD[1,4], Martijn J Schuemie, PhD[1], Sabine MJM Straus, MD, PhD[1,2], Rosa Gini, PhD[5], Ron Herings, PhD[1,6], Giampiero Mazzaglia, MD, PhD[7], Gino Picelli[8], Arianna Ghirardi, MSc[9], Lars Pedersen, PhD[10], Bruno HC Stricker, MB,PhD[1,11,12], Johan van der Lei, MD, PhD[1], and Miriam CJM Sturkenboom, PharmD,PhD[1,11] on behalf of the EU-ADR consortium

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

ABSTRACT

Background. Traditional pharmacovigilance activities do not focus specifically on children and medicines in children are frequently being prescribed off-label based on extrapolating experience in adults to children. In Europe, the EU-ADR Project (Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge) aims to use information from various electronic healthcare record (EHR) databases to produce a computerized integrated system for the early detection of drug safety signals. This might also prove to be a useful tool in pediatric pharmacovigilance.

Objectives. To provide estimates on the number of drugs and incidence rates (IRs) of adverse events that can be monitored in children and adolescents in the EU-ADR database network.

Methods. Demographic, clinical events, and outpatient drug prescription/dispensing data were obtained for children and adolescents (0 to 18 years) from seven population-based EHR databases of the EU-ADR network from Denmark, Italy, and the Netherlands. Data were analyzed for the period January 1st 1996 through December 31st 2008. We estimated the number and types of drugs for which specific adverse events can be monitored as a function of actual drug use, minimally detectable relative risk (RR), and empirically-derived incidence rates for 10 events deemed to be important in pharmacovigilance. The same was done for adverse events frequently reported in children, using age-dependent IRs described in literature.

Results. The pediatric population (0-18 years) of the EU-ADR network comprised 4 838 146 individuals contributing 25 575 132 person years (PYs) of follow-up during the study period. Within this population, a total of 2 170 drugs (i.e., distinct chemical substances) were prescribed during the study period with a total drug exposure of 1 610 631 PYs. Eighteen of the 2 170 drugs (0.8%) comprised half of the total drug exposure while 90% of the total drug exposure in PYs was represented by 158 drugs (7.3%). For a relatively frequent event such as upper gastrointestinal bleeding (IR= 14.4/100 000 PYs), there were 39 drugs (comprising 66% of total exposure in PY) for which an association with a RR≥4, if present, can be investigated. For rare events such as anaphylactic shock and bullous eruptions, there were 8 drugs (comprising 35% of total exposure) and 9 drugs (comprising 37% of total exposure) respectively, for which an association of same magnitude can be investigated. Based on literature-derived IR, there is a higher number of drugs that can be monitored for the events febrile convulsions, suicide attempt, and epilepsy at the same magnitudes of risk.

Conclusion. Drug use in children is rare and shows little variation; only 18 out of the total 2 170 prescribed drugs make up half of the total exposure time to drugs in the pediatric population. The number of drugs with enough exposure to detect safety signals for rare events in children and adolescents using EHRs is limited. Mining within EHR databases seems especially promising for events that have a high background incidence in the pediatric population and for drugs with a large amount of exposure. Inter-continental collaboration will be necessary to gain overall adequate statistical power for pediatric drug safety signal detection.

INTRODUCTION

Medicines in children are frequently being prescribed off-label as little information is available from clinical trials and data is extrapolated from adults to children.[1] The number of clinical trials is expected to increase with the introduction of new legislation with respect to the approval of drugs used in children in the United States (US) (2002) and in Europe (2007).[2-4] However, clinical trials are primarily designed to assess therapeutic efficacy and have well known limitations for the assessment of risks since the number of included children is often too limited to draw firm conclusions with respect to safety. These limitations underline the importance of monitoring potential safety signals of drugs with a pediatric indication during the post-marketing phase. Currently, spontaneously reported adverse drug reactions (ADRs) and post-marketing safety studies are the most important sources for identifying such safety signals both in children and adults.[5-6]

Although there is a fair amount of experience with using spontaneous reporting systems (SRS) to study vaccine safety in children,[7-12] the usefulness of such systems for routine safety surveillance of non-vaccine medicines in children is limited. Studies on ADR reporting primarily focus on the number of ADR-related hospital admissions or are descriptive in nature.[13-16] Little is known on the capability for signal detection in the pediatric population. And although safety signal detection in SRS have proven their value, mainly as a hypothesis-generating tool in adults, there are well-recognized limitations and biases such as selective underreporting, stimulated reporting and the lack of exposure data.[17-19]

To complement SRS and other traditional pharmacovigilance systems, several initiatives, both in the US and in Europe, have set up population-based surveillance systems that make use of longitudinal healthcare data.[20-22] 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 funded by the European Commission. The EU-ADR project is a collaboration of 18 public and private institutions representing academic research, general practice, healthcare services administration, and the pharmaceutical industry. EU-ADR aims to exploit information from various electronic healthcare records (EHR) and other biomedical databases in Europe to produce a computerized integrated system for the early detection of drug safety signals. [20] In a previous study using data from the EU-ADR project, Coloma *et al.*,[23] provided estimates of the number and types of drugs that can be monitored for safety surveillance in the general population using EHR databases. It was concluded that signal detection within these data was possible, but that the statistical power might be low for infrequently used drugs or for rare outcomes. Although this study included pediatric data, no specific analyses were performed on the pediatric sub-population. As disease prevalence, pharmacokinetics and pharmacodynamics, as well as drug exposure are different for children compared to adults,[24-26] pediatric ADR data should be analyzed separately. Many EHR databases contain data on large numbers of children, which make such databases a good source for safety monitoring.

In this study, we first aimed to provide estimates of the number of drugs that have enough exposure to be monitored in children and adolescents based on actual drug use, minimally detectable relative risk (RR), and empirically-derived incidence rates (IRs) for 10 adverse events currently being investigated within the EU-ADR project. Second, the same estimation of the number of drugs to be monitored was done for adverse events reported as frequently occurring in children, using age-dependent IRs described in literature. Third, we aimed to provide information on the range of IRs that can be monitored in children and adolescents based on the actual drug-exposure in the pediatric cohort and the minimal detectable RR.

METHODS

Data sources and setting

We used data from the EU-ADR network, of which a detailed description has been published earlier.[20, 27] In summary, the EU-ADR platform currently comprises data from eight EHR databases in four European countries. For the current study we used pediatric data from seven of the databases (from three European countries: Denmark; Italy; and the Netherlands). Health Search/CSD LPD (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 and Tuscany (ARS) 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 and Denmark and covered by obligatory health insurance in the Netherlands. In all of these countries general practitioners function as gatekeepers of the healthcare system. Children aged 0 to 18 years included in these databases were included in the current study.

The study period ran from January 1st 1996 to December 31st 2008. Follow up started after a run-in period of 365 days. This run-in period was required to determine if an event was incident. The run-in period was omitted for children younger than one 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 on, 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 the data into a common format followed by local aggregation using custom-built software, Jerboa©.[20]

Drug exposure

Drug use was categorized using the World Health Organization's (WHO) Anatomical Therapeutic Chemical (ATC) classification system.[28] Drug exposure was measured in terms of person-years (PYs). We further analyzed drug use by anatomical main groups (ATC 1st level) and by chemical substances (ATC 5th level). Drug exposure was stratified according to one year age-categories.

Drugs were subsequently categorized based on the total amount of drug exposure in PYs as follows: <10 PYs; >10 - ≤50 PYs; >50 - ≤100 PYs; >100 - ≤500 PYs; >500 - ≤1 000 PYs; >1 000 - ≤5 000 PYs; >5 000 - ≤10 000 PYs; and >10,000 PYs. Furthermore, the number of drugs (distinct chemical substances) that accounted for 50% and 90% of the total drug exposure in the population were calculated.

Events

The identification of the events of interest in EU-ADR has been described in detail in an earlier publication.[20] Only those events considered to be most serious and most relevant (generally within the context of pharmacovigilance in adults) were included. In summary, events of interest were identified 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. Since the databases included in EU-ADR use four nomenclature systems to code the events, these different terminologies were first mapped using the Unified Medical Language System1 (UMLS1), a biomedical terminology integration system handling more than 150 terminologies.[29] The processes of terminology mapping, harmonization, and benchmarking of event extractions from the various databases have been described in more detail in other publications.[27, 30-31]

The following events have been mapped and harmonized in the EU-ADR platform: (1) acute liver injury; (2) acute myocardial infarction; (3) acute renal failure; (4) anaphylactic shock; (5) bullous eruptions; (6) cardiac valve fibrosis; (7) hip fractures; (8) neutropenia; (9) acute pancreatitis; (10) pancytopenia; (11) progressive multifocal leukoencephalopathy; (12) rhabdomyolysis; and (13) upper gastrointestinal bleeding. Not all events selected within the EU-ADR platform were considered relevant to study in children and adolescents. Therefore, only events occurring in children with an annual incidence rate of >1/100 000 PYs were included. For children and adolescents the following 10 events were considered relevant and were included in this study: (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.

The events as currently monitored in the EU-ADR project are not pediatric-specific and therefore the analyses were extended to include events that are recognized as posing risk in children and adolescents. The following serious events were chosen based on studies on ADR-related hospital admissions in children and reviews of spontaneous reported adverse events in children:[13, 32-33] (1) completed suicide; (2) suicide attempt; (3) febrile convulsions; and (4) epilepsy.

Statistical analysis

Required amount of drug exposure to detect safety signals

Given the pooled population-based IR of the 10 events that were evaluated 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 (weak signal), 4 (moderate signal), and 6 (strong signal) 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 and discussed previously by Coloma *et al.*[23] We subsequently determined the number of drugs for which there would be sufficient data for safety monitoring. The number of drugs was expressed as the number of unique chemical substances (i.e., ATC 5th level). For the drugs with enough exposure to detect the RR of interest, the proportion of the PYs of exposure to these drugs, compared to the total PYs of exposure for all drugs, was calculated.

Based on the actual exposure and hypothetical incidences of adverse events it was also calculated for how many drugs within the anatomical main groups (ATC 1st level) was there enough exposure to detect associations with varying magnitudes. The following (hypothetical) incidences were considered: 1/100 000 PYs; 10/100 000 PYs; 50/100 000 PYs; 100/100 000 PYs; and 500/100 000 PYs.

Age-specific incidence rates for the following additional events were obtained from the literature: (1) completed suicide; (2) suicide attempt; (3) febrile convulsions; and (4) seizures and convulsions (see Appendix).

Required incidence based on the actual drug exposure to detect safety signals

Analogous to the analysis described under the heading 'Required amount of drug exposure to detect safety signals,' we calculated the range of IRs of events that can be monitored to detect weak (RR= 2), moderate (RR= 4) or strong (RR= 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 by age.

Stratification by age

Results were stratified in one-year age categories and according to four age categories based on the guidelines of the International Conference of Harmonization (ICH): 0-<2 years, 2-≤5 years; 6-≤11 years and 12-<18 years.[34]

RESULTS

The pediatric population of the EU-ADR network comprised 4 838 146 children and adolescents (0 to 18 years) contributing 25 575 132 PYs of follow-up between 1996 and 2008. Of these PYs of follow-up, 12.8% were for children aged 0 to<2years, 22.2% for children aged 2 to ≤5 years, 32.7% for children aged 6 to ≤11 years and 32.3% for adolescents aged 12 to <18 years.

A total of 2 170 drugs (i.e., distinct chemical substances) were prescribed or dispensed to this population during the study period with a total exposure of 1 610 631 PYs. An overview of drug exposure, at the anatomical level of the ATC classification across different age categories is illustrated in Figure 1. Up to 12 years of age, the drug classes with the highest exposure are respiratory drugs (6.5-9.3 PYs of exposure/100 000 PYs of follow up) and anti-infective drugs (2.1-5.2 PYs of exposure/100 000 PYs of follow up). Likewise, from 12 years on the genitourinary drugs were increasingly prescribed up to 10.2 PYs of exposure/100 000 PYs of follow up.

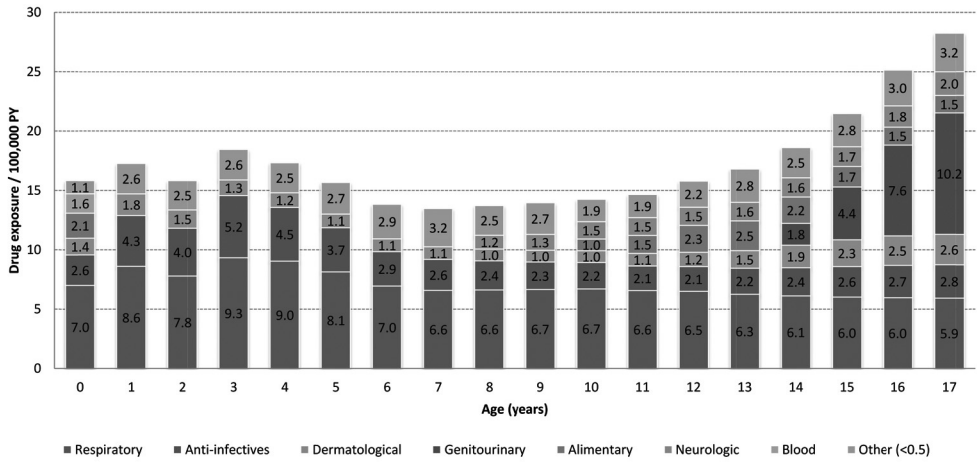


Figure 1 | Drug exposure in person-years/ 100 000 person-years by age ***(see page 308 for colour figure).***

NOTE: Drug exposure is aggregated on the 1st ATC level (anatomical main group). 'Other' represents all groups with an exposure of < 1.0 / 100 000 PYs.

Required exposure for monitoring of pre-defined events within EU-ADR and for events frequently occurring in children

The number of drugs on a chemical substance level (5th ATC level) that have enough exposure to detect weak (RR= 2), moderate (RR= 4) or strong (RR= 6) associations for the 10 EU-ADR events are presented in Table 1. Since the numbers are low, these results were not further stratified by age. The stronger the association to be studied, the higher is the number of drugs that can be studied, which is expected from the power calculations. Conversely, the higher the IR of the event the higher the number of drugs that can be studied. Considering the IR of upper gastrointestinal bleeding (UGIB) within our pediatric population of 14.4/100 000 PYs, which was relatively high compared to the other events included in the study, a minimal exposure of 55 725 PYs was required to detect a weak association (RR≥2). Within our population five drugs fulfilled this criterion. These five drugs made up 26.2% of the total drug exposure in PYs. To detect a moderate association (RR≥4) with UGIB, a minimal exposure of 8 532 PYs was required; 39 drugs, covering 66.3% of the total drug exposure had this minimal exposure. Finally, to assess a strong

Table 1 | Amount of required amount drug exposure to identify potential safety signals

Event Type	IR/ 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	15.31	52 501	6	29.5	8 039	42	67.8	3 589	81	80.4
Upper GI bleeding	14.42	55 725	5	26.2	8 532	39	66.3	3 810	79	79.9
Neutropenia	8.10	99 259	2	13.0	15 198	25	56.9	6 786	48	70.5
Acute liver injury	3.96	202 733	0	0	31 041	9	37.3	13 860	26	57.8
Pancytopenia	3.73	215 469	0	0	32 991	9	37.3	14 730	25	56.9
Bullous eruption	3.58	224 394	0	0	34 358	9	37.3	15 341	24	56.0
Anaphylactic shock	3.23	248 526	0	0	38 053	8	35.0	16 990	20	52.1
Cardiac valve fibrosis	2.91	275 840	0	0	42 235	8	35.0	18 858	15	46.6
Acute renal failure	1.55	517 050	0	0	79 168	3	17.9	35 348	9	37.3
Acute pancreatitis	1.55	519 664	0	0	79 568	3	17.9	35 527	9	37.3

Drugs (N): Number of drugs at 5th ATC, chemical substance level that have enough PY of exposure to detect a potential signal (total= 2 170).
% of Exp: Proportion of PYs of exposure of the drugs with enough exposure compared to the total PYs of exposure for all drugs.

association (RR≥6), a total of 79 drugs (79.9% of the total exposure) had enough exposure. The IR of pancreatitis was low, 1.6/100 000 PYs. Since the IR was low, none of the drugs had enough exposure to detect a weak signal (RR≥2), 3 drugs (17.9% of the total exposure) had enough exposure to detect a moderate signal (RR≥4), and 9 drugs (37.3% of the total exposure) had enough exposure to detect a strong signal (RR≥6).

The number of drugs, stratified at the anatomical level of the ATC classification, with enough exposure to study hypothetical IRs to detect weak, moderate or strong associations is given in Table 2. For drugs rarely prescribed in the study population (such as antineoplastic, antiparasitic, and cardiovascular drugs), no drug had enough exposure to monitor an association with RR≥2 for any of the hypothetical incidences ranging from 1 to 500/100 000 PYs. Respiratory drugs and anti-infectives were, however, included in those drugs having enough exposure to monitor associations of RR≥2, RR≥4 and RR≥6 for events having a (hypothetical) incidence of 10/100 000 PYs and higher.

As illustrated in Figure 2, only a small number of the drugs have a high exposure in the pediatric population, 53% of the drugs have a total exposure of less than 10 PYs. This is most pronounced in the youngest children, for which 75% of the drugs have a total exposure of less than 10 PYs. In the table accompanying Figure 2, the minimal detectable IRs for the exposure-categories for each RR is given: for drugs with an exposure of less than 10 PYs, IRs of maximal 765/1 000 PY can be detected for RR≥2, maximal 12/1 000 PY for RR≥4 and maximal 5.2/1 000 PY for RR≥6. The proportion of the drugs with an exposure of more than 1 000 PYs is less than 5% for all age categories, and is only 8.4% for the total pediatric population. An exposure of more than 1 000 PYs is necessary to detect IRs of more than 1.6/1 000 PYs with a RR≥2.

We estimated the power of the EU-ADR system to detect events frequently occurring in children based on the literature-derived IRs for 'completed suicide', 'suicide attempt', febrile seizures', and 'seizures and convulsions' (see Appendix). For events with a high incidence rates like febrile convulsion, a large number of drugs have enough exposure to detect a potential safety signal. However, for very rare events such as Reye's syndrome, only one drug has enough exposure to detect a strong association (RR=6) in children aged 2-≤ 5 years.

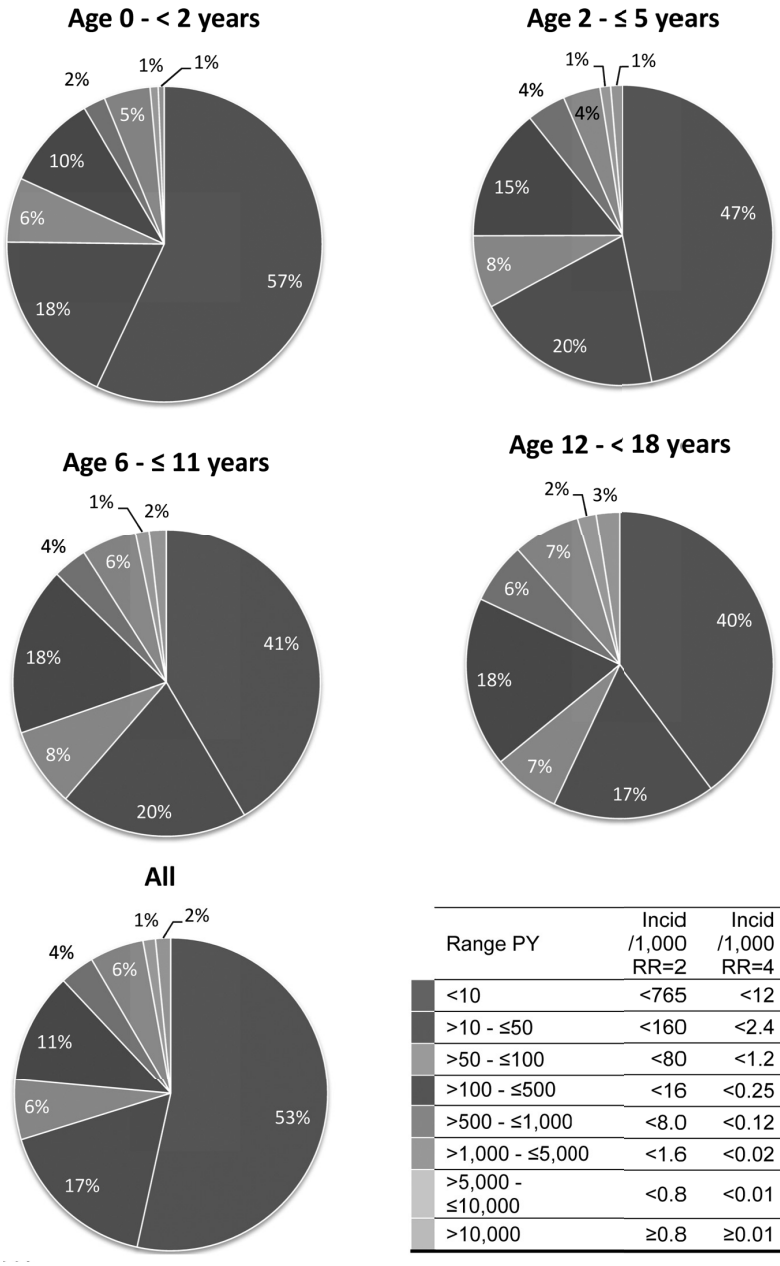
Range of incidence rates of events that can be monitored within the network

Eighteen of the 2 170 drugs (0.8%) made up 50% of the total drug-exposure (Table 3). For 0 to <2 years, 2 to ≤5 years; 6 to ≤11 years and 12 to <18 years there were 8 (0.6%), 8 (0.5%), 14 (0.9%) and 20 (1.0%) drugs prescribed/dispensed. These drugs have corresponding exposures of ≥7,024 PY (0 to <2 years), ≥10 951 PYs (2 to ≤5 years), ≥6 822 PYs (6 to ≤11 years) and ≥7 227 PYs (12 to <18 years). Based on these exposure data, for the age 0 to <2 years, events with IR>114/100 000 PYs (at RR≥2), IR>18/100 000 PYs (at RR≥4) and IR>7.8/100 000 PYs (at RR≥6) can be detected. For the age 2 to ≤5 years, events with IR>73/100 000 PYs (atRR≥2), IR>11/100 000 PYs (at RR≥4) and IR>5.4/100 000 PYs (at RR≥6) can be detected. For the age range 6 to ≤11 years, events with IR>118/100 000 PYs (at RR≥2), IR>18/100 000 PYs (at RR≥4) and IR>8.1/100 000 PYs (at RR≥6) can be detected. Finally for the age 12 to <18 years, events with IR of >111/100 000 PYs (at RR≥2),

Table 2 | Number of drugs (5th ATC level) by class with enough exposure to study the given incidences with the given RRs

1st ATC class *		≤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)	37 (23.1)	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)	35 (17.2)	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)	28 (12.1)	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)	27 (10.0)	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)	19 (11.2)	11 (6.5)	34 (20.1)	42 (24.9)
Cardiovascular [N=192]	N (%)	–	–	–	–	–	–	–	–	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)	11 (7.2)	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)	14 (19.4)	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)	18 (17.3)
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)	16 (16.5)
Antiparasitic [N=35]	N (%)	–	–	–	–	–	–	–	–	1 (2.9)	–	1 (2.9)	1 (2.9)	–	5 (14.3)	11 (31.4)
Antineoplastics [N=76]	N (%)	–	–	–	–	–	–	–	–	2 (2.6)	–	2 (2.6)	5 (6.6)	–	6 (7.9)	10 (13.2)
Total [N=2 170]	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)	133 (6.1)	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.



Range PY	Incid /1,000 RR=2	Incid /1,000 RR=4	Incid /1,000 RR=6
<10	<765	<12	<5.2
>10 - ≤50	<160	<2.4	<1.1
>50 - ≤100	<80	<1.2	<0.55
>100 - ≤500	<16	<0.25	<0.11
>500 - ≤1,000	<8.0	<0.12	<0.055
>1,000 - ≤5,000	<1.6	<0.02	<0.011
>5,000 - ≤10,000	<0.8	<0.01	<0.0055
>10,000	≥0.8	≥0.01	≥0.0055

Figure 2 | Distribution of exposure in PY by age-groups (5th ATC level, chemical subgroup) ***(see page 309 for colour figure)*****.**

The range in PY is given with the corresponding incidence rates of events that can be monitored. PY=person-years; incid=incidence

IR>17/100 000 PYs (at RR≥4) and IR>7.6/100 000 PYs (at RR≥6) can be detected (*data not shown)*. Data from 90% of the total drug exposure (represented by 158 drugs) will allow detection of events with IR>387/100 000 PYs (at RR≥2 and RR≥4) and IR>173/100000 PYs (at RR≥6) for 0 to <2 years; IR>281/100 000 PYs (at RR≥2 and RR≥4) and IR>125/100 000 PYs (at RR≥6) for 2 to ≤5 years; IR>313/100 000 PYs (at RR≥2 and RR≥4) and IR>140/100 000 PYs (at RR≥6) for 6 to ≤11 years; and IR>258/100 000 PYs (at RR≥2 and RR≥4) and IR>115/100 000 PYs (at RR≥6) for the age 12 to <18 years (*data not shown)*.

Table 3 | Drugs that cover 50% of the total drug exposure in person years by age-categories.

Age 0 to < 2 years *	**Age 2 to ≤ 5 years***	**Age 6 to ≤ 11 years***	**Age 12 to < 18 years***	**Total ***
Beclometasone [R03BA01] (13.1)	Beclometasone [R03BA01] (12.5)	Salbutamol [R03AC02] (6.8)	Levonorgestrel and estrogen [G03AA07] (11.9)	Beclometasone [R03BA01] (6.8)
Salbutamol [R03AC02] (10.5)	Salbutamol [R03AC02] (9.1)	Beclametasone [R03BA01] (6.4)	Sodium fluoride [A01AA01] (3.9)	Salbutamol [R03AC02] (6.2)
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)
Amoxicillin / Clavulanic acid [J01CR02] (4.9)	Amoxicillin [J01CA04] (5.0)	Fluticasone [R03BA05] (4.8)	Salbutamol [R03AC02] (3.2)	Levonorgestrel and estrogen [G03AA07] (4.6)
Phytomenadione (vitamin K) [B02BA01] (4.4)	Fluticasone [R03BA05] (5.0)	Cetirizine [R06AE07] (4.1)	Cyproterone and estrogen [G03HB01] (3.2)	Amoxicillin [J01CA04] (3.6)
Fluticasone [R03BA05] (3.8)	Budesonide [R03BA02] (4.3)	Budesonide [R03BA02] (3.5)	Cetirizine [R06AE07] (2.5)	Fluticasone [R03BA05] (3.4)
Budesonide [R03BA02] (3.6)	Clarithromycin [J01FA09] (3.8)	Amoxicillin [J01CA04] (3.4)	Beclametasone [R03BA01] (2.4)	Budesonide [R03BA02] (2.9)
Flunisolide [R03BA03] (3.6)	Flunisolide [R03BA03] (3.1)	Methylphenidate [N06BA04] (3.2)	Amoxicillin [J01CA04] (2.3)	Cetirizine [R06AE07] (2.6)
		Salmeterol and other drugs for obstructive airway diseases [R03AK06] (2.7)	Ferrous sulfate [B03AA07] (2.1)	Clarithromycin [J01FA09] (2.2)
		Clarithromycin [J01FA09] (2.7)	Methylphenidate [N06BA04] (1.8)	Sodium fluoride [A01AA01] (1.9)
		Desmopressin [H01BA02] (2.3)	Salmeterol and other drugs for obstructive airway diseases [R03AK06] (1.7)	Fluisolide [R03BA03] (1.7)
		Montelukast [R03DC03] (1.7)	Desloratadine [R06AX27] (1.6)	Methylphenidate [N06BA04] (1.6)
		Fluticasone (nasal) [R01AD08] (1.7)	Budesonide [R03BA02] (1.6)	Salmeterol and other drugs for obstructive airway diseases [R03AK06] (1.6)

Age 0 to < 2 years *	Age 2 to ≤ 5 years*	Age 6 to ≤ 11 years*	Age 12 to < 18 years*	Total *
		Terbutaline [R03AC03] (1.6)	Fluticasone [R03BA05] (1.6)	Terbutaline [R03AC03] (1.5)
			Levocetirizine [R06AE09] (1.4)	Cyproterone and estrogen [G03HB01] (1.2)
			Gestodene and estrogen [G03AA10] (1.4)	Fluticasone [R01AD08] (1.1)
			Clarithromycin [J01FA09] (1.3)	Montelukast [R03DC03] (1.1)
			Fluticasone (nasal) [R01AD08] (1.3)	Salbutamol and other drugs for obstructive airway diseases [R03AK04] (1.1)
			Terbutaline [R03AC03] (1.2)	
			Mometasone [R01AD09] (1.1)	

DISCUSSION

There is a growing number of initiatives evaluating the use of EHR databases as a source for safety signal detection.[21-22, 35-36] Some of these include data on children and adolescents; however, we are not aware of any specific analyses that have been carried out regarding the pediatric population. To our knowledge, this is the first study that explores the feasibility of using EHR databases as a source for safety signal detection in children and adolescents.

Despite the large number of children and adolescents included in the EU-ADR system, the number of drugs that have enough exposure to study weak, moderate or strong associations with the events currently monitored in EU-ADR network is limited. For a rare event like anaphylactic shock (incidence rate=3.2/100 000 PY) there were no drugs with enough exposure to study a weak association (RR≥2); there were 8 drugs to study a moderate association (RR≥4), and 20 drugs to study a strong association (RR≥6). These numbers are low compared to the total of 2 170 different drugs prescribed in the pediatric population. It is mainly for drugs that are known to be chronically used in children[25] (e.g., anti-infective drugs, respiratory drugs and hormones) that there is 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 (CNS)-drugs: ADHD (attention deficit-hyperactivity disorder)-drugs, anti-epileptics, antidepressants analgesic drugs.[33] Methylphenidate was the only neurological drug within the group of drugs that covered half of the total drug exposure in PYs.

This study showed that within the pediatric population of the EU-ADR database network, drug exposure is low and that a limited number of drugs cover the majority of the prescriptions.

The 1.6 million PYs of exposure were distributed over 2 170 individual drugs, compared to 2 289 for the overall population (all ages) in EU-ADR (95%). Of these 2 170 drugs, only 18 represent 50% of the entire exposure time while 158 drugs cover 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. In view of the total exposure time in PYs of all drugs, the number of drugs that have enough exposure to study anaphylactic shock is therefore relatively high. The 20 drugs that have enough exposure to study a strong association with anaphylactic shock (at RR≥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 incidence of 10/100 000 and up (29 drugs, covering 60% of the total exposure), while for events having an incidence of 50/100 000 and up weak associations can also be studied (20 drugs, covering 52% of the total exposure). It should be noted that these results have not been corrected for multiple testing.

The study was also limited by the low IRs of the 10 adverse events as directly derived from the EU-ADR data within the pediatric population. The low IRs are expected since the events were chosen based on safety issues which were more relevant in adults. Furthermore, the mechanisms of action of certain adverse events differ between children and adults. For example, hip fractures, which have the highest IR within this population (15.3/ 100 000 PYs), is caused by a high-energy trauma in 85-90% of the cases in children and are likely unrelated to drug use,[37-38] while in adults the main causes are falls and osteoporosis, which may be associated with use of certain drugs. The causal pathway for this particular event makes it less important to study in children. Also, since the symptoms of the same condition can differ between children and adults, there is a higher chance of misclassification if this is not accounted for in the selection of cases. In future initiatives to set up drug surveillance systems for the pediatric population using EHRs, it is very important to choose age-appropriate events with age-appropriate symptoms because as we have demonstrated, events with a higher incidence in children (such as febrile convulsions) require less PYs of exposure to study associations.

As emphasized in the recently published CIOMS VIII report, an important unaddressed question is whether the positive predictive value of mining longitudinal healthcare database as a source for signal detection will be higher than data mining in spontaneous reporting systems.[18] Trifirò and colleagues have tried to address this issue in a study where potential signals derived from the EU-ADR network were compared with signals derived from spontaneous reporting systems (SRSs).[39] The SRSs were more likely to detect potential signals for events with a low incidence in the general population and commonly regarded as drug-induced like bullous eruptions and anaphylactic shock. At the same time, it was noted that systems like EU-ADR may complement traditional SRS in the detection of adverse events that are frequent in the general population and are not commonly regarded to be drug-induced. This is in line with the results we obtained specifically for the pediatric population. For events with a low incidence and a high probability to be drug-induced only a small number of drugs had enough exposure to detect potential safety

signals. For events with a high incidence, a larger number of drugs could be studied. It is also important to note that although the number of drugs that can be studied for rare events is low, the drugs that can be studied have a relatively large exposure within the population and, hence, EHR databases appear to be able to detect associations for drugs that are frequently used. It is known that ADRs have the highest chance 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 effects. Because of the longitudinal nature of the data collection in EHR databases, signals may also be detected after long-term use of drugs, possibly even for rare diseases, and may thereby further complement SRSs.

Limitations and future directions

Our study illustrates that the capacity of EHR databases as a source for safety signal detection in children and adolescents is not only limited by the size of the population, but is mainly hampered by the fact that the majority of the drugs are prescribed very rarely in this population and the variation is small; 53% of the drugs had an exposure of less than 10 PY and 88% of the drugs had an exposure of less than 500 PY. We emphasize that the results should be interpreted within the context of the data sources which gave rise to these results. Since the databases are primary care-based, specialist prescriptions (e.g., for antineoplastic drugs) are only captured in the system if continued by the general practitioner. Expansion of the database network to include other populations would be necessary to capture *all* drugs prescribed in the population, not only to increase the size of the studied population, but also to increase the variation in prescribing patterns.

Other possible sources of pediatric data include the Mini-Sentinel and OMOP.[22, 35] The pediatric population within Mini-Sentinel comprises ~27 million children and adolescents up to 19 years (21.6% of total)[40] OMOP comprises ~39.5 children and adolescents up to 18 years (*personal communication)*. If it would be possible to combine these data sources altogether, the current study population will be enlarged with a factor 14.7. 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 ~23.7 million PYs. Consequently, for anaphylactic shock the number of drugs with enough exposure to study a moderate association (RR≥4) will increase from 8 to 100 and for a more frequent event like upper gastrointestinal bleeding 242, instead of 39, drugs could be investigated to study a moderate association. Global collaboration will be necessary for further development of pediatric drug safety monitoring systems using EHRs, although such collaborations may still be incapable of studying the majority, if not all, drugs used in children and adolescents.

CONCLUSION

Drug use in children is rare and shows little variation; only 18 out of the total 2 170 prescribed drugs make up half of the total exposure to drugs in the pediatric population in EU-ADR. The number of drugs with enough exposure to detect safety signals using EHR databases for rare events in children and adolescents is limited. Mining within EHR databases seems especially promising for events with a high background incidence in the pediatric population and for drugs with a large amount of exposure. Inter-continental collaboration will be necessary to gain enough statistical power for pediatric safety signal detection.

APPENDIX

Incidence rates of important events in terms of drug safety in children and adolescents and number of drugs at 5th ATC (chemical substance level) that have enough PY of exposure to study a potential signal.

Event	Estimated incidence based on literature	RR≥2	RR≥4	RR≥6
Suicide				
6-≤11 years	2 / 100 000 PY	0	0	1
12-<18 years	10 / 100 000 PY	0	9	25
Suicide attempt				
6-≤11 years	150 / 100 000 PY	15	80	129
12-<18 years	750 / 100 000 PY	188	288	423
Febrile convulsions				
0-<2 years	1 400 / 100 000 PY	50	132	188
2-≤5 years	1 400 / 100 000 PY	70	206	294
Epilepsy				
0-<2 years	75 / 100 000 PY	3	24	44
2-≤5 years	75 / 100 000 PY	8	33	59
6-≤11 years	75 / 100 000 PY	10	50	83
12-<18 years	75 / 100 000 PY	10	73	121

REFERENCES

1. Schirm E, Tobi H, de Jong-van den Berg LT. Unlicensed and off label drug use by children in the community: cross sectional study. *BMJ (Clinical research ed.* Jun 1 2002;324(7349):1312-1313.
2. Dunne J. The European Regulation on medicines for paediatric use. *Paediatric respiratory reviews.* Jun 2007;8(2):177-183.
3. Mathis LL, Iyasu S. Safety monitoring of drugs granted exclusivity under the Best Pharmaceuticals for Children Act: what the FDA has learned. *Clin Pharmacol Ther.* Aug 2007;82(2):133-134.
4. Breslow LH. The Best Pharmaceuticals for Children Act of 2002: the rise of the voluntary incentive structure and congressional refusal to require pediatric testing. *Harvard J Legis.* Winter 2003;40(1):133-193.
5. Hartmann K, Doser AK, Kuhn M. Postmarketing safety information: how useful are spontaneous reports? *Pharmacoepidemiology and drug safety.* Apr 1999;8 Suppl 1:S65-71.
6. Harmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. *European journal of clinical pharmacology.* Aug 2008;64(8):743-752.
7. 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.* Sep 14 2001;19(32):4627-4634.
8. Aagaard L, Hansen EW, Hansen EH. Adverse events following immunization in children: retrospective analysis of spontaneous reports over a decade. *European journal of clinical pharmacology.* Nov 16 2010.
9. 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.* Feb 26 2008;26(9):1185-1194.
10. Zeinoun Z, Seifert H, Verstraeten T. Quantitative signal detection for vaccines: effects of stratification, background and masking on GlaxoSmithKline's spontaneous reports database. *Human vaccines.* Sep 2009;5(9):599-607.
11. 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(8):667-674.
12. Varricchio F, Iskander J, Destefano F, et al. Understanding vaccine safety information from the Vaccine Adverse Event Reporting System. *The Pediatric infectious disease journal.* Apr 2004;23(4):287-294.
13. Star K, Noren GN, Nordin K, Edwards IR. Suspected Adverse Drug Reactions Reported For Children Worldwide: An Exploratory Study Using VigiBase. *Drug Saf.* May 1 2011;34(5):415-428.
14. Aagaard L, Weber CB, Hansen EH. Adverse drug reactions in the paediatric population in Denmark: a retrospective analysis of reports made to the Danish Medicines Agency from 1998 to 2007. *Drug Saf.* Apr 1 2010;33(4):327-339.
15. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. *British journal of clinical pharmacology.* Jul 2001;52(1):77-83.
16. Star K. Detecting unexpected adverse drug reactions in children. *Paediatric drugs.* Apr 1 2011;13(2):71-73.
17. Goldman SA. Limitations and strengths of spontaneous reports data. *Clinical therapeutics.* 1998;20 Suppl C:C40-44.
18. *Practical Aspects of Signal Detection in Pharmacovigilance. Report of CIOMS Working Group VIII.* Geneva: CIOMS;2010.
19. 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. *Pharmacoepidemiology and drug safety.* Jun 2003;12(4):271-281.
20. Coloma PM, Schuemie MJ, Trifiro G, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiology and drug safety.* Jan 2011;20(1):1-11.
21. Platt R, Wilson M, Chan KA, Benner JS, Marchibroda J, McClellan M. The new Sentinel Network--improving the evidence of medical-product safety. *The New England journal of medicine.* Aug 13 2009;361(7):645-647.

22. Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Annals of internal medicine.* Nov 2 2010;153(9):600-606.

23. Coloma PM, Trifirò G, Schuemie MJ, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? *Pharmacoepidemiology and drug safety.* 2012 Feb 8. doi: 10.1002/pds.3197. [Epub ahead of print]

24. Stephenson T. How children's responses to drugs differ from adults. *British journal of clinical pharmacology.* Jun 2005;59(6):670-673.

25. Sturkenboom MC, Verhamme KM, Nicolosi A, et al. Drug use in children: cohort study in three European countries. *Bmj.* 2008;337:a2245.

26. Lasky T. Estimates of pediatric medication use in the United States: current abilities and limitations. *Clinical therapeutics.* Feb 2009;31(2):436-445.

27. 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;148:43-49.

28. WHO Collaborating Centre for Drug Statistics Methodology. Guidelines for ATC classification and DDD assignment. http://www.whocc.no/, 02-02-2012.

29. Unified Medical Language System® (UMLS®). http://www.nlm.nih.gov/research/umls/. Accessed 02-02-2012.

30. Trifiro G, Pariente A, Coloma PM, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiology and drug safety.* Dec 2009;18(12):1176-1184.

31. Avillach P, Mougin 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.

32. Martinez-Mir I, Garcia-Lopez M, Palop V, Ferrer JM, Rubio E, Morales-Olivas FJ. A prospective study of adverse drug reactions in hospitalized children. *British journal of clinical pharmacology.* Jun 1999;47(6):681-688.

33. Clavenna A, Bonati M. Adverse drug reactions in childhood: a review of prospective studies and safety alerts. *Arch Dis Child.* Sep 2009;94(9):724-728.

34. International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use. Clinical Investigation Of Medicinal Products In The Pediatric Population E11. *ICH Harmonised Tripartite Guideline* 2000; http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E11/Step4/E11_Guideline.pdf. Accessed 02-02-2012.

35. Curtis LH, Weiner MG, Boudreau DM, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiology and drug safety.* Jan 2012;21 Suppl 1:23-31.

36. 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.* Mar 1 2010;33(3):223-232.

37. Beaty JH. Fractures of the hip in children. *Orthop Clin North Am.* Apr 2006;37(2):223-232, vii.

38. Bimmel R, Bakker A, Bosma B, Michielsen J. Paediatric hip fractures: a systematic review of incidence, treatment options and complications. *Acta Orthop Belg.* Feb 2010;76(1):7-13.

39. Trifiro G, Patadia V, Schuemie MJ, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Stud Health Technol Inform.* 2011;166:25-30.

40. Mini-Sentinel. Mini-Sentinel Distributed Database "At A Glance". December 2011; http://www.mini-sentinel.org/about_us/MSDD_At-a-Glance.aspx.

4.3

A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection using Electronic Healthcare Records (EHR) Databases

Preciosa M. Coloma, MD, MSc[1], Paul Avillach, MD, PhD[2,3], Francesco Salvo, MD, PhD[4-6], Martijn J. Schuemie, PhD[1], Carmen Ferrajolo, PharmD, MSc[1], Antoine Pariente, MD, PhD[4-6], Annie Fourrier-Réglat, PharmD, PhD[4-6], Mariam Molokhia, MD, PhD[7], Vaishali Patadia, MSc[1,8], Johan van der Lei, MD, PhD [1], Miriam Sturkenboom, PharmD, PhD[1,9], Gianluca Trifirò, MD, PhD [1,10], on behalf of the EU-ADR consortium

[1]Department of Medical Informatics, Erasmus Medical Center, Rotterdam, the Netherlands; [2]LESIM, ISPED, University Bordeaux 2, Bordeaux, France; [3]LERTIM, EA 3283, Faculté de Médecine, Université Aix Marseille 2, Marseille, France; [4]Inserm U 657, Pharmacology Department, Bordeaux, France; [5]CHU Bordeaux, France; [6]Department of Pharmacology, University of Bordeaux, France; [7]Primary Care & Population Sciences, Kings College, London, UK; [8]Astellas Pharmaceuticals, Deerfield, Illinois, USA; [9]Department of Epidemiology, Erasmus University Medical Center, Rotterdam; and [10]Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Messina, Italy

Drug Safety 2012 [In press]

ABSTRACT

Background. The growing interest in using electronic healthcare record (EHR) databases for drug safety surveillance has spurred the development of new methodologies for signal detection. Although several drugs have been withdrawn post-marketing by regulatory authorities after scientific evaluation of harms and benefits, there is no definitive list of confirmed signals (i.e., list of all known adverse reactions and which drugs can cause them). As there is no true gold standard, prospective evaluation of signal detection methods remains a challenge.

Objective. Within the context of methods development and evaluation in the EU-ADR Project, we propose a surrogate reference standard of drug-adverse event associations based on existing scientific literature and expert opinion.

Methods. The reference standard was constructed for 10 top-ranked events judged as important in pharmacovigilance. A stepwise approach was employed to identify which among a list of drug-event associations are well-recognized (known positive associations) or highly unlikely ('negative controls') based on MEDLINE-indexed publications, drug product labels, spontaneous reports made to the World Health Organization's pharmacovigilance database, and expert opinion. Only drugs with adequate exposure in the EU-ADR database network (comprising ≈60 million person-years of healthcare data) to allow detection of an association were considered. Manual verification of positive associations and negative controls was independently performed by two experts proficient in clinical medicine, pharmacoepidemiology, and pharmacovigilance. A third expert adjudicated equivocal cases and arbitrated any disagreement between evaluators.

Results. 94 drug-event associations comprised the reference standard, which included 44 positive associations and 50 negative controls for the 10 events of interest: bullous eruptions; acute renal failure; anaphylactic shock; acute myocardial infarction; rhabdomyolysis; aplastic anemia/pancytopenia; neutropenia/agranulocytosis; cardiac valve fibrosis; acute liver injury; and upper gastrointestinal bleeding. For cardiac valve fibrosis, there was no drug with adequate exposure in the database network that satisfied the criteria for a positive association.

Conclusion. A strategy for the construction of a reference standard to evaluate signal detection methods that use EHR has been proposed. The resulting reference standard is by no means definitive, however, and should be seen as dynamic. As knowledge on drug safety evolves over time and new issues in drug safety arise, this reference standard can be re-evaluated.

INTRODUCTION

The growing interest in the utility of electronic healthcare records (EHR) for drug safety surveillance has spurred the development of new methodologies for quantitative and automated signal detection. Timely detection of safety signals remains a challenge because no single technique ensures identification of all drug-related adverse events, whether signal detection is done using spontaneous reports[1] or using healthcare records.[2] Generation of false alarms similarly constitutes a public health hazard, not only overwhelming regulatory agencies and diverting already scarce resources, but also triggering unwarranted warnings or even drug market withdrawals.[3] Thus, proper evaluation of signal detection methodologies calls for the creation of a reference standard, the purpose of which is to better define the predictive value of these new techniques, as well as their added value to the current pharmacovigilance armamentarium.

Signal detection in the context of pharmacovigilance

The World Health Organization (WHO) has defined 'signal' as 'reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented.'[4] An updated and more encompassing definition has been proposed recently based on a systematic review of how the term is being applied in current pharmacovigilance: a signal represents information that arises from one or multiple sources which suggests a new potentially causal association, or a new aspect of a known association, between an intervention and an event or set of related events, either adverse or beneficial, and is judged to be of sufficient likelihood to justify verificatory and remedial actions.[5] Although a 'gold standard' of confirmed signals, i.e., causal drug-adverse event associations, does not exist, a reference standard of recognized associations based on existing published scientific literature, regulatory actions (e.g., labeling changes or withdrawal of marketing authorization), as well as expert opinion may serve as suitable surrogate. In this paper we describe a reference standard that was put together in the context of methods development within the EU-ADR Project ('Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge,' http://www.euadr-project.org), which aims to exploit information from various EHR databases in Europe to produce a computerized integrated system for the early detection of drug safety signals.[6] This reference standard was developed for the primary purpose of evaluating performance of methods for signal detection using EHR.

METHODOLOGY

The EU-ADR network currently comprises anonymous healthcare data from eight established European databases located in four countries (Denmark, Italy, the Netherlands, and United Kingdom).[7] Clinical and drug dispensing/prescription data used for this paper represent data from 19 647 445 individuals with 59 929 690 person-years (PYs) of follow-up.

Adverse Events

In the EU-ADR Project we have chosen an event-based approach to active drug safety surveillance, focusing on events considered to be important from a pharmacovigilance and public health perspective. For the construction of this reference standard, we considered the following top 10 events which have been selected from a list of 23 events ranked on the basis of importance in pharmacovigilance using predefined criteria: (1) bullous eruptions; (2) acute renal failure; (3) anaphylactic shock; (4) acute myocardial infarction; (5) rhabdomyolysis; (6) aplastic anemia/pancytopenia; (7) neutropenia/agranulocytosis; (8) cardiac valve fibrosis; (9) acute liver injury; and (10) upper gastrointestinal bleeding.[8]

Drug selection

The procedure employed in the construction of the reference standard is outlined in Figure 1. It was first necessary to ensure that the drug-event associations to be included in the reference standard are identifiable in clinical practice and could be investigated in the EU-ADR network. That is, there should be adequate exposure to the drugs to permit detection of an association with the adverse event of interest, if present. In another publication we described the sample size calculations used to derive the total amount of PYs of drug exposure required to detect an association between a drug and a particular event over varying magnitudes of relative risk (RR), using one-sided significance level $\alpha = 0.05$ and power of 80%, given pooled population-based incidence rates (IR) estimated directly within the EU-ADR network.[2] For this reference standard we employed in the calculations RR of at least 2 for all events except for rhabdomyolysis, bullous eruptions, and anaphylactic shock, where we used RR of at least 4. The latter was done to account for the very low background incidence rates of these events in the population (2.5/100 000 PYs for rhabdomyolysis, 5.7/100 000 PYs for anaphylactic shock, and 5.9/100 000 PYs for bullous eruptions). A series of steps was subsequently employed to select the positive drug-event associations and 'negative controls' among those potentially eligible (i.e., drugs with adequate amount of exposure to detect the association of interest) (see Figure 1).

Information retrieval from published literature

To streamline the scientific literature search, we utilized a tool developed within the EU-ADR Project that automatically searches MEDLINE-indexed publications concerning adverse drug reaction (ADRs).[9] A subset of MEDLINE was downloaded (via PubMed) and imported in a database including all the citations from December 1952 to February 2010 with the 'adverse effects' MeSH subheading. For each citation the PubMed identification (PMID), MeSH descriptors, major/minor subheadings, substances, date of creation of the citation, as well as publication type, were obtained. Co-occurrence of the drug (from 'substances' OR 'MeSH heading' fields) and the event (under the subheading 'adverse effects') in a citation were noted. Drug codes in WHO Anatomical Therapeutic Chemical (ATC) classification were first mapped to MeSH headings or supplementary concept records using standardized concept unique identifiers from the Unified Medical Language System (UMLS).[10] Drugs from the 'substances' field were taken into account only if their pharmacological action was qualified by the subheading 'adverse effects.' Taking the

pharmacological action as an additional element for consideration was an attempt to establish a link between the adverse event of interest and the drug in the context of drug safety and not just a co-occurrence in a MEDLINE citation. This becomes particularly important when more than one drug is mentioned in the citation.[10]

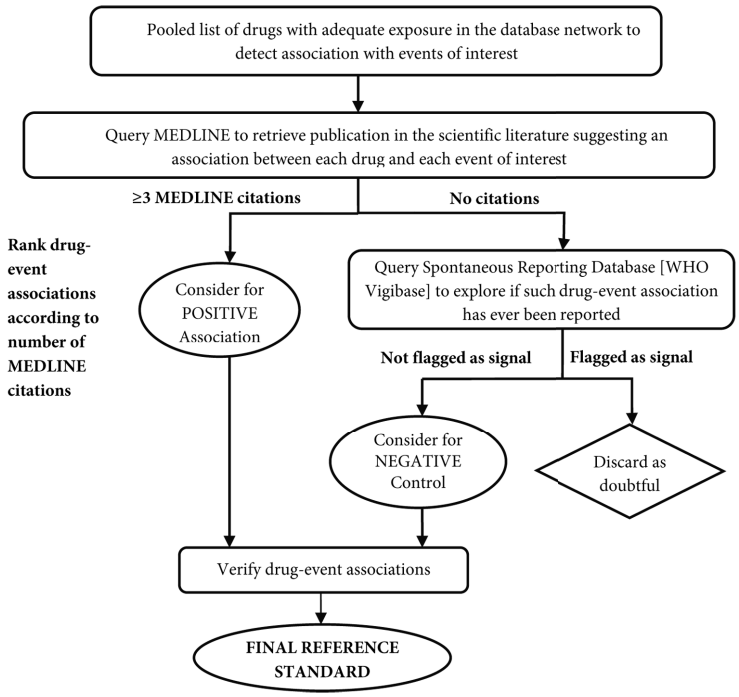


Figure 1 | Flowchart showing the process of the construction of the reference standard.

Selection of known positive drug-event associations

The drug-event associations were ranked according to the number of MEDLINE citations with co-occurrence of the drug and the adverse event of interest. For the pool of positive drug-event associations, we considered those with the highest number of citations. This meant that more published evidence was available on these associations. Citations may refer to case reports, observational studies, clinical trials, reviews, or meta-analyses. The type of publication was taken into account in the evaluation of the evidence regarding each drug-adverse event association, as subsequently described. Supplementary information was obtained from the Summary of Product Characteristics (SPCs) or product labels.[11-16] The aim was to select five drugs which are positively associated with each event of interest. Whenever possible, drugs belonging to different classes were included in the pool. However, the need for minimizing ambiguity (i.e., by selecting strong and well-substantiated drug-adverse event associations) took precedence over the need for diversity in terms of drug class. Except for fixed-dose combinations, drug preparations with more than one active substance were excluded from the pool.

Selection of 'negative controls'

A drug-event association was considered for the pool of 'negative controls' if there were no MEDLINE citations with co-occurrence of the drug and the event of interest and if there was no explicit mention of such adverse event in the drug product label. The pool of 'negative controls' was further evaluated using the WHO spontaneous reporting database (VigiBase) to exclude associations flagged as potential signal using standard data mining methodology. The list of potential signals from VigiBase (including data up to the fourth quarter of 2010) was generated using Oracle Health Sciences EmpiricaTM Signal tool (courtesy of Astellas Pharmaceuticals). Bayesian disproportionality analysis was performed using preferred terms (PTs) mapped to the events of interest.[17] A value greater than 2 for the lower bound of the 90% confidence interval of the Empirical Bayes Geometric Mean (EB05) and the presence of at least one report were used as the criteria for flagging a signal.[18] The aim was to likewise obtain five drug-event associations as 'negative controls' per event of interest.

Evaluation of the evidence from literature

Table 1 shows the scheme that was used as guide to evaluate evidence from the literature. Manual verification of the positive associations and 'negative controls' was conducted by two physicians with proficiency in clinical medicine, epidemiology, and pharmacovigilance. A third expert arbitrated any disagreement between evaluators. The following indices of agreement between evaluators were assessed: (1) proportion of overall agreement; (2) proportion of specific agreement; and (3) kappa statistic, κ for chance-corrected agreement. The earliest date of MEDLINE citation was also noted for each drug-event association.

Table 1 | Levels of evidence used in the evaluation of drug safety information from the literature.

Level of Evidence	Description
Level I	Evidence from at least one (properly designed) randomized controlled trial or meta-analysis.
Level II	Evidence from at least one observational study (e.g., cohort/case–control/case-cross over/self-controlled case series) OR from at least three published case reports from different sources and concerning different patients
Level III	Evidence from not more than two published case reports OR from unpublished reports in pharmacovigilance databases and no further substantiation in the literature
Level IV	Included in drug label (Summary of Product Characteristics, SPC) but no case reports or published studies
Level V	No evidence from published literature or from WHO spontaneous reporting database and not mentioned in SPC

Recommendations:

Levels I and II → positive association
Levels III and IV → cannot be determined → disregard as doubtful
Level V → 'negative control'

RESULTS

The amount of drug exposure required to detect a potential signal in the EU-ADR database network for each of the events of interest is shown in Table 2. Overall, there were 893 drugs (i.e., unique ATC codes, 5th level – chemical substance) with enough exposure to permit detection of an association with at least one of the 10 events of interest. Out of the 893 drugs, the following are the number (i.e., count) of drugs for which there were at least three MEDLINE citations with co-occurrence of the drug and the corresponding event: acute liver injury, 21; acute myocardial infarction, 52; acute renal failure, 51; anaphylactic shock, 26; bullous eruptions, 47; cardiac valve fibrosis, 2; neutropenia/agranulocytosis, 30; aplastic anemia/pancytopenia, 21; rhabdomyolysis, 8; upper gastrointestinal bleeding, 54. Close to 1 200 abstracts and, when necessary, also the full-text journal articles pertaining to all 10 events were reviewed to arrive at a shortlist of potential positive associations and 'negative controls.' Specific citations in drug product labels concerning 'undesirable effects,' 'warnings,' and 'adverse reactions' were used to further restrict the shortlist of associations. Table 3 shows how the manual evaluation of a positive association for acute liver injury with valproic acid and for upper gastrointestinal bleeding with indometacin were done. The complete evaluation for all the positive drug-adverse event associations of interest can be found in the Appendix.

Table 2 | Amount of drug exposure required to detect a potential signal in the EU-ADR database network for the events of interest.

Event	Required exposure (person-years)	No. of drugs with sufficient exposure to detect association and with ≥3 MEDLINE citations
Acute liver injury	32 769	21
Acute myocardial infarction	4 706	52
Acute renal failure	30 397	51
Anaphylactic shock	21 733	26
Bullous Eruptions	20 823	47
Cardiac valve fibrosis	13 604	2
Neutropenia/agranulocytosis	82 697	30
Aplastic anemia/pancytopenia	77 192	21
Rhabdomyolysis	49 593	8
Upper gastrointestinal bleeding	12 028	54

The final reference standard consisted of 94 drug-event associations, which included 44 positive associations and 50 'negative controls' related to the 10 events of interest. Table 4 lists the positive associations, including the corresponding level of evidence. Majority of the positive associations were based on Level II evidence. The associations for which there was Level I evidence include that of nonsteroidal anti-inflammatory drugs (NSAIDs) and of heparin with upper gastrointestinal

bleeding, the association of the statins with rhabdomyolysis, and the association of coxibs and of rosiglitazone with acute myocardial infarction. All 'negative controls,' by definition, have Level V evidence and are listed in Table 5. Both positive and 'negative control' associations comprised 68 unique drugs (i.e., ATC 5th level) belonging to 42 different pharmacological subgroups (i.e., ATC 3rd level). Only four drugs having sufficient exposure in the database network satisfied the criteria for a positive association with rhabdomyolysis, all of them being HMG-CoA reductase inhibitors (statins). Fibrates, as a class (ATC 4th level, chemical subgroup), comprised enough exposure to detect an association with rhabdomyolysis, but the individual drugs did not. For cardiac valve fibrosis, no drug with adequate exposure met the criteria for a positive association after review of the literature.

Table 3 | Example summary of manual evaluation of positive drug-event associations for Indometacin and Valproic acid.

ATC	Drug Name	Event type	No. of MEDLINE notices	Labeled as AE in SPC?YES/NO [Source and Label Section]
N03AG01	Valproic acid	Acute liver injury	Total no. of citations = 31	YES DailyMed *[Boxed warning, Adverse reactions]* eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse Reactions]*
			Review† = 1	
			Clinical trial = 1 [RCT]	
			Epidemiologic study = 1 [cohort study]	
			Case reports‡ = 28 [1 citation involving 3 cases, 1 citation involving 5 cases, 1 citation reviewing 31 cases, 2 other citations with literature review]	
M01AB01	Indometacin	Upper gastrointestinal bleeding	Total no. of citations = 45	YES eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
			Review = 13	
			Clinical trial = 16 [9 RCTs]	
			Epidemiologic study = 5 [1 case control and 4 cohort studies]	
			Case reports = 11	

NOTE: † Review refers to both systematic and narrative reviews

‡ Case reports involve only one case pertinent to the drug of interest, unless specified

Legend: SPC: Summary of Product Characteristics

Inter-evaluator agreement

The indices for agreement were computed across all drug-event pairs evaluated (179 drug-event pairs), including those which eventually did not get included in the final reference standard. The proportion of overall agreement (the proportion of cases for which both evaluators agreed across all evaluation categories) was 0.93 (95% confidence interval (CI) 0.89 to 0.97). The proportions of

Table 4 | Positive associations.

Event	ATC	Name	Level of Evidence
Acute Liver Injury	N03AF01	Carbamazepine	II
	N03AG01	Valproic acid	II
	M01AX17	Nimesulide	II
	J01CR02	Amoxicillin and clavulanic acid	II
	A07EC01	Sulfasalazine	II
Acute Myocardial Infarction	M01AH02	Rofecoxib	I
	A10BG02	Rosiglitazone	I
	G03AA07	Levonorgestrel and estrogen	II
	N02CC01	Sumatriptan	II
	M01AH03	Valdecoxib	I
Acute Renal Failure	C09AA01	Captopril	II
	M01AE01	Ibuprofen	II
	N02BE01	Paracetamol	II
	J01MA02	Ciprofloxacin	II
	N05AN01	Lithium	II
Anaphylactic Shock	B01AC06	Acetylsalicylic acid	II
	N02BE01	Paracetamol	II
	J01CA04	Amoxicillin	II
	J01MA02	Ciprofloxacin	II
	M01AB05	Diclofenac	II
Bullous Eruptions	N03AF01	Carbamazepine	II
	J01EE01	Sulfamethoxazole and trimethoprim	II
	N03AX09	Lamotrigine	II
	M04AA01	Allopurinol	II
	C03CA01	Furosemide	II
Cardiac Valve Fibrosis	No drug with sufficient exposure that satisfies criteria for True Positive		
Neutropenia/ agranulocytosis	H03BB02	Thiamazole	II
	B01AC05	Ticlopidine	II
	C09AA01	Captopril	II
	N03AF01	Carbamazepine	II
	N03AG01	Valproic acid	II
Aplastic anemia/ Pancytopenia	B01AC05	Ticlopidine	II
	N03AF01	Carbamazepine	II
	H03BB02	Thiamazole	II
	M04AA01	Allopurinol	II
	C09AA01	Captopril	II
Rhabdomyolysis	C10AA07	Rosuvastatin	I
	C10AA05	Atorvastatin	I
	C10AA03	Pravastatin	I
	C10AA01	Simvastatin	I
Upper Gastrointestinal Bleeding	N02BA01/ B01AC06	Acetylsalicylic acid	I
	M01AB01	Indometacin	I
	B01AB01	Heparin	I
	H02AB06	Prednisolone	II
	M01AE01	Ibuprofen	I

Table 5 | 'Negative control' associations*

Event	ATC	Name
Acute Liver Injury	R03AC13	Formoterol
	S01ED05	Carteolol
	G04CA03	Terazosin
	N04BA02	Levodopa and decarboxylase inhibitor
	C01DA02	Glyceryl trinitrate
Acute Myocardial Infarction	A10AD01	Insulin (human)
	B03AA07	Ferrous sulfate
	J01CR02	Amoxicillin and clavulanic acid
	J05AB11	Valaciclovir
	C10AB04	Gemfibrozil
Acute Renal Failure	R01AD09	Mometasone
	H03AA01	Levothyroxine sodium
	R06AX26	Fexofenadine
	N04BA02	Levodopa and decarboxylase inhibitor
	B03AA07	Ferrous sulfate
Anaphylactic Shock	N06AX11	Mirtazapine
	H03AA01	Levothyroxine sodium
	C02AC01	Clonidine
	C02CA04	Doxazosin
	N05BA04	Oxazepam
Bullous Eruptions	C01BC03	Propafenone
	C07AB03	Atenolol
	R03BB01	Ipratropium bromide
	R03BB04	Tiotropium bromide
	C08CA02	Felodipine
Cardiac Valve Fibrosis	N06AB08	Fluvoxamine
	L04AX03	Methotrexate
	C09CA04	Irbesartan
	C03CA01	Furosemide
	G03CA03	Estradiol
Neutropenia/Agranulocytosis	C07AA07	Sotalol
	H03AA01	Levothyroxine sodium
	C10AA05	Atorvastatin
	C01DA14	Isosorbide mononitrate
	G04CA02	Tamsulosin
Aplastic anemia/ Pancytopenia	C09CA04	Irbesartan
	C10AA04	Fluvastatin
	S01EE01	Latanoprost
	S01ED01	Timolol
	R06AX27	Desloratadine
Rhabdomyolysis	G03CA03	Estradiol
	C02CA04	Doxazosin
	A10BB12	Glimepiride
	S01ED01	Timolol
	C01DA02	Glyceryl trinitrate
Upper Gastrointestinal Bleeding	R06AX26	Fexofenadine
	C10AA01	Simvastatin
	S01EC03	Dorzolamide
	L02AE03	Goserelin
	N05CF01	Zopiclone

* All negative control associations have Level V evidence

specific agreement were as follows: (1) 'positive' agreement 0.96 (95% CI 0.93 to 0.98); (2) 'negative' agreement 0.90 (95% CI 0.89 to 0.90). There were three instances where one evaluator considered a drug-event association 'undetermined' while the other considered it a positive association (paracetamol-anaphylactic shock, bromocriptine-acute myocardial infarction, and acetylsalicylic acid-bullous eruptions). Of these three instances only one was eventually included in the reference standard after arbitration (paracetamol-anaphylactic shock). There was a single case where one evaluator marked the association 'undetermined' while the other marked it as 'negative control' (prednisone-neutropenia/agranulocytosis). Arbitration was done by a third expert. There was no disagreement between evaluators in the final list of 'negative control' associations. The chance-corrected agreement kappa coefficient, κ, was 0.83 (unweighted, 95% CI 0.74 to 0.92).

DISCUSSION

In this paper we present a novel approach to identify a surrogate 'gold standard' for drug safety signal detection using a systematic and rigorous methodology, applied across various data sources and which could be extended to examine other drug-event associations. We put together a list of drug-adverse event associations known to be true and drug-event associations considered to be unlikely based on current published scientific literature, drug product labels, spontaneous ADR reports, and expert opinion. Although the rationale for creating this reference standard is to have one single index against which signal detection methods (as applied to EHR data) can be tested, this reference standard can be re-evaluated and adapted to different settings as needed.

In evaluating the evidence from the literature we only considered associations that were reported with use of the drug in therapeutic doses, which is consistent with the definition of an adverse drug reaction.[19] For acetylsalicylic acid, citations referring to both cardiovascular prophylactic (low-dose) and analgesic doses were considered. We considered, aside from case reports that described the clinical characteristics leading to suspicion of an ADR, publications that proposed (or elucidated) biologic mechanisms for the associations. Such publications came in the form of both narrative reviews and systematic reviews. We likewise considered associations that were described in the context of drug-drug interactions (e.g., aplastic anemia resulting from the synergistic interaction between azathioprine and allopurinol).[20] For the event acute renal failure, we disregarded associations which arose from rhabdomyolysis leading to renal failure, but considered the reverse situation (i.e., associations for rhabdomyolysis which resulted in renal failure). While randomized controlled trials (RCTs) and meta-analyses are considered supreme with respect to level of evidence, this is more true for evidence regarding efficacy, not so much for safety, of interventions.[21-24] This is apparent in Table 5, where most of the evidence pertaining to the positive associations came from observational studies and case reports (or reviews). The associations with Level I evidence are those that are well-known (e.g., association of the NSAIDs and of heparin with upper gastrointestinal bleeding), or well-investigated – either because of controversy or public health impact (e.g., the association of the statins with rhabdomyolysis, and

the association of coxibs and of rosiglitazone with acute myocardial infarction). Interestingly, but perhaps not surprisingly, the most widely-investigated association was that between aspirin and upper gastrointestinal bleeding (259 MEDLINE citations overall, see Appendix). Most of the publications related to this association, including clinical trials, described the drug as a comparator to other drugs that are presumed (and proven) to confer a lower risk of the event.

There have been previous attempts to develop a reference standard with which data mining methods for safety signal detection can be evaluated, 'rules of evidence' being devised *ad hoc*.[25-27] In the creation of this reference standard we employed a systematic approach incorporating various sources of drug safety information, the process designed to be transparent and reproducible, thus also making it easier to update. Different sources have varying comprehensiveness and accuracy with regards to documenting drug-adverse event associations. Because RCTs may be restricted to specific populations and lack statistical power to detect rare events, they must be supplemented by non-experimental studies and other types of evidence, including case reports.[21-24] Rare or idiosyncratic events (e.g., bullous eruption-Stevens Johnson syndrome) and events occurring after chronic exposure (e.g., cardiac valvulopathy) are unlikely to be identified in clinical trials, but rather in case reports or observational studies.

There was only one disagreement between evaluators in the final list of positive associations ('undetermined' vs. 'positive' for the association paracetamol-anaphylactic shock; arbitration resulted in positive association). There was no disagreement between evaluators in the final list of 'negative control' associations. Although this high overall agreement between evaluators indicates that the resulting reference standard fulfills the pre-determined criteria, as the definitions of positive associations and 'negative controls' are based on existing knowledge at the time of this review, these associations (especially the 'negative controls') may be refuted as new data come along.[28] Hence, this reference standard should be considered dynamic and will need periodic re-evaluation. Adoption of this reference standard for use by other investigators can validate its applicability in other settings and will facilitate its further improvement.

While a reference standard, however rigorously constructed, may be able to permit evaluation and comparison of methods for signal detection, a method shown to successfully detect known drug-adverse events associations is not a guarantee that such method will also be able to detect signals, i.e., new, currently unknown, drug-event associations (problem of contemporary comparison).[29]

Limitations

Since the selection of drugs for the reference standard was dependent on the presence of adequate exposure to detect an association within the EU-ADR network (i.e., drugs that are more frequently used in the population were more likely to be chosen), this reference standard may not be as useful for evaluation in situations where the drug use patterns are expected to be different. In particular, the EU-ADR database network is unable to capture information on drugs which are primarily used in hospitals or specialist centers (e.g., anti-cancer drugs) and for this reason such drugs have not been included in the reference standard. This criterion also precluded the

inclusion of known associations with drugs which have been withdrawn from the market for a long time before the accrual of healthcare data in the databases. Because of this, there was no drug that could be used as a positive reference for the event cardiac valve fibrosis; the use of the appetite suppressants fenfluramine and phentermine, as well as the dopamine agonists pergolide and cabergoline, were inadequately documented or no longer captured in the databases because of the decline in use (or eradication in practice) of these drugs.[30] The choice as to which drug-event pairs can be considered for the positive associations was primarily established on the basis of the number of publications (i.e., number of MEDLINE citations with co-occurrence of the drug and the event of interest). This meant that drugs that have been on the market longer – or were involved in high-profile or controversial issues – had a higher chance of being included in the reference standard.

Finally, the availability of a surrogate 'gold standard' is only one component of the evaluation process for signal detection methodologies.[3,31] Other issues that need to be considered in performance evaluation of these methods include standardization of event definitions, establishment of reliable and consistent criteria for adjudicating causality and expectedness of adverse events, as well as understanding variations in database content and quality.

CONCLUSION

A unique strategy for the construction of a reference standard to evaluate drug safety signal detection methodologies using EHR has been proposed. This reference standard should be considered dynamic and as knowledge on drug safety evolves over time and new issues in drug safety arise, this reference standard can be periodically re-evaluated. Our proposed strategy represents a novel contribution to pharmacovigilance, with opportunities for adaptation to evaluate harms and benefits for other suspected ADRs.

ACKNOWLEDGMENT

The authors thank the reviewers and Anders Ottosson of Astra Zeneca for the valuable comments and insights.

REFERENCES

1. Amery WK. Signal generation from spontaneous adverse event reports. *Pharmacoepidemiol Drug Saf* 1999 Mar; 8(2):147-50.
2. Coloma PM, Trifirò G, Schuemie MJ et al., on behalf of the EU-ADR Consortium. Electronic healthcare databases for active drug safety surveillance: Is there enough leverage? *Pharmacoepidemiol Drug Saf* 2012 Feb 8. doi: 10.1002/pds.3197. [Epub ahead of print]
3. Hauben M, Reich L. Drug-induced pancreatitis: lessons in data mining. *Br J Clin Pharmacol* 2004 Nov; 58 (5): 560-2.
4. World Health Organization. Safety of medicines: a guide to detecting and reporting adverse drug reactions online]. Available from URL: http://whqlibdoc.who.int/hq/2002/WHO_EDM_QSM_2002.2.pdf [Last accessed 10 September 2011]
5. Hauben M, Aronson JK. Defining 'Signal' and its Subtypes in Pharmacovigilance Based on a Systematic Review of Previous Definitions. *Drug Saf* 2009; 32 (2): 99-110.
6. 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.
7. Coloma PM, Schuemie MJ, Trifiró G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom M, 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 S*af 2011; 20 (1):1-11. doi: 10.1002/pds.2053. Epub 2010 Nov 8.
8. 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.
9. Avillach P, Dufour JC, Diallo G, et al. Design and Validation of an Automated Method to Detect Known Adverse Drug Reactions in MEDLINE: a Contribution to the European EU-ADR Project. [Submitted, AMIA 2010 Annual Symposium].
10. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004 Jan 1;32(Database issue): D267-70.
11. European Medicines Agency (EMA). http://www.ema.europa.eu/ema/index.jsp?curl=/pages/medicines/landing/epar_search.jsp&murl=menus/medicines/medicines.jsp&mid=WC0b01ac058001d125 [Last accessed 20 September 2011]
12. DailyMed. http://dailymed.nlm.nih.gov/dailymed/about.cfm [Last accessed 20 September 2011]
13. electronic Medicines Compendium (for drugs licensed in the United Kingdom). http://www.medicines.org.uk/emc/ [Last accessed 20 September 2011]
14. Micromedex. https://www.thomsonhc.com/hcs/librarian/ [Last accessed 20 September 2011]
15. RxList. http://www.rxlist.com/ [Last accessed 20 September 2011]
16. Drugbank. URL: http://www.drugbank.ca/ [Last accessed 20 September 2011]
17. Trifirò G, Patadia V, Schuemie MJ, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Stud Health Technol Inform* 2011; 166:25-30.
18. DuMouchel W, Smith ET, Beasley R, et al. Association of asthma therapy and Churg-Strauss syndrome: an analysis of postmarketing surveillance data. *Clin Ther* 2004 Jul;26(7):1092-1104.
19. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet* 2000 Oct 7;356(9237): 1255-9.
20. Kennedy DT, Hayney MS, Lake KD. Azathioprine and allopurinol: the price of an avoidable drug interaction. *Ann Pharmacother* 1996 Sep;30(9):951-4.
21. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312 (7040):1215-1218.
22. Papanikolaou PN CG, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies *CMAJ* 2006;74:635–641.

23. Stricker BH PB. Detection, verification, and quantification of adverse drug reactions. *BMJ* 2004 Jul 3;329(7456):44-47.

24. Ray W. Population-based studies of adverse drug effects. *New Engl J Med* 2003;349(1592-1594).

25. Lindquist M, Ståhl M, Bate A, et al. 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.

26. Hauben M, Reich L. Safety related drug-labelling changes: findings from two data mining algorithms. *Drug Saf* 2004; 27(10): 735-44. Erratum in: Drug Saf. 2006;29(12):1192.

27. Hochberg AM, Hauben M, Pearson RK, et al. An evaluation of three signal-detection algorithms using a highly inclusive reference event database. *Drug Saf* 2009; 32(6): 509-25. doi: 10.2165/00002018-200932060-00007.

28. Waller P. Dealing with uncertainty in drug safety: lessons for the future from sertindole. *Pharmacoepidemiol Drug Saf* 2003 Jun; 12(4):283-7; discussion 289-90.

29. Bate A, Edwards IR. Data Mining Techniques in Pharmacovigilance. In: Hartzema AG, Tilson HH, Chan KA, eds. Pharmacoepidemiology and Therapeutic Risk Management. Cincinnati: Harvey Whitney, 2008: 239-272.

30. Bhattacharyya S, Schapira AH, Mikhailidis DP, et al. Drug-induced fibrotic valvular heart disease. *Lancet* 2009; 374 (9689): 577-85.

31. Levine JG, Tonning JM, Szarfman A. Reply: The evaluation of data mining methods for the simultaneous and systematic detection of safety signals in large databases: lessons to be learned. *Br J Clin Pharmacol* 2006 Jan;61(1): 105-13; author reply 115-7.

APPENDIX

Summary of manual evaluation of positive associations for 9 adverse events of interest*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Acute Liver Injury	N03AF01	Carbamazepine	Total no. of citations = 55	YES
			Review† = 4	DailyMed *[Adverse reactions]*
			Clinical trial = 1 [RCT]	eMC *[Undesirable effects]*
			Case reports‡ = 50 [5 citations involving 2 cases, 1 citation involving 3 cases, 1 other citation involving 2 cases with literature review, 1 other citation with literature review]	Micromedex *[Adverse reactions]*
Acute Liver Injury	N03AG01	Valproic acid	Total no. of citations = 31	YES
			Review = 1	DailyMed *[Boxed warning, Adverse reactions]*
			Clinical trial = 1 [RCT]	eMC [*Special warnings and precautions for use, Undesirable effects]*
			Epidemiologic study = 1 [cohort study]	Micromedex *[Adverse reactions]*
			Case reports = 28 [1 citation involving 3 cases, 1 citation involving 5 cases, 1 citation reviewing 31 cases, 2 other citations with literature review]	
Acute Liver Injury	M01AX17	Nimesulide	Total no. of citations = 38	YES
			Review = 2	Micromedex *[Adverse reactions]*
			Epidemiologic study = 5 [1 cohort, 1 case control, 2 case/non-case studies, 1 case series]	
			Case reports = 31 [3 citations involving 3 cases, 2 citations involving 6 cases, 2 other citations with literature review]	

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Acute Liver Injury	J01CR02	Amoxicillin and clavulanic acid	Total no. of citations = 37 Review = 2 Epidemiologic study = 4 [2 case series, 2 cohort studies] Case reports = 31 [2 citations involving 2 cases, 1 citation involving 15 cases, 1 citation involving 9 cases, 1 citation involving 8 cases, 1 citation involving 5 cases, 4 other citations with review of literature]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Acute Liver Injury	A07EC01	Sulfasalazine	Total no. of citations = 39 Review = 1 Epidemiologic study = 1 [case series] Case reports = 37 [4 citations involving 2 cases, 1 other citation involving 2 cases with review of the literature, 3 other citations with review of the literature]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Acute Myocardial Infarction	M01AH02	Rofecoxib	Total no. of citations = 46 Meta-analysis = 4 Review = 11 Clinical trial = 3 [all RCTs] Epidemiologic study = 27 [16 cohort and 11 case control studies] Case report = 1	Withdrawn from the market
Acute Myocardial Infarction	A10BG02	Rosiglitazone	Total no. of citations = 14 Meta-analysis = 3 Review = 5 Epidemiologic study = 6 [3 cohort and 3 case control studies]	YES Suspended by EMA because benefit does not outweigh risk *[increased risk of cardiovascular disorders including heart attacks, heart failure and stroke]* DailyMed *[FDA Boxed warning, Adverse Reactions]* eMC *[Contraindications, Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse Reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Acute Myocardial Infarction	G03AB03	Levonorgestrel and estrogen	Total no. of citations = 44 Review = 33 Epidemiologic study = 8 [3 cohort and 5 case control studies] Case report = 3	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Contraindications, Special warnings and precautions for use]* Micromedex *[Adverse reactions]*
Acute Myocardial Infarction	N02CC01	Sumatriptan	Total no. of citations = 17 Epidemiologic study = 1 [cohort study] Case report = 16 [1 citation involving 4 cases with literature review, 2 other citations with literature review]	YES DailyMed *[Adverse reactions]* eMC *[Contraindications, Undesirable Effects]* Micromedex *[Adverse reactions]*
Acute Myocardial Infarction	M01AH03	Valdecoxib	Total no. of citations = 11 Meta-analysis = 2 Review = 5 Epidemiologic study = 4 [2 cohort and 2 case control studies]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Undesirable Effects, under Parecoxib]* Micromedex *[Adverse reactions]*
Acute Renal Failure	C09AA01	Captopril	Total no. of citations = 61 Review = 5 Clinical trial = 2 Epidemiologic study = 2 [both cohort studies] Case reports = 52 [4 citations involving 2 cases, 2 citations involving 6 cases, 1 citation involving 7 cases, 1 citation involving 5 cases,1 citation with literature review]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Acute Renal Failure	M01AE01	Ibuprofen	Total no. of citations = 35 Review = 1 Epidemiologic study = 2 [1 cohort, 1 case control study] Case reports = 32 [4 citations involving 4 cases and with literature review, 1 citation involving 3 cases, 4 other citations with literature review]	YES eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Acute Renal Failure	N02BE01	Paracetamol	Total no. of citations = 16	YES
			Case reports = 16 [4 citations involving 2 cases, 1 citation involving 4 cases, 1 other citation with literature review]	eMC *[Special warnings and precautions for use, Overdose]* Micromedex *[Adverse reactions]*
Acute Renal Failure	J01MA02	Ciprofloxacin	Total no. of citations = 29	YES
			Case reports = 29 [2 citations involving 2 cases, 1 citation involving 3 cases, 1 citation involving 2 cases with literature review, 1 citation involving 3 cases with literature review, 1 citation involving 5 cases with literature review, 5 other citations with literature review]	DailyMed *[Warnings, Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
Acute Renal Failure	N05AN01	Lithium	Total no. of citations = 16	YES
			Review = 5	
			Epidemiologic study = 1 [cohort study]	DailyMed [Warnings] eMC *[Special warnings and precautions for use, Undesirable effects]*
			Case reports = 10 [2 citations involving 2 cases, 1 other citation with literature review]	Micromedex *[Adverse reactions]*
Anaphylactic shock	B01AC06/ N02BA01	Acetylsalicylic acid	Total no. of citations = 27	YES
			Review = 16	
			Epidemiologic study = 1 [cohort study]	DailyMed *[Warnings]* eMC *[Contraindications, Undesirable effects]*
			Case reports = 10 [1 citation involving 2 cases]	Micromedex *[Adverse reactions]*
Anaphylactic shock	N02BE01	Paracetamol	Total no. of citations = 16	YES
			Review = 2	
			Epidemiologic study = 1 [cohort study]	Micromedex *[Adverse reactions]*
			Case reports = 13 [1 citation involving 3 cases, 1 other citation involving 4 cases]	

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Anaphylactic shock	J01CA04	Amoxicillin	Total no. of citations = 12 Clinical trial = 1 Case reports = 11 [1 citation involving 3 cases, 1 citation involving 8 cases with literature review, 2 other citations with literature review]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special Warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Anaphylactic shock	J01MA02	Ciprofloxacin	Total no. of citations = 17 Case reports = 17 [1 citation involving 3 cases, 2 other citations with review of the literature]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special Warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Anaphylactic shock	M01AB05	Diclofenac	Total no. of citations = 16 Review = 2 Case reports = 14 [1 citation with literature review]	YES DailyMed *[Warnings, Contraindications, Adverse Reactions]* eMC *[Special Warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Bullous Eruptions	N03AF01	Carbamazepine	Total no. of citations = 34 Review = 3 Epidemiologic study = 10 [5 case-control, 4 case series, and 1 cohort study] Case reports = 21 [1 citation involving 4 cases, 1 other citation with literature review]	YES DailyMed *[Warnings, Adverse Reactions]* eMC *[Special Warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Bullous Eruptions	J01EE01	Sulfame-thoxazole and trimethoprim	Total no. of citations = 19 Review =2 Epidemiologic study = 9 [5 cohort, 2 case control, and 2 case series] Case reports = 8 [1 citation involving 8 cases]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special Warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Bullous Eruptions	N03AX09	Lamotrigine	Total no. of citations = 18 Review = 2 Epidemiologic study = 7 [4 case control, 3 cohort study] Case reports = 9	YES DailyMed *[FDA Boxed warning, Adverse reactions]* eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Bullous Eruptions	M04AA01	Allopurinol	Total no. of citations = 22 Review = 2 Epidemiologic study = 3 [all case control studies] Case reports = 17 [1 citation involving 3 cases]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
Bullous Eruptions	C03CA01	Furosemide	Total no. of citations = 7 Review = 3 Case reports = 4	YES DailyMed *[Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
Neutropenia/ Agranulocytosis	H03BB02	Thiamazole	Total no. of citations = 56 Review = 4 Epidemiologic study = 11 [5 cohort, 3 case series, and 3 case control studies] Case reports = 41 [1 citation involving 7 cases, 1 citation involving 4 cases, 1 citation involving 3 cases, 1 citation involving 2 cases, 1 citation involving 7 cases with literature review, 1 citation involving 2 cases with literature review, 2 other citations with literature review]	YES DailyMed *[Adverse reactions]* Micromedex *[Adverse reactions]*
Neutropenia/ Agranulocytosis	B01AC05	Ticlopidine	Total no. of citations = 34 Review = 5 Clinical trial = 3 [all RCTs] Epidemiologic study = 1 [cohort study] Case reports = 25 [1 citation involving 2 cases with literature review]	YES DailyMed *[FDA Boxed warning, Adverse reactions]* Micromedex *[Adverse reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Neutropenia/ Agranulocytosis	C09AA01	Captopril	Total no. of citations = 33 Review = 7 Case reports = 26 [2 citations involving 2 cases, 1 other citation involving 3 cases]	YES DailyMed *[Warnings, Adverse reactions]* eMC [*Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Neutropenia/ Agranulocytosis	N03AF01	Carbamazepine	Total no. of citations = 18 Review = 4 Clinical trial = 1 [RCT] Epidemiologic study = 4 [1 case series and 3 cohort studies] Case reports = 9 [2 citations involving 2 cases, 1 other citation with literature review]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Neutropenia/ Agranulocytosis	N03AG01	Valproic acid	Total no. of citations = 15 Review = 1 Epidemiologic study =3 [all cohort studies] Case reports = 11 [1 citation with literature review]	YES DailyMed *[Adverse Reactions]* eMC *[Precautions/Interaction with other medicinal products, Undesirable effects]* Micromedex *[Adverse reactions]*
Aplastic anemia/ Pancytopenia	B01AC05	Ticlopidine	Total no. of citations = 29 Review = 2 Clinical trial = 1 [RCT] Case reports = 26 [1 citation involving 2 cases, 1 citation involving 3 cases with literature review, 3 other citations with literature review]	YES DailyMed *[Warnings, Adverse reactions]* Micromedex *[Adverse reactions]*
Aplastic anemia/ Pancytopenia	N03AF01	Carbamazepine	Total no. of citations = 28 Review = 11 Epidemiologic study = 5 [3 cohort and 2 case control studies] Case reports = 12	YES DailyMed *[Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Aplastic anemia/ Pancytopenia	H03BB02	Thiamazole	Total no. of citations = 16 Review = 1 Epidemiologic study = 1 [case series] Case reports = 14 [1 citation with literature review]	YES DailyMed *[Adverse reactions]* Micromedex *[Adverse reactions]*
Aplastic anemia/ Pancytopenia	M04AA01	Allopurinol	Total no. of citations = 15 Review = 1 Epidemiologic study = 1 [cohort study] Case reports = 13	YES DailyMed *[Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
Aplastic anemia/ Pancytopenia	C09AA01	Captopril	Total no. of citations = 8 Review =1 Case reports = 7	YES DailyMed *[Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
Rhabdomyolysis	C10AA07	Rosuvastatin	Total no. of citations = 19 Meta-analysis = 1 Review =7 Epidemiologic study = 4 [all cohort studies] Case reports = 7 [1 citation involving 8 cases, 1 other citation with literature review]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Rhabdomyolysis	C10AA05	Atorvastatin	Total no. of citations = 24 Meta-analysis = 2 Review = 4 Epidemiologic study = 2 [1 case series and 1 cohort study] Case reports = 16	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Rhabdomyolysis	C10AA03	Pravastatin	Total no. of citations = 18 Meta-analysis = 1 Review = 7 Case reports = 10	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Rhabdomyolysis	C10AA01	Simvastatin	Total no. of citations = 86 Review = 11 Epidemiologic study = 1 [case series] Case reports = 75 [4 citations involving 4 cases, 3 other citations with literature review]	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Upper Gastrointestinal Bleeding	N02BA01/ B01AC06	Acetylsalicylic acid	Total no. of citations = 259 Meta-analysis = 3 Review = 96 Clinical trial = 53 [24 RCTs] Epidemiologic study = 56 [3 case series, 21 case control, and 32 cohort studies] Case reports = 51 [1 citation involving 4 cases]	YES eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Upper Gastrointestinal Bleeding	M01AB01	Indometacin	Total no. of citations = 45 Review = 13 Clinical trial = 16 [9 RCTs] Epidemiologic study = 5 [1 case control and 4 cohort studies] Case reports = 11	YES eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*
Upper Gastrointestinal Bleeding	B01AB01	Heparin	Total no. of citations = 25 Review = 11 Clinical trial = 2 [1 RCT] Epidemiologic study = 8 [all cohort studies] Case reports = 4	YES DailyMed *[Warnings, Adverse reactions]* eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*
Upper Gastrointestinal Bleeding	H02AB06	Prednisolone	Total no. of citations = 13 Review = 3 Epidemiologic study = 2 [1 case series, 1 cohort study] Case reports = 8	YES DailyMed *[Adverse reactions]* eMC *[Undesirable effects]* Micromedex *[Adverse reactions]*

Event type	ATC	Drug Name	No. of MEDLINE citations	Labeled as AE in SPC? YES/NO [Source and Label Section]
Upper Gastrointestinal Bleeding	M01AE01	Ibuprofen	Total no. of citations = 43 Meta-analysis = 2 Review = 7 Clinical trial = 16 [10 RCTs] Epidemiologic study = 10 [5 case control and 5 cohort studies] Case reports = 8 [1 citation involving 4 cases]	YES DailyMed *[Warnings]* eMC *[Special warnings and Precautions for use, Undesirable effects]* Micromedex *[Adverse reactions]*

NOTE: *There was no drug with adequate exposure in the database network that satisfied the criteria for a positive association with the event cardiac valve fibrosis.

† Review refers to both systematic and narrative reviews

‡ Case reports involve only one case pertinent to the drug of interest, unless specified

Legend:
SPC: Summary of Product Characteristics
RCT: Randomized Controlled Trial
DailyMed: website for drugs currently marketed and approved by the United States Food and Drug Administration, URL: http://dailymed.nlm.nih.gov/
eMC: electronic Medicines Compendium, for drugs licensed in the United Kingdom, URL: http://www.medicines.org.uk
Micromedex: family of international databases providing full-text drug and substance information, URL: http://www.thomsonhc.com/micromedex2/

4.4

Using electronic healthcare records for drug safety signal detection: A comparative evaluation of statistical methods

Martijn J. Schuemie, PhD[1], Preciosa M Coloma, MD, MSc[1], Huub Straatman, MSc[2], Ron MC Herings, PhD[2], Gianluca Trifirò, MD, PhD[1,3], Justin N. Matthews, MSc[4], David Prieto-Merino, PhD[4], Mariam Molokhia, PhD[5], Lars Pedersen, PhD[6], Rosa Gini, PhD[7], Francesco Innocenti, MSc[8], Giampiero Mazzaglia, MD, PhD[8], Gino Picelli, PhD[9], Lorenza Scotti, MSc[10], Johan van der Lei, MD, PhD[1], Miriam C.J.M. Sturkenboom, PharmD, PhD[1], on behalf of the EU-ADR consortium

[1]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands; [2]PHARMO Institute, Utrecht, the Netherlands; [3]Department of Clinical and Experimental Medicine and Pharmacology, University of Messina, Messina, Italy; [4]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, UK; [5]Department of Primary Care and Public Health Sciences, Kings College London, UK; [6]Department of Clinical Epidemiology, Aarhus University Hospital, Århus Sygehus, Denmark; [7]Agenzi Regionali di Sanità della Toscana, Italy; [8]Health Search, Italian College of General Practitioners, Italy; [9]Pedianet, Societa' Servizi Telematici SRL, Italy; [10]Department of Statistics, Università di Milano-Bicocca, Italy

Submitted

ABSTRACT

Background and Objective. Safety monitoring of drugs post-marketing relies primarily on spontaneous reporting, but mining of electronic healthcare records (EHR) offers a possible complement in the detection of safety signals. In this paper we evaluate the relative performance of different statistical methods in the detection of potentially drug-induced adverse events from EHR data.

Research Design. Data from seven EHR databases across three countries in Europe comprising over 20 million subjects were used to compute relative risk estimates for drug-event pairs using 10 different methods, including those developed for spontaneous reporting systems, cohort methods such as the Longitudinal Gamma Poisson Shrinker (LGPS), and case-based methods like case control. The newly developed method 'Longitudinal Evaluation of Observational Profiles of Adverse events Related to Drugs' (LEOPARD) was used to remove associations likely to be due to protopathic bias. Data from the different databases were combined by simple data pooling and by meta-analysis for random effects. Method performance was evaluated using a surrogate reference standard of known positive drug-event associations and 'negative controls.'

Measures. Area under the receiver operator characteristic curve (AUC) was calculated for each method, both with and without LEOPARD filtering.

Results. The highest AUC (0.83) was achieved by the combination of either LGPS or case-control with LEOPARD filtering, although the difference in performance between methods was marginal. LEOPARD increased overall performance for all methods, but flagged several known drug-event associations as due to protopathic bias.

Conclusions. Combinations of methods demonstrate good performance in distinguishing known associations from 'negative controls.' These methods can be further explored to detect new, previously unknown, drug safety signals.

INTRODUCTION

Modern drug legislation on post-marketing drug safety monitoring was prompted over 40 years ago as a result of the devastating teratogenic effects of thalidomide.[1] Since then, the mainstay of drug safety surveillance has been the collection of spontaneous reports of suspected adverse drug reactions (ADRs).[2,3] However, a number of recent high-impact issues (e.g., cardiovascular risk with rofecoxib and rosiglitazone) require re-thinking of the way safety monitoring is conducted.[4] Spontaneous Reporting Systems (SRS) have inherent limitations that hamper safety signal detection[5] such as under-reporting and biases due to selective reporting.[6] The percentage of ADRs being reported by health professionals varies between 1 and 10% of those actually occurring in clinical practice[7-13], and this problem occurs both in the primary care and hospital settings.[14,15] A recent study in Spain showed that less than two-thirds of ADRs recorded in electronic medical records were actually reported to the Spanish Pharmacovigilance System.[14] As a consequence, SRS may not always guarantee timely signal detection[16] and additional data sources have to be explored.

One possible complementary approach is mining of electronic healthcare record (EHR) databases for drug safety signal detection. These databases, which include both electronic medical record and administrative claims databases, have been most commonly used to confirm or refute potential signals flagged by spontaneous reporting or other surveillance systems. However, appropriate use of these databases has enormous potential for earlier detection of drug safety signals because of the availability of large amounts of routinely collected, time-stamped data from actual clinical practice.[17,18] Several international initiatives have recently embarked on developing such post-marketing surveillance systems: in the US the Sentinel Initiative was established with the Mini-Sentinel (http://mini-sentinel.org) and Observational Medical Outcomes Partnership (OMOP, http://omop.fnih.org) as pilot initiatives, and in Europe the PROTECT (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium, http://imi-protect.eu) and EU-ADR (Exploring and Understanding Adverse Drug Reactions. http://euadr-project.org) projects were launched.

The EU-ADR project aims to develop a safety surveillance system using EHR databases within a distributed database architecture.[19] In this paper we apply a wide range of statistical methods to the EHR data in EU-ADR and compare their relative performances on the task of distinguishing drug-adverse event associations that are well-recognized from those which are highly unlikely ('negative controls').

METHODS

Data sources

Data used in this analysis come from seven EHR databases in three European countries: Health Search/CSD LPD (HSD, Italy), Integrated Primary Care Information (IPCI, 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 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 pharmacy services, are provided for, or subsidized by, the state in Italy, Denmark, and the UK and covered by obligatory health insurance in the Netherlands. The databases in EU-ADR and their characteristics are listed in Table 1. The database network has been described in more detail in an earlier publication.[19] All databases used the World Health Organization's (WHO) Anatomical Therapeutic Chemical (ATC) classification system for coding the drugs. Events of interest were extracted using a variety of structured disease coding algorithms and free text queries.[19] Although all statistical analyses were performed at a central site, data aggregation was performed locally in the databases using a Java-based tool developed in the EU-ADR project, called Jerboa®, before combining data across databases. This ensured that patient-level data remain protected within the confines of the individual databases.

Table 1 | Participating databases and their characteristics.

	Pedianet	**Health Search Lombardy**		**ARS**	**IPCI**	**PHARMO**	**Aarhus**
	(Italy)	(Italy)	(Italy)	(Italy)	(Netherlands)	(Netherlands)	(Denmark)
Subjects	140 000 children	1 000 000	10 000 000	4 000 000	750 000	1 280 000	2 000 000
Type of database	General Practice pediatric database	General Practice database	Record linkage system with:	Record linkage system with:	General Practice database	Record linkage system with:	National health registry with:
			1) Registry inhabitants	1) Registry inhabitants		1) Registry inhabitants	1) Registry inhabitants
			2) Regional Drug dispensation records	2) Regional Drug dispensation records		2) Regional Drug dispensation records	2) Regional Drug dispensation records
			3) Hospital claims database	3) Hospital claims database		3) Hospital claims database	3) Hospital claims database
				4) Death registry		4) Lab values	4) Lab values
							5) Death registry

Statistical methods

Table 2 | Overview of the methods evaluated here.

SRS methods	Bayesian	Ranking criteria for AUC
Proportional Reporting Ratio (**PPR**)		PRR
Reporting Odds Ratio (**ROR**)		ROR
Gamma Poisson Shrinker (**GPS**)	Yes	Point estimate of the RR
Bayesian Confidence Propagation Neural Network (**BCPNN**)	Yes	Information Component (IC)
Cohort methods		
Incidence Rate Ratio (**IRR**)		IRR
Longitudinal GPS (**LGPS**)	Yes	Point estimate of the IRR
Bayesian Hierarchical Model (**BHM**)	Yes	Point estimate of the IRR
Case-based methods		
Matched case-control (**CC**)		Beta (odds ratio estimate)
Self-Controlled Case Series (**SCCS**)		Beta (relative risk estimate)
Elimination of protopathic bias		
LEOPARD		Eliminate if $p< 0.5$

Spontaneous Reporting System methods

Signal detection methods originally developed for SRS can also be used on EHR data by transforming the data to a format suitable for these methods. For this transformation we assume that whenever one of the events of interest occurs during a period associated with the exposure to a drug, that this will lead to a 'report' describing a potential ADR involving the drug and event. The number of reports for a particular drug-event pair can then be used as if it is a report count from an SRS database, as shown in Table 3. In this table, w_{00} is the number of events A that occurred during exposure to drug X, w_{01} is the number of events of a different type than A that occurred during exposure to X, w_{10} is the number of events A that occurred during exposure to drugs other than X, and w_{11} is the number of events of a different type than A, that occurred during exposure to drugs other than X.

Table 3 | Overview of report counts generated for event A and drug X.

	Event A	Not Event A
Drug X	w_{00}	w_{01}
Not drug X	w_{10}	w_{11}

Based on this table, the different metrics described below can be calculated. The major disadvantage of these disproportionality methods is that they do not use all the information

available in the longitudinal healthcare records but focus only on the cases. In addition, it is difficult to adjust for many confounding factors with these methods. They can be regarded as easy screening methods and can be scaled easily to large EHR databases as they are not computationally intensive. The following disproportionality methods were included in the evaluation:

- **Proportional Reporting Ratio (PRR)** is the ratio of the proportion of all reported cases of the event of interest among people exposed to a particular drug compared with the corresponding proportion among people exposed to all drugs.[23]
- **Reporting Odds Ratio (ROR)** is the reformulation of the PRR as an odds ratio.[24]
- **Gamma Poisson Shrinker (GPS)** also determines the disproportionality of reports for a particular drug compared to all exposure, but uses a Bayesian model to shrink relative risk estimates when less data is available.[25] The prior distribution is established empirically using data of all drug-event pairs.
- **Bayesian Confidence Propagation Neural Network (BCPNN)** works similarly to GPS, in that it also uses a Bayesian model to shrink estimates of risk. Typically, the output of a BCPNN is expressed as the Information Component (IC); the logarithm of the ratio between observed and expected number of reports for a particular drug-event pair.[26]

Cohort methods

One of the limitations of data from SRS (and consequently, methods developed for SRS) is that only numerator data is available, i.e., the number of people who are exposed to drugs and have the event of interest. What is missing is the denominator data: the number of people that are exposed to the drugs. In EHR databases this information is readily available, as well as the duration of drug exposure. It is this information which is used in the cohort methods.

- **Incidence Rate Ratio (IRR)** is the ratio between the incidence rate during exposure to the drug and a background incidence rate. A Mantel-Haenszel test is used to test the differences between the incidence rates, correcting for age and sex.
- **Longitudinal GPS (LGPS)** is an adaptation of the GPS to longitudinal data, and applies Bayesian shrinkage to the IRR. It was developed in the EU-ADR project.[27]
- **Bayesian Hierarchical Model (BHM)** uses a full Bayesian approach to perform shrinkage of risk estimates, but instead of using a single prior distribution for all drugs, priors are also created for classes and super classes of drugs. The BHM combines statistical models for the observations given the parameters (likelihood) and the parameters themselves (priors). In the application reported here, the incidence rate is modeled using a Poisson process, and the priors as a hierarchy (using guidance from Gelman[28]). The groupings forming the hierarchy are decided a priori based on criteria of similarity between drugs; in this case we have used ATC coding levels based on organ/systems and therapeutic or chemical characteristics. Berry and Berry[29] used a similar hierarchical approach, with a hierarchy based on related outcomes rather than drugs. The BHM shrinks the original estimates to give an updated posterior distribution of each individual drug to the group mean and reduces its variance. This is because the posterior considers both the data provided by the drug and by the other

drugs in the same group. Shrinkage is stronger for drugs with an initial large variance (less information) and larger effects. These novel methods can offer key advantages by reducing the likelihood of false positive or false negative results obtained from the data. Although the BHM is grouped here with the cohort methods, it can also be applied to other types of relative risk estimates.

Case based methods

Several analytical epidemiological methods start with persons with the disease or event of interest (cases) and compare these with a sample of the population that gives rise to the cases (i.e., the controls) to evaluate differences in exposure status. Since the case-based methods are more efficient in terms of data needs (exposure assessed only at one point in time), they allow for easier adjustments for confounding factors.

- **Matched Case Control (CC)** starts with all cases and finds for every case a predefined number of controls (in our experiments two controls per case), where controls have the same age and sex as the case. For both cases and controls, the exposure to drugs is determined at the calendar time of the event (also known as the index date). A logistic regression, conditioned on the case sets, is performed to determine the effect size (odds ratio) of exposure to a drug. To adjust for comorbidity and overall patient health status we included the variable **drug count**[30] in the regression, which we defined as the number of different drugs (distinct ATC codes) the subject was exposed to in one year prior to the event date until one month prior to the event date.
- **Self-Controlled Case Series (SCCS)** investigates the association between acute outcomes and transient exposures, whereby cases are used as their own controls. In essence, the SCCS is a Poisson regression conditioned on the patient.[31] Only information of cases is used in this analysis, all other persons are ignored.

Other types of methods

One other method not categorized elsewhere remains:

- **Longitudinal Evaluation of Observational Profiles of Adverse events Related to Drugs (LEOPARD)**[27] attempts to detect protopathic bias. Protopathic bias occurs when a drug is prescribed for an early manifestation of a disease that has not yet been diagnosed. For every drug-event combination, the number of prescriptions initiated in the 25 days prior to the event is compared to the number or prescriptions starting in the 25 days after the event. An increase in the number of prescriptions after the event date, relative to the number of prescriptions before the event date, is taken to be an indication that the drug is used to treat the event or a precursor of the event, rather than cause it. This is assessed using a binomial test. Appendix I shows as an example the association between the proton pump inhibitor Omeprazole and upper gastrointestinal bleeding. A signal is considered to be caused by protopathic bias if the p-value is below 0.5.

Settings common to all methods

For all methods, these specifications were used to define exposures and outcomes:

- *Incident events*. Only the first occurrence of an event was considered; patient time after an event was censored.
- *Run-in period of 365 days*. In order to determine that an event is incident, some patient time has to be available before the event occurred. Hence, during the first year of observation subjects were not considered for events or exposure counts, but events during this so-called run-in period were used to determine whether later events were truly incident events. This run-in period was omitted for children younger than one year at the start of observation.
- *Exposure window definition*. Exposure to a drug was defined as the estimated length of the drug prescription in days, excluding the first day of the prescription. The duration was calculated based on the prescribed daily dosage in some databases (IPCI, PHARMO, Pedianet), and on the defined daily dose and the quantity prescribed in other databases (Lombardy, Aarhus, Health Search, ARS). If two prescriptions of the same drug overlapped in time, the exposure was assumed to start the day after the first day of the first prescription, and end on the last day of the last prescription. Exposure carry-over and prescription gaps were not considered in these analyses.
- *Age stratification*. Whenever appropriate, age was stratified in 5-year age intervals.
- *Independence of drug risks*. Currently, every drug-event pair is evaluated separately; co-medication is not taken into account.

LEOPARD was considered to be potentially complementary to all methods, and was therefore applied as a filter to the results derived from each method. LEOPARD can be applied at the level of the individual drug, but it can also be applied at the drug class level. LEOPARD appears to be better at detecting protopathic bias when drugs are grouped within the same pharmacological subgroup (i.e., ATC 3rd level, which usually means the same indication) (unpublished results). Signals that are flagged by LEOPARD, either at individual or at group level, were ranked lower in the list of signals than signals that were not flagged when calculating the AUC.

Combination of databases

The information from the different databases was combined to generate a single score per drug-event pair per method. In principle there are two approaches: pooling of the data as if the databases together form one large database, or computing the score per database and using meta-analytic techniques to combine the scores. We have tested both data pooling and meta-analysis for risk estimates assuming random effects. Meta-analysis used weighting by inverse variance (both within and between database variance).

Performance metrics

Typically, the output of a signal detection method is turned into a binary decision (positive or negative) by employing a threshold, for instance that the relative risk be greater than 2. By comparing the binary outcomes of the method to the reference standard, sensitivity and specificity

can be computed. However, sensitivity and specificity can be traded off by varying the threshold, and comparing individual values of sensitivity and specificity is therefore not informative. Typically for method comparison, the Receiver Operator Characteristics (ROC) curve is plotted, showing all values of sensitivity and specificity. (Note that this implies varying the threshold from the smallest relative risk to the largest relative risk found for the associations in the reference standard.) Such a curve can subsequently be summarized into one statistic: the area under the ROC Curve (AUC). The AUC indicates the overall performance of a method, independent of any threshold. An AUC of 0.5 indicates random performance; an AUC of 1.0 indicates a perfect performance. The measure used to calculate the AUC for each method is specified in Table 2.

RESULTS

A total of 146 830 906 patient-years of follow-up data concerning 20 042 652 subjects from three European countries from 1997 until 2010 were included in the study.

Overall performance of methods

Figure 1 shows the performance of the different methods on the reference standard, using meta-analysis for random effects and data pooling.

All methods performed better than random baseline and LEOPARD filtering for protopathic bias always improved overall performance, but less so for methods that are already performing well. The differences in the performance of the methods were marginal. In general, LGPS and the case-control method (adjusting for drug count) slightly outperformed the other methods and the SRS-based methods have lower performance, although this was not statistically significant.

Signal detection using LGPS and LEOPARD

Figure 2 shows that most of the known positive drug-event associations have an estimated IRR greater than 1. Unfortunately, this is also the case for a large number of the negative controls, but many of these are flagged by LEOPARD as protopathic bias. This reduction in false positives comes at a price: several of the known positive associations are also flagged as protopathic bias. Notably, Ciprofloxacin and ARF, which is a known association, has an estimated IRR of 13.98, but because there are more prescriptions starting in the 25 days following ARF than in the 25 days before (631 and 574 prescriptions, respectively) this signal is rejected by LEOPARD. The strongest false positive not discarded by LEOPARD is Fexofenadine and ARF. **Figure 3** shows that most other methods indicate a low relative risk instead.

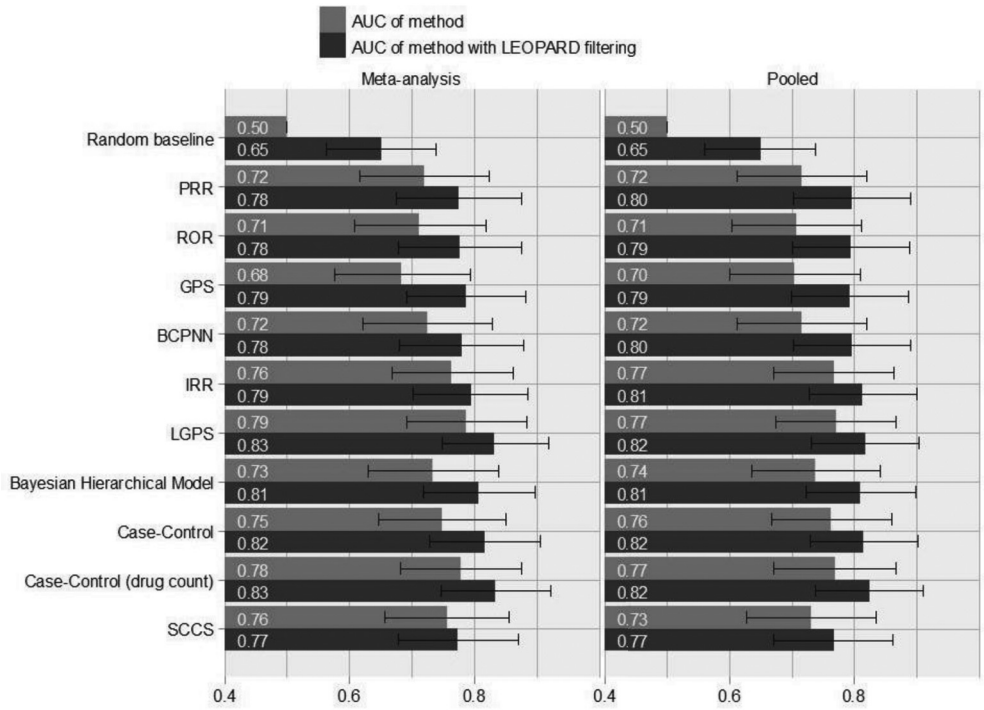


Figure 1 | Area Under the ROC Curve (AUC) for all methods, with and without LEOPARD filtering. Combination across databases was performed using meta-analysis for random effects (left panel) and pooling (right panel). Error bars indicate the 95% confidence interval. See Appendix for area under ROC curves ***(see page 310 for colour figure)*****.**

Legend: **PRR** = Proportional Reporting Ratio **ROR** = Reporting Odds Ratio
GPS = Gamma Poisson Shrinker **BCPNN** = Bayesian Confidence Propagation Neural Network
IRR = Incidence Rate Ratio **LGPS** = Longitudinal GPS
SCCS = Self-Controlled Case Series

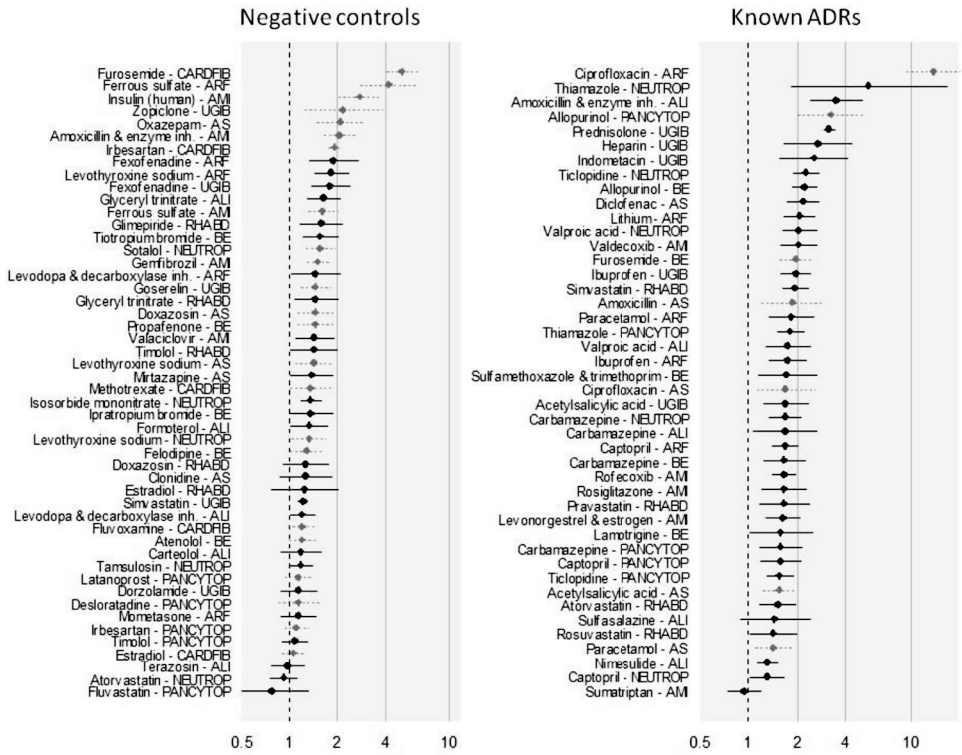


Figure 2 | Estimates of the risk estimates for all drug-outcome combinations in the reference set using LGPS (Longitudinal Gamma Poisson Shrinker) and using meta-analysis for random effects to combine estimates across databases. Error bars indicate the 95% confidence interval. Gray markers and dashed lines indicate a drug-event pair has been flagged as protopathic bias by LEOPARD.

Legend: **AMI** = acute myocardial infarction; **ALI** = acute liver injury; **ARF** = acute renal failure; **AS** = anaphylactic shock; **BE** = bullous eruption; **CARDFIB** = cardiac valve fibrosis; **NEUTROP** = neutropenia; **PANCYTOP** = pancytopenia/aplastic anemia; **RHABD** = rhabdomyolysis; **UGIB** = upper gastrointestinal bleeding

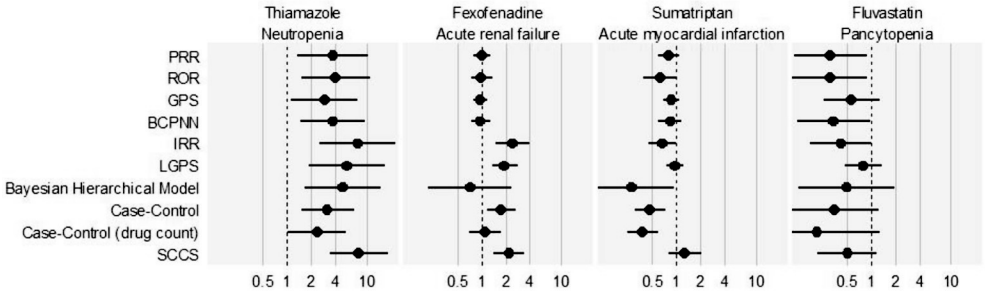


Figure 3 | Estimates and 95% confidence intervals of the risk estimates as calculated using the different methods for four drug-event combinations, using meta-analysis for random effects. Thiamazole and neutropenia is the known positive association with the highest overall LGPS estimate (true positive). Fexofenadine and ARF is the negative control with the highest overall LGPS estimate (false positive). Sumatriptan and AMI is the known ADR with the lowest LGPS estimate (false negative), and Fluvastatin and pancytopenia is the negative control with the lowest LGPS estimate (true negative).

If we were to use a threshold value of IRR > 1.5, and were to remove signals flagged by LEOPARD, LGPS using meta-analysis for random effects would achieve a sensitivity of 0.73, and a specificity of 0.88.

DISCUSSION

In general the methods showed good performance, with the best performing method achieving an area under the ROC curve of 0.83. When using a default threshold of IRR >1.5, a sensitivity and specificity of 0.73 and 0.88 respectively, are achieved. On the one hand this is not surprising, since the reference standard was limited to drugs with a large amount of exposure. This probably explains why the Bayesian methods did not perform much better than the non-Bayesian ('frequentist') methods, as these methods are designed to deal with sparse data. On the other hand, the overall good performance is surprising given the fact that most of these methods are very simple: the best performing methods do not take any potential confounding factors into account other than age and sex. It is expected that by including potential confounders in the analysis, method performance could be improved further, but the data needed for this is not uniformly available across databases. For example: diabetes is a risk factor for myocardial infarction and diabetics will tend to be exposed to anti-diabetic drugs; diabetes is therefore a confounder between myocardial infarction and anti-diabetics, and should ideally be included in the analysis. However, diabetes will be coded differently in the different databases in EU-ADR and harmonization of database queries to identify diabetes (as well as other diseases that can be potential confounders) would be needed to extract this data in a uniform way. Since for every drug-event combination

there can be different potential confounders, many such variables would need to be extracted, which is currently not feasible. One possible solution could be the development of techniques for adjusting for confounding without the use of harmonized covariates, for instance by using summary statistics such as propensity scores instead.[32] Another simplification is that the methods currently consider each drug independently, which could lead to other drugs being implicated because of frequent co-prescribing with drugs that do cause an ADR (so called 'innocent bystander'). Since data on most drug prescriptions is available, this problem could potentially be solved by adapting the methods to include all drugs in one analysis.

Filtering signals for protopathic bias using LEOPARD has a positive effect on overall performance, but some of the known positive associations are incorrectly flagged as protopathic bias. For example, Ciprofloxacin is known to be associated with ARF,[33-35] and we indeed find that Ciprofloxacin users have an increased risk of ARF. However, the data also shows that subjects are more likely to receive a prescription after ARF when compared to the period preceding the event, which indicates potential protopathic bias. One possible explanation for this is that patients with ARF are at risk of developing various infections and may therefore be treated with Ciprofloxacin, which is a broad-spectrum antibiotic. In fact, for patients with ARF that require dialysis, Ciprofloxacin is used in the treatment of dialysis-associated peritonitis.[36] It seems that the results derived using LEOPARD cannot be used to rule out drug safety signals, but only as indication that protopathic bias *might* be present. The strongest false positive for LGPS not eliminated by LEOPARD is Fexofenadine and ARF, but interestingly this signal is not detected by the majority of other methods. It appears that methods need to be tailored to certain types of drugs or events to achieve better results; this is something that needs to be explored in future investigations.

We also identified some positive drug-event associations that are not picked up by the methods, even before LEOPARD filtering. For example Sumatriptan is known to be associated with AMI[37] but none of the methods find an increased risk for this drug. Sumatriptan is used for the treatment of migraine headaches and may be taken by patients only intermittently and as needed, which can be a long time after the drug is prescribed. Since our current definition of the exposure window assumes exposure starts on the day of prescribing, this could explain why this association was not detected. In general, we expect that signal detection using EHRs will not capture well the associations involving drugs that are taken as needed because of uncertainty about the drug exposure window.

These results are based on data from seven different databases, but we did not investigate the method performance per database because none of the databases by itself has enough data to detect all drug-event pairs in the reference standard. The need for sufficient data on drug exposure is one important consideration when using EHR for drug safety surveillance,[38] and combining databases can overcome this problem to some extent. We investigated two methods of combining data: pooling and meta-analysis, and the differences in performance were negligible.

CONCLUSION

The results from this study indicate that there are several combinations of statistical methods that show good performance in distinguishing known drug-adverse event associations from 'negative controls.' These methods can be further explored to detect new, previously unknown, drug safety signals or to re-appraise signals previously identified in SRS. The methods and the approach described here may contribute to, and further improve, current methodologies in post-marketing drug safety surveillance.

ACKNOWLEDGMENTS

The authors would like to acknowledge Stephen Evans, Annie Fourrier-Reglat, and Silvana Romio for their participation in the many discussions about this research.

APPENDIX I. LEOPARD and protopathic bias

For every suspect drug-event combination, the number of prescriptions starting in a 51-day window around the event date is counted, as shown in the figure below for the drugs Omeprazole and Naproxen for the event upper gastrointestinal bleeding (data from the IPCI database). The number of prescriptions in the 25 days prior to the event is compared to the number or prescriptions starting in the 25 days after the event. An increase in the number of prescriptions after the event date, relative to the number of prescriptions before the event date, is taken to be an indication that the drug is used to treat the event or a precursor of the event, rather than cause it. This is tested using a binomial test. The increased risk with Omeprazole (IRR= 3.9) is likely due to protopathic bias. 978 prescriptions of Omeprazole were initiated in the 25 days before the event while 3 459 prescriptions were initiated in the 25 days after. For Omeprazole, p< 0.001, indicating the signal is probably caused by protopathic bias, while Naproxen has p = 1.00, indicating that the signal is probably not caused by protopathic bias. A signal is considered to be caused by protopathic bias if the p-value is below 0.5.

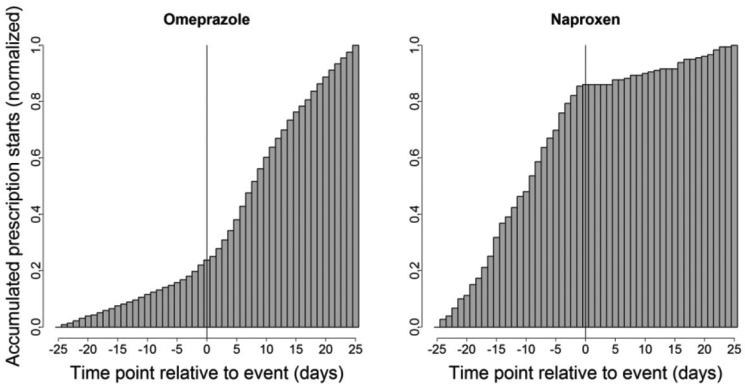
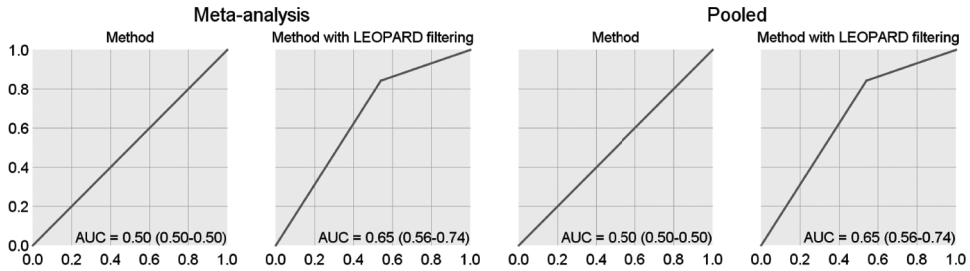
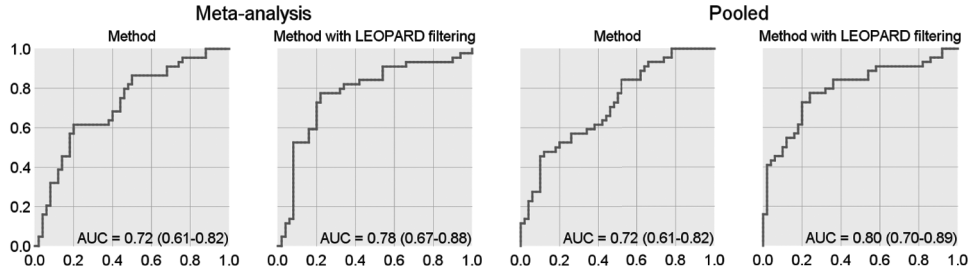


Figure A1 | Empirical cumulative distribution functions of prescription starts in a window around upper gastrointestinal bleeding occurrences for Omeprazole and Naproxen ***(see page 310 for colour figure).***

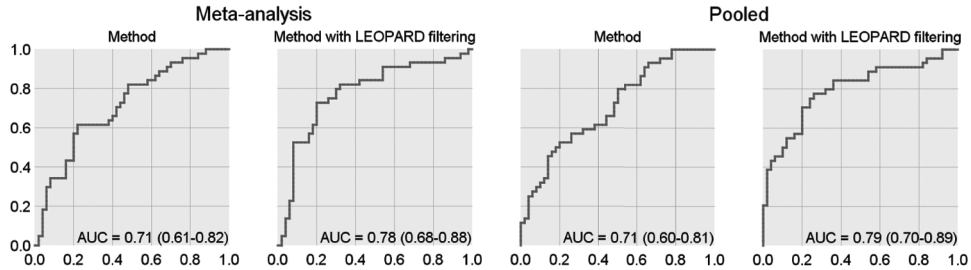
APPENDIX II. Receiver Operator Characteristic (ROC) Curves



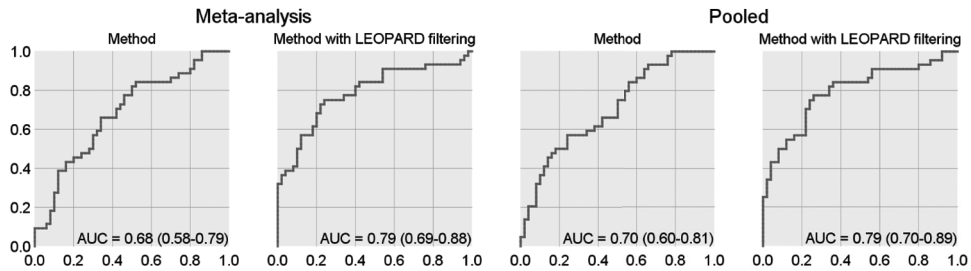
Proportional Reporting Ratio (PRR)

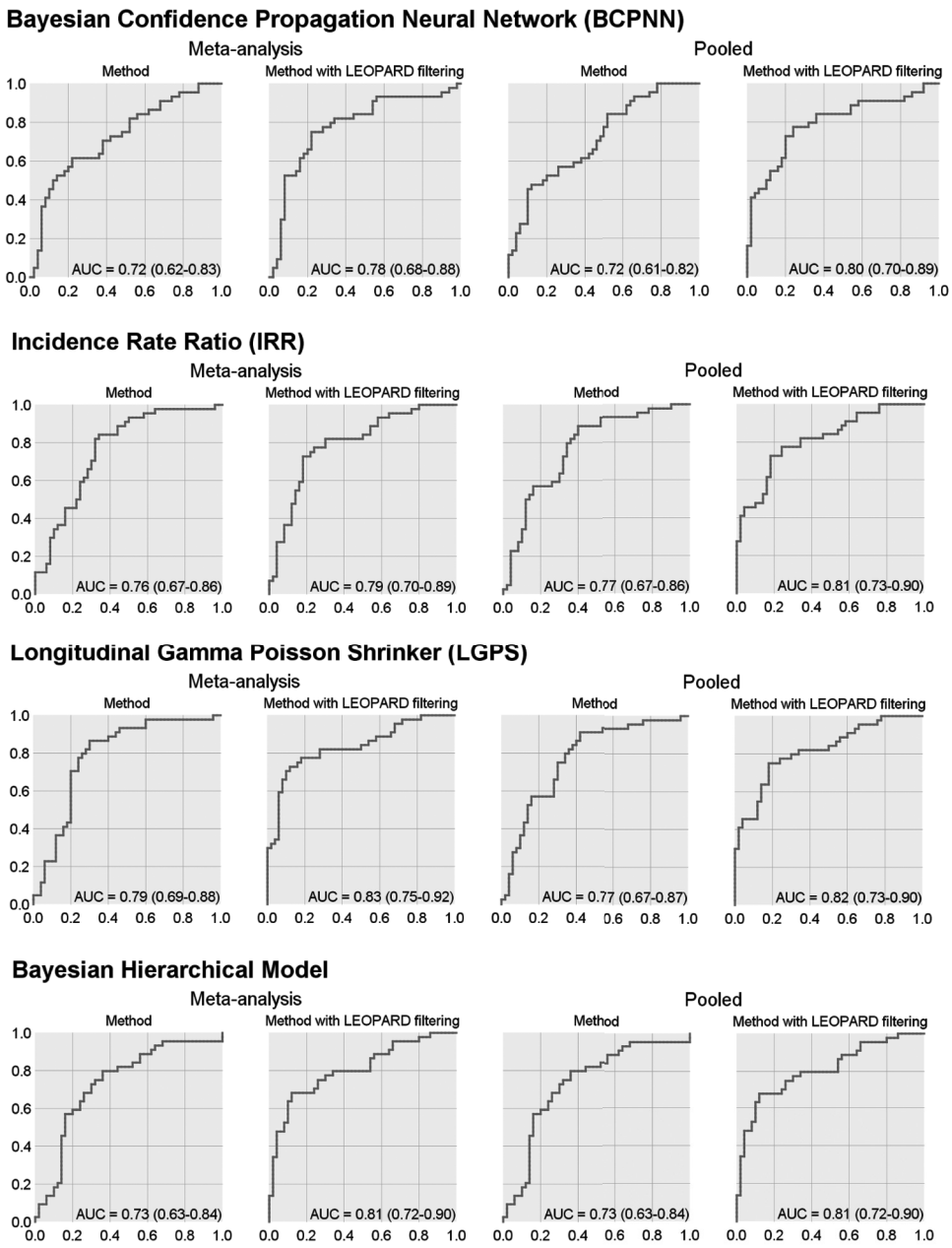


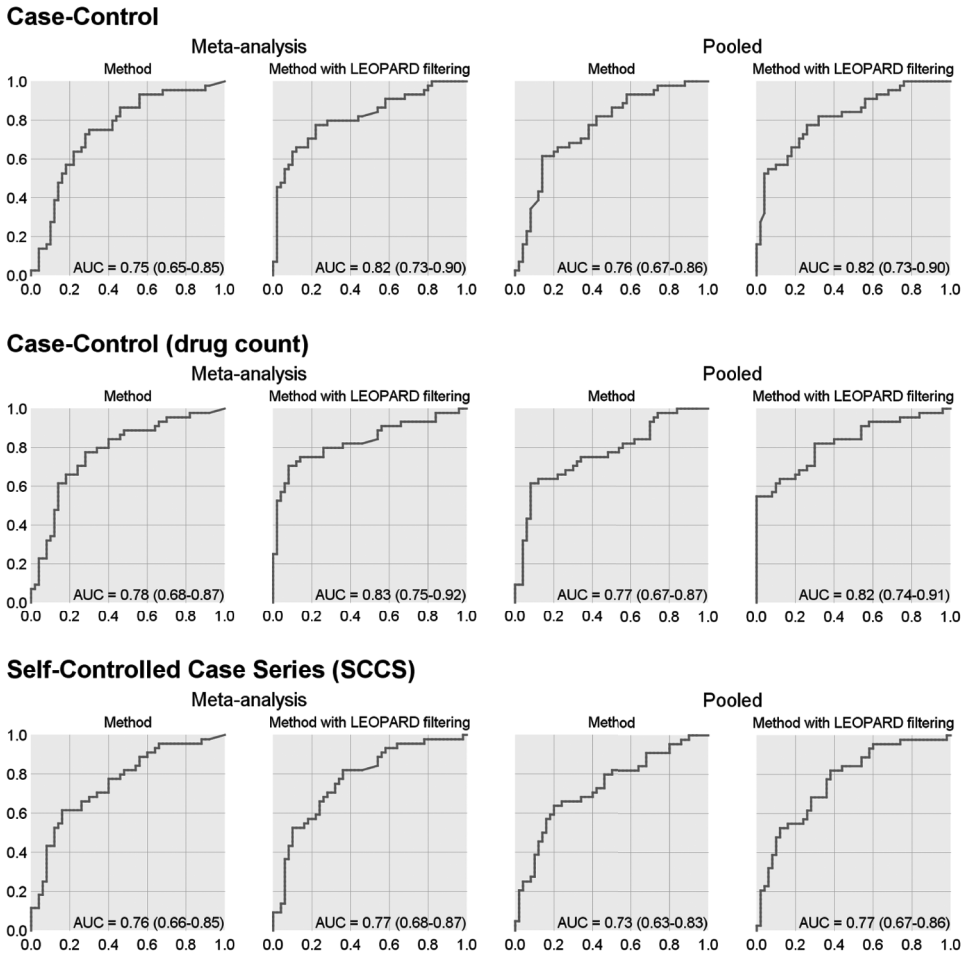
Reporting Odds Ratio (ROR)



Gamma Poisson Shrinker (GPS)




Bayesian Confidence Propagation Neural Network (BCPNN)
Meta-analysis
Pooled
Method
Method with LEOPARD filtering
AUC = 0.72 (0.62-0.83)
AUC = 0.78 (0.68-0.88)
AUC = 0.72 (0.61-0.82)
AUC = 0.80 (0.70-0.89)
Incidence Rate Ratio (IRR)
AUC = 0.76 (0.67-0.86)
AUC = 0.79 (0.70-0.89)
AUC = 0.77 (0.67-0.86)
AUC = 0.81 (0.73-0.90)
Longitudinal Gamma Poisson Shrinker (LGPS)
AUC = 0.79 (0.69-0.88)
AUC = 0.83 (0.75-0.92)
AUC = 0.77 (0.67-0.87)
AUC = 0.82 (0.73-0.90)
Bayesian Hierarchical Model
AUC = 0.73 (0.63-0.84)
AUC = 0.81 (0.72-0.90)
AUC = 0.73 (0.63-0.84)
AUC = 0.81 (0.72-0.90)


Case-Control
Meta-analysis
Pooled
Method
Method with LEOPARD filtering
AUC = 0.75 (0.65-0.85)
AUC = 0.82 (0.73-0.90)
AUC = 0.76 (0.67-0.86)
AUC = 0.82 (0.73-0.90)
Case-Control (drug count)
AUC = 0.78 (0.68-0.87)
AUC = 0.83 (0.75-0.92)
AUC = 0.77 (0.67-0.87)
AUC = 0.82 (0.74-0.91)
Self-Controlled Case Series (SCCS)
AUC = 0.76 (0.66-0.85)
AUC = 0.77 (0.68-0.87)
AUC = 0.73 (0.63-0.83)
AUC = 0.77 (0.67-0.86)
0.0 0.2 0.4 0.6 0.8 1.0

REFERENCES

1. Mann R, Andrews E, eds. Pharmacovigilance. John Wiley & Sons; 2002
2. Ahmad SR. Adverse drug event monitoring at the Food and Drug Administration. *J Gen Intern Med* 2003;18:57-60
3. Olsson S. The role of the WHO programme on International Drug Monitoring in coordinating worldwide drug safety efforts. *Drug Saf* 1998;19:1-10
4. Avorn J. Evaluating drug effects in the post-Vioxx world: there must be a better way. *Circulation* 2006;113:2173-2176
5. Meyboom RH, Egberts AC, Edwards IR, et al. Principles of signal detection in pharmacovigilance. *Drug Saf* 1997;16:355-365
6. Belton KJ. Attitude survey of adverse drug-reaction reporting by health care professionals across the European Union. The European Pharmacovigilance Research Group. *Eur J Clin Pharmacol* 1997;52:423-427
7. Alvarez-Requejo A, Carvajal A, Begaud B, et al. Under-reporting of adverse drug reactions. Estimate based on a spontaneous reporting scheme and a sentinel system. *Eur J Clin Pharmacol* 1998;54:483-488
8. Eland IA, Belton KJ, van Grootheest AC, et al. Attitudinal survey of voluntary reporting of adverse drug reactions. *Br J Clin Pharmacol* 1999;48:623-627
9. De Bruin ML, van Puijenbroek EP, Egberts AC, et al. Non-sedating antihistamine drugs and cardiac arrhythmias -- biased risk estimates from spontaneous reporting systems? *Br J Clin Pharmacol* 2002;53:370-374
10. Begaud B, Martin K, Haramburu F, et al. Rates of spontaneous reporting of adverse drug reactions in France. *JAMA* 2002;288:1588
11. Hazell L, Shakir SA. Under-reporting of adverse drug reactions : a systematic review. *Drug Saf* 2006;29:385-396
12. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. *Drug Saf* 2009;32:19-31
13. Perez Garcia M, Figueras A. The lack of knowledge about the voluntary reporting system of adverse drug reactions as a major cause of underreporting: direct survey among health professionals. *Pharmacoepidemiol Drug Saf* 2011;20:1295-1302
14. Gonzalez-Rubio F, Calderon-Larranaga A, Poblador-Plou B, et al. Underreporting of recognized adverse drug reactions by primary care physicians: an exploratory study. *Pharmacoepidemiol Drug Saf* 2011;20:1287-1294
15. Pushkin R, Frassetto L, Tsourounis C, et al. Improving the reporting of adverse drug reactions in the hospital setting. *Postgrad Med* 2010;122:154-164
16. Meyboom RH, Hekster YA, Egberts AC, et al. Causal or casual? The role of causality assessment in pharmacovigilance. *Drug Saf* 1997;17:374-389
17. McClellan M. Drug safety reform at the FDA--pendulum swing or systematic improvement? *N Engl J Med* 2007;356:1700-1702
18. Platt R. Challenges for the FDA: The Future of Drug Safety, Workshop Summary.: National Academy of Sciences; 2007
19. Coloma PM, Schuemie MJ, Trifiro G, et al. 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
20. Trifiro G, Pariente A, Coloma PM, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf* 2009;18:1176-1184
21. Coloma PM, Avillach P, Schuemie MJ, et al. A Reference Standard for Evaluating Methods for Drug Safety Signal Detection using Electronic Healthcare Record (EHR) Databases. 11th ISOP Annual Meeting. Istanbul; 2011
22. Coloma PM, Avillach P, Salvo F, et al. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection using Electronic Healthcare Records (EHR) Databases. *Drug Safety* 2012 [In press]
23. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf* 2001;10:483-486

24. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol Drug Saf* 2004;13:519-523
25. DuMouchel W. Bayesian Data Mining in Large Frequency Tables, with an Application to the FDA Spontaneous Reporting System. *Am Stat* 1999;53:177-190
26. Norén GN, Bate A, Orre R, et al. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events. *Stat Med* 2006;25:3740-3757
27. Schuemie MJ. Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD. *Pharmacoepidemiol Drug Saf* 2011;20:292-299
28. Gelman A. Prior distributions for variance parameters in hierarchical models. Bayesian Analysis 2006;1: 515-534
29. Berry SM, Berry DA. Accounting for multiplicities in assessing drug safety: a three-level hierarchical mixture model. *Biometrics* 2004;60:418-426
30. Schneeweiss S, Seeger JD, Maclure M, et al. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol 2001;154:854-864
31. Whitaker HJ, Farrington CP, Spiessens B, et al. Tutorial in biostatistics: the self-controlled case series method. *Stat Med* 2006;25:1768-1797
32. Rassen JA, Solomon DH, Curtis JR, et al. Privacy-maintaining propensity score-based pooling of multiple databases applied to a study of biologics. *Med Care* 2010;48:S83-89
33. Gerritsen WR, Peters A, Henny FC, et al. Ciprofloxacin-induced nephrotoxicity. *Nephrol Dial Transplant* 1987;2:382-383
34. Allon M, Lopez EJ, Min KW. Acute renal failure due to ciprofloxacin. *Arch Intern Med* 1990;150:2187-2189
35. Hootkins R, Fenves AZ, Stephens MK. Acute renal failure secondary to oral ciprofloxacin therapy: a presentation of three cases and a review of the literature. *Clin Nephrol* 1989;32:75-78
36. Goffin E, Herbiet L, Pouthier D, et al. Vancomycin and ciprofloxacin: systemic antibiotic administration for peritoneal dialysis-associated peritonitis. *Perit Dial Int* 2004;24:433-439
37. Ottervanger JP, Wilson JH, Stricker BH. Drug-induced chest pain and myocardial infarction. Reports to a national centre and review of the literature. *Eur J Clin Pharmacol* 1997;53:105-110
38. Coloma PM, Trifirò G, Schuemie MJ, et al. Electronic healthcare databases for active drug safety surveillance: Is there enough leverage? *Pharmacoepidemiology and Drug Safety* 2012. Feb 8. doi: 10.1002/pds.3197. [Epub ahead of print]

4.5

Positive predictive value of different coding-based algorithms for identification of patients with upper gastrointestinal bleeding in four European electronic healthcare record databases

Vera E. Valkhoff, MD, MSc[1,2], Preciosa M. Coloma, MD, MSc[1], Rosa Gini,PhD[3], Francesco Lapi[4], Mariam Molokhia, MD, PhD[5], Bartholemeus Mosseveld[1], Malene Schou Nielsson[6], Martijn Schuemie, PhD[1], Frantz Thiessard, PhD[7], Miriam Sturkenboom, PharmD, PhD[1,8], Gianluca Trifiro,MD, PhD [1,9] on behalf of the EU-ADR Consortium

[1]Department of Medical Informatics, Erasmus University Medical Centre, Rotterdam, the Netherlands; [2]Department of Gastroenterology and Hepatology, Erasmus University Medical Centre, Rotterdam, the Netherlands; [3]Agenzia Regionale di Sanità della Toscana, Florence, Italy; [4]Health Search, Italian College of General Practitioners, Florence, Italy; [5]Department of Primary Care & Public Health Sciences, Kings College London, United Kingdom; [6]Department of Clinical Epidemiology, Aarhus University Hospital, Denmark; [7]LESIM, ISPED, University Bordeaux 2, France; [8]Department of Epidemiology, Erasmus University Medical Centre, Rotterdam, the Netherlands; [9]Department of Clinical and Experimental Medicine and Pharmacology, University of Messina, Italy.

ABSTRACT

Purpose. Validation of outcome events is warranted when using electronic healthcare records (EHR) for drug safety surveillance. We evaluated and compared the accuracy of various coding algorithms to identify patients with upper gastrointestinal bleeding (UGIB) from four European health care databases (DBs).

Methods. We conducted a validation study in the following EHR databases: 1) IPCI (Netherlands); 2) HSD (Italy); 3) ARS Regional DB (Italy); and 4) Aarhus (Denmark). The first two are medical record databases from primary care while the latter two are administrative/claims DBs. Three diagnosis coding systems were used in the identification of patients with UGIB: 1) International Classification of Diseases (ICD-) 9th revision (HSD and ARS); 2) ICD-10th revision (Aarhus); and 3) International Classification of Primary Care (ICPC) (IPCI). Additionally, free text search was conducted in IPCI and HSD. We identified patients using UGIB-specific codes (ICD-10: K25-29, K92.0, K92.1, K92.2; ICD-9: 531-535, 578; ICPC: D14, D15, D85, D86) or key words as identified by UMLS mapping. A random sample of 200 cases from all identified UGIB cases were selected from each DB (400 for IPCI) and reviewed manually by medically trained assessors. Positive predictive values (PPV) were calculated for each coding algorithm in each database.

Results. For IPCI, 70% of cases were identified via free text search using keywords. Upon review of the medical charts, 84 patients were confirmed as UGIB cases, resulting in a PPV of 24% (95% CI: 20%-29%). PPV was 23% (95% CI: 17%-30%) and 25% (95% CI: 19%-31%) for free text and ICPC codes respectively. For HSD, 70% of cases were identified via ICD9 codes. The overall PPV was 78% (95% CI: 72%-84%); the PPV was 90% (95% CI: 85%-95%) for codes and 50% (95% CI: 37%-63%) for free text only. The overall PPV for the ICD9-based algorithm in ARS was 77% (95% CI: 71%-83%). The estimated PPV was 66% for the ICD10-based coding algorithm (95% CI: 59%-73%).

Conclusions. EHR databases present a potentially good source of identifying patients with UGIB. However, there are differences in the accuracy of automated case identification using various databases (hospitalization claims vs. medical records) that stem from differences in both the coding systems used and in the type of data collected. Database-specific factors need to be factored in when evaluating events identified using disparate databases.

INTRODUCTION

Electronic healthcare record (EHR) databases are frequently used data sources when conducting pharmacoepidemiological studies.[1] The advantages of using EHR databases rely on the fact that real-world data are captured on a large scale, allowing for cost-efficiency and flexibility in study design to study the risk of adverse drug events with multiple drugs. Both medical record databases and administrative/claims databases have been used to characterize healthcare utilization patterns as well as monitor patient outcomes.[2-3] The use of such databases for proactive drug safety surveillance is gaining worldwide interest.[4-6]

One of the important drug-induced adverse events that need to be monitored is upper gastrointestinal bleeding (UGIB), largely associated with use of non-steroidal anti-inflammatory drugs (NSAIDs) and a reason for many adverse drug reaction (ADR)-related hospitalizations. The accuracy of the diagnosis codes for the identification of UGIB has been previously assessed in several studies,[7-12] but only for two disease coding terminologies: International Classification of Diseases (ICD) 9th revision and READ codes. To our knowledge, no validation studies have been conducted so far to evaluate the accuracy of International Classification of Primary Care (ICPC) and ICD-10 codes for ascertainment of UGIB in EHR.

This study was conducted within the context of the EU-ADR Project (*Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge*; http://www.euadr-project.org), a computerized integrated system of European EHR databases and biomedical databases geared towards the early detection of drug safety signals.[6] UGIB is one of the events considered in EU-ADR to be of particular interest from a pharmacovigilance and public health perspective.[13] Using the Unified Medical Language System (UMLS)[6] as guide, harmonization of event definitions was carried out in EU-ADR to adapt clinically sound criteria to the specific database terminologies.[14] To assess the accuracy of the harmonized coding algorithms, we determined the positive predictive value (PPV) of ICD-9, ICD-10, and ICPC coding-based algorithms for the identification of patients with UGIB in EHR databases. We also estimated the PPV of free text search algorithms employed by databases with large amounts of narrative data.

METHODS

Data sources

The EU-ADR platform comprises anonymized healthcare data from eight European databases. Four databases agreed to participate in this validation study by performing hospitalization chart/electronic medical records review. The four databases are located in three European countries: two general practice (GP) databases (Health Search/CSD Patient DB (HSD, Italy);[15] Integrated Primary Care Information (IPCI, Netherlands);[16] and two administrative databases (Aarhus University Hospital Database (Aarhus, Denmark),[17] and ARS Tuscany (Italy)),[18] which are comprehensive record-linkage systems in which drug dispensing data of a well-defined population are linked to a

registry of hospital discharge diagnoses and various other registries. In these countries, all citizens are registered with a primary care practice, which maintains computerized records of all relevant medical information and drug prescriptions. The Italian databases (HSD and ARS) recorded medical events using the ICD- 9th Revision, Clinical Modification (ICD-9-CM), the Aarhus DB used the ICD-10th revision,[19] and IPCI, the International Classification for Primary Care (ICPC).[20] HSD and IPCI are primary care databases that also register clinical notes from general practitioners as narratives which may be used for the identification of medical events. These four databases cover overall a population of 8 million European inhabitants with a follow-up ranging from 1995 to 2010 (2003-2009 for HSD, 2002-2008 for ARS, 1995-2011 for IPCI, and 1999-2008 for Aarhus). All databases have been extensively used for pharmacoepidemiological research.[6]

UGIB case identification

Potential case patients with UGIB were identified through automated retrieval using diagnosis codes and free text which are listed in Tables 1 and 2. Only HSD and IPCI performed the free text searches. The process of mapping and harmonization of event data extraction from different EHR databases in the EU-ADR project has been described in more detail in previous publications.[6, 14]

UGIB case validation

For each DB a random (unweighted) sample of 200 potential UGIB cases were selected. Since the use of free text search was extensive in IPCI, an additional 200 potential cases were drawn from those identified via free text.

To estimate the overall and code/free text-specific positive predictive value (PPV), all the medical records in primary care DBs (IPCI and HSD) and the medical charts in claims DBs (ARS and Aarhus) of the randomly selected UGIB cases were retrieved and abstracted to perform the case validation. To ensure uniformity in the process of data abstraction for UGIB case validation across the four DBs, a custom-built electronic data entry software Chameleon© was provided and locally installed in each DB. A standardized questionnaire containing all relevant questions eliciting medical information was implemented in Chameleon©. Medically trained assessors provided the required data for the validation of each of these potential cases according to the questionnaire. In particular, clinically relevant information was captured regarding signs/symptoms (hematemesis, melena, peptic ulcer disease), bleeding site (as documented by endoscopy, surgery, or autopsy), cause of death, endoscopy findings, and potential alternative explanations, such as black stools due to ingestion of iron/ferrous-containing medications. The etiology of UGIB, if known, was likewise noted: peptic ulcer bleeding, Mallory-Weiss (MW) syndrome, angiodysplasia, Dieulafoy lesion, iatrogenic UGI bleeding caused by procedure with endoscopy (i.e. endoscopic retrograde cholangiopancreatography; papillotomy), variceal bleeding, or bleeding caused by a gastrointestinal malignancy.

Table 1 | List of diagnostic codes (plus related label) used in the coding algorithm for UGIB case identification in electronic medical records stratified by coding system.

ICD-9 (ARS, HSD)		ICD-10 (for Aarhus)		ICPC (for IPCI)	
530.21	Ulcer of esophagus with bleeding				
530.82	Esophageal hemorrhage				
531.00/ 531.01	Gastric ulcer, Acute with hemorrhage	K25.0	Gastric ulcer, Acute with hemorrhage		
531.10	Gastric ulcer, Acute with perforation	K25.1	Gastric ulcer, Acute with perforation		
531.20/531.21	Gastric ulcer, Acute with hemorrhage and perforation	K25.2	Gastric ulcer, Acute with both hemorrhage and perforation		
532.00/532.01	Duodenal ulcer, Acute with hemorrhage	K26.0	Duodenal ulcer, Acute with hemorrhage	D85	Duodenal ulcer
532.10	Duodenal ulcer, Acute with perforation	K26.1	Duodenal ulcer, Acute with perforation		
532.20	Duodenal ulcer, Acute with hemorrhage and perforation	K26.2	Duodenal ulcer, Acute with both hemorrhage and perforation		
533.00	Peptic ulcer, site unspecified, Acute with hemorrhage	K27.0	Peptic ulcer, site unspecified, Acute with hemorrhage	D86	Peptic ulcer, other
533.10	Peptic ulcer, site unspecified, Acute with perforation	K27.1	Peptic ulcer, site unspecified, Acute with perforation		
533.20	Peptic ulcer, site unspecified, Acute with hemorrhage and perforation	K27.2	Peptic ulcer, site unspecified, Acute with both hemorrhage and perforation		
534.00/534.01	Gastrojejunal ulcer, Acute with hemorrhage	K28.0	Gastrojejunal ulcer, Acute with hemorrhage		
534.10	Gastrojejunal ulcer, Acute with perforation	K28.1	Gastrojejunal ulcer, Acute with perforation		
534.20/534.21	Gastrojejunal ulcer, Acute with hemorrhage and perforation	K28.2	Gastrojejunal ulcer, Acute with both hemorrhage and perforation		
535.01	Acute gastritis, with hemorrhage	K29.0	Acute hemorrhagic gastritis		
535.11	Atrophic gastritis, with hemorrhage				
535.41	Other specified gastritis, with hemorrhage				
535.51	Unspecified gastritis and gastroduodenitis, with hemorrhage				
578.0	Hematemesis, Vomiting of blood	K92.0	Hematemesis	D15	Hematemesis
578.1	Blood in stool, Melena	K92.1	Melena	D14	Melena
578.9	Hemorrhage of gastrointestinal tract, unspecified	K92.2	Gastrointestinal hemorrhage, unspecified		

Table 2 | Free text search strings employed in the identification of cases of upper gastrointestinal bleeding in general practitioner (GP) databases (search done both for complaints/symptoms and diagnoses).

IPCI (Dutch)	HSD (Italian)
(bloed AND maag) OR (bloed AND mela) OR (loed AND braken)OR (bloed AND overgev) OR (zwart AND ontlast) OR haematcemesis OR melaena OR hematemesis OR melena OR heamatemesis OR meleana)	*'Emorragia gastr' OR 'Emorragia tratto gastr' OR 'Emorragia intest' OR 'Vomito Sangue' OR 'Ematemesi' OR 'Melena' OR 'feci nere' OR 'feci scure' OR 'feci color pece' OR 'feci picee' OR 'Sanguinamento gastr' OR 'Sanguinamento tratto gastr' OR 'Sanguinamento intest' OR 'Sanguinamento esof' OR 'Sanguinamento du' OR 'Sanguinamento digiu' OR 'Sanguinamento prepi' OR 'Sanguinamento pil' OR 'Ulcera con emo' OR 'Ulcera sang' OR 'Esofagite con emo' OR 'Gastrite con emo' OR 'Duodenite con emo' OR 'Antrite con emo' OR 'Esofagite emo' OR 'Gastrite emo' OR 'Duodenite emo' OR 'Antrite emo' OR 'Perforazione con sanguinamento' OR 'Perforazione con emo'*
Refinement done*: Discarded records identified by code D86 unless a specificity procedure involving free text and diagnostic procedures (gastroscopy or duodenoscopy) within 30 days is successful. Discarded records which contain the string 'geen' ('no/none'). Discarded records with reference to chronic UGIB, UGIB in father/mother/sister/brother, prior history of UGIB*	Refinement done*: Discarded records that contain the string 'rettor'.*

Additionally, the exact date of onset of bleeding (i.e. date of first symptom/occurrence of UGIB), if documented, was retrieved ('reviewed date') to compare it with the date that was automatically detected using the coding algorithm ('coded date'). For GP databases (IPCI and HSD) it was indicated also whether the UGIB diagnosis was made by a GP or medical specialist.

Two types of validation were performed: assessor judgement and algorithm based on structured feature extraction. Based on the information collected in the questionnaire, the potential UGIB cases were classified as: (1) definite case; (2) non-case; or (3) non-assessable case, if the available information was judged as not sufficient to perform the case validation (i.e., a diagnosis of gastrointestinal bleeding from unspecified location without further clinical or endoscopic information). A patient with UGIB was defined as a confirmed case if: (1) melena or hematemesis was described and bleeding site was located in esophagus/stomach/duodenum; or (2) gastrointestinal bleeding was mentioned in the records as being consistent with peptic ulcer disease, gastric erosions or gastritis and no source of bleeding outside the UGI tract. Drug safety surveillance being the primary concern in the EU-ADR project, we were interested in potentially drug-induced UGIB and therefore in one of the two algorithms for case validation applied, we have excluded the following conditions as case: Mallory Weiss syndrome, angiodysplasia, Dieulafoy lesion, iatrogenic UGI bleeding, variceal bleeding, or bleeding caused by a gastrointestinal malignancy ('algorithm narrow'). As primary outcome of the study, we used the PPV derived via the algorithm for UGIB definition in general ('algorithm broad') to ensure uniformity across DBs

in the case definition. The validation questionnaire implemented in the data extraction software Chameleon© is given in Appendix 1.

Statistical Analysis

For the primary analysis, we calculated the overall and code/free text-specific PPVs, based on both automatic ('broad' and 'narrow') and assessor-based validation. The overall PPV was calculated as the probability of the number of UGIB confirmed cases and the total number of randomly selected potential cases, together with 95% confidence intervals (95% CIs). PPV was likewise calculated for every code and free text search used. Non-assessable cases were not included in either the numerator or the denominator for the PPV calculation.

Secondly, we calculated the difference in days between the UGIB date of the registration in the DB ('coded date') and the true date of onset of UGIB ('reviewed date'). This was done for the confirmed UGIB cases only.

RESULTS

Overall, the four databases covered a population of 8 044 637 European inhabitants with 41 086 273 person-years of follow-up in the period 1995-2011. Within this population, a total of 39 870 potential cases of UGIB were identified. The medical/hospitalization charts for 985 of the 1 000 randomly selected potential cases were retrieved for validation: ARS: 185; HSD: 200; Aarhus: 200; IPCI: 400, subdivided in 200 records for text and 200 records for codes. Case validation in IPCI via review of electronic medical records maintained by general practitioners (GPs), while for ARS and Aarhus, validation was done using chart review from hospitals in the corresponding catchment area. In all databases but IPCI, there were more male patients with potential UGIB than females (male sex: ARS 55%; HSD 53%; Aarhus: 59%; IPCI: 43%). Mean age ranged between 52.5 and 69.3 years old. Melena was more frequently reported than hematemesis as sign of UGIB. Four fatal cases of UGIB were identified in Aarhus and one in IPCI; there were no fatal cases of UGIB documented in the other two databases. Peptic ulcer bleeding (PUB) was reported for 24% to 59% of confirmed cases, followed by esophagitis/gastritis (3% to 15%) and erosive gastrointestinal disease (0% to 8%). The diagnosis was made by a specialist in two thirds of the confirmed cases in IPCI while only 3 out of 156 confirmed cases were diagnosed by a specialist in HSD (the rest were diagnosed by the GP, see Table 3).

Of the 979 potential UGIB cases which were eventually reviewed for validation, 514 (ARS: 142; HSD: 156; ARS: 132; IPCI: 84) were confirmed cases. The positive predictive value (PPV) for the coding algorithms used in ARS and Aarhus was high, with PPVs of 77% (95% CI: 71%-83%) and 66% (95% CI: 59%-73%), respectively. For IPCI, the PPV (both ICPC coding and text) was 24% (95% CI: 20%-29%) and for HSD (ICD-9 coding and text) it was 78% (95% CI: 72%-84%).

Table 3 | Characteristics of random sample of potential UGIB case patients

	ARS		HSD		Aarhus		IPCI	
	Total N:185 (%)	Confirmed cases N: 142 (%)	Total N: 200 (%)	Confirmed cases N: 156 (%)	Total N: 200 (%)	Confirmed cases N: 132 (%)	Total N: 400 (%)	Confirmed cases N: 84 (%)
Male sex	101 (55)	84 (59)	105 (53)	83 (53)	117 (59)	80 (61)	172 (43)	44 (52)
Female sex	84 (45)	58 (41)	95 (48)	73 (47)	83 (42)	52 (40)	228 (57)	40 (48)
Mean age (years)	67.4	69.3	65.0	64.9	69.3	70.4	52.5	57.4
Age categories:								
0 - 20 years	14 (8)	8 (6)	2 (1)	2 (1)	2 (1)	1 (1)	48 (12)	8 (10)
20 - 40 years	8 (4)	7 (5)	21 (11)	18 (9)	13 (7)	6 (5)	69 (17)	8 (10)
40 - 60 years	26 (14)	18 (10)	50 (25)	37 (24)	39 (20)	26 (20)	120 (30)	24 (29)
60 - 80 years	82 (44)	61 (33)	85 (43)	66 (42)	87 (44)	55 (42)	101 (25)	28 (33)
> 80 years	55 (30)	48 (34)	42 (21)	33 (21)	59 (30)	44 (33)	62 (16)	16 (19)
Sign of UGIB:*								
melena	98 (53)	93 (65)	95 (48)	87 (56)	94 (47)	92 (70)	83 (21)	40 (48)
hematemesis	42 (23)	38 (27)	25 (13)	25 (16)	78 (39)	76 (58)	60 (15)	42 (50)
Etiology of UGIB:								
Mallory-Weiss syndrome	1 (1)	1 (1)	0 (0)	0 (0)	1 (1)	1 (1)	11 (3)	6 (7)
Angiodysplasia	3 (2)	3 (2)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)	1 (1)
Dieulafoy	1 (1)	1 (1)	0 (0)	0 (0)	5 (3)	5 (4)	0 (0)	0 (0)
Iatrogenic	2 (1)	1 (1)	0 (0)	0 (0)	1 (1)	1 (1)	1 (0)	1 (1)
Varices	3 (2)	3 (2)	3 (2)	1 (0.6)	2 (1)	2 (2)	2 (1)	2 (2)
Cancer	10 (5)	7 (5)	1 (1)	1 (0.6)	6 (3)	6 (4)	12 (3)	4 (5)
Peptic ulcer	83 (45)	72 (51)	47 (24)	41 (26.3)	78 (39)	78 (59)	24 (6)	20 (24)
Erosion	14 (8)	12 (8)	0 (0)	0 (0)	6 (3)	6 (4)	4 (1)	4 (5)
Esophagitis/ Gastritis	16 (9)	13 (9)	5 (3)	4 (2.6)	13 (7)	12 (9)	16 (4)	13 (15)
Undocumented	52 (28)	29 (20)	144 (72)	109 (70)	88 (44)	20 (15)	329 (82)	33 (39)
Diagnosis made by:**								
GP	NA	NA	197 (99)	153 (98)	NA	NA	75 (19)	27 (32)
Specialist	100%	100%	3 (2)	3 (2)	100%	100%	79 (20)	55 (65)

* Can sum up to more than 100%
** Only for GP databases of relevance (HSD and IPCI)

For ARS, 142 patients were defined as confirmed UGIB case. Among the 18 ICD-9 codes used for UGIB identification code 532.00 (22%) was the most frequently used code, followed by 578.9 (19%) and 531.00 (16%) with respective PPVs equal to 95%, 67%, and 83%. For HSD a similar list of ICD-9 codes was retrieved and seven codes were available in the random sample, besides free text. In particular, UGIB cases were more frequently identified using 578.1 code (melena) (39%)

and free text (30%). Apart from code 531.10 (n= 3) and code 533.00 (n= 3), the PPVs for all codes were above 85%, while the PPV for free text was much lower: 50% (95% CI: 37%-63%).

As for the extraction of UGIB events in Aarhus, the following codes had the greatest contribution: K26.0 (25%), K25.0 (21%), and K92.0 (18%), with corresponding PPVs of 82%, 85%, and 72%. Only in Aarhus was the code for hematemesis more frequently reported than the code for melena (36 vs. 19 case patients). When comparing the PPVs for the ICD-9 based algorithms with the ICD10-based algorithm, code 578.9 for ICD-9 (unspecified hemorrhage of gastrointestinal tract) had a higher PPV for Italian DBs (ARS: 67%; HSD: 100%) compared to the PPV for its corresponding code for ICD-10 in Aarhus (K92.2; 43%).

The overall PPV for the algorithms used in IPCI was 24% (95% CI: 20%-29%). When stratified according to free text or ICPC codes, the PPV was 23% (95% CI: 16%-29%) and 25% (95% CI: 19%-32%), respectively. All four codes based on the harmonized ICPC list in **Table 1** were retrieved in the random sample of 200 patients, with the code for melena being the most frequent and the code for hematemesis being the least frequent.

Table 4 shows the PPV for each of the codes used in the search algorithm for each database. Of the codes referring to physical sign of UGIB, hematemesis had a higher PPV than melena for UGIB identification. The PPVs calculated according to the case validation as manually performed by the assessor were not different from those obtained automatically by the algorithm, except for the ICPC code D14 for IPCI (algorithm: PPV 71% versus assessor: 33%).

Table 4 | Positive predictive value of diagnostic codes for upper gastrointestinal bleeding, stratified by coding system and database.

Codes	Charts reviewed N (%)	PPV Broad algorithm (95% CI)	PPV Assessor (95% CI)	PPV Narrow algorithm (95% CI)
ICD-9 ARS:	185 (100)	76.8 (70.7 - 82.8)	72.3 (65.8 - 78.8)	71.9 (65.4 - 78.4)
530.82 Esophageal hemorrhage	1 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
531.00 Gastric ulcer, Acute with hemorrhage	29 (16)	82.8 (69 - 96.5)	79.3 (64.6 - 94.1)	79.3 (64.6 - 94.1)
531.01 Gastric ulcer, Acute with hemorrhage, with obstruction	1 (1)	100 (100 - 100)	100 (100 - 100)	0 (0 - 0)
531.10 Gastric ulcer, Acute with perforation	1 (1)	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
532.00 Duodenal ulcer, Acute with hemorrhage	40 (22)	95 (88.2 - 101.8)	92.5 (84.3 - 100.7)	95 (88.2 - 101.8)
532.01 Duodenal ulcer, Acute with hemorrhage, with obstruction	2 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
532.10 Duodenal ulcer, Acute with perforation	9 (5)	44.4 (12 - 76.9)	50 (15.4 - 84.6)	44.4 (12 - 76.9)
533.00 Peptic ulcer, site unspecified, Acute with hemorrhage	5 (3)	100 (100 - 100)	100 (100 - 100)	80 (44.9 - 115.1)
533.20 Peptic ulcer, site unspecified, Acute with hemorrhage and perforation	1 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
534.00 Gastrojejunal ulcer, Acute with hemorrhage	4 (2)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)

Codes	Charts reviewed N (%)	PPV Broad algorithm (95% CI)	PPV Assessor (95% CI)	PPV Narrow algorithm (95% CI)
534.01 Gastrojejunal ulcer, Acute with hemorrhage, with obstruction	1 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
534.21 Gastrojejunal ulcer, Acute with hemorrhage and perforation	1 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
535.01 Acute gastritis with hemorrhage	17 (9)	64.7 (42 - 87.4)	64.7 (42 - 87.4)	64.7 (42 - 87.4)
535.11 Atrophic gastritis with hemorrhage	1 (1)	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
535.41 Other specified gastritis with hemorrhage	3 (2)	33.3 (-20 - 86.7)	33.3 (-20 - 86.7)	33.3 (-20 - 86.7)
578.0 Hematemesis	12 (6)	83.3 (62.2 - 104.4)	83.3 (62.2 - 104.4)	66.7 (40 - 93.3)
578.1 Melena	21 (11)	66.7 (46.5 - 86.8)	57.1 (36 - 78.3)	61.9 (41.1 - 82.7)
578.9 Hemorrhage of gastrointestinal tract, unspecified	36 (19)	66.7 (51.3 - 82.1)	52.8 (36.5 - 69.1)	58.3 (42.2 - 74.4)
ICD-9 HSD:	200 (100)	78 (72.3 - 83.7)	80.7 (74.7 - 86.8)	75.5 (69.5 - 81.5)
All codes:	140 (70)	90 (85 - 95)	89.1 (83.5 - 94.7)	87.9 (82.4 - 93.3)
531.10 Gastric ulcer, Acute with perforation	3 (2)	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
531.00 Gastric ulcer, Acute with hemorrhage	19 (10)	100 (100 - 100)	94.7 (84.7 - 104.8)	100 (100 - 100)
532.00 Duodenal ulcer, Acute with hemorrhage	4 (2)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
533.00 Peptic ulcer, site unspecified, Acute with hemorrhage	3 (2)	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
578.0 Hematemesis	17 (9)	100 (100 - 100)	94.1 (82.9 - 105.3)	94.1 (82.9 - 105.3)
578.1 Melena	78 (39)	89.7 (83 - 96.5)	85.7 (77.9 - 93.5)	88.5 (81.4 - 95.6)
578.9 Hemorrhage of gastrointestinal tract, unspecified	16 (8)	100 (100 - 100)	100 (100 - 100)	93.8 (81.9 - 105.6)
Text	60 (30)	50 (37.3 - 62.7)	57.1 (42.2 - 72.1)	46.7 (34 - 59.3)
ICD-10 Aarhus:	200 (100)	66 (59.4 - 72.6)	72.9 (65.9 - 79.9)	60.5 (53.7 - 67.3)
K25.0 Gastric ulcer, Acute with hemorrhage	41 (21)	85.4 (74.5 - 96.2)	84.2 (72.6 - 95.8)	78 (65.4 - 90.7)
K25.1 Gastric ulcer, Acute with perforation	10 (5)	10 (-8.6 - 28.6)	0 (0 - 0)	10 (-8.6 - 28.6)
K25.2 Gastric ulcer, Acute with both hemorrhage and perforation	2 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
K26.0 Duodenal ulcer, Acute with hemorrhage	50 (25)	82 (71.4 - 92.6)	92.9 (85.1 - 100.6)	82 (71.4 - 92.6)
K26.1 Duodenal ulcer, Acute with perforation	7 (4)	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
K26.2 Duodenal ulcer, Acute with both hemorrhage and perforation	2 (1)	100 (100 - 100)	100 (100 - 100)	100 (100 - 100)
K27.0 Peptic ulcer, site unspecified, Acute with hemorrhage	2 (1)	50 (-19.3 - 119.3)	100 (100 - 100)	50 (-19.3 - 119.3)
K28.0 Gastrojejunal ulcer, Acute with hemorrhage	2 (1)	50 (-19.3 - 119.3)	50 (-19.3 - 119.3)	50 (-19.3 - 119.3)
K29.0 Acute hemorrhagic gastritis	8 (4)	62.5 (29 - 96)	80 (44.9 - 115.1)	62.5 (29 - 96)
K92.0 Hematemesis	36 (18)	72.2 (57.6 - 86.9)	70.4 (53.1 - 87.6)	58.3 (42.2 - 74.4)
K92.1 Melena	19 (10)	47.4 (24.9 - 69.8)	58.3 (30.4 - 86.2)	42.1 (19.9 - 64.3)
K92.2 Gastrointestinal hemorrhage, unspecified	21 (11)	42.9 (21.7 - 64)	42.9 (16.9 - 68.8)	33.3 (13.2 - 53.5)

Codes	Charts reviewed N (%)	PPV Broad algorithm (95% CI)	PPV Assessor (95% CI)	PPV Narrow algorithm (95% CI)
ICPC IPCI:	400 (100)	24.1 (19.6 - 28.6)	16.3 (12.7 - 19.9)	21.5 (17.2 - 25.8)
All codes:	200 (50)	24.9 (18.5 - 31.2)	15.6 (10.5 - 20.6)	21.5 (15.4 - 27.5)
D14 Hematemesis/ vomiting blood	40 (10)	71 (55 - 86.9)	32.5 (18 - 47)	58.1 (40.7 - 75.4)
D15 Melena	62 (16)	38.8 (25.1 - 52.4)	24.6 (13.8 - 35.4)	34.7 (21.4 - 48)
D85 Duodenal ulcer	49 (12)	2.1 (-2 - 6.1)	2 (-1.9 - 6)	2.1 (-2 - 6.1)
D86 Peptic ulcer other	49 (12)	4.1 (-1.5 - 9.6)	4.1 (-1.5 - 9.6)	4.1 (-1.5 - 9.6)
Free text	200 (50)	22.5 (16.2 - 28.8)	17 (11.8 - 22.2)	21.5 (15.4 - 27.7)

Validation of date of Upper Gastrointestinal Bleeding occurrence

For ARS, 39 out of 142 (27%) confirmed cases had a different actual date of onset ('reviewed date') than the automatically determined UGIB date ('coded date'). For the majority (82%) the date of onset upon review was earlier than the recorded diagnosis date (range -29 day to +19 day). For Aarhus, 21% of confirmed cases had a different reviewed date of event as recorded; for 79% the reviewed date was one or two days earlier than the coded date (range -2 days to +2922 days). For IPCI, of the 82 confirmed cases, 25 had a reviewed date earlier than the coded event date and one had a reviewed date later than coded date (range -1200 days to + 1096 days) (Figure 1). For HSD, the coded date for all confirmed cases corresponded to the reviewed date

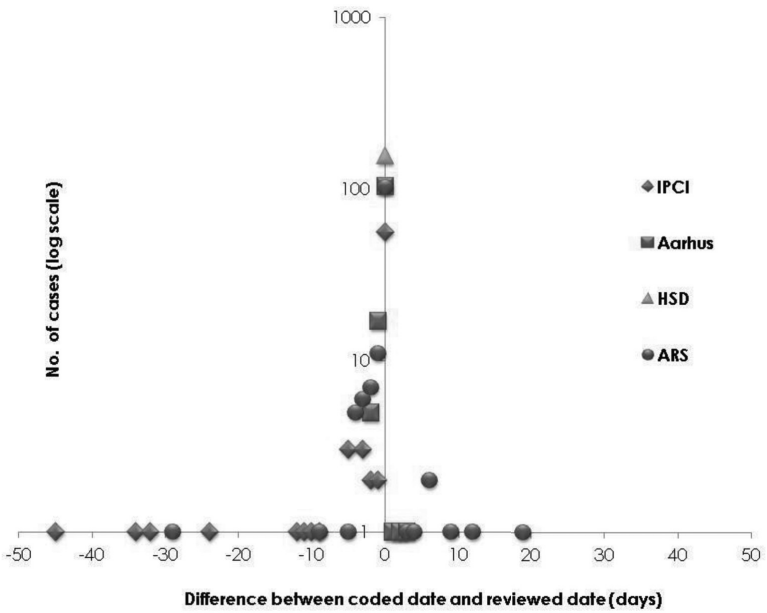


Figure 1 | Distribution of difference in days between coded date and reviewed date of event in confirmed UGIB cases only ***(see page 311 for colour figure)*****.**

DISCUSSION

In this validation study we determined the positive predictive value (PPV) of search algorithms employing diagnostic codes and free text for identification of patients with UGIB in different European EHR databases. We analyzed the accuracy of three well-described disease coding systems (ICD-9th and 10th revision and ICPC), free text versus code search, databases originating from three different countries (Denmark, Italy, and the Netherlands), and within the same country (Italy). In addition, we explored the difference in PPV for the same codes between a primary care (HSD) versus hospital setting (ARS), while using the same coding scheme (ICD-9). This is a strength and novelty of our study.

In general the PPV of the more extensive coding system of ICD-9 and ICD-10 outperformed the less granular coding system of ICPC. This is explainable by the fact that none of the four available ICPC codes refers specifically to UGIB, but rather to uncomplicated peptic ulcers or signs of UGIB only. Moreover, UGIB case misclassification in primary care databases such as IPCI capturing the less severe events may be more likely than hospital discharge databases containing the more severe hospitalized cases. For two databases we had the opportunity to explore the PPV of free text search which used the label of the selected codes as key words for the identification of UGIB events. For HSD, 30% of cases were identified via free text with a PPV of 50%, indicating that the PPV for free text search was significantly lower than for diagnosis codes. In IPCI, up to 70% of the cases were identified by free text search alone. The PPV for free text was very low (23%) but surprisingly, was equal to PPV for codes alone (25%). This shows that for complete information regarding the cases, the free text search is of importance, but it should not be used in an automatic search without manual validation so as not to increase the false positive rate. In the IPCI database manual review is always conducted, and currently advanced text mining activities are implemented to reduce the burden of manual inspection.[21]

Our findings of the PPVs originating from the ICD-9 diagnoses codes (ARS, HSD) can be compared with other studies. One study, also originating from Italy, studied UGIB (including perforation) and the PPVs are in line with the PPVs we obtained.[8] The PPV for hematemesis (ICD-9: 578.0) we reported was lower (59% versus ARS 83% and HSD 100%), but our results on the PPV concerning code 578 were similar to the study by Raiford *et al.*[10] Two other studies with ulcer disease and UGIB as endpoint had lower PPVs (around 60%) as reported in the current study,[7, 11] especially the PPV for unspecific codes (578) was reported low (9% only) in a study with eight Health Maintenance Organizations.[7] We are unaware of studies reporting the PPV for UGIB codes for ICPC or ICD-10. However, one study is available reporting a PPV of 98% for ulcer disease in general for ICD-10 codes in Aarhus.[22]

There are several limitations to our study. First, despite a standardized event definition for UGIB provided and a common tool (Chameleon©) with which to enter the available information in a structured manner, the UGIB case ascertainment process relied, to some degree, on the assessor's judgement. Therefore, apart from the judgement of the assessor, an automated algorithm was developed to allocate UGIB cases to avoid subjectivity in the case definition across EHRs. This

automated algorithm, however, was dependent on the completeness of the available data and thus had limited value. The PPVs derived from the automated algorithm and from the assessor were similar, except for hematemesis in IPCI, which might be explained by a conservative validation approach by the medical assessor in IPCI. Moreover, in this study we have limited the event definition to UGIB only and to potentially drug-induced UGIB, whereas other studies included perforation or uncomplicated ulcer disease in general, leading to conservative estimates. The reason that the PPVs for the narrow concept of UGIB were not much lower than the PPVs for the broad concept of UGIB is understandable as peptic ulcer bleeding and erosive disease are responsible for the great majority of UGIB cases as reported in literature,[23] and similar to our results.

In conclusion, EHR databases present a potentially good source of identifying patients with UGIB. However, there are differences in the accuracy of automated case identification using various coding-based algorithms in databases that stem from differences in both the coding systems used and in the type of data collected (hospitalization claims vs. medical records) which are, in turn, a consequence of the primary purpose for which the data are being collected in the first place. Database-specific factors need to be factored in when evaluating events identified using disparate databases.

REFERENCES

1. Strom BL, Carson JL. Automated data bases used for pharmacoepidemiology research. *Clinical pharmacology and therapeutics.* Oct 1989;46(4):390-394.
2. Schneeweiss S AJ. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005;58(4):323-337.
3. Suissa S GE. Primer: administrative health databases in observational studies of drug effects—advantages and disadvantages. . *Nat Clin Pract Rheum* 2007;3(12):725-732.
4. 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.* Aug 13 2009;361(7):645-647.
5. 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.* Nov 2 2010;153(9):600-606.
6. Coloma PM, Schuemie MJ, Trifiro G, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiol Drug Saf.* Jan 2011;20(1):1-11.
7. Andrade SE, Gurwitz JH, Chan KA, et al. Validation of diagnoses of peptic ulcers and bleeding from administrative databases: a multi-health maintenance organization study. *J Clin Epidemiol.* Mar 2002;55(3):310-313.
8. Cattaruzzi C, Troncon MG, Agostinis L, Garcia Rodriguez LA. Positive predictive value of ICD-9th codes for upper gastrointestinal bleeding and perforation in the Sistema Informativo Sanitario Regionale database. *J Clin Epidemiol.* Jun 1999;52(6):499-502.
9. Margulis AV, Garcia Rodriguez LA, Hernandez-Diaz S. Positive predictive value of computerized medical records for uncomplicated and complicated upper gastrointestinal ulcer. *Pharmacoepidemiology and drug safety.* Oct 2009;18(10):900-909.
10. Raiford DS, Perez Gutthann S, Garcia Rodriguez LA. Positive predictive value of ICD-9 codes in the identification of cases of complicated peptic ulcer disease in the Saskatchewan hospital automated database. *Epidemiology (Cambridge, Mass.* Jan 1996;7(1):101-104.
11. Wahl PM, Rodgers K, Schneeweiss S, et al. Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population. *Pharmacoepidemiology and drug safety.* Jun 2010;19(6):596-603.
12. Lopushinsky SR, Covarrubia KA, Rabeneck L, Austin PC, Urbach DR. Accuracy of administrative health data for the diagnosis of upper gastrointestinal diseases. *Surgical endoscopy.* Oct 2007;21(10):1733-1737.
13. Trifiro G, Pariente A, Coloma PM, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf.* Dec 2009;18(12):1176-1184.
14. Avillach P, Joubert M, Thiessard F, et al. Design and evaluation of 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.* 2010;160(Pt 2):1085-1089.
15. Cricelli C, Mazzaglia G, Samani F, et al. Prevalence estimates for chronic diseases in Italy: exploring the differences between self-report and primary care databases. *Journal of public health medicine.* Sep 2003;25(3):254-257.
16. Vlug AE, van der Lei J, Mosseveld BM, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf Med.* Dec 1999;38(4-5):339-344.
17. Christensen S, Riis A, Norgaard M, Thomsen RW, Sorensen HT. Introduction of newer selective cyclo-oxygenase-2 inhibitors and rates of hospitalization with bleeding and perforated peptic ulcer. *Aliment Pharmacol Ther.* Apr 15 2007;25(8):907-912.
18. 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)]. *Epidemiologia e prevenzione.* May-Jun 2006;30(3):161-168.
19. Hauben M MD, Gerrits CM, Walsh L, Van Puijenbroek EP. The role of data mining in pharmacovigilance. *Expert Opin Drug Saf* 2005;4(5):929-948.
20. Lamberts H, Wood M, Hofmans-Okkes IM. International primary care classifications: the effect of fifteen years of evolution. *Family practice.* Sep 1992;9(3):330-339.

21. Schuemie MJ, Sen E, t Jong GW, van Soest EM, Sturkenboom MC, Kors JA. Automating classification of free-text electronic health records for epidemiological studies. *Pharmacoepidemiology and drug safety.* Jan 24.

22. Thygesen SK, Christiansen CF, Christensen S, Lash TL, Sorensen HT. The predictive value of ICD-10 diagnostic coding used to assess Charlson comorbidity index conditions in the population-based Danish National Registry of Patients. *BMC medical research methodology.* 2011;11:83.

23. van Leerdam ME, Vreeburg EM, Rauws EA, et al. Acute upper GI bleeding: did anything change? Time trend analysis of incidence and outcome of acute upper GI bleeding between 1993/1994 and 2000. *The American journal of gastroenterology.* Jul 2003;98(7):1494-1499.

APPENDIX 1. Upper Gastro-Intestinal Bleeding (UGIB)

Questionnaire for Case Validation

1. Questionnaire for assessors

Based on the information reported in the medical definition, assessors should fill in the following questionnaire to be able to apply subsequently a validation algorithm:

Database: _____
ID Patient: _____
Gender: M/F
Birth date: dd/mm/yy
Date of event as reported in Jerboa output: dd/mm/yy
Age:_____years (calculated automatically with reference to the date of event)

Note: *a specific timeframe is not indicated* for *the search of the information, as reported in the following questionnaire. The assessors should consider all the information that is reasonably time-related to the occurrence of the event*

A) Information on characteristics and detection of UGIB

1. Was there any mention of gastrointestinal bleeding? Yes/No/Not assessable (NA)

2. Was there any specific mention of **upper** gastrointestinal bleeding (UGIB)? Yes/No/Not assessable (NA)

3. Was there any mention/evidence of melena/black stools/tarry stools? Yes/No

4. Was there any mention/evidence of hematemesis/blood vomiting? Yes/No

5. If the patient died: is there an explicit mention of UGIB as a cause of death? Yes/No/NA

6. If an endoscopy was performed, was bleeding detected? Yes/No/NA

7. Was it possible to identify the exact bleeding location? Yes/No

8. If yes, please specify:
    1. oral cavity
    2. esophagus
    3. stomach
    4. duodenum

5. jejunum
6. ileum
7. cecum
8. colon
9. rectum
10. anus
11. nasopharyngeal
12. pulmonary
13. pancreatico-biliary tract

9. Was there any endoscopic/clinical evidence of any the following gastrointestinal diseases potentially associated with UGIB? Yes/No/NA
   If yes, please mark the correct one(s):
   - ☐ Dyspepsia
   - ☐ Gastro-Esophageal Reflux Disease (GERD)
   - ☐ Peptic ulcer
   - ☐ Esophageal/Gastric/Duodenal ulcer
   - ☐ Esophageal/Gastric/Duodenal erosion
   - ☐ Acute gastritis/duodenitis/esophagitis

10. FOR GP DATABASES: Indicate if the UGIB diagnosis was made by:
- ☐ General Practitioner
- ☐ Specialist/Hospital discharge diagnosis

B) Information about potential alternative explanations for Upper GI bleeding

1. Was there any mention/evidence of any of the following diseases? Yes/No/NA;

If yes, please mark the correct one(s):
- ☐ black stools due to ingestion of iron/ferrous-containing medications
- ☐ Mallory-Weiss Syndrome or gastro-esophageal laceration syndrome
- ☐ Osler-Weber-Rendu Syndrome
- ☐ Angiodysplasia of the upper Gastrointestinal tract
- ☐ Dieulafoy lesion
- ☐ Iatrogenic Upper GI bleeding caused by procedure with endoscopy
- ☐ Varices bleeding
- ☐ Gastrointestinal cancer (esophageal, gastric, intestinal)

C) <u>In addition to this specific episode, was there any mention of a previous episode of UGIB?</u> Yes/No/NA

D) <u>ONLY FOR GP DATABASES:</u> **Was there sufficient information available for validation? Yes/No**

E) **Only for hospital databases**

Based on the available information, when did the bleeding (including melena or blood vomiting) start**?**

- ☐ During the hospital stay
- ☐ On the exact day of hospital admission
- ☐ Within the 7 days preceding the hospital admission
- ☐ Within one month preceding the hospital admission
- ☐ More than one month preceding the hospital admission
- ☐ Based on the available information, it is not possible to establish the exact date of the start of the bleeding (including melena or blood vomiting)

<u>For both hospital and GP databases</u>

F) **If it was clearly reported, please indicate the exact date of the start of bleeding (including melena or blood vomiting): DD/MM/YY**

Other comments:

_____

_____

_____

4.6

Accuracy of coding-based algorithms in identification of acute myocardial infarction in multi-country electronic healthcare record (EHR) databases

Preciosa M. Coloma, MD, MSc[1], Vera E Valkhoff, MD, MSc[1,2], Giampiero Mazzaglia, MD, PhD[3], Malene Schou Nielsson[4], Lars Pedersen, PhD[4], Mariam Molokhia, MD, PhD[5], Bartholemeus Mosseveld[1], Paolo Morabito[3], Martijn J Schuemie, PhD[1], Miriam Sturkenboom, PharmD, PhD[1,6], Gianluca Trifiro, MD, PhD [1,7] on behalf of the EU-ADR Consortium

[1]Department of Medical Informatics, Erasmus University Medical Centre, Rotterdam, the Netherlands; [2]Department of Gastroenterology and Hepatology, Erasmus University Medical Centre, Rotterdam, the Netherlands; [3]Health Search, Italian College of General Practitioners, Florence, Italy; [4]Department of Clinical Epidemiology, Aarhus University Hospital, Denmark; [5]Primary Care & Population Sciences, Kings College, London, UK; [6]Department of Epidemiology, Erasmus University Medical Centre, Rotterdam, the Netherlands; [7]Department of Clinical and Experimental Medicine and Pharmacology, University of Messina, Italy.

ABSTRACT

Background and Objective. Accuracy of event ascertainment is crucial to ensure validity when mining multiple electronic healthcare record (EHR) databases for drug safety signal detection. In this paper we aimed to evaluate and compare the accuracy of various coding algorithms used to identify cases of acute myocardial infarction (AMI) in three European EHR databases.

Methods. We conducted a validation study in three databases (DBs) of the EU-ADR network: (1) IPCI (general practitioner (GP) DB, Netherlands); (2) HSD (GP DB, Italy): and 3) Aarhus (claims DB, Denmark). We identified cases of AMI from GP medical records, primary hospital discharge diagnoses, and death registries using coding algorithms which included three different disease terminology schemes: (1) International Classification of Primary Care (ICPC); (2) International Classification of Diseases-9th revision-clinical modification (ICD9-CM); (3) ICD-10th revision. We also employed free text using key words consistent with AMI. A random sample of 200 cases per DB was obtained from all potential cases identified. Additional 200 potential cases identified by free text search were obtained in IPCI. Manual review of medical records and hospitalization charts was performed using a standardized questionnaire implemented as computerized data entry using custom-built software Chameleon,© locally installed in each database. Positive predictive values (PPV) were calculated overall and for each code and free text query.

Results. The study population for this analysis comprised healthcare data from 4 034 232 individuals with 22 428 883 person-years (PYs) of follow-up during the period 1995-2010. Within this population, a total of 42 774 potential cases of AMI were identified. From the random sample of 800 potential cases of AMI selected for validation, 748 records were retrieved (93.5%) and reviewed. All ICD-10 codes used (I21.0, I21.1, I21.2, I21.3, I21.4, and I21.9) had 100% PPV. Overall the ICD9-CM codes had very good PPV, with 410.9/410.90, the most frequently occurring code having a PPV of 96.5% (95%CI 93.5-100.4). The ICPC code K75 had a PPV of 75% (67.4-82.6). Use of free text had a lower PPV: 60% (95% CI 17.1-102.9) in HSD and 19.7% (95%CI 12.9 -26.5) in IPCI.

Conclusion. The results obtained in this study are consistent with the PPV estimates for ICD-9CM and ICD-10 cited in the literature. Strategies are necessary to further optimize the value of free text search in the identification of AMI in EHR databases.

INTRODUCTION

Clinical and drug use information derived from electronic healthcare record (EHR) databases are increasingly being used for pro-active drug safety surveillance.[1-3] Drug-induced acute myocardial infarction (AMI) is an important safety consideration, as shown by the rofecoxib and rosiglitazone stories.[4-7] AMI may be identified using population-based EHR databases, which provide data on how drugs are being used in real-world settings. As in any association study, however, accurate case definitions for AMI are essential to the proper evaluation of increased risk for such an event in association with use of certain drugs.[8] Innovations in recent years have brought about discovery, and subsequent clinical use, of novel cardiac biomarkers that allow earlier recognition of disease as well as therapeutic interventions that reduce the extent of myocardial injury and mortality. Such developments have led to revisions in the definition of AMI and consequently, changes in the diagnosis and prognosis.[9-11] Studies that estimate AMI incidence from EHR data must also then consider the implications of new diagnostic criteria on disease coding practices in EHR databases.[12-14]

The accuracy of AMI identification in EHR using specific disease coding schemes has been evaluated in several studies. Most of these studies have been performed in administrative claims databases and have calculated the positive predictive value (PPV) of International Classification of Diseases – 9th revision – Clinical Modifications (ICD-9-CM) codes as well as diagnosis- related groups (DRG) codes.[15-18] A recent study evaluated the PPV of ICD-10th revision diagnostic codes used to assess Charlson comorbidity index conditions, including myocardial infarction, in the population-based Danish National Registry of Patients.[19] To date there is no study that evaluates the validity of International Classification of Primary Care (ICPC) codes and free text search, or the combination of diagnosis codes and free text search, in the identification of AMI in EHR data.

Within the context of the EU-ADR project, we conducted a validation study to evaluate and compare the accuracy of coding-based search algorithms from three different terminologies (International Classification of Primary Care (ICPC), ICD-9CM, and ICD-10) used to identify cases of AMI in population-based EHR databases from Denmark, Italy, and the Netherlands. We further assessed the accuracy of free text search using key words in the identification of AMI.

METHODS

Data sources

The EU-ADR network currently comprises anonymized clinical and drug prescription/ dispensing information of about 20 million individuals from eight population-based electronic healthcare databases in four European countries. The data are pooled using a distributed network approach that allows data holders to maintain control over their protected data. Validation of AMI identification was performed in three of these databases: (1) Integrated Primary Care Information (IPCI, Netherlands); (2) Health Search/CSD Patient DB (HSD, Italy); and (3) Aarhus University

Hospital Database (Aarhus, Denmark). IPCI and HSD are both general practice (GP) databases documenting patient consults, including referrals for hospitalization or specialist care as well as prescriptions for medications. Aarhus is a comprehensive record-linkage database system in which drug dispensing data are linked to a registry of hospital discharge diagnoses and various other registries. A more detailed description of the characteristics of the databases has been previously published.[3,20] The three databases employ different disease coding terminologies: IPCI uses the International Classification for Primary Care (ICPC); HSD uses the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM); and Aarhus uses the ICD 10th revision (ICD10). Clinical notes from GP records in both HSD and IPCI also use free text narratives which can be used to identify medical events. All databases have considerable experience in using patient data for research purposes.[21-24]

Identification of acute myocardial infarction

Potential cases of AMI were identified using harmonized and database-specific coding algorithms which included the ICPC code K75 (IPCI), ICD9-CM codes 410/410.x/ 410.x0 (HSD), and the ICD-10 codes I21.x (Aarhus). IPCI and HSD also performed free text search using specific key words. The free text search strings employed in IPCI and HSD are given in Appendix 1. The process of mapping and the harmonization of event data extraction from different EHR databases in the EU-ADR project is based on medical concepts derived from the Unified Medical Language System (UMLS).[25-26]

Case validation

A random sample of 200 potential AMI cases was selected in each of the three databases. Additional 200 potential cases identified by free text search were obtained in IPCI. Manual review of GP records and hospitalization charts was performed by medically trained assessors using a standardized questionnaire that was pilot-tested in the databases and reviewed by a panel of experts. Diagnostic criteria for AMI as prescribed in current guidelines[8-9] were incorporated in the questionnaire, as well as information regarding cardiovascular risk factors and potential alternative diagnoses that could explain findings suggestive of AMI. For GP databases (IPCI and HSD) it was also determined whether the AMI diagnosis was made by a GP or medical specialist. The standardized questionnaire was then implemented as computerized data entry algorithm using custom-built software Chameleon©. The software was installed locally in each database, allowing the data holders to keep patient-level data within their protected environment. The data entry algorithm is shown in Figure 1 and the questionnaire shown in Appendix 2. Based on the information collected in the questionnaire, the potential AMI cases were classified as: (1) definite case; (2) non-case; or (3) non-assessable case, if the available information was deemed insufficient to validate the case.

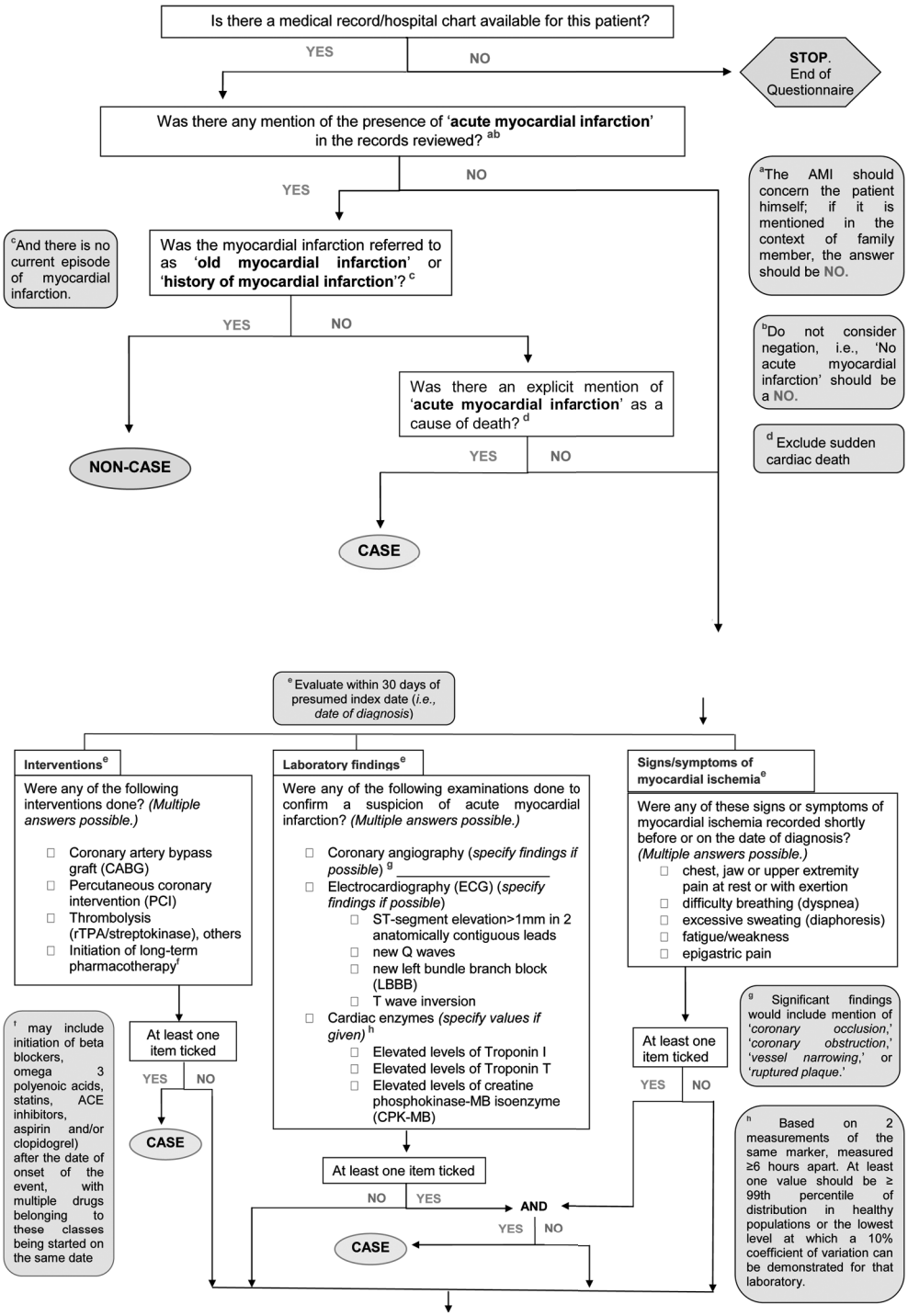
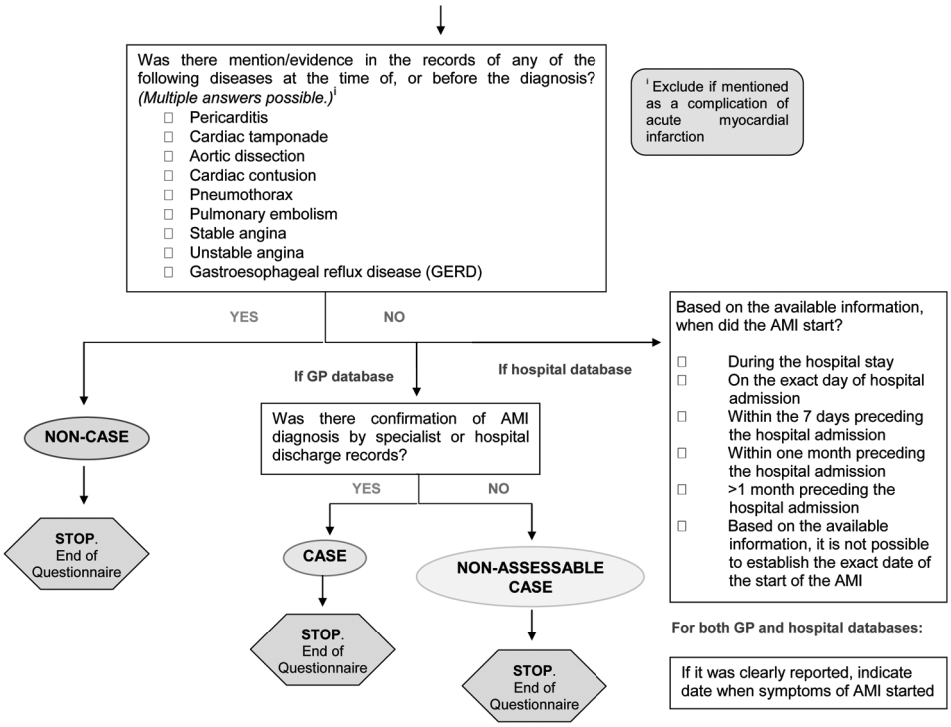


Figure 1 | Data entry algorithm implemented based on a standardized questionnaire.

(Continued on next page)

(Continued from previous page)



Consistency of index date

We determined how the coded date of the event (which is detected automatically, but may differ among databases) was related to the actual date of diagnosis of AMI as well as to the date of onset of first symptoms. In addition, for the claims database Aarhus we evaluated the coded date with respect to the day of hospital admission.

Statistical analysis

Positive predictive values (PPV) and corresponding 95% confidence intervals (95%CI) were calculated overall in each database and for each code or free text query, using the medical charts as gold standard. Potential cases judged as non-assessable were not included in the PPV calculation.

RESULTS

The three EU-ADR databases considered for this analysis comprised healthcare data from 4 034 232 individuals with 22 428 883 person-years (PYs) of follow-up during the period 1995-2010. Within this population, a total of 42 774 potential cases of AMI were identified. From the random sample of 800 potential cases of AMI (200 cases per database plus additional 200 cases for free text-identified cases in IPCI) selected for validation, 748 records were retrieved (93.5%) and reviewed. The hospital medical charts of 52 potential cases in Aarhus could not be accessed. The demographic and clinical characteristics of the potential cases from the random sample are shown in Table 1. The mean age was 67 years across all the databases and the patients were predominantly male (62-70% overall). Chest pain at rest or with exertion was the most frequently reported symptom of AMI cases across all the databases (more than 50% of confirmed cases in both IPCI and Aarhus and 11% in HSD). Diabetes, hypertension, and dyslipidemia were the most frequently recorded cardiovascular risk factors.

All 148 potential cases of AMI identified in Aarhus were confirmed by manual chart review. 93 (46.5%) out of the total 200 potential ICPC-coded cases were confirmed in IPCI, 31(15.5%) were judged as non-cases, and 76 (38%) judged as non-assessable. From the 200 potential cases identified by free text search, 26 (13%) cases were confirmed and 68 (34%) considered non-assessable from the data available, while the remaining (106, 53%) were non-cases. For HSD, 115 (57.5%) cases were confirmed and 79 (39.5%) declared non-assessable. Table 2 shows the overall PPV for the coding algorithms used for AMI identification in each database. In Table 3, the percentage distribution and PPV for the specific diagnosis codes from each terminology and free text search are given. All of the ICD-10 codes used in Aarhus had 100% PPV. Overall the ICD9-CM codes had very good PPV, with 410.9/410.90, the most frequently reported code, having a PPV of 96.5 (95%CI: 93.5 to 100.0). The PPV of free text search alone in HSD was 60% (95%CI 17.1-100.0). In the IPCI database, the ICPC code K75 had a PPV of 75 (95%CI 67.4 - 82.6), while free text search alone had a PPV of 19.7 (95%CI 12.9-26.5). All validated cases of AMI in IPCI and HSD were supported with diagnosis of a medical specialist.

The relationship between the coded date and the date of onset of symptoms across the three databases is shown in Figure 2. There was not enough information to enable inference on such relationship for 25 of the cases in Aarhus (16.9%). The date of symptom onset ranged from 1 day before to more than 60 days before the automatically detected event date. The coded date for the majority of cases coincided with the onset of symptoms in all database: Aarhus=72 cases (48.6%); HSD=110 cases (95.6%); and IPCI= 67 cases (56.3%). In Aarhus there were three instances where the coded date for AMI actually preceded the onset of symptoms by two days and there was one case each in Aarhus and IPCI where the coded date preceded symptoms by one day.

For the claims database Aarhus, the characterization of the coded index date with respect to hospitalization is as follows: (1) date of hospital admission= 100 cases (67.6%); (2) during hospital stay= 9 cases (6.1%); (3) ≥7 days preceding hospitalization= 23 (15.5%); (4) not possible to establish= 16 (10.8%).

Table 1 | Characteristics of patients in the random sample of potential AMI cases.

	IPCI		HSD		Aarhus	
	Total N: 400 (%)	Confirmed cases N:93 (%)	Total N: 200(%)	Confirmed cases N:115 (%)	Total N: 148 (%)	Confirmed cases N:148 (%)
Male sex (%)	246 (61.5)	87 (93.5)	132 (66.0)	80 (69.6)	103 (69.6)	82 (55.4)
Mean age (years)	66	65	68	67	67	67
Cardiovascular risk factors						
Family history of coronary heart disease	86 (21.5)	14 (15.1)	11 (5.5)	8 (7.0)	19 (12.8)	19 (12.8)
Dyslipidemia	73 (18.2)	19 (20.4)	98 (49.0)	58 (50.4)	19 (12.8)	19 (12.8)
Diabetes mellitus	62 (15.5)	13 (14.0)	66 (33.0)	42 (36.5)	17 (11.4)	17 (11.4)
Hypertension	126 (31.5)	36 (38.7)	113 (56.5)	76 (66.0)	44 (29.7)	44 (29.7)
Obesity	79 (19.8)	23 (24.7)	21 (10.5)	16 (13.9)	7 (4.7)	7 (4.7)
Cigarette smoking	111 (27.8)	36 (38.7)	39 (19.5)	22 (19.1)	50 (33.8)	50 (33.8)
Clinical manifestations						
chest, jaw or upper extremity pain at rest or with exertion	102 (25.5)	52 (55.9)	23 (11.5)	13 (11.3)	118 (79.7)	118 (79.7)
difficulty breathing (dyspnea)	38 (9.5)	15 (20.4)	10 (5.0)	6 (5.2)	22 (14.8)	22 (14.8)
excessive sweating (diaphoresis)	30 (7.5)	19 (20.4)	2 (1.0)	0 (0)	5 (3.4)	5 (3.4)
fatigue/weakness	7 (1.8)	0 (0)	2 (1.0)	1 (0.9)	5 (3.4)	5 (3.4)
Diagnostic workup done						
Coronary angiography	48 (12.0)	39 (41.9)	39 (19.5)	38 (33.0)	117 (79.1)	117 (79.1)
Electrocardiography						
ST-segment elevation	27 (6.7)	25 (30.1)	7 (3.5)	5 (4.4)	59 (39.9)	59 (39.9)
new Q waves	7 (1.8)	4 (4.3)	1 (0.50)	1 (0.9)	9 (6.1)	9 (6.1)
new left bundle branch block (LBBB)	3 (0.8)	0 (0)	0 (0)	0 (0)	2 (1.4)	2 (1.4)
T wave inversion	17 (4.2)	9 (9.7)	8 (4.0)	6 (5.2)	19 (12.8)	19 (12.8)
Other (ST segment depression, etc.)	19 (7.2)	19 (20.4)	7 (3.5)	7 (6.1)	9 (6.1)	9 (6.1)
Cardiac enzymes						
Elevated cardiac troponin I	5 (1.2)	41 (44.1)	2 (1.0)	2 (1.7)	0 (0)	0 (0)
Elevated cardiac troponin T	6 (1.5)	6 (9.7)	0 (0)	0 (0)	123 (83.1)	123 (83.1)
Elevated creatine phosphokinase (MB isoenzyme)	23 (5.8)	23 (24.7)	5 (2.5)	5 (4.4)	86 (58.1)	86 (58.1)
Other	8 (2.0)	5 (5.4)	3 (1.5)	3 (2.6)	5 (3.4)	5 (3.4)
Interventions done						
Coronary artery bypass graft (CABG)	15 (3.8)	10 (10.8)	17 (8.5)	17 (14.8)	7 (4.7)	7 (4.7)
Percutaneous coronary intervention (PCI)	78 (19.5)	67 (72.0)	41 (20.5)	41 (35.6)	93 (62.8)	93 (62.8)

	IPCI		HSD		Aarhus	
	Total N: 400 (%)	Confirmed cases N:93 (%)	Total N: 200(%)	Confirmed cases N:115 (%)	Total N: 148 (%)	Confirmed cases N:148 (%)
Thrombolysis (rTPA/ streptokinase, others)	6 (1.5)	5 (5.4)	4 (2.0)	4 (3.5)	9 (6.1)	9 (6.1)
Initiation of long-term pharmacotherapy	116 (29.0)	93 (100)	121 (60.5)	90 (77.6)	81 (56.8)	81 (56.8)
Deaths	7 (1.8)	7 (7.5)	2 (1.0)	2 (1.7)	2 (1.4)	2 (1.4)
Diagnosis confirmed by medical specialist**	113 (28.2)	113 (100)	1 (0.5)	1 (0.9)	NA	NA

* Can add up to more than 100% ** Only relevant for GP databases (HSD and IPCI)

Table 2 | PPV of EHR-based definitions for acute myocardial infarction (AMI) in three databases from four countries.

Source	Coding system	No. of cases sampled	No of cases retrieved	No. of cases confirmed	No. of cases considered non-assessable	Overall PPV (95%CI)
General practitioner (GP)/specialist diagnoses (IPCI, Netherlands)	ICPC, free text	400	400	119	144	46.5 (40.4 -52.6)
GP/specialist diagnoses (HSD, Italy)	ICD9-CM, free text	200	200	115	79	95 (91.2 - 98.9)
Primary hospital discharge diagnoses (Aarhus, Denmark)	ICD-10	200	148	148	0	100 (100 - 100)

DISCUSSION

Proactive surveillance of potentially drug-induced outcomes using multi-country EHR databases is a promising avenue to overcome some of the limitations of traditional spontaneous reporting systems. The validity of such surveillance activities depends, however, on the accuracy of the definitions of the adverse events being investigated. Therefore, it is important to validate cases identified from hospital discharge diagnoses in claims databases as well as GP-recorded diagnoses.

We examined the positive predictive value (PPV) of primary hospital discharge diagnosis codes and GP-recorded diagnoses for AMI in three European EHR databases (Denmark, Italy, and the Netherlands). The overall PPV for the coding scheme-based diagnoses was good, ranging from 75% (IPCI, ICPC coding) to 95% (HSD, ICD9-CM) to 100% (Aarhus, ICD10). The use of free text search was more extensive in IPCI compared to HSD, largely due to the lesser granularity of the ICPC coding system. The use of free text alone had a low accuracy, ranging from 19.7% in IPCI to 60% in HSD.

Table 3 | Number and distribution of confirmed AMI cases by diagnostic code or free text

Database/Code	Code description	No. of records reviewed	No. of cases confirmed	% of all potential cases in sample	PPV (95%CI)
IPCI					
ICPC K75	Acute myocardial infarction	200	93	100	75.0 (67.4 - 82.6)
Free text	Specific key words*	200	26	100	19.7 (12.9 - 26.5)
HSD					
ICD9-CM					
410 or 410.0	Acute myocardial infarction (AMI) of anterolateral wall	12	6	6.0	85.7 (59.8-111.6)
410.1 or 410.10	Acute MI of other anterior wall	4	4	2.0	100 (100 – 100)
410.20	Acute MI of inferolateral wall	1	1	0.5	100 (100 – 100)
410.3	Acute MI of inferoposterior wall	1	1	0.5	100 (100 – 100)
410.7	Subendocardial infarction	3	2	1.5	100 (100 – 100)
410.9 or 410.90	Acute MI, unspecified site	157	94	78.5	96.9 (93.5 - 100.4)
410.9+Free text	Acute MI, unspecified site	8	4	4.0	100 (100 – 100)
411.81+Free text	Acute coronary occlusion without MI	1	0	0.5	0
Free text	Specific key words*	13	3	6.5	60 (17.1 - 102.9)
Aarhus					
ICD-10					
I21.0	Acute transmural MI of anterior wall	20	20	13.5	100 (100 – 100)
I21.1	Acute transmural MI of inferior wall	17	17	11.5	100 (100 – 100)
I21.2	Acute transmural MI of other sites	2	2	1.4	100 (100 – 100)
I21.3	Acute transmural MI of unspecified site	26	26	17.6	100 (100 – 100)
I21.4	Acute subendocardial MI	56	56	37.8	100 (100 – 100)
I21.9	AMI, unspecified	27	27	18.2	100 (100 – 100)

*see Appendix 1 for key words used

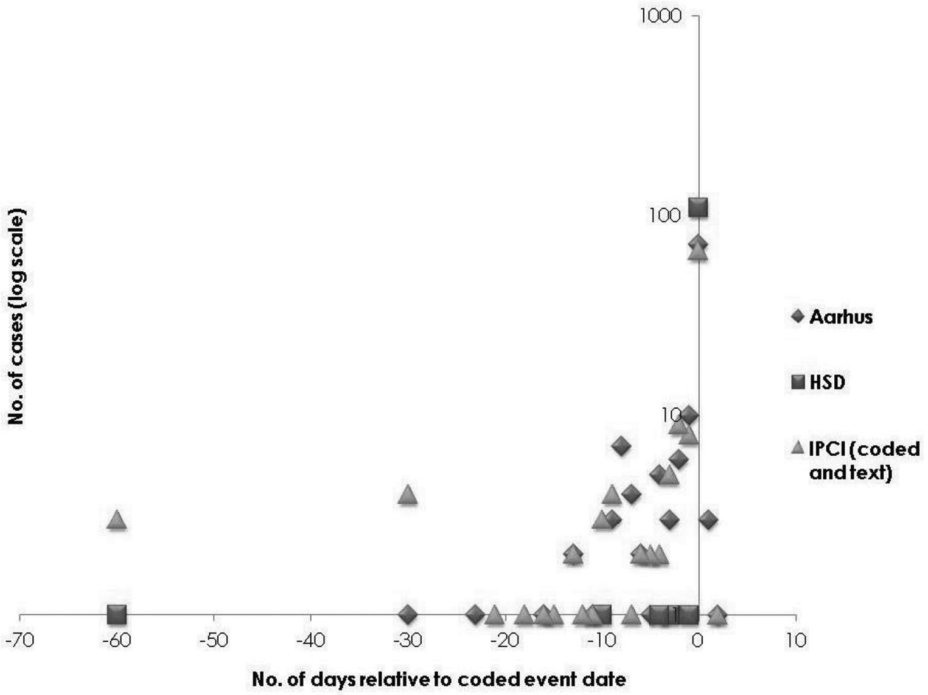


Figure 2 | Differences in the coded event date and date of onset of symptoms across the three databases ***(see page 311 for color figure).***

The results we obtained in this study are consistent with the PPV estimates for ICD-10 and for ICD-9CM cited in the literature. The ICD10 codes I21, I22, and I23 were found to have 98% PPV in a Danish study evaluating the accuracy of ICD10-coded myocardial infarction as a component of the Charlson comorbidity index.[19] Previous studies evaluating earlier versions of ICD have also demonstrated the accurate coding practices in Danish administrative registries, including the Danish MONICA (Monitoring Trends and Determinants in Cardiovascular Disease) study where 93.5% of the patients in the Danish National Patient Register were found to have definite or possible AMI.[27] The PPV of the ICD9-CM code 410 to identify cases of AMI among records with a prior primary hospital discharge code in the Saskatchewan Hospital Automated Database was 97%.[16] In another study using Medicare claims, the PPV of several ICD9-CM codes (410.01, 410.11, 410.21, 410.31, 410.41, 410.51, 410.61, 410.71, 410.81, or 410.91) for identifying AMI in either primary or secondary hospital discharge diagnoses was 94.1%.

While ICPC codes are often used to estimate incidence or prevalence of various clinical outcomes,[28-30] we are not aware of any published studies that have assessed the accuracy of ICPC codes in the identification of AMI in EHR data. A study in the Netherlands evaluated ICPC-coded

diagnoses in GP records in the context of cardiovascular risk factor assessment after pre-eclampsia, but only the validity of ICPC-coded pre-eclampsia diagnosis was determined.[31]

The available knowledge regarding the value of free-text mining in identifying outcomes from EHR data is still limited, but this is an area of research that is gaining a lot of interest.[32-33] Our findings show that there is potential for the use of free text search in identification of AMI from EHR databases, but that appropriate combination of key words needs to be further evaluated and optimized.

The use of a standardized questionnaire implemented in an automated data entry validation algorithm facilitated harmonized data collection and analysis across the different databases without compromising data protection. The procedure also enabled us to document recorded database information on cardiovascular risk factors such as family history of coronary artery disease, hypertension, diabetes, dyslipidemia, smoking, and obesity. Such information may be useful in evaluating the potential confounder effects when performing signal detection activities. Our evaluation of the automatically detected index date shows that most of the time the coded event date coincided with the date of onset of first symptoms (and with date of hospital admission for claims database) although there can be a wide range between these two dates. The possible implications of this finding on the exposure definitions used in the estimation of risk for AMI during drug use needs to be further investigated.

CONCLUSION

We have shown that multi-country EHR databases with different disease coding schemes can accurately identify patients with acute myocardial infarction. The results obtained in this study are consistent with the PPV estimates for ICD-9CM and ICD-10 cited in the literature. Strategies are necessary to optimize the use of free text search in the identification of AMI in EHR databases.

APPENDIX 1 | Free text search strings employed in the identification of cases of acute myocardial infarction in General Practitioner (GP) Databases

IPCI (Dutch)	HSD (Italian)
" (myocard AND infarct) OR (hart AND infar) OR (latera AND infar) OR (posterior AND infar) OR (anterior AND infar) OR (septal AND infar) OR (hart AND aanval) OR (Q- AND wave) OR ('ST' AND verhog) OR (wand AND infar) OR (parde AND wave) OR 'VWI' OR 'AMI'"	*"infarto miocardico acuto OR infarto%miocard%acut OR necrosi%miocard%acut OR infarto%lateral OR infarto%infer% OR %infarto%transmur OR infarto%anter OR infarto%poster OR infarto%subendocard"*

APPENDIX 2 | Standardized questionnaire for the validation of potential cases of acute myocardial infarction identified from EHR databases

Questionnaire for assessors

Based on the information reported in the medical definition, assessors should answer the following questions to be able to apply subsequently a validation algorithm:

Database: _____

ID Patient: _____

Sex: M/F

Birthdate: dd/mm/yyyy

Date of event as reported in the automated search: dd/mm/yyyy

Age:_____years (calculated automatically with reference to the date of event)

Is there a medical record available for this patient? Yes/No

A) Information on characteristics and detection of AMI

1. Was there any mention of the presence of '**acute myocardial infarction**' in the records reviewed? Yes/No
2. If (*Answer to 1 is*) YES, was the myocardial infarction referred to as '**old myocardial infarction**' or '**history of myocardial infarction**'? Yes/No
3. Was there an explicit mention of '**acute myocardial infarction**' as a cause of death? Yes/No

For Questions 4-6, evaluate within 30 days of presumed index date [i.e., date of diagnosis]

4. Were any of the following interventions done? Multiple answers are possible:
   - ☐ Coronary artery bypass graft (CABG)
   - ☐ Percutaneous coronary intervention (PCI)
   - ☐ Thrombolysis (rTPA/streptokinase, others)
   - ☐ Initiation of long-term pharmacotherapy
   - ☐ None of the above

5. Were any of the following examinations done to confirm a suspicion of acute myocardial infarction? Multiple answers are possible:
   - ☐ Coronary angiography (specify findings if possible)
     - ☐ coronary occlusion
     - ☐ coronary obstruction
     - ☐ vessel narrowing
     - ☐ ruptured plaque
     - ☐ other, please specify _____
   - ☐ Electrocardiography (ECG) (specify findings if possible)
     - ☐ ST-segment elevation>1mm in 2 anatomically contiguous leads
     - ☐ new Q waves
     - ☐ new left bundle branch block (LBBB)
     - ☐ T wave inversion
     - ☐ Other, please specify _____
   - ☐ Cardiac enzymes (specify values and units, if given)
     - ☐ Elevated levels of Troponin I_____
     - ☐ Elevated levels of Troponin T_____
     - ☐ Elevated levels of creatine phosphokinase-MB
     - ☐ isoenzyme (CPK-MB) _____
     - ☐ Other, specify if possible_____
   - ☐ None of the above

6. Were any of these signs or symptoms of myocardial ischemia recorded shortly on or before the date of diagnosis? Multiple answers are possible:
   - ☐ chest, jaw or upper extremity pain at rest or with exertion
   - ☐ difficulty breathing (dyspnea)
   - ☐ excessive sweating (diaphoresis)
   - ☐ fatigue/weakness
   - ☐ epigastric pain
   - ☐ none of the above
   - ☐ other, please specify _____

B) Information about cardiovascular risk factors

1. Was there mention/evidence in the records of any of the following risk factors for acute myocardial infarction? Multiple answers are possible.
   - ☐ Family history of myocardial infarction/cardiovascular disease
   - ☐ Dyslipidemia
   - ☐ Diabetes mellitus
   - ☐ Hypertension
   - ☐ Obesity
   - ☐ Cigarette smoking
   - ☐ None of the above

C) Information about potential alternative explanations for the signs/symptoms/ laboratory findings

1. Was there mention/evidence in the records of any of the following diseases at the time of/before the diagnosis? Multiple answers are possible.
   - ☐ Pericarditis and/or Cardiac Tamponade _____
   - ☐ Myocarditis _____
   - ☐ Aortic dissection _____
   - ☐ Cardiac contusion _____
   - ☐ Pneumothorax _____
   - ☐ Pulmonary embolism _____
   - ☐ Stable angina _____
   - ☐ Unstable angina _____
   - ☐ Gastroesophageal reflux disease (GERD) _____
   - ☐ None of the above

Only for GP databases: Indicate if there was a confirmation of the diagnosis of acute myocardial infarction by:

- ☐ Specialist
- ☐ Hospital discharge records

Was there sufficient information available for validation? Yes/No

D) Date of the event:

For both hospital and GP databases: If it was clearly reported, indicate the date when the symptoms of AMI started (index date)

_____

Only for hospital databases:

Based on the available information, when did the AMI start**?**

- ☐ During the hospital stay
- ☐ On the exact day of hospital admission
- ☐ Within the 7 days preceding the hospital admission
- ☐ Within one month preceding the hospital admission
- ☐ More than one month preceding the hospital admission
- ☐ Based on the available information, it is not possible to establish the exact date of the start of the AMI

E) In your opinion, was this a:

- ☐ CASE?
- ☐ NON-CASE?
- ☐ NON-ASSESSABLE CASE?

Other comments

_____

_____

REFERENCES

1. 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.* Nov 2 2010;153(9):600-606.
2. 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.* Aug 13 2009;361(7):645-647.
3. Coloma PM, Schuemie M, Trifiró G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom MCJM, 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 Jan;20(1):1-11.
4. Bresalier RS, Sandler R, Quan H, Bolognese JA, Oxenius B, Horgan K, Lines C, Riddell R, Morton D, Lanas A, Konstam MA, Baron JA; Adenomatous Polyp Prevention on Vioxx (APPROVe) Trial Investigators. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005 Mar 17;352(11):1092-1102.
5. Topol E. Failing the public health--rofecoxib, Merck, and the FDA. *N Engl J Med.* 2004 Oct 21;351(17):1707-1709.
6. Loke YK, Kwok CS, Singh S. Comparative cardiovascular effects of thiazolidinediones: a systematic review and meta-analysis of observational studies. *BMJ.* 2011 Mar 17;342.
7. Trelle S RS, Wandel S, Hildebrand P, Tschannen B, Villiger PM, Egger M, Juni P. Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. *BMJ* Jan 11;342:c7086. doi: 10.1136/bmj.c7086.
8. Luepker RV, Apple FS, Christenson RH, et al. Case definitions for acute coronary heart disease in epidemiology and clinical research studies: a statement from the AHA Council on Epidemiology and Prevention; AHA Statistics Committee; World Heart Federation Council on Epidemiology and Prevention; the European Society of Cardiology Working Group on Epidemiology and Prevention; Centers for Disease Control and Prevention; and the National Heart, Lung, and Blood Institute. *Circulation.* Nov 18 2003;108(20):2543-2549.
9. Thygesen K, Alpert JS, White HD, et al. Universal definition of myocardial infarction. *Circulation.* Nov 27 2007;116(22):2634-2653.
10. Alpert JS, Thygesen K, White HD, Jaffe AS. Implications of the universal definition of myocardial infarction. *Nat Clin Pract Cardiovasc Med.* Nov 2008;5(11):678-679.
11. Malasky BR, Alpert JS. Diagnosis of myocardial injury by biochemical markers: problems and promises. *Cardiol Rev.* Sep-Oct 2002;10(5):306-317.
12. Graven T, Kruger O, Bronstad G. Epidemiological consequences of introducing new biochemical markers for detection of acute myocardial infarction. *Scand Cardiovasc J.* Sep 2001;35(4):233-237.
13. Kavsak PA, MacRae AR, Lustig V, et al. The impact of the ESC/ACC redefinition of myocardial infarction and new sensitive troponin assays on the frequency of acute myocardial infarction. *Am Heart J.* Jul 2006;152(1):118-125.
14. Sanfilippo FM, Hobbs MS, Knuiman MW, Hung J. Impact of new biomarkers of myocardial damage on trends in myocardial infarction hospital admission rates from population-based administrative data. *Am J Epidemiol.* Jul 15 2008;168(2):225-233.
15. Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. *Am Heart J.* Jul 2004;148(1):99-104.
16. Varas-Lorenzo C, Castellsague J, Stang MR, Tomas L, Aguado J, Perez-Gutthann S. Positive predictive value of ICD-9 codes 410 and 411 in the identification of cases of acute coronary syndromes in the Saskatchewan Hospital automated database. *Pharmacoepidemiol Drug Saf.* Aug 2008;17(8):842-852.
17. Pladevall M, Goff DC, Nichaman MZ, et al. An assessment of the validity of ICD Code 410 to identify hospital admissions for myocardial infarction: The Corpus Christi Heart Project. *Int J Epidemiol.* Oct 1996;25(5):948-952.
18. Petersen LA, Wright S, Normand SL, Daley J. Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database. *J Gen Intern Med.* Sep 1999;14(9):555-558.

19. Thygesen SK, Christiansen CF, Christensen S, Lash TL, Sorensen HT. The predictive value of ICD-10 diagnostic coding used to assess Charlson comorbidity index conditions in the population-based Danish National Registry of Patients. *BMC Med Res Methodol.* 2011;11:83.

20. Coloma PM, Trifiro G, Schuemie MJ, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? *Pharmacoepidemiol Drug Saf.* Feb 8 2012.

21. Vlug AE, van der Lei J, Mosseveld BM, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf Med.* 1999;38:339–344.

22. Verhamme K, Mosis G, Dieleman J, Stricker B, Sturkenboom M. Spironolactone and risk of upper gastrointestinal events: population based case-control study. *BMJ.* 2006;333(7563).

23. Christensen S, Riis A, Nørgaard M, Thomsen W, Sørensen HT. Introduction of newer selective cyclooxygenase-2 inhibitors and rates of hospitalization with bleeding and perforated peptic ulcer. *Aliment Pharmacol Ther* 2007;25:907–912.

24. Sacchetti E, Trifiro G, Caputi A, et al. Risk of stroke with typical and atypical antipsychotics: a retrospective cohort study including unexposed subjects. *J Psychopharmacol.* 2008;22(1):39-46.

25. Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med.* Aug 1993;32(4):281-291.

26. Avillach P, Joubert M, Thiessard F, et al. Design and evaluation of 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.* 2010;160(Pt 2):1085-1089.

27. Madsen M, Davidsen M, Rasmussen S, Abildstrom SZ, Osler M. The validity of the diagnosis of acute myocardial infarction in routine statistics: a comparison of mortality and hospital discharge data with the Danish MONICA registry. *J Clin Epidemiol.* Feb 2003;56(2):124-130.

28. Hak E, Rovers MM, Kuyvenhoven MM, Schellevis FG, Verheij TJ. Incidence of GP-diagnosed respiratory tract infections according to age, gender and high-risk co-morbidity: the Second Dutch National Survey of General Practice. *Fam Pract.* Jun 2006;23(3):291-294.

29. Britt H, Meza RA, Del Mar C. Methodology of morbidity and treatment data collection in general practice in Australia: a comparison of two methods. *Fam Pract.* Oct 1996;13(5):462-467.

30. Esteban-Vasallo MD, Dominguez-Berjon MF, Astray-Mochales J, et al. Epidemiological usefulness of population-based electronic clinical records in primary care: estimation of the prevalence of chronic diseases. *Fam Pract.* Dec 2009;26(6):445-454.

31. Nijdam ME, Timmerman MR, Franx A, et al. Cardiovascular risk factor assessment after pre-eclampsia in primary care. *BMC Fam Pract.* 2009;10:77.

32. Warrer P, Hansen EH, Juhl-Jensen L, Aagaard L. Using text-mining techniques in electronic patient records to identify ADRs from medicine use. *Br J Clin Pharmacol.* 2011 Nov 29;Epub ahead of print

33. Schuemie MJ, Sen E, t Jong GW, van Soest EM, Sturkenboom MC, Kors JA. Automating classification of free-text electronic health records for epidemiological studies. *Pharmacoepidemiol Drug Saf.* Jan 24 2012.

4.7

Epidemiology of Progressive Multifocal Leukoencephalopathy: population-based study from three European countries

Preciosa M Coloma, MD, MSc[1], Gianluca Trifirò, MD, PhD[1,2], Martijn J Schuemie, PhD[1], Rosa Gini, PhD[3], Ron Herings, PhD[4], Giampiero Mazzaglia, MD[5], PhD, Lorenza Scotti, MSc[6], Lars Pedersen, PhD [7], Sabine Straus, MD, PhD[1,8], Johan van der Lei, MD, PhD[1], Miriam Sturkenboom, PharmD, PhD[1,9] on behalf of the EU-ADR consortium

[1]Department of Medical Informatics, Erasmus Medical Center, Rotterdam, the Netherlands; [2]Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Messina, Italy; [3]Agenzia Regionale di Sanità della Toscana, Florence, Italy; [4]PHARMO Institute, Utrecht, the Netherlands; [5]Società Italiana di Medicina Generale, Florence, Italy; [6]Department of Statistics, Università di Milano-Bicocca, Milan, Italy; [7]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; [8]Dutch Medicines Evaluation Board, The Hague, the Netherlands, and [9]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands

ABSTRACT

Background. Progressive multifocal leukoencephalopathy (PML) is a rare central nervous system demyelinating disease resulting from reactivation of the JC virus. PML occurs almost exclusively in immunosuppressed individuals, including patients with HIV/AIDS, leukemia, tumors, or those receiving immunosuppressive therapy such as patients undergoing organ transplants. The disease has recently generated widespread concern after reports of PML allegedly developing after treatment with several new biologic agents. Because of its rarity, few data are available regarding incidence of PML in non-immunocompromised populations.

Objective. To provide estimates of the incidence of PML in the general population using electronic healthcare data from three countries in Europe.

Methods. Demographic and clinical data were obtained from six population-based electronic healthcare record (EHR) databases of three European countries within the EU-ADR network (Denmark, Italy, and the Netherlands). Data were analyzed for the years 1995–2011. Cases of PML were identified from primary and secondary hospital discharge diagnoses and death registries (claims databases) and from general practitioner or specialist diagnoses (medical records databases) using diagnostic codes and free text. Case validation by medical chart/records review was performed in a subset of cases.

Results. A total of 1 150 cases of PML were identified in the database network representing 21 171 291 individuals with 154 474 063 person-years (PYs) of follow-up. About half of the cases were males (603, 52.4%). The largest proportion of male cases were in the age group 30-50 years old (267 cases, 23.2% of total), while the largest proportion of female cases came from those 70 years and older (330 cases, 28.7% of total). The age-standardized rate across the databases ranged from 0.03 – 2.29 per 100 000 PYs. The overall incidence rate (IR) was 0.75 per 100 000 PYs.

Conclusion. PML is a rare disease among the general European population and its incidence is largely associated with HIV disease, malignancy, and rheumatologic conditions. An observed increase in the incidence of PML in the elderly requires further investigation. Electronic healthcare data from multiple countries may be able to provide estimates of background rates of diseases such as PML in the general population.

INTRODUCTION

Progressive multifocal leukoencephalopathy (PML) is a rare central nervous system demyelinating disease resulting from reactivation of the JC virus. Healthy individuals may harbor the virus and seroprevalence increases with age, although acquisition of the virus per se is not associated with a clinical syndrome.[1-2] Studies indicate that 50-90% of adults have been exposed to the JC virus, with 19-27% of these people shedding the virus in their urine.[3-4] A diagnosis of PML is usually made on the basis of neurological features at presentation, characteristic brain imaging changes and the presence of JC virus DNA in cerebrospinal fluid.[5] The understanding of the epidemiology of PML has evolved and paralleled key developments in medicine: the Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome (HIV/AIDS) epidemic; the introduction of highly active antiretroviral therapy (HAART); and the advancement in novel biologic therapies. Before the emergence of HIV, PML developed mostly in patients with lymphoma, other malignancies, and rare forms of immunosuppression. Now, majority of PML cases are reported in patients with HIV and PML is considered an AIDS-defining illness, affecting about 5% of HIV-infected patients.[6-8] There is no specific treatment for PML and although HIV-positive patients with PML are now living longer, the disease still has a grave prognosis. In the pre-AIDS era the mean survival was approximately 6 months and mortality was 80% within 9 months of disease onset; with the advent of combination antiretroviral therapy, the one-year survival rate has increased to 50%.[5,9] The disease has recently generated widespread concern after reports of PML allegedly developing after treatment with some biologic agents. These agents include the monoclonal antibodies natalizumab (used in the treatment of multiple sclerosis and Crohn's disease), efalizumab (used in the treatment of severe forms of plaque psoriasis), and rituximab (used for hematologic malignancies and rheumatoid arthritis).[10-18]

The growing concern that PML may be a devastating adverse effect of drug therapy prompted drug regulatory authorities as well as physicians and other healthcare practitioners to obtain more information on this disease, including its background frequency in the general population. Because of its rarity and the complexity of the diagnosis, very few population-based data are available regarding incidence of PML in patients without HIV, malignancies, or rheumatologic conditions. In response to a request made by the Dutch Medicines Evaluation Board, we provide in this paper estimates of the incidence of PML in the general population using multiple electronic healthcare data from three countries in Europe (Denmark, Italy, and the Netherlands).

METHODS

Data sources

Demographic and clinical data were obtained from six population-based electronic healthcare records (EHR) databases of the EU-ADR network:[19] (1) Health Search/CSD LPD (HSD, Italy); (2) Interdisciplinary Processing of Clinical Information (IPCI, formerly known as

Integrated Primary Care Information, the Netherlands);[20] (3) PHARMO Network (PHARMO, the Netherlands);[21] (4) Aarhus University Hospital Database (Aarhus, Denmark); (5) Lombardy regional database (Lombardy, Italy); and (6) Tuscany regional database (Tuscany, Italy).[19] The first two (HSD and IPCI) are population-based general practice (GP) databases containing patients' clinical information as well as prescriptions for medications; the last four databases are all comprehensive record-linkage systems of a regional/national catchment area wherein registries of claims for hospitalization and other healthcare services are linked to drug dispensing data. 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. A more detailed description of the databases can be found in earlier publications.[19, 22-23] Data in EU-ADR are combined using a distributed network approach: standardized input files are created locally by each database; the files are then processed using custom-built software (Jerboa©), which aggregates, de-identifies and sends the results in encrypted format to a central repository for evaluation and further analysis.

Outcome identification and validation

PML was identified in the EHR databases through a multi-step process consisting of:

a) Event definition based on clinical criteria established from literature;
b) Development of harmonized database-specific coding algorithms, using diagnosis codes and free text,[21] which reflected the event definition. Because each database uses one of three nomenclature systems to code the events, the different terminologies were first mapped using the Unified Medical Language System1 (UMLS1) to ensure that events are identified using a common language.[25] Cases of PML were identified from primary and secondary hospital discharge diagnoses as well as death registries (claims databases) and from general practitioner or specialist diagnoses (medical records databases) using diagnosis codes and free text (see Table 1).
c) Search in EHR databases using the database-specific coding algorithm;
d) Case validation. Potential cases identified via broad search were subsequently manually validated through review of medical records and hospitalization charts in those databases where validation was possible (HSD and IPCI).

Statistical analysis

Custom-built software Jerboa© manages and aggregates the data over all patients locally, producing as output the amount of person-time of follow-up and the number of events per sex and age group. We only considered the first recorded occurrence of the event of interest after a one year run-in period. We calculated age- and sex-specific incidence rates (IRs) within each database and performed direct standardization using the WHO World Standard Population as reference to account for age differences when comparing the overall rates.[24] Incidence rates were expressed as number of cases per 100 000 person-years (PYs) together with 95% Confidence Interval (95%CI).

Table 1 | Codes and free text used in the database queries to identify PML cases.

UMLS Concept	Concept Description	ICD 10	ICD 9CM	ICPC	Free text
C0023524	Leukoencephalopathy, Progressive Multifocal	A81.2	046.3	MTHU044827 MTHU050534	– Leukoencephalopathy, Progressive Multifocal – Multifocal leukoencephalopathy, NOS – Multifocal leucoencephalopathy – PMLE - Progressive multifocal leukoencephalopathy – PML - Progressive multifocal leucoencephalopathy – Progressive multifocal leukoencephalopathy (disorder) – progressieve multifocale leukoencefalopathie (Dutch) – Leucoencefalopatia multifocale progressive (Italian)
C0022369	JC Virus	–	–	–	– JC Virus – JC polyomavirus – polyomavirus JC – Human polyomavirus type JC – Human polyomavirus JCV – polyoma virus JC JCV – JC virus (organism) – JC polyomavirus JCV

Legend: ICPC – International Classification of Primary Care

RESULTS

A total of 1 150 cases of PML were identified from demographic and clinical data in three European countries representing 21 171 291 individuals with 154 474 063 person-years (PYs) of follow-up. The overall incidence rate (IR) was 0.75 per 100 000 PYs. Figure 1 shows the IRs (pooled from all the databases) stratified by sex and the different age groups. About half of the PML cases were males (603 cases, 52.4%). The largest proportion of male cases were in the age group 30-50 years old (267 cases, 23.2% of total), while the largest proportion of female cases came from those 70 years and older (330 cases, 28.7% of total). The majority of cases overall were also in the age group of 70 years and older (46.7%). The increased IR in the elderly, as seen in Figure 2, was due primarily to the cases in HSD and ARS. Table 2 shows the number of PML cases identified in each of the six databases and the corresponding IRs. The actual codes and search criteria that identified the cases are given in Table 3. The highest IRs were observed in the Italian databases, with the GP database HSD having a crude IR of 4.97/100 000 PYs (standardized 2.29/100 000 PYs), followed by the regional administrative databases of ARS and Lombardy at 1.96/100 000 PYs (standardized 0.73/100 000 PYs) and 0.49/100 000 PYs (standardized 0.39/100 000 PYs), respectively. The IRs for the databases in the Netherlands and in Denmark were quite similar, the standardized rates

ranging from 0.03/100 000 PYs (PHARMO) to 0.04/100 000 PYs (Aarhus) and 0.09/100 000 PYs (IPCI).

A subset analysis of the identified cases in ARS, done *post-hoc*, showed that human immune deficiency virus (HIV) disease was the most common co-morbid condition, being the most frequently reported primary hospital discharge diagnosis in those with a secondary diagnosis of PML. An overview of the most frequently reported primary hospital discharge diagnoses associated with a secondary hospital discharge diagnosis of PML is shown in Table 4.

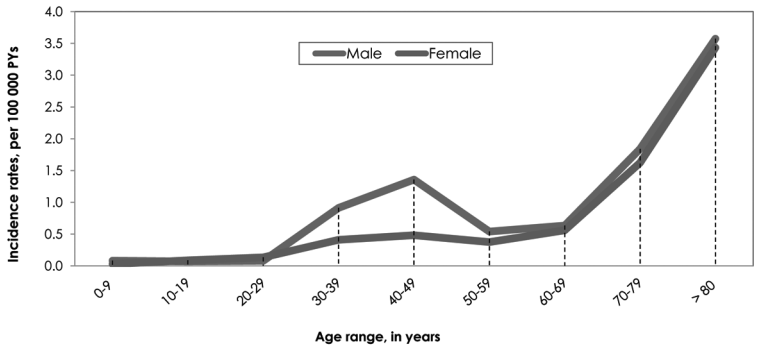


Figure 1 | Incidence rates of PML across age groups and sex (all databases) ***(see page 313 for colour figure).***

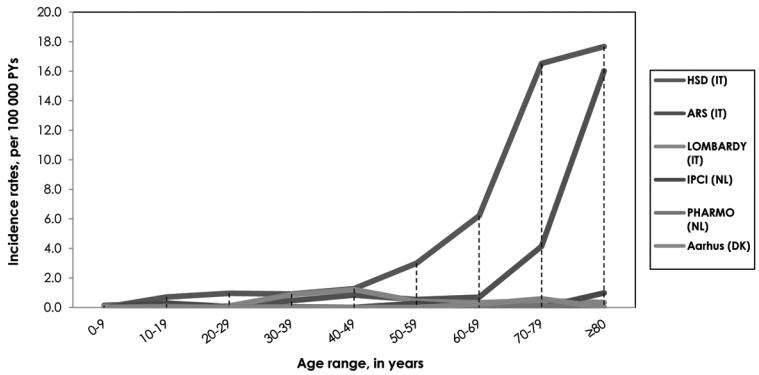


Figure 2 | Age-specific incidence rates of PML across the six databases from four countries ***(see page 312 for colour figure).***

Table 2 | Incidence Rates of PML across six databases in three countries

Database *(type, country of origin)*	**No. of cases identified**	**Person-time observed** *(person-years, PYs)*	**Incidence rate (95%CI)** *(per 100 000 PYs)*	**Standardized Incidence rate*** *(per 100 000 PYs)*
ARS (claims, Italy)	362	18 493 714	1.96 (1.76, 2.17)	0.73
Lombardy (claims, Italy)	484	98 905 832	0.49 (0.45, 0.53)	0.39
HSD/Thales (GP records, Italy)	289	5 814 074	4.97 (4.42, 5.57)	2.29
IPCI (GP records, Netherlands)	3	2 862 113	0.10 (0.03, 0.28)	0.09
PHARMO (claims, Netherlands)	4	13 951 825	0.03 (0.01, 0.07)	0.03
Aarhus (claims, Denmark)	8	13 914 643	0.06 (0.03, 0.11)	0.04
Total	1 150	153 942 201	0.75 (0.70, 0.79)	

*age-standardized to the World Health Organization's Standard Population

Table 3 | Distribution of cases according to the database-specific coding scheme.

Database	**Code/Criteria**	**Source**	**Number of cases** *(% total in database)*
Aarhus	ICD10 : A812	Primary or secondary hospital discharge diagnoses	6 *(75)*
		Death registry	2 *(25)*
ARS	ICD9CM: 046.3	Primary hospital discharge diagnosis	149 *(41)*
		Secondary hospital discharge diagnosis	209 *(58)*
		Death registry	4 *(1)*
Lombardy	ICD9CM: 046.3	Primary or secondary hospital discharge diagnoses	484 *(100)*
HSD	ICD9CM: 046.3 and free text	GP or specialist diagnosis	289 *(100)*
PHARMO	ICD9CM: 046.3	Primary or secondary hospital discharge diagnoses	4 *(100)*
IPCI	Free text	GP or specialist diagnosis	3 *(100)*

Table 4 | Most frequently reported primary hospital discharge diagnoses in patients with secondary discharge diagnosis of PML in the Italian database ARS.

ICD9 Code	Description	No. of times reported
042	Human immunodeficiency virus (HIV) disease	20
435.9	Unspecified transient cerebral ischemia	12
507.0	Pneumonitis due to inhalation of food or vomitus	11
518.81	Acute respiratory failure	10
427.5	Cardiac arrest	7
345.1	Generalized convulsive epilepsy	5
434.01	Cerebral thrombosis with infarct	5
276.5	Volume depletion	4
290.41	Vascular dementia with delirium	3
571.5	Cirrhosis of liver without mention of alcohol	3
414.0	Coronary atherosclerosis of unspecified type of vessel, native or graft	3
038.4	Gram-negative septicemia not specified	3
293.0	Delirium due to conditions classified elsewhere	3
485	Bronchopneumonia, organism unspecified	3
585	Chronic kidney disease (CKD)	3
584.9	Acute renal failure, unspecified	3
854	Intracranial injury of other and unspecified nature	2
431	Intracerebral hemorrhage	1
162.2	Malignant neoplasm of the main bronchus	1
V10.06	History of malignant neoplasm of the rectum, rectosigmoid junction, and anus	1
185	Malignant neoplasm of the prostate	1
197	Secondary malignant neoplasm of the respiratory and digestive systems	1
203.00	Multiple myeloma, without mention of remission	1
203.8	Other immunoproliferative neoplasm	1
710.1	Systemic sclerosis	1
714	Rheumatoid arthritis and other inflammatory polyarthropathies	1
438	Late effects of cerebrovascular disease	1

DISCUSSION

The advent of promising biologic therapies has raised hopes for long-suffering patients, but has also raised concerns for PML as a devastating possible consequence of drug treatment.[25] A lot of uncertainty remains about this disease, but what is clear is that physicians hitherto unaware or unfamiliar with the disease must now consider PML in their differential diagnoses and marketing authorization holders and drug regulators alike must ensure that the possible risk of PML is addressed in the risk management plan, and carefully monitored after drug approval, of

certain drugs. In this paper we have attempted to address some of the uncertainty by providing an overview of the epidemiology of PML in the general population of three European countries using electronic healthcare record databases. It has been suggested that non-immunosuppressed, healthy individuals account for <1% of all cases of PML.[7] A study using health insurance claims in the US estimated the prevalence of PML in the general population at 4.4 cases per 100 000 individuals.[26] To date, there is no study describing the epidemiology of PML in the general European population against which we can compare our results, but our findings are compatible with this US estimate, albeit slightly higher (1 150 PML cases out of the database network population of 21 171 291 individuals translates to 5.5 cases per 100 000 individuals). An increased awareness about the disease in the last six years since the publication of this US study may account for this discrepancy.

The incidence rates of PML observed in all three Italian databases were higher than those observed in both Denmark and the Netherlands, which were comparable with one another. We investigated whether this would be explained by a higher prevalence of HIV/AIDS in Italy compared to the other two countries. However, latest estimates from the Joint United Nations Programme on HIV/AIDS (UNAIDS) show that in 2009 the prevalence among adults aged 15-49 was the same in the three countries: 0.3% for Italy and 0.2% for both Denmark and the Netherlands (140 000 people living with HIV/AIDS in Italy compared to 22 000 in the Netherlands and 5 300 in Denmark).[27] It might be possible that there is a higher incidence of HIV/AIDS among those 50 years and older in Italy.[28-32] Another possible explanation is that while incidence and prevalence of HIV/AIDS may be the same in all three countries, access to healthcare resources (including antiretroviral therapy) may be different. The ICD9-CM code 046.3 is a very specific code for PML and it seems unlikely to be recorded unless there is a high index of suspicion. Review of the medical records of the PML cases identified in HSD showed that PML was mostly diagnosed in patients with cerebrovascular ischemia (transient ischemic attack, stroke), dementia of various etiologies, autoimmune diseases (multiple sclerosis and rheumatoid arthritis), and cancer. This is probably due to the fact that these conditions share the same initial clinical manifestations as those of PML. Records also showed that about 50% of the cases underwent neuroimaging studies (MRI, CT scan) although the actual results were not always available. While the majority of these GP-recorded diagnoses (both for HSD and IPCI) were confirmed by a neurologist or other specialist, the possibility of misclassification or misdiagnosis cannot be disregarded. With respect to claims reimbursements, PML does not translate to higher incentive according to the Drug-Related Groups (DRG) coding system and would not explain an eagerness for recording the diagnosis.

Because of the widespread nature of JCV infection and the association of PML with immune suppression, it is generally believed that viral reactivation from sites of latency triggers the initiating event that results in the development of PML.[18] The observed peak of the IR in middle age most likely parallels the peak in HIV/AIDS incidence, especially in homosexual males. The increasing incidence of PML observed from age 50 years onwards could be a sign that the prevalence of HIV/AIDS in the elderly is underestimated, age being an important predictor of progression in HIV infections.[33-34] Early symptoms of HIV, including fatigue and physical and neurocognitive deterioration, could be mistaken for other conditions frequently seen in the elderly such as

Parkinson's disease and Alzheimer's, or even attributed to 'normal aging.'[35] Such scenarios can lead to a delay in seeking medical care and delayed diagnosis. Another, perhaps more probable, explanation is that the increasing incidence of PML in the elderly reflects the decline in the functioning of the immune system that comes with age and a greater risk for developing various malignancies, organ failure requiring transplant, as well as complications of autoimmune diseases.

Aside from HIV/AIDS, other conditions that have been described in the literature to occur with, or predispose to, PML include systemic lupus erythematosus (SLE), inflammatory myositis, dermatomyositis-polymyositis, rheumatoid arthritis, other autoimmune vasculitis, the leukemias, lymphomas, and bone marrow or solid organ transplants.[8] Our findings reveal some of these conditions as well as others which have not previously been described in the literature such as reticulosarcoma and malignant neoplasms of the lung, prostate, and bone and multiple myeloma. Our findings also show that patients with PML have multiple co-morbidities including cardiac and cerebrovascular disease, which may mask the symptoms of PML and delay diagnosis, as well as advanced liver and kidney disease which may predispose to immune suppression.

Limitations

Despite the wide population coverage of our study, there are some limitations. The EU-ADR database network was primarily designed for drug safety surveillance and although incidence rates of various diseases are useful side products of the system, there remain caveats in the interpretation of such results. We only used diagnosis codes in identifying the cases of PML. Although corroborating evidence from neuroimaging, JCV DNA studies, or brain biopsy were sought during case validation, such data were not uniformly available in all the databases. Administrative databases may provide information on whether – and how many -such procedures were performed, but the nature of claims databases is such that the actual results are not routinely documented. While a patient with PML may have his or her first clinical encounter with a GP, subsequent work-up requiring specialist care may not always be documented in the GP records. It was not feasible to conduct case validation by medical chart review in all the databases and this was done only in a subset of the cases.

CONCLUSION

We have described the epidemiology of progressive multifocal leukoencephalopathy among the general population in three European countries and we have confirmed that PML is a rare disease that is not only associated with HIV, but also with malignancy and autoimmune conditions. An observed increase in the incidence of PML in the elderly requires further investigation. With proper case validation, EHR databases may provide estimates of background incidence rates in the general population for rare events such as PML.

REFERENCES

1. White MK, Khalili K. Pathogenesis of progressive multifocal leukoencephalopathy--revisited. *J Infect Dis.* Mar 1 2011;203(5):578-586.
2. Walker DL, Padgett BL. The epidemiology of human polyomaviruses. *Prog Clin Biol Res.* 1983;105:99-106.
3. Markowitz RB, Thompson HC, Mueller JF, Cohen JA, Dynan WS. Incidence of BK virus and JC virus viruria in human immunodeficiency virus-infected and -uninfected subjects. *J Infect Dis.* Jan 1993;167(1):13-20.
4. Egli A, Infanti L, Dumoulin A, et al. Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors. *J Infect Dis.* Mar 15 2009;199(6):837-846.
5. Berger JR. The clinical features of PML. *Cleve Clin J Med.* Nov 2011;78 Suppl 2:S8-12.
6. Saribas AS, Ozdemir A, Lam C, Safak M. JC virus-induced Progressive Multifocal Leukoencephalopathy. *Future Virol.* May 2010;5(3):313-323.
7. Simpson DM. HIV-associated PML: changing epidemiology and clinical approach. *Cleve Clin J Med.* Nov 2011;78 Suppl 2:S24-27.
8. Amend KL, Turnbull B, Foskett N, Napalkov P, Kurth T, Seeger J. Incidence of progressive multifocal leukoencephalopathy in patients without HIV. *Neurology.* Oct 12 2010;75(15):1326-1332.
9. Tan CS, Koralnik IJ. Progressive multifocal leukoencephalopathy and other disorders caused by JC virus: clinical features and pathogenesis. *Lancet Neurol.* Apr 2010;9(4):425-437.
10. Carson KR, Focosi D, Major EO, et al. Monoclonal antibody-associated progressive multifocal leucoencephalopathy in patients treated with rituximab, natalizumab, and efalizumab: a Review from the Research on Adverse Drug Events and Reports (RADAR) Project. *Lancet Oncol.* Aug 2009;10(8):816-824.
11. Langer-Gould A, Atlas SW, Green AJ, Bollen AW, Pelletier D. Progressive multifocal leukoencephalopathy in a patient treated with natalizumab. *N Engl J Med.* Jul 28 2005;353(4):375-381.
12. Kleinschmidt-DeMasters BK, Tyler KL. Progressive multifocal leukoencephalopathy complicating treatment with natalizumab and interferon beta-1a for multiple sclerosis. *N Engl J Med.* Jul 28 2005;353(4):369-374.
13. Van Assche G, Van Ranst M, Sciot R, et al. Progressive multifocal leukoencephalopathy after natalizumab therapy for Crohn's disease. *N Engl J Med.* Jul 28 2005;353(4):362-368.
14. Korman BD, Tyler KL, Korman NJ. Progressive multifocal leukoencephalopathy, efalizumab, and immunosuppression: a cautionary tale for dermatologists. *Arch Dermatol.* Aug 2009;145(8):937-942.
15. Herold T, Seiler T, Egensperger R, et al. Progressive multifocal leukoencephalopathy after treatment with rituximab, fludarabine and cyclophosphamide in a patient with chronic lymphocytic leukemia. *Leuk Lymphoma.* Jan 2012;53(1):169-172.
16. Clifford DB, De Luca A, Simpson DM, Arendt G, Giovannoni G, Nath A. Natalizumab-associated progressive multifocal leukoencephalopathy in patients with multiple sclerosis: lessons from 28 cases. *Lancet Neurol.* Apr 2010;9(4):438-446.
17. Tavazzi E, Ferrante P, Khalili K. Progressive multifocal leukoencephalopathy: an unexpected complication of modern therapeutic monoclonal antibody therapies. *Clin Microbiol Infect.* Dec 2011;17(12):1776-1780.
18. Major EO. Progressive multifocal leukoencephalopathy in patients on immunomodulatory therapies. *Annu Rev Med.* 2010;61:35-47.
19. Coloma PM SM, Trifiró G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom MCJM, 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 Jan;20(1):1-11.
20. Interdisciplinary Processing of Clinical Information. http://www.ipci.nl.
21. PHARMO Institute for Drug Outcomes Research. http://www.pharmo.nl.
22. Trifiro G, Patadia V, Schuemie MJ, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Stud Health Technol Inform.* 2011;166:25-30.
23. Coloma PM, Trifiro G, Schuemie MJ, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? *Pharmacoepidemiol Drug Saf.* Feb 8 2012.
24. World Health Organization. WHO Standard Population. http://www.who.int/healthinfo/paper31.pdf.

25. Koralnik IJ. Progressive multifocal leukoencephalopathy revisited: Has the disease outgrown its name? *Ann Neurol.* Aug 2006;60(2):162-173.
26. Eng PM TB, Cook SF, Davidson JE, Kurth T, Seeger JD. Characteristics and antecedents of progressive multifocal leukoencephalopathy in an insured population. *Neurology.* 2006 Sep 12;67(5):884-886.
27. UNAIDS. Joint United Nations Programme on HIV/AIDS UNAIDS. 2009; http://www.unaids.org/en/. Accessed Last accessed 20 January 2012.
28. Verdecchia A, Mariotto A, Capocaccia R, Mariotti S. An age and period reconstruction of the HIV epidemic in Italy. *Int J Epidemiol.* Oct 1994;23(5):1027-1039.
29. Suligoi B, Pezzotti P, Boros S, Urciuoli R, Rezza G. Epidemiological changes in AIDS and HIV infection in Italy. *Scand J Infect Dis Suppl.* 2003;106:12-16.
30. Madeddu G, Rezza G, Mura MS. Trends in the European HIV/AIDS epidemic: a perspective from Italy. *Expert Rev Anti Infect Ther.* Feb 2009;7(1):25-36.
31. Muller V, Maggiolo F, Suter F, et al. Increasing clinical virulence in two decades of the Italian HIV epidemic. *PLoS Pathog.* May 2009;5(5):e1000454.
32. Barcherini S, Cantoni M, Grossi P, Verdecchia A. Reconstruction of human immunodeficiency virus (HIV) sub-epidemics in Italian regions. *Int J Epidemiol.* Feb 1999;28(1):122-129.
33. Adler WH, Baskar PV, Chrest FJ, Dorsey-Cooper B, Winchurch RA, Nagel JE. HIV infection and aging: mechanisms to explain the accelerated rate of progression in the older patient. *Mech Ageing Dev.* Jun 1997;96(1-3):137-155.
34. Inelman EM, Gasparini G, Enzi G. HIV/AIDS in older adults: a case report and literature review. *Geriatrics.* Sep 2005;60(9):26-30.
35. Butt AA, Dascomb KK, DeSalvo KB, Bazzano L, Kissinger PJ, Szerlip HM. Human immunodeficiency virus infection in elderly patients. *South Med J.* Apr 2001;94(4):397-400.

Chapter 5

Separating the Big Fish from the Small Fry

'The world is richer in associations than meanings, and it is part of wisdom to differentiate the two.'

– John Barth, American novelist

Triage and Evaluation of Potential Safety Signals Identified from Electronic Healthcare Record (EHR) Databases

Preciosa M Coloma, MD, MSc[1], Martijn J Schuemie, PhD[1], Gianluca Trifirò, MD, PhD[1,2], Erik van Mulligen, PhD[1], Paul Avillach, MD, PhD3[3,4], Anna Bauer-Mehren, PhD[5], Laura Furlong, PhD[5], Jan Kors,[1] PhD[1], Ferran Sanz,[5] PhD[1], Jodi Mestres,[5] PhD[1], José Luis Oliveira, PhD[6], Scott Boyer, PhD[7], Ernst Ahlberg Helgee, MSc[7], Mariam Molokhia, MD, PhD[8], Justin Matthews, MSc[9], David Prieto-Merino, PhD[9], Rosa Gini, PhD[10], Ron Herings, PhD[11], Giampiero Mazzaglia, MD, PhD[12], Gino Picelli, PhD[13], Lorenza Scotti, MSc[14], Lars Pedersen, PhD[15], Johan van der Lei, MD, PhD[1], Miriam Sturkenboom, PharmD, PhD[1,16], on behalf of the EU-ADR consortium

[1]Department of Medical Informatics, Erasmus Medical Center, Rotterdam, the Netherlands; [2]Department of Clinical and Experimental Medicine and Pharmacology, Section of Pharmacology, University of Messina, Messina, Italy; [3]LESIM, ISPED, University Bordeaux Segalen, Bordeaux, France; [4]LERTIM, EA 3283, Faculté de Médecine, Université Aix Marseille 2, Marseille, France; [5]Research Programme on Biomedical Informatics (GRIB), IMIM-Hospital del Mar - Universitat Pompeu Fabra-PRBB, Barcelona, Spain; [6]IEETA, Universidade de Aveiro, Portugal; [7]Safety Assessment, AstraZeneca Research & Development, Mölndal, Sweden; [8]Department of Primary Care & Public Health Sciences, Kings College London, United Kingdom; [9]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, UK; [10]Agenzia Regionale di Sanità della Toscana, Florence, Italy; [11]PHARMO Institute, Utrecht, The Netherlands; [12]Società Italiana di Medicina Generale, Florence, Italy; [13]Pedianet-Società Servizi Telematici SRL, Padova, Italy; [14]Department of Statistics, Università di Milano-Bicocca, Milan, Italy; [15]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; [16]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands

ABSTRACT

Background. There is huge potential for mining electronic healthcare record (EHR) databases to augment current systems in pharmacovigilance. Like any signal detection system, there is a need to establish 'rules' how to trigger an alert, when to consider a signal likely enough to be true to warrant follow-up or even to require immediate health policy intervention.

Objectives. To describe the process of prioritization of drug-adverse event associations derived from signal detection using EHR databases in the EU-ADR Project.

Methods. Association measures between drug use and acute myocardial infarction (AMI) were generated by first applying various statistical methods on healthcare data from seven databases of the EU-ADR network. Association estimates were ranked based on the best performing method (Longitudinal Gamma Poisson Shrinker). Matched case-control and self-controlled case series methods were additionally conducted to deal with temporality and confounding effects, while the LEOPARD method was applied to specifically detect protopathic bias. Consistency of the association among drugs of the same class and the number of excess cases attributable to the drug exposure were further assessed to prioritize the list of potential signals. Finally, signal filtering and signal substantiation were done using different bioinformatics workflows to determine the novelty and plausibility of the identified signals.

Results. Demographic, clinical and prescription/dispensing data from seven databases in three European countries were obtained, representing 21 171 291 individuals with 154 474 063 person-years of follow-up in the period 1995-2011. Overall, 163 potential signals for AMI were identified based on preliminary screening for statistical associations. Of these, 72 signals were flagged by LEOPARD as likely due to protopathic bias. Further signal refinement to reduce possible confounding decreased the number of signals to 39. The following nine signals remained after applying the criteria for novelty and plausibility: (1) metoclopramide; (2) cisapride; (3) domperidone; (4) betamethasone; (5) erythromycin; (6) roxythromycin; (7) azithromycin; (8) fluconazole; and (9) megestrol.

Conclusion. We propose a prioritization strategy for drug safety signal detection using EHR by taking into account, in addition to statistical associations, also public health relevance, novelty, and plausibility. This strategy needs to be further tested using other EHR data sources and other adverse events.

INTRODUCTION

The scale-up and automation of drug safety surveillance efforts has become both a science[1] and a crusade and has brought to the fore the huge potential for mining electronic healthcare record (EHR) databases to augment current systems in pharmacovigilance. While the advantages of automated surveillance and quantitative signal detection are obvious, there are growing concerns that such data mining may generate more signals than can be followed up effectively with currently available resources. This concern is not entirely unfounded, considering that the annual volume of reports received in spontaneous reporting systems (SRS), database systems primarily designed for signal detection, has become enormous and unmanageable.[2-3] The problem is likely to be worse with the use of EHR data which have been intended for other purposes and which can be mined for associations without routine human evaluation of potential alternative explanations.

Signal detection is only the initial step in the long and complex process of post-marketing safety surveillance. The evaluation of a signal may take years, from the earliest suspicion of a potential risk to an established mechanism of causation and fully understood phenomenon.[4] While signals derived from EHR data may give a good snapshot of how drugs are being used in real-world settings, there remains the need to establish guidelines as to when - and how - to consider a signal likely to be substantial enough to warrant verification and follow-up. Various strategies for signal prioritization have been proposed in many publications, although most of these refer to signals derived from SRS.[4-8] These strategies focus consistently on signals with serious adverse effect, strong supporting evidence, and greatest public health impact. With the new pharmacovigilance legislation adopted by the European Parliament and European Council[9-10] and strengthened legal basis for additional post-authorization monitoring (consequently increasing the work and accountability of both drug manufacturers and regulators), there is an even more pressing need to streamline the process of surveillance.

In this paper we describe the process of triage and evaluation of drug-event associations derived from signal detection using EHR databases based on preliminary results from the EU-ADR Project. Taking the event acute myocardial infarction (AMI) as an example, we propose a strategy for combining evidence from different data sources and modalities to prioritize signals that may represent genuine risk and, hence, necessitate further investigation through traditional hypothesis testing studies or action from regulatory agencies.

METHODS

Data sources

The EU-ADR network currently comprises anonymized demographic, clinical and drug prescription/dispensing information from eight population-based electronic healthcare records (EHR) databases in four European countries. Signal detection was performed in the following seven databases for the period 1995-2011: (1) Health Search/CSD LPD (HSD, Italy); (2) Interdisciplinary

Processing of Clinical Information (IPCI, formerly known as Integrated Primary Care Information, the Netherlands); (3) Pedianet (Italy); (4) PHARMO Network (PHARMO, the Netherlands); (5) Aarhus University Hospital Database (Aarhus, Denmark); (6) Lombardy regional database (Lombardy, Italy); and (7) Tuscany regional database (Tuscany, Italy). HSD, IPCI, and Pedianet are general practice (GP) databases, where both clinical information and drug prescriptions are recorded. Aarhus, PHARMO, Lombardy, and Tuscany are all comprehensive record-linkage systems in which drug dispensing data of a well-defined population are linked to a registry of hospital discharge diagnoses and various other registries. 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.[11-24] Most 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 and turnover is low. In all of the countries with GP databases, GPs function as 'gatekeepers' of the healthcare system. A more detailed description of the database network can be found in earlier publications.[25-27] Data in EU-ADR are combined using a distributed network approach: standardized input files are created locally by each database; the files are then processed using custom-built software (Jerboa©), which aggregates, de-identifies and sends the results in encrypted format to a central repository for evaluation and further analysis.

Identification and validation of cases of acute myocardial infarction

Potential cases of acute myocardial infarction (AMI) in the database network were identified using harmonized and database-specific coding algorithms that utilize three disease coding terminologies: (1) International Classification of Primary Care (ICPC) for IPCI; (2) International Classification of Diseases 9th revision-Clinical Modification (ICD-9CM) for ARS, HSD, Lombardy, and PHARMO; and (3) ICD 10th revision for Aarhus. IPCI, HSD, and Pedianet also performed free text search using specific key words. The process of mapping and the harmonization of event data extraction from different EHR databases in the EU-ADR project was based on medical concepts derived from the Unified Medical Language System (UMLS)[28] and has been described in other publications.[29-31] Case validation using manual review of hospitalization records and GP records was done in a subset of the cases. Details of the search algorithms used as well as the case validation performed are described in another paper.[32]

Drug exposure

Drug prescription and/or dispensing data were used to estimate event rates during drug exposure. Drug prescriptions and dispensings in EU-ADR are locally coded using the national product codes, which differ among countries but these product codes are linked to the Anatomical Therapeutic Chemical (ATC) classification system of the World Health Organization (WHO) system.[33] The ATC code is used as the drug code in the EU-ADR input files. Each database holder estimates the duration covered by each prescription or dispensing according to the legend duration (if dosing regimen is available), or is otherwise based on the defined daily dose (DDD).[34]

Overlapping treatment episodes with the same drug (same ATC code) are combined into a single episode of drug use that starts when the first prescription begins and stops when the last prescription ends. When a patient uses more than one drug at a time, the corresponding person-time is labeled accordingly. Using individual data on start date and end date of prescription or dispensing, Jerboa© determines and marks as unexposed those periods during which an individual is included in the study but is not using any drug. Events are then assigned to the episodes (drug use/non-use) in which they occurred.

Signal Detection

The EU-ADR system comprises three automated steps that generate a list of potential safety signals. For signal triage we describe additional measures that could be used for final prioritization if the interest would be to create a list of signals that requires further follow-up.

EU-ADR Step 1: Data mining techniques for signal detection

In a previous paper we described in detail the various statistical methods currently employed for signal detection in EU-ADR.[35] The following statistical methods were applied to data pooled from seven databases representing three countries in the EU-ADR network (Denmark, Italy, and the Netherlands): (1) SRS-based methods, including the proportional reporting ratio (PRR), reporting odds ratio (ROR), Gamma Poisson Shrinker (GPS), and Bayesian Confidence Propagation Neural Network (BCPNN); (2) cohort-based methods, including estimation of the incidence rate ratio (IRR), longitudinal GPS, and Bayesian hierarchical modeling; and (3) case-based methods, including case-control and self-controlled case series (SCCS). Another method, LEOPARD (Longitudinal Evaluation of Observational Profiles of Adverse events Related to Drugs), developed in EU-ADR, attempts to single out signals which may occur because the drug is used to treat the event, or a precursor of the event, rather than cause it. For every suspect drug-event pair, LEOPARD compares the rates of prescription starts within a fixed window (±25 days) prior to and after the occurrence of the event. An increase in the number of prescriptions after an event relative to the number of prescriptions prior to the event is taken to be an indication of protopathic bias.[36] This is assessed using a binomial test.

EU-ADR Step 2: Automated signal filtering

The drug-adverse event associations from *Step 1* were further evaluated to discriminate among potentially relevant new signals and already known signals. This process of signal filtering was carried out by mining of published literature (via MEDLINE) as well as publicly available pharmacological databases and repositories.[37-38]

EU-ADR Step 3: Automated signal substantiation

The process of signal substantiation requires that the signal be placed in the context of current knowledge of pharmacological mechanisms that might explain it.[37, 39] To substantiate the likelihood of certain signals based on biological plausibility, automatic linkage of biomedical entities (drugs,

proteins and their genetic variants, biological pathways and clinical events) via customized bioinformatics methods was done to find supporting biological evidence for the potential signals. All signals were processed by a computational framework that identifies pair-wise relationships between the drug and the event based on *in silico* prediction of drug targets, analysis of drug metabolites and gene-disease associations.[39, 40] **Figure 1** shows a schematic representation of the substantiation set-up. Both signal filtering and signal substantiation workflows were developed as discrete web services and are currently implemented in a web-based platform that allows automatic processing of the signals.[41-42]

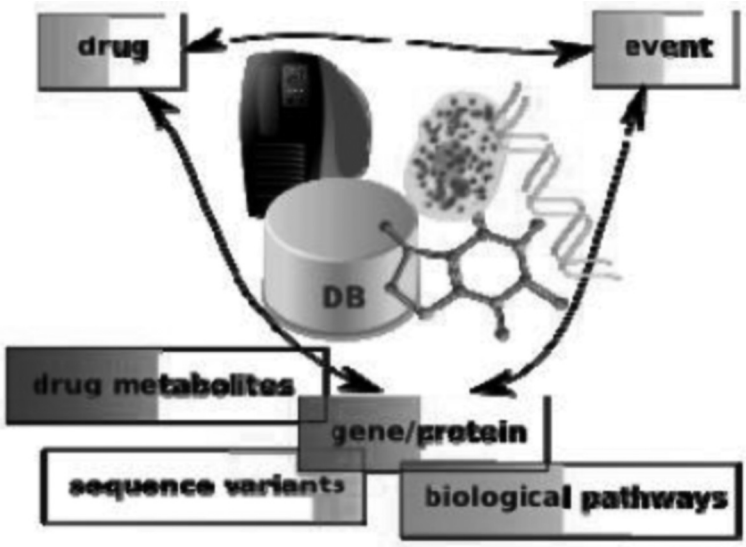


Figure 1 | Schematic representation of the substantiation workflows that can guide causal inference for a signal.

Triage of signals for further follow up

Prioritization of signals for further follow-up was done based on a sequence of additional methods that mimic the usual steps in the triage of signals generated from SRS and based on the Bradford-Hill criteria for causality.

Strength of statistical association. The Longitudinal Gamma Poisson Shrinker (LGPS) method was identified as the best-performing among the methods currently being used in EU-ADR, i.e., it had the highest area under the receiver operator characteristic curve (AUC) when tested against a reference standard of known positive drug-event associations and 'negative controls.'[35, 43-44] We used the risk estimate obtained by this method (relative risk, RR_{LGPS}) to initially rank the list of potential signals from *Step 1*. A value of $RR_{LGPS} \geq 2.0$ and a lower 95% CI of $RR_{LGPS} > 1$ were used as threshold values for further processing.

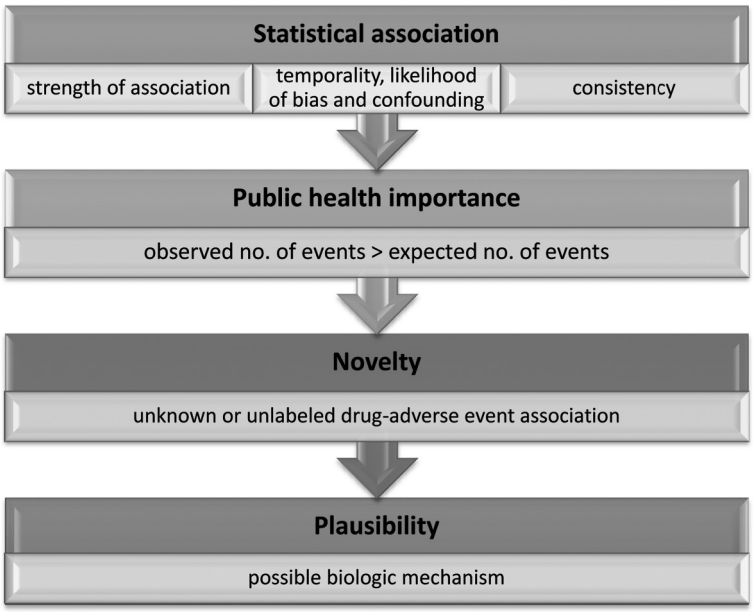


Figure 2 | Framework for the triage of signals identified from EHR data..

Alternative explanation: protopathic bias and confounding. All potential signals flagged by LEOPARD as likely due to protopathic bias were eliminated from the list. To account for possible confounding, we further sorted out the list and considered only those signals which had significant increased risk estimates based on the case-control method (lower 95% CI of exposure odds ratio (OR)>1) or the SCCS (lower 95% CI of incidence rate ratio (IRR_{SCCS})>1). In the case control method, each case was matched to two controls of the same age, sex, and index date (date of event). To adjust for comorbidity and global patient health status we used as proxy the number of different drugs (i.e., distinct Anatomical Therapeutic Chemical Classification (ATC) codes)[45] an individual was exposed to within the period included between one year and one month prior to the index date.[46] The SCCS method controls for time-fixed confounders such as genetic factors, socio-economic status, individual frailty, and severity of underlying disease.[47]

Consistency of association among drugs of the same class. While it may not always be the case, drugs belonging to the same class usually have a similar pharmacological mechanism of therapeutic action and safety profile.[48-50] Based on this assumption, signals that involve drugs belonging to the same class may require more thorough investigation.

Public health importance. To quantify the public health impact of a potential signal, we used as a surrogate the number of excess cases of patients exposed to the drug relative to the background unexposed population (observed – expected). At the outset, we already focused our signal

detection efforts in EU-ADR on events that are considered important from a pharmacovigilance and public health perspective.[51]

Novelty. The very nature of a signal requires it to be a phenomenon describing a 'new potentially causal association, or a new aspect of a known association.'[52] We incorporate the criterion for novelty in the triage framework by making use of the signal filtering workflows described in *Step 2*. A drug-adverse event association that is flagged as known, or previously reported, in more than one of three biomedical databases (MEDLINE: URL http://www.nlm.nih.gov/bsd/pmresources.html; DrugBank: URL http://www.drugbank.ca/; or DailyMed: URL http://dailymed.nlm.nih.gov/dailymed/) is taken off the signals list.

Plausibility. To identify meaningful signals, signal detection must go beyond statistical association to demonstration of a possible mechanism for the drug-adverse event association via pathophysiologic pathways and metabolic targets. This process of signal substantiation can steer the signal triage process into more sound decisions as to which signals are worth following up. As a final assessment procedure, we retained in the signal list only those associations for which the substantiation workflow could find a possible biologic mechanism.

Using the substantiation qualification may preclude finding those associations which cannot be predicted from the drug's pharmacological action. There are, however, many known ADRs the mechanisms of which remain unclear or have yet to be elucidated, including the so-called idiosyncratic reactions.[53,54-55] To account for this type of ADRs, we determined how many and which signals would remain if we keep those signals for which the substantiation workflow did not find anything, but passed the novelty requirement. A manual search of the literature was performed to determine whether there could be some explanation for these additional signals.

RESULTS

The results described below are based on healthcare data from 21 171 291 individuals with 154 474 063 person-years (PYs) of follow-up within the period 1995-2011. A total of 235 283 cases of AMI were identified in the entire database network, with pooled background incidence rate=153.7/100 000 PYs (no AMI cases were identified in Pedianet). There were 163 potential signals (i.e., drug-event associations) identified for AMI through mining of EHR using LGPS as a screening method. Of these, 72 signals were flagged and discarded by LEOPARD as likely due to protopathic bias. Table 1 shows the therapeutic classification of drugs comprising the remaining 91 signals. Further refinement of the signals to reduce possible confounding by means of the case control and SCCS methods reduced the number of signals to 39. The classes of drugs associated with more than one signal are shown in Table 2.

The number of excess cases attributable to drug exposure ranged from 18 for the systemic antibacterial rokitamycin to 2 445 for fixed-dose combinations containing metformin. After applying the criteria for novelty and plausibility, only 9 signals remained. These potential signals are listed in **Table 3**. To consider the possibility of idiosyncratic reactions or associations which

cannot be explained by the drug pharmacology, signals not substantiated by the workflows were included back in the list. This increased the number of potential signals to 27.

Table 1 | Therapeutic classification of drugs comprising potential signals for acute myocardial infarction*.

Therapeutic class	No. of drugs
Systemic antibacterials	22
Opioid analgesics	10
Antiemetic/gastric prokinetic agents	9
Antineoplastics or immune modulators	9
Systemic corticosteroids	5
Antihypertensives	4
Oral hypoglycemic agents	4
Antiretrovirals	4
β-adrenergic agonists	4
Antihemorrhagic and anti-anemic preparations	3
Systemic antifungals	2
Antipsychotics	2
Cough and cold preparations	2
Antacids	2
Nonsteroidal anti-inflammatory drugs	2
Disease-modifying antirheumatic drugs	1
Muscle relaxants	1
Antidepressants	1
Hormonal contraceptives	1
Antimycobacterials	1
Antidote/ion-exchange resins	1
Phosphate binders	1
Total	91

* based on $RR_{LGPS} \geq 2$, lower 95% CI of $RR_{LGPS} > 1$, and unlikely due to protopathic bias

Table 2 | Drugs comprising potential signals*

Therapeutic class	Drug (ATC code)	RR_{LGPS} (95% CI)	OR (95% CI)	IRR_{SCCS} (95% CI)	No. of excess cases
Oral hypoglycemic agent	Metformin and sulfonamides (A10BD02)	2.5 (2.4, 2.6)	1.9 (1.8, 2.0)	1.5 (1.4, 1.6)	2 445
Antihypertensive	Nifedipine (C08CA05)	2.1 (2.0, 2.2)	1.6 (1.6, 1.7)	1.8 (1.7, 2.0)	2 097
Systemic corticosteroid	Prednisone (H02AB07)	2.5 (2.4, 2.6)	1.5 (1.4, 1.6)	2.2 (1.9, 2.6)	1 261
β-adrenergic agonist	Salbutamol (systemic) (R03AC02)	2.1 (2.0, 2.2)	1.2 (1.2, 1.3)	1.9 (1.6, 2.2)	1 017
Systemic corticosteroid	Methylprednisolone (H02AB04)	2.3 (2.2, 2.4)	1.5 (1.3, 1.6)	2.0 (1.7, 2.3)	832
Opioid analgesic	Tramadol (N02AX02)	2.1 (2.0, 2.2)	1.3 (1.2, 1.4)	2.2 (1.7, 2.8)	736
Oral hypoglycemic agent	Glibenclamide (A10BB01)	2.2 (2.1, 2.4)	1.6 (1.6, 1.8)	1.3 (1.1, 1.6)	686
Antihypertensive	Clonidine (C02AC01)	2.9 (2.7, 3.1)	1.8 (1.6, 1.9)	2.5 (1.9, 3.2)	650
Systemic antibacterial	Clarithromycin (J01FA09)	3.5 (3.2, 3.7)	2.4 (2.2, 2.6)	3.3 (2.8, 3.8)	645
β-adrenergic agonist	Fenoterol (inhaled) (R03AK03)	2.5 (2.3, 2.6)	1.4 (1.3, 1.5)	1.6 (1.1, 2.3)	588
β-adrenergic agonist	Salbutamol (inhaled) (R03AK04)	2.4 (2.2, 2.6)	1.3 (1.2, 1.4)	1.7 (1.4, 2.2)	510
Systemic antibacterial	Amoxicillin (J01CA04)	2.2 (2.0, 2.3)	1.6 (1.5, 1.8)	2.0 (1.8, 2.4)	497
Systemic corticosteroid	Betamethasone (H02AB01)	2.9 (2.7, 3.2)	1.7 (1.5, 2.0)	3.3 (2.6, 4.3)	365
Antacid	Magaldrate (A02AD02)	2.8 (2.5, 3.0)	1.9 (1.7, 2.2)	4.8 (3.9, 5.9)	365
Systemic antibacterial	Phenoxymethylpenicillin (J01CE02)	3.6(3.3, 4.0)	2.6 (2.3, 2.9)	3.8 (3.0, 4.9)	335
Systemic corticosteroid	Dexamethasone (H02AB02)	3.2 (2.9, 3.5)	1.9 (1.7, 2.2)	5.4 (4.1, 7.2)	285
Antacid	Combinations of aluminum, magnesium, or calcium salts (A02AD01)	3.1 (2.8, 3.5)	1.9 (1.6, 2.2)	4.4 (3.3, 5.7)	265
Opioid analgesic	Fentanyl (N02AB03)	2.5 (2.3, 2.8)	1.2 (1.1, 1.4)	2.1 (1.2, 3.9)	249
Antiemetic/gastric prokinetic	Metoclopramide (A03FA01)	5.7 (5.1, 6.4)	2.6 (2.2, 3.1)	8.9 (5.1, 15.6)	236
Antiemetic/gastric prokinetic	Domperidone (A03FA03)	2.8 (2.5,3.1)	1.6 (1.4, 1.8)	3.1 (2.4, 4.0)	229
Systemic antibacterial	Azithromycin (J01FA10)	2.8 (2.5, 3.2)	1.7 (1.5, 2.1)	2.5 (1.8, 3.5)	159
Systemic antibacterial	Pivampicillin (J01CA02)	4.5 (3.9, 5.2)	3.1 (2.6, 3.7)	3.6 (2.1, 6.1)	156
Systemic antibacterial	Ceftriaxone (J01DD04)	8.2 (7.0, 9.4)	5.2 (2.1, 13.1)	5.6 (2.8, 11.0)	154

Therapeutic class	Drug (ATC code)	RR_{LGPS} (95% CI)	OR (95% CI)	IRR_{SCCS} (95% CI)	No. of excess cases
Nonsteroidal anti-inflammatory drug	Ketorolac (M01AB15)	4.6 (3.9, 5.3)	2.7 (2.1, 3.4)	2.8 (1.7, 4.7)	135
Other anti-anemic	Darbepoetin alfa (B03XA02)	3.3 (2.8, 3.8)	1.7 (1.4, 2.1)	3.2 (1.8, 5.6)	126
Systemic antibacterial	Cefixime (J01DD08)	3.1 (2.6, 3.6)	2.4 (1.9, 3.1)	4.4 (3.1, 6.2)	104
Systemic antibacterial	Roxithromycin (J01FA06)	3.3 (2.8, 3.9)	2.4 (1.9, 3.0)	2.9 (1.8, 4.9)	89
Opioid analgesic	Ketobemidone and antispasmodics (N02AG02)	2.2 (1.9, 2.6)	1.2 (1.0, 1.5)	2.6 (1.2, 5.7)	78
Systemic antibacterial	Dicloxacillin (J01CF01)	2.6 (2.2, 3.1)	1.8 (1.5, 2.2)	2.5 (1.2, 5.2)	73
Antiemetic/gastric prokinetic	Cisapride (A03FA02)	2.1 (1.8, 2.5)	1.2 (1.0, 1.5)	2.4 (1.6, 3.6)	69
Antineoplastic/ immunomodulator	Azathioprine (L04AX01)	2.1 (1.7, 2.4)	1.2 (1.1, 1.5)	3.4 (1.9, 6.1)	69
Oral hypoglycemic agent	Gliquidone (A10BB08)	2.7 (2.2, 3.2)	2.1 (1.6, 2.7)	2.2 (1.2, 4.0)	66
Systemic antibacterial	Erythromycin (J01FA01)	3.7 (3.0, 4.6)	2.6 (1.9, 3.4)	2.4 (1.1, 5.1)	63
Systemic antifungal	Fluconazole (J02AC01)	2.7 (2.2, 3.3)	1.5 (1.2, 2.0)	2.2 (1.18, 4.4)	53
Phosphate binder	Polystyrene sulfonate (V03AE01)	4.8 (3.6,6.4)	2.2 (1.5, 3.1)	3.3 (1.0, 10.3)	48
Antiemetic/gastric prokinetic	Butylscopolamine (A03BB01)	5.8 (4.2, 7.7)	2.1 (1.4, 3.4)	11.3 (5.0, 25.8)	45
Antineoplastic/ immunomodulator	Megestrol (L02AB01)	3.2 (2.5, 4.0)	2.5 (1.8, 3.4)	4.0 (1.8, 9.3)	44
Systemic antibacterial	Ceftibuten (J01DD14)	2.3 (1.8, 3.0)	1.9 (1.3, 2.7)	3.0 (1.7, 5.2)	31
Systemic antibacterial	Rokitamycin (J01FA12)	2.6 (1.8, 3.7)	1.8 (1.1, 3.0)	4.3 (2.3, 8.0)	18

*Associations with: (i) with RRLGPS≥2 and lower 95% CI of RRLGPS>1; (ii) unlikely to be due to protopathic bias; (iii) lower 95% CI of OR>1; (iv) lower 95% CI of IRRSCCS>1

Table 3 | Drugs comprising potential signals for increased risk of acute myocardial infarction which passed the signal filtering (novelty criterion) and signal substantiation (plausibility criterion) procedures.

Drugs that satisfied both novelty and plausibility criteria	Drugs that satisfied only novelty criterion
Metoclopramide	Combinations of aluminium, magnesium, and calcium salts
Cisapride	Magaldrate
Domperidone	Butylscopolamine
Betamethasone	Gliquidone
Erythromycin	Metformin combinations with sulfonamides
Roxithromycin	Methylprednisolone
Azithromycin	Pivampicillin
Fluconazole	Phenoxymethylpenicillin
Megestrol	Dicloxacillin
	Ceftriaxone
	Cefixime
	Ceftibuten
	Rokitamycin
	Azathioprine
	Ketobemidone and antispasmodics
	Fenoterol (inhaled)
	Salbutamol (inhaled)
	Polystyrene sulfonate

DISCUSSION

We have described a strategy for prioritizing potential safety signals identified from EHR data. We attempted to simulate, using this strategy, how a physician or a regulator would go about evaluating suspected drug-induced adverse events. This is the first signal triage strategy developed for use – and tested – in electronic healthcare data. In this strategy, we take into account public health relevance, novelty, and plausibility in addition to statistical association. In a concept paper sent for public consultation regarding the implementation of the new pharmacovigilance legislation, the EC proposed to determine the evidence contained in a signal using a common methodology that takes into account the 'quantitative strength of the association, consistency of the data, exposure response relationship, biological plausibility, experimental findings, possible analogies and the nature and quality of the data.[56] Similar attributes have been deemed important in evaluating associations, both within and outside pharmacovigilance.[57-58]

Stepwise exclusion of alternative causes for the potential signal is part of an etiology-based approach for the assessment of ADRs and while this is usually inherent in physician-reported ADRs this is not the case with associations obtained from EHR data (particularly with claims

data), which are inferred outside the actual physician-patient encounter. We have tried to simulate this process by assessing how protopathic bias and confounding (due to time-invariant factors as well as comorbidity and overall health status) could explain the signals identified. Although the mechanisms behind most ADRs are still not completely understood, accumulating evidence over the years indicate the interplay of various factors and an increasing role of inter-individual genetic variants, most notably single nucleotide polymorphisms (SNPs), in genes encoding drug metabolizing enzymes and drug target genes.[17] The triage strategy we developed takes into account these various pathways which can lead to a plausible explanation of the identified drug-adverse event associations.

Potential signals identified by the triage strategy

Systemic antibacterials and AMI. Several studies have shown that β-lactam antibiotics may cause allergic reactions and initiate acute coronary syndrome (ACS) in hypersensitive individuals. ACS refers to a spectrum of clinical presentations that include ST-segment elevation myocardial infarction (STEMI), non–ST-segment elevation myocardial infarction (NSTEMI), and unstable angina. The coincidental occurrence of chest pain and allergic-anaphylactic reaction accompanied by clinical and laboratory findings of classical angina pectoris, caused by inflammatory mediators released during an allergic insult, constitutes the so-called Kounis syndrome.[59-60] This allergic angina can progress to AMI, also referred to as 'allergic myocardial infarction.' The clinical manifestations of Kounis syndrome, including electrocardiographic findings, are similar to AMI. The proposed mechanism behind the Kounis syndrome is via the action of cardiac mast cells found in the coronary artery intimal layer and atherosclerotic plaques. It has been demonstrated that the density of mast cells in the culprit atheroma of patients who died from AMI was 200 times higher than the density in normal coronary vessels from the same patients.[61] These mast cells become activated during the allergic reaction and release endogenous mediators, including histamine, leukotrienes, thromboxane, platelet activation factor (PAF), tryptase, chymase, and rennin – all of which affect different receptors on the coronary vessel wall that may result in AMI.[62] Histamine, the main amine released during allergic reactions, increases the movement of calcium ions into the myocardial cells and accelerates phase 4 spontaneous depolarization of the ventricular action potential. The effects of histamine on cardiac function, including increased cardiac contractility and heart rate as well as coronary vasospasm, are mediated via H1- and H2- receptors situated on the four cardiac chambers and coronary arteries. Activation of H1 receptors on endothelial cells stimulates the release of prostaglandins, interleukin 6, and interleukin 8 into the circulation which causes inflammation. Histamine also induces tissue factor expression and contributes to thrombus formation.[63] In addition to direct coronary vasoconstriction and thrombus generating effects, histamine also potentiates the platelet aggregating response to adrenaline. Kounis syndrome has been described with use of penicillin, ampicillin, amoxicillin, cefuroxime, cefoperazone, and cefoxitin.[62] To date, there have been no reports in the literature associating macrolide antibiotics (which include erythromycin, roxithromycin, and azithromycin) with the Kounis syndrome. It is, however, possible that the macrolides produce coronary vasopasm via the same mechanism

as that of the β-lactam antibiotics; after all, macrolides also induce hypersensitivity reactions, although much less frequently than the β-lactams.[64-65] Both immediate-type hypersensitivity (i.e., anaphylaxis) non-immediate reactions like fixed drug eruptions, toxic epidermal necrolysis and leukocytoclastic vasculitis have been reported with the use of macrolide antibiotics.[64, 66] Another possible, perhaps more likely explanation, is related to channeling: macrolide antibiotics are used preferentially in those patients who may be at higher risk for developing hypersensitivity to β-lactam antibiotics, and consequently at risk for developing Kounis syndrome.

Recreational drug use with cocaine and oral contraceptive use in women are the main culprits usually implicated when AMI occurs in a young patient with no clinically evident coronary disease or other known cardiovascular risk factors.[67-68] The foregoing results suggest that systemic antibiotics may have an increasing, or previously unrecognized, role in drug-induced AMI; further studies are needed to explore this.

Gastric prokinetic agents and AMI. Among the gastric prokinetic agents, cisapride has the most well-characterized cardiac adverse effect profile, which includes ventricular arrhythmia, QT prolongation and *torsades de pointes.*[69-71] However, both metoclopramide and domperidone have also been reported to have arrythmogenic potential.[72] Their effects on the cardiovascular system are related to their action on dopaminergic and 5-HT receptors; this could be the same mechanism that predisposes to myocardial ischemia or infarction, although how this may happen is unclear.[73] Most of the literature associating gastric prokinetic agents with AMI are related to the use of these agents in those patients who already have AMI. These anti-emetics are also given to patients as pre-anesthesia medication in procedures such as coronary angiography in patients who are at risk for, or may already have, AMI.[74] In addition, prokinetic agents may be co-administered with morphine in patients presenting with chest pain and suspected AMI.[75] These two scenarios present possible explanations for detecting the association with AMI, although if this were the case this signal would have been likely to be flagged by LEOPARD as being caused by protopathic bias.

Systemic antifungals and AMI. Fluconazole has been associated with cardiac adverse effects including QT prolongation and *torsades de pointes*,[76-77] but not with myocardial ischemia or infarction (at least not directly). Another drug belonging to the same class (itraconazole), has been described as causing a negative inotropic effect resulting in congestive heart failure, with the characteristic triad of hypertension, hypokalemia, and edema.[78-79] The product label of itraconazole has been changed to include a warning to avoid administration to patients with evidence, or history, of heart failure.[80] Itraconazole was filtered out in the initial signal screening, because it had an RR_{LGPS} for AMI of 1.5 (95% CI 1.2, 1.9). There were 17 excess AMI cases attributable to Itraconazole exposure and although it was not flagged by LEOPARD for protopathic bias, the risk estimates obtained using case-control (OR 1.1 with 95% CI 0.8, 1.6) and SCCS (IRR 1.6 with 95%CI 0.9, 3.1) were not considered significant. This, however, does not discount the possibility that the azole antifungals may trigger AMI in those already at risk by modifying lipid profile, an important determinant of cardiovascular risk. The product label of fluconazole indicates that there have been

post-marketing reports of both hypercholesterolemia and hypertriglyceridemia with fluconazole use.[81] There have been no published assessment or validation of such reports,[82] however, and it seems worthwhile to investigate this association further and explore the possible mechanisms for the dyslipidemia. Drug-drug interactions may also play a role in the development of AMI, especially in high-risk patients who are taking multiple cardiac drugs. Because of their mechanism of action, all the azole antifungals inhibit CYP450 enzymes to some degree and may predispose to adverse cardiac complications, including rhythm problems and ischemia or infarction.[83-84]

Systemic corticosteroids and AMI. Lipodystrophy, weight gain, and hypertension are known corticosteroid-induced adverse effects.[85] Hyperlipidemia is usually associated with long-term corticosteroid use. Cases of AMI with use of systemic corticosteroids have also been reported.[86] In a Danish study of 12 089 patients with an out-of-hospital cardiac arrest (OHCA), pharmacotherapy with systemic corticosteroids, bronchodilators, and antipsychotics were found to have the strongest association up to 30 days before OHCA.[87] Similar results were found for associations up to one year before OHCA. The association of AMI with systemic corticosteroid use may also be a reflection of the underlying population of corticosteroid users rather than a direct adverse effect of the drug itself. Systemic corticosteroids are used in patients with systemic lupus erythemaytosus (SLE) and other rheumatologic diseases. Accelerated atherosclerosis and premature coronary artery disease (CAD) are recognized complications of SLE, psoriasis, and other rheumatologic disorders, although the exact etiology remains unclear and is likely to be multifactorial.[88-89] In a UK-wide study involving 53 patients with either AMI or angina pectoris, SLE patients with clinical CAD were more likely to have been treated with corticosteroids (OR 2.46; 95% CI 1.03, 5.88).[90] Patients on corticosteroids are also likely to have associated chronic kidney disease (CKD), which may further predispose them to develop dyslipidemia.[91]

Progesterone derivatives and AMI. Megestrol (marketed as megestrol acetate) is a progesterone derivative that is used for hot flushes, for palliative treatment of some hormone-dependent malignant neoplasms such as advanced carcinoma of the breast, endometrium, and prostate, as well as treatment of anorexia and cachexia in patients with AIDS or cancer. There is not much data in the literature regarding cardiovascular risk of progestogens, except within the context of combination with estrogens in hormonal replacement therapy,[92-95] although some data suggest that progestogens having higher androgenic potency may attenuate the beneficial effects of estrogen on lipid profile and vasomotor response of the coronary arteries.[96-101] Megestrol may predispose to AMI via its effects on known cardiovascular risk factors. Weight gain, hypertension, and hyperglycemia or diabetes mellitus may occur with use of megestrol via glucocorticoid action-mediated increased peripheral insulin resistance (especially with long-term use).[102-105]

New associations but unsubstantiated by automatic procedures

We looked at the signals that were new (i.e., not previously described in the literature) but were not automatically substantiated to recover associations which may not be explained by the drug's pharmacology.[106-108] Doing away with the substantiation requirement, however, yielded drugs that are similar to those described above. Butylscopolamine is another prokinetic agent, methylprednisolone is another corticosteroid, while pivampicillin, phenoxymethylpenicillin, dicloxacillin, ceftriaxone, cefixime, ceftibuten, and rokitamycin are all systemic antibacterials (and all β-lactams except for the last one which is a macrolide). In the same UK study that looked at CAD in SLE patients, azathioprine exposure was more likely in patients who developed CAD compared to those who did not have CAD (OR 2.33; 95% CI 1.16, 4.67).[90] The bronchodilators Fenoterol and Salbutamol may trigger AMI via their action on β- receptors, primarily increased heart rate and cardiac contractility, resulting in increased myocardial oxygen demand and consequent ischemia and infarction. Some newer antidiabetic drugs have been found to increase the risk for AMI, but it seems more likely that the association with gliquidone and metformin reflect the underlying condition in such patients, diabetes being a major risk factor in the development of CAD and AMI (confounding by indication). No plausible mechanisms could be found for the associations involving the antacids, the opioid ketobemidone, and the phosphate binder polysterene sulfonate.

Limitations and next steps

Although we tried to incorporate into this signal triage strategy most of the criteria considered pertinent and relevant, there is still room for some improvement. Refinement of signal detection methods, including optimization of parameter settings to determine which method (or methods) can be best suited for particular types of adverse events, is currently ongoing.[11] Although we tried to take into account global health status and co-morbidities, residual confounding cannot be ruled out. Dose-response relationships, exposure duration and carryover effects, and the effect of concomitant use of other drugs (including drug-drug interactions) were not considered in this triage strategy. Such factors are difficult, but not impossible, to assess in aggregated healthcare data when *all* drugs are being studied (as how it is in signal detection). Having a 'priority list' of signals should make this problem of multidimensionality more manageable. Continuous update of the knowledge/biomedical databases is necessary; many new molecular entities are introduced into the market every year and databases that catalog the pharmacology and toxicology of these drugs (including information on molecular targets and gene associations) have not been able to keep up. Furthermore, although information on these drugs may be available, many of these bioinformatics databases may not be publicly available. We used the ATC drug coding system to harmonize the identification of drugs in the EHR databases within EU-ADR (which have their own drug coding systems). However, there are some drugs, particularly new drugs, which could not be found in the bioinformatics databases using the ATC-based algorithms developed. This issue of attribution is also a problem when looking at fixed drug combinations, because drug information is usually catalogued per individual active drug ingredient. While the automated signal filtering and signal substantiation streamlined the triage and greatly reduced manual work, full automation of the

signal prioritization process is still not possible at this time. Manual verification of the output produced by these workflows, in terms of both accuracy and completeness, remains a crucial step.

Finally, signal detection is, by definition, a hypothesis-generating exercise. Formal pharmacoepidemiologic studies to investigate the associations identified by the triage algorithm as necessitating follow-up are obvious and necessary next steps. It is not surprising that the results we obtained in this study led to more questions rather than answers. Some questions appear to deserve further attention, such as the issue of Kounis syndrome occurring in other drugs that elicit hypersensitivity reactions.

CONCLUSION

We have proposed a signal triage strategy for use in signal detection using EHR that goes beyond statistical association to demonstration of public health relevance, novelty, and plausibility. This strategy needs to be further tested using other EHR data sources and other adverse events.

REFERENCES

1. 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.* Nov 2 2010;153(9):600-606.
2. US Food and Drug Administration. Reports Received and Entered into AERS by Year. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm070434.htm. Accessed Last accessed 10 December 2011.
3. The Uppsala Monitoring Centre. WHO Global Database Reaches 7 Million. http://www.who-umc.org/DynPage.aspx?id=105196&mn1=7347&newsid=10596. Accessed [Last accessed 30 December 2011].
4. Meyboom RH, Lindquist M, Egberts AC, Edwards IR. Signal selection and follow-up in pharmacovigilance. *Drug Saf.* 2002;25(6):459-465.
5. Waller P, Heeley E, Moseley J. Impact analysis of signals detected from spontaneous adverse drug reaction reporting data. *Drug Saf.* 2005;28(10):843-850.
6. Heeley E, Waller P, Moseley J. Testing and implementing signal impact analysis in a regulatory setting: results of a pilot study. *Drug Saf.* 2005;28(10):901-906.
7. Levitan B, Yee CL, Russo L, Bayney R, Thomas AP, Klincewicz SL. A model for decision support in signal triage. *Drug Saf.* 2008;31(9):727-735.
8. Lindquist M. Use of triage strategies in the WHO signal-detection process. *Drug Saf.* 2007;30(7):635-637.
9. European Medicines Agency. 2010 pharmacovigilance legislation. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000492.jsp&mid=WC0b01ac058033e8ad&jsenabled=true. Accessed [Last accessed 15 January 2012].
10. Borg JJ, Aislaitner G, Pirozynski M, Mifsud S. Strengthening and rationalizing pharmacovigilance in the EU: where is Europe heading to? A review of the new EU legislation on pharmacovigilance. *Drug Saf.* Mar 1 2011;34(3):187-197.
11. Mazzaglia G, Mantovani LG, Sturkenboom MC, et al. Patterns of persistence with antihypertensive medications in newly diagnosed hypertensive patients in Italy: a retrospective cohort study in primary care. *J Hypertens.* Nov 2005;23(11):2093-2100.
12. Sacchetti E, Trifiro G, Caputi A, et al. Risk of stroke with typical and atypical anti-psychotics: a retrospective cohort study including unexposed subjects. *J Psychopharmacol.* Jan 2008;22(1):39-46.
13. Vlug AE, van der Lei J, Mosseveld BM, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf Med.* Dec 1999;38(4-5):339-344.
14. Verhamme K, Mosis G, Dieleman J, Stricker B, Sturkenboom M. Spironolactone and risk of upper gastrointestinal events: population based case-control study. *BMJ.* Aug 12 2006;333(7563):330.
15. Sturkenboom M, Nicolosi A, Cantarutti L, et al. Incidence of mucocutaneous reactions in children treated with niflumic acid, other nonsteroidal antiinflammatory drugs, or nonopioid analgesics. *Pediatrics.* Jul 2005;116(1):e26-33.
16. Giaquinto C, Van Damme P, Huet F, et al. Clinical consequences of rotavirus acute gastroenteritis in Europe, 2004-2005: the REVEAL study. *J Infect Dis.* May 1 2007;195 Suppl 1:S26-35.
17. Christensen S, Riis A, Norgaard M, Thomsen RW, Sorensen HT. Introduction of newer selective cyclo-oxygenase-2 inhibitors and rates of hospitalization with bleeding and perforated peptic ulcer. *Aliment Pharmacol Ther.* Apr 15 2007;25(8):907-912.
18. Dalton SO, Johansen C, Poulsen AH, et al. Cancer risk among users of neuroleptic medication: a population-based cohort study. *Br J Cancer.* Oct 9 2006;95(7):934-939.
19. Penning-van Beest FJ, Termorshuizen F, Goettsch WG, Klungel OH, Kastelein JJ, Herings RM. Adherence to evidence-based statin guidelines reduces the risk of hospitalizations for acute myocardial infarction by 40%: a cohort study. *Eur Heart J.* Jan 2007;28(2):154-159.
20. Goettsch WG, de Jong RB, Kramarz P, Herings RM. Developments of the incidence of osteoporosis in The Netherlands: a PHARMO study. *Pharmacoepidemiol Drug Saf.* Feb 2007;16(2):166-172.
21. Corrao G, Zambon A, Conti V, et al. Menopause hormone replacement therapy and cancer risk: an Italian record linkage investigation. *Ann Oncol.* Jan 2008;19(1):150-155.

22. Corrao G, Botteri E, Bertu L, Zambon A, Favilli S. Exploring the effect of transient exposure on the risk of acute events by means of time-window designs: an application to fluoroquinolone antibacterials and arrhythmia. *Pharmacoepidemiol Drug Saf.* Jan 2006;15(1):31-37.

23. 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.* May-Jun 2006;30(3):161-168.

24. Monami M, Balzi D, Lamanna C, et al. Are sulphonylureas all the same? A cohort study on cardiovascular and cancer-related mortality. *Diabetes Metab Res Rev.* Sep 2007;23(6):479-484.

25. 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;148:43-49.

26. Coloma PM, Schuemie MJ, Trifiro G, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiol Drug Saf.* Jan 2011;20(1):1-11.

27. Coloma PM, Trifiro G, Schuemie MJ, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? *Pharmacoepidemiol Drug Saf.* Feb 8 2012.

28. Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med.* Aug 1993;32(4):281-291.

29. Avillach P, Mougin 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.

30. Avillach P, Joubert M, Thiessard F, et al. Design and evaluation of 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.* 2010;160(Pt 2):1085-1089.

31. Avillach P, Coloma PM, Gini R, Schuemie MJ, Mougin F, Dufour JC, Mazzaglia G, Giaquinto C, Fornari C, Herings R, Hippisley-Cox J, Molokhia M, Pedersen L, Fourrier-Réglat A, Fieschi M, Sturkenboom M, van Der Lei J, Pariente A, Trifirò G, on behalf of the EU-ADR consortium. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. Submitted.

32. Coloma PM, Valkhoff V, Mazzaglia G, Nielsson MS, Pedersen L, Molokhia M, Mosseveld B, Morabito P, Schuemie MJ, Sturkenboom MCJM, Trifiro G, on behalf of the EU-ADR Consortium. Accuracy of coding-based algorithms in identification of acute myocardial infarction in multi-country electronic healthcare records (EHR) databases. Unpublished.

33. WHO Collaborating Centre for Drug Statistics Methodology. WHO ATC Classification system. http://www.whocc.no/atc/structure_and_principles/.

34. WHO Collaborating Centre for Drug Statistics Methodology. WHO DDD Classification. http://www.whocc.no/ddd/definition_and_general_considera/.

35. Schuemie MJ, Coloma PM, Straatman H, Herings RMC, Trifirò G, Matthews JN, Prieto-Merino D, Molokhia M, Pedersen L, Gini R, Innocenti F, Mazzaglia G, Picelli G, Scotti L, van der Lei J, Sturkenboom MCJM. Using electronic healthcare records for drug safety signal detection: a comparative evaluation of statistical methods. *Submitted.* 2012.

36. Schuemie MJ. Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD. *Pharmacoepidemiol Drug Saf.* Mar 2011;20(3):292-299.

37. Avillach P, Coloma PM, Bauer-Mehren A, Gini R, Thiessard F, Trifirò G, Schuemie MJ, Sturkenboom MCJM, Oliveira JL, Sanz F, Molero E, Díaz C, and van der Lei J, on behalf of the EU-ADR Group. Design, development and validation of a computerised system that exploits data from electronic health records and biomedical information for the early detection of adverse drug reactions. The EU-ADR project: Preliminary Results. *13th World Congress on Medical and Health Informatics MEDINFO*. Cape Town, South Africa2010.

38. van Mulligen EM AF-R, D Gurwitz, et al. The EU-ADR Corpus: Annotated Drugs, Diseases, Targets, and their Relationships. *J Biomed Inform.* 2012;Accepted.

39. Bauer-Mehren A, van Mulligen E, Avillach P, Coloma PM, Garcia-Serna R, Herings R, Sturkenboom M, Boyer S, Sanz F, Mestres J, Kors J, Furlong L. Substantiation of Drug Safety Signals: Taking Automated Signal Detection a Step Further. *Presented at the 26th International Conference on Pharmacoepidemiology and Therapeutic Risk Management*. Brighton, United Kingdom2010.

40. Bauer-Mehren A, van Mulligen E, Avillach P, et al. . Automatic filtering and substantiation of drug safety signals. *PLoS Comp Biol.* 2012;Accepted.
41. Oinn T, Addis M, Ferris J, Marvin D, Senger M, Greenwood M, Carver T, Glover K, Pocock, MR WA, et al. Taverna: a tool for the composition and enactment of bioinformatics workflows. *Bioinformatics* 2004;20(17):3045-3054.
42. The EU-ADR Project. EU-ADR Web Platform. https://bioinformatics.ua.pt/euadr/.
43. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourrier-Réglat A, Molokhia M, Patadia V, van der Lei J, Sturkenboom M, and Trifirò G, on behalf of the EU-ADR consortium. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection using Electronic Healthcare Records (EHR) Databases. *Drug Safety.* 2012;Accepted.
44. Coloma PM, Schuemie MJ, Trifiro G, et al. Comparison of methods for drug safety signal detection using electronic healthcare record (EHR) databases: the added value of longitudinal, time-stamped patient information. Paper presented at: Presented at the 27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management2011; Chicao, Illinois, USA.
45. WHO Collaborating Centre for Drug Statistics Methodology. http://www.whocc.no/atc/structure_and_principles/. Accessed 10 November 2011.
46. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *Am J Epidemiol.* Nov 1 2001;154(9):854-864.
47. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. *Stat Med.* May 30 2006;25(10):1768-1797.
48. Stricker BH, Psaty BM. Detection, verification, and quantification of adverse drug reactions. *BMJ.* Jul 3 2004;329(7456):44-47.
49. Evans M, Rees A. Effects of HMG-CoA reductase inhibitors on skeletal muscle: are all statins the same? *Drug Saf.* 2002;25(9):649-663.
50. McGettigan P, Henry D. Cardiovascular risk with non-steroidal anti-inflammatory drugs: systematic review of population-based controlled observational studies. *PLoS Med.* Sep 2011;8(9):e1001098.
51. Trifiro G, Pariente A, Coloma PM, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf.* Dec 2009;18(12):1176-1184.
52. Hauben M, Aronson JK. Defining ‚signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions. *Drug Saf.* 2009;32(2):99-110.
53. Pirmohamed M, Breckenridge AM, Kitteringham NR, Park BK. Adverse drug reactions. *BMJ.* Apr 25 1998;316(7140):1295-1298.
54. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet.* Oct 7 2000;356(9237):1255-1259.
55. Park BK, Kitteringham NR, Powell H, Pirmohamed M. Advances in molecular toxicology-towards understanding idiosyncratic drug toxicity. *Toxicology.* Nov 16 2000;153(1-3):39-60.
56. European Commission Health and Consumers Directorate General. Implementing measures in order to harmonise the performance of the pharmacovigilance activities provided for in Directive 2001/83/EC and Regulation (EC) No. 726/2004. The concept paper submitted for public consultation. 08 September 2011 http://ec.europa.eu/health/files/pharmacovigilance/2011-09_concept-paper.pdf. Accessed [Last accessed 30 December 2011]
57. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med.* May 1965;58:295-300.
58. Caster O, Edwards IR. Reflections on attribution and decisions in pharmacovigilance. *Drug Saf.* Oct 1 2010;33(10):805-809.
59. Biteker M. Current understanding of Kounis syndrome. *Expert Rev Clin Immunol.* Sep 2010;6(5):777-788.
60. Kounis NG, Zavras GM. Histamine-induced coronary artery spasm: the concept of allergic angina. *Br J Clin Pract.* Summer 1991;45(2):121-128.
61. Kovanen PT, Kaartinen M, Paavonen T. Infiltrates of activated mast cells at the site of coronary atheromatous erosion or rupture in myocardial infarction. *Circulation.* Sep 1 1995;92(5):1084-1088.
62. Ridella M, Bagdure S, Nugent K, Cevik C. Kounis syndrome following beta-lactam antibiotic use: review of literature. *Inflamm Allergy Drug Targets.* Mar 2009;8(1):11-16.

63. Steffel J, Akhmedov A, Greutert H, Luscher TF, Tanner FC. Histamine induces tissue factor expression: implications for acute coronary syndromes. *Circulation.* Jul 19 2005;112(3):341-349.

64. Araujo L, Demoly P. Macrolides allergy. *Curr Pharm Des.* 2008;14(27):2840-2862.

65. Thong BY. Update on the management of antibiotic allergy. *Allergy Asthma Immunol Res.* Apr 2010;2(2):77-86.

66. Demoly P, Gomes ER. Drug hypersensitivities: definition, epidemiology and risk factors. *Eur Ann Allergy Clin Immunol.* Jun 2005;37(6):202-206.

67. Rezkalla SH, Kloner RA. Cocaine-induced acute myocardial infarction. *Clin Med Res.* Oct 2007;5(3):172-176.

68. Rubin JB, Borden WB. Coronary heart disease in young adults. *Curr Atheroscler Rep.* Apr 2012;14(2):140-149.

69. Quigley EM. Cisapride: what can we learn from the rise and fall of a prokinetic? *J Dig Dis.* Jun 2011;12(3):147-156.

70. Keller GA, Di Girolamo G. Prokinetic agents and QT prolongation: a familiar scene with new actors. *Curr Drug Saf.* Jan 2010;5(1):73-78.

71. Hennessy S, Leonard CE, Newcomb C, Kimmel SE, Bilker WB. Cisapride and ventricular arrhythmia. *Br J Clin Pharmacol.* Sep 2008;66(3):375-385.

72. Collins KK, Sondheimer JM. Domperidone-induced QT prolongation: add another drug to the list. *J Pediatr.* Nov 2008;153(5):596-598.

73. Tonini M, De Ponti F, Di Nucci A, Crema F. Review article: cardiac adverse effects of gastrointestinal prokinetics. *Aliment Pharmacol Ther.* Dec 1999;13(12):1585-1591.

74. Lencz L. Role of domperidone in improving intestinal activity in acute myocardial infarction patients. *Ther Hung.* 1991;39(3):145-147.

75. Simpson PM, Bendall JC, Middleton PM. Review article: Prophylactic metoclopramide for patients receiving intravenous morphine in the emergency setting: a systematic review and meta-analysis of randomized controlled trials. *Emerg Med Australas.* Aug 2011;23(4):452-457.

76. Khazan M, Mathis AS. Probable case of torsades de pointes induced by fluconazole. *Pharmacotherapy.* Dec 2002;22(12):1632-1637.

77. McMahon JH, Grayson ML. Torsades de pointes in a patient receiving fluconazole for cerebral cryptococcosis. *Am J Health Syst Pharm.* Apr 1 2008;65(7):619-623.

78. Ahmad SR, Singer SJ, Leissa BG. Congestive heart failure associated with itraconazole. *Lancet.* Jun 2 2001;357(9270):1766-1767.

79. Sharkey PK, Rinaldi MG, Dunn JF, Hardin TC, Fetchick RJ, Graybill JR. High-dose itraconazole in the treatment of severe mycoses. *Antimicrob Agents Chemother.* Apr 1991;35(4):707-713.

80. US National Library of Medicine. DailyMed. FDA-approved label Itraconazole. http://dailymed.nlm.nih.gov/dailymed/lookup.cfm?setid=a4d555fa-787c-40fb-bb7d-b0d4f7318fd0. Accessed [Last accessed 10 January 2012].

81. US National Library of Medicine. DailyMed. FDA-approved label Fluconazole. http://dailymed.nlm.nih.gov/dailymed/lookup.cfm?setid=f694c617-3383-416c-91b6-b94fda371204. Accessed [Last accessed 10 January 2012.]

82. eHealthMe. Real world drug outcomes. Does fluconazole cause hypercholesterolemia? ; http://www.ehealthme.com/ds/fluconazole/hypercholesterolemia. Accessed [Last accessed 30 January 2012.].

83. Groll AH, Piscitelli SC, Walsh TJ. Clinical pharmacology of systemic antifungal agents: a comprehensive review of agents in clinical use, current investigational compounds, and putative targets for antifungal drug development. *Adv Pharmacol.* 1998;44:343-500.

84. Brüggemann RJ, Alffenaar J, Blijlevens NM, Billaud EM, Kosterink JG, Verweij PE, Burger DM. Clinical relevance of the pharmacokinetic interactions of azole antifungal drugs with other coadministered agents. *Clin Infect Dis.* 2009 May 15;48(10):1441-1458.

85. Stone N. Secondary causes of hyperlipidemia. *Med Clin North Am.* 1994;78(1):117-141.

86. Sarnes E, Crofford L, Watson M, Dennis G, Kan H, Bass D. Incidence and US costs of corticosteroid-associated adverse events: a systematic literature review. *Clin Ther.* Oct 2011;33(10):1413-1432.

87. Weeke P, Folke F, Gislason GH, et al. Pharmacotherapy and hospital admissions before out-of-hospital cardiac arrest: a nationwide study. *Resuscitation.* Dec 2010;81(12):1657-1663.

88. Gisondi P, Girolomoni G. Psoriasis and atherothrombotic diseases: disease-specific and non-disease-specific risk factors. *Semin Thromb Hemost.* Apr 2009;35(3):313-324.

89. Szekanecz Z, Kerekes G, Der H, et al. Accelerated atherosclerosis in rheumatoid arthritis. *Ann N Y Acad Sci.* Jun 2007;1108:349-358.

90. Haque S, Gordon C, Isenberg D, et al. Risk factors for clinical coronary heart disease in systemic lupus erythematosus: the lupus and atherosclerosis evaluation of risk (LASER) study. *J Rheumatol.* Feb 2010;37(2):322-329.

91. Lacquaniti A, Bolignano D, Donato V, Bono C, Fazio MR, Buemi M. Alterations of lipid metabolism in chronic nephropathies: mechanisms, diagnosis and treatment. *Kidney Blood Press Res.* 2010;33(2):100-110.

92. Kim J, Evans S, Smeeth L, Pocock S. Hormone replacement therapy and acute myocardial infarction: a large observational study exploring the influence of age. *Int J Epidemiol.* Jun 2006;35(3):731-738.

93. Bernstein P, Pohost G. Progesterone, progestins, and the heart. *Rev Cardiovasc Med.* Fall 2010;11(4):228-236.

94. Schneider C, Jick SS, Meier CR. Risk of cardiovascular outcomes in users of estradiol/dydrogesterone or other HRT preparations. *Climacteric.* Oct 2009;12(5):445-453.

95. Sare GM, Gray LJ, Bath PM. Association between hormone replacement therapy and subsequent arterial and venous vascular events: a meta-analysis. *Eur Heart J.* Aug 2008;29(16):2031-2041.

96. Keeling D. Combined oral contraceptives and the risk of myocardial infarction. *Ann Med.* 2003;35(6):413-418.

97. Chakhtoura Z, Canonico M, Gompel A, Scarabin PY, Plu-Bureau G. Progestogen-only contraceptives and the risk of acute myocardial infarction: a meta-analysis. *J Clin Endocrinol Metab.* Apr 2011;96(4):1169-1174.

98. Rosendaal FR, Van Hylckama Vlieg A, Tanis BC, Helmerhorst FM. Estrogens, progestogens and thrombosis. *J Thromb Haemost.* Jul 2003;1(7):1371-1380.

99. Nilsson SE, Fransson E, Brismar K. Relationship between serum progesterone concentrations and cardiovascular disease, diabetes, and mortality in elderly Swedish men and women: An 8-year prospective study. *Gend Med.* Sep 2009;6(3):433-443.

100. Al-Azzawi F, Thompson J, Stevenson J. Which progestogen is more likely to increase the risk of fatal myocardial infarction: a combination of epidemiological and trial evidence. *Maturitas.* May 20 2006;54(2):154-163.

101. Sitruk-Ware R. Progestogens in hormonal replacement therapy: new molecules, risks, and benefits. *Menopause.* Jan-Feb 2002;9(1):6-15.

102. Panwalker AP. Hyperglycemia induced by megestrol acetate. *Ann Intern Med.* May 15 1992;116(10):878.

103. Henry K, Rathgaber S, Sullivan C, McCabe K. Diabetes mellitus induced by megestrol acetate in a patient with AIDS and cachexia. *Ann Intern Med.* Jan 1 1992;116(1):53-54.

104. Leinung MC, Liporace R, Miller CH. Induction of adrenal suppression by megestrol acetate in patients with AIDS. *Ann Intern Med.* Jun 1 1995;122(11):843-845.

105. Kilby JM, Tabereaux PB. Severe hyperglycemia in an HIV clinic: preexisting versus drug-associated diabetes mellitus. *J Acquir Immune Defic Syndr Hum Retrovirol.* Jan 1 1998;17(1):46-50.

106. Uetrecht J. Idiosyncratic drug reactions: past, present, and future. *Chem Res Toxicol.* Jan 2008;21(1):84-92.

107. Uetrecht J. Idiosyncratic drug reactions: current understanding. *Annu Rev Pharmacol Toxicol.* 2007;47:513-539.

108. Uetrecht J. Evaluation of which reactive metabolite, if any, is responsible for a specific idiosyncratic reaction. *Drug Metab Rev.* 2006;38(4):745-753.

Chapter 6

Summary Discussion and Perspectives

The ultimate test for pharmacovigilance systems is the demonstration of public health benefit and it is this test which signal detection methodologies need to meet if expectations of all stakeholders are to be fulfilled.

– CIOMS Working Group VIII

General Discussion

The concept of a signal, from a drug surveillance point of view has evolved from its definition by the World Health Organization (WHO) in 2002[1] to a more synthesized and comprehensive definition proposed by Hauben and Aronson:[2] (1) It is based on information from one or more sources (including observations and experiments), suggesting an association (either adverse or beneficial) between a drug or intervention and an event or set of related events (e.g., a syndrome); (2) It represents an association that is new and important, or a new aspect of a known association, and has not been previously investigated and refuted; and (3) It demands investigation, being judged to be of sufficient likelihood to justify verificatory and, when necessary, remedial actions. It is thus evident that a signal in pharmacovigilance may – and will – arise from various data sources. It has been posited that electronic healthcare record (EHR) databases represent an important resource for safety signal detection and active surveillance and can augment existing pharmacovigilance systems.[3] Various public-private initiatives worldwide have been launched, and great investments have been made, to explore the secondary use of EHR for this purpose. In this thesis, we have drawn from the experience of the EU-ADR network (Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge, http://www.euadr-project.org/), a federation of eight EHR databases in four countries in Europe, to demonstrate the feasibility of combining diverse and differently structured data and pave the way for pro-active signal detection. This database network represents a resource of unprecedented size for monitoring of drug safety in Europe, containing healthcare data from more than 21 million individuals with over 154 million person-years (PYs) of follow-up. In the following discussion we give a summary of the opportunities and challenges that come with heterogeneity in database structure, with differences in language and coding of both drugs and diseases, and with the diversity in the organization of European healthcare systems.

MAIN FINDINGS

In our review of existing pharmacovigilance systems we reiterated the important contribution of spontaneous reporting systems (SRS) and prescription event monitoring systems (PEM) to post-marketing surveillance. Despite their well-described limitations, spontaneous reports have unquestionably played a major role in highlighting potential safety signals post-marketing and have provided supporting evidence for various regulatory actions. SRS gather real-life data on marketed drugs and, when review of individual case reports or case-series analysis is possible, may permit the identification of potential safety concerns. There are numerous examples of signals that have been generated or reinforced through SRS.[4-9] PEM has, likewise, identified some important signals.[10-11] Active ascertainment with PEM supplements SRS by providing prescription follow-up information that can serve as a denominator for calculating event rates, but important gaps remain. One of the principal objectives of exploring EHR databases for safety surveillance is to be

able to fill some of these gaps. EHR-based signal detection systems are not subject to many of the limitations of SRS, including the lack of 'denominator,' underreporting, variable and incomplete reporting, or lack of information regarding overall health status prior to and after a suspected ADR. Potential risk associated with drug use should be measured both in terms of risk to the individual and the population frequency, which requires knowledge of the level and duration of exposure. Time-stamped, population-based healthcare data enables calculation of background incidence rates and allows comparison of event rates before and after drug exposure – information which spontaneous reports are not able to provide. With the use of proper methodologies, the longitudinal nature of the data may allow identification of adverse events that have a long delay between exposure and clinical manifestations (e.g., cardiac valvulopathy or cancer), especially in databases with long patient follow-up and low turnover. While most spontaneous reports usually involve newly marketed drugs, EHR data can highlight new risks associated with old drugs (as a consequence of new indications of use or new generation of users), as well as adverse events that have high background incidence rates (such as acute myocardial infarction) and events which are less likely to be suspected as drug-induced, thus less likely to be reported. Furthermore, data from EHR provide greater detail regarding patient demographics, drug use, and utilization of healthcare services which permit evaluation of the benefit-risk profile of drugs, hence putting safety issues in a broader perspective and fostering sound regulatory decisions.

There have been many efforts to evaluate the usefulness of EHR databases for drug safety signal detection, including the adaptation of signal detection algorithms developed in SRS for use in EHR and the development of novel signal detection methods that can be employed specifically within the context of EHR.[12-15] Within the last five years international collaborations have been forged to put together multiple EHR databases and develop EHR-based drug safety signal detection systems.[16-17]

Where and how to focus efforts in safety surveillance using electronic healthcare data

Active safety surveillance using EHR databases being an emerging science still in its infancy, there are important issues that are being encountered for the first time. When using data mining to detect signals in EHR databases, a decision needs to be made about the type of approach, which can be drug-based or event-based. In a drug-based approach, a set of specific drugs are monitored for their association with all possible events. In an event-based approach, a set of specific events are evaluated for their association with all possible drugs. In the EU-ADR project, we adopted an event-based approach because the definition of drugs across databases in various countries can more easily be harmonized compared to the definition of events. Moreover, the event-based approach allows us to focus on those events that are considered important from a public health perspective irrespective of the drug associated with the event. There is, however, no literature to date which prescribes what would be the primary events of interest for intensive monitoring in pharmacovigilance when applying data mining techniques. We laid the groundwork in a study where we created such a list of events ranked by importance. Peer-reviewed publications, medical

textbooks, and websites of regulatory agencies were reviewed to create a preliminary list of events that are deemed important in pharmacovigilance. Two teams of pharmacovigilance experts independently rated each event on five criteria: 'trigger for drug withdrawal'; 'trigger for black box warning'; 'leading to emergency department visit or hospital admission'; 'probability of event to be drug-related'; and 'likelihood of death'. An initial list comprising 23 adverse events was identified and a ranked list was subsequently established. The five top-ranked events were: (1) bullous eruptions (collectively consisting of Stevens Johnson syndrome, toxic epidermal necrolysis, and erythema multiforme); (2) acute renal failure; (3) anaphylactic shock; (4) acute myocardial infarction; and (5) rhabdomyolysis. This shortlist of events served as the initial focus of signal detection in the EU-ADR project. Although the prioritization of adverse events for drug safety monitoring done in this study was based on thorough evaluation of evidence from various sources of information by pharmacovigilance experts, this list of events and their ranking is by no means definitive and will need to be updated as new data comes along.

Data from EHR can be used to monitor drug safety, but in order to compare and pool data from various databases of different countries and constructs, the extraction of potential adverse events must be harmonized. Each of the eight databases participating in EU-ADR has unique characteristics depending on its primary objective and local function (i.e., administrative/claims, medical records) and contains medical information coded according to different languages and terminologies. For these reasons, queries for data extraction concerning potential adverse events had to be created based on local expertise. Due to structural, syntactic and semantic heterogeneity of the databases, it was not possible to construct a single query for data extraction that could be used as such in all databases. In the context of large-scale drug safety monitoring using EHRs, the event data extraction from different databases required harmonization, i.e., a process geared towards reaching a common definition of events which is both clinically sound and agreeable to all stakeholders. Such a process would also facilitate transparency in the extraction of the events of interest and understanding of differences between databases. We described in another study the iterative harmonization process for the data extraction concerning the five top-ranked events mentioned above. We used the Unified Medical Language System (UMLS) to identify concepts and corresponding codes in each terminology. We utilized a common database model, implemented via custom-built software Jerboa©, to share and pool data and to verify the semantic basis of the event identification queries. Feedback interaction with the database owners was employed at various stages to refine the extraction queries. We used age-adjusted incidence rates to support harmonization of data extraction processes across the databases. We obtained the following age- and sex-adjusted incidence rates for the five events of interest across databases: (1) acute myocardial infarction 60-148/100 000 PYs; (2) acute renal failure 3-49/100 000 PYs; (3) anaphylactic shock 2-12/100 000 PYs; (4) bullous eruptions 2-17/100 000 PYs; and (5) rhabdomyolysis 0.1-8/100 000 PYs. This study showed how event extractions may differ across databases and how different choices impact on the estimated incidence of the events. The knowledge described in the various terminologies, which are included in the UMLS, was inadequate to define all the clinical aspects of an event for our purposes, and so expert knowledge and experience from the database holders

were necessary to build a more comprehensive definition of the event. This study further reiterated that use of EHR databases requires an understanding of how the healthcare data are generated from the initial patient encounter all the way to completion of the database entry.

In a proof-of-concept study we described the framework and process of combining data from eight European EHR databases of different countries and origins (medical records, administrative registries, record-linkage databases) to scale up safety signal detection. In combining the data of the various databases it was crucial to take into account ethical issues regarding the processing of anonymized healthcare data. The databases in EU-ADR have well-developed safeguard mechanisms ensuring compliance with the European directives and national regulations as well as database governance rules. Since no new data are collected, other than those made available by the participating databases, the cornerstone of ensuring proper ethical and legal conduct is set down in the rules and regulations that govern each database. Rather than imposing a one-size-fits-all approach and compel people to change their data, we leveraged on the diversity of the databases to use local expertise and maximize extraction of relevant information. Within this distributed network system, databases retain ownership of their respective data, extraction being done locally, and only the aggregated, non-identifiable data are shared with the rest of the network. Revisiting the known association of upper gastrointestinal bleeding (UGIB) and nonsteroidal anti-inflammatory drug (NSAID) use, we have shown that data sharing can take place within a distributed networking system to provide consistent incidence rates and detect a known drug-adverse event association. Crude incidence rates of UGIB varied between 38.8 and 109.5/100 000 person-years (PYs), depending on country and type of database, while age-standardized rates ranged from 25.1 to 65.4/100 000 person-years. A statistically significant age- and sex-adjusted association between use of any NSAID and increased risk for UGIB was confirmed in all databases, incidence rate ratio (IRR) from 2.0 to 4.3, which is consistent with what has been described in the literature.[18]

Is there enough leverage?

Data mining on healthcare databases generates statistical associations between drug exposure and adverse outcomes and may be able to augment the current passive–reactive system and facilitate earlier detection of potential safety issues. We have shown in the proof-of-concept study that such an approach for safety surveillance is promising; now 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. 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.[19] 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. We then estimated 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 incidence rates of events as observed in EU-ADR. Data from almost 20 million individuals with 60 million PYs of follow-up who used 2 289 drugs in the EU-ADR network showed 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% of total) for which an association of same magnitude can be investigated. Our findings showed that (1) 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 (2) such a system has more power to detect signals with lower strength of association for relatively frequent events such as acute myocardial infarction and upper gastrointestinal bleeding. 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. 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 about 50% for relatively frequent events such as acute myocardial infarction and upper gastrointestinal bleeding and less than 10% for rare events such as rhabdomyolysis. 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.

Because the prevalence and nature of many diseases are not the same in children as in adults and because of age-related variation in drug pharmacology and drug utilization,[20-24] we performed a separate study to determine the capabilities of the system for pediatric drug safety signal detection. Our findings reveal that drug use in children is rare, <1% of the total drugs prescribed or dispensed make up 50% of the total exposure time to drugs in the pediatric population. For a more frequent event such as upper gastrointestinal bleeding (IR= 14.4/100 000 PYs), there were 39 drugs (66% of total exposure in PYs) for which an association with RR≥ 4, if present, can be investigated. For rare events such as anaphylactic shock and bullous eruptions, there were 8 drugs (35% of total exposure) and 9 drugs (37% of total exposure) respectively, for which an association of same magnitude can be investigated. Based on literature-derived incidence rates, there is a higher number of drugs that can be monitored at the same magnitudes of risk for the following events known to be more frequently occurring in children: febrile convulsions; suicide attempt; and epilepsy.

How do we know the system works?

Although several drugs have been withdrawn post-marketing by regulatory authorities after scientific evaluation of harms and benefits, there is no definitive list of confirmed signals (i.e., list of all known adverse reactions and which drugs can cause them). As there is no true gold standard, prospective evaluation of signal detection methods remains a challenge. Within the context of methods development and evaluation in the EU-ADR Project, we proposed a surrogate reference standard of drug-adverse event associations based on existing scientific literature and

expert opinion. The reference standard was constructed for 10 top-ranked events judged as important in pharmacovigilance. A stepwise approach was employed to identify which among a list of drug-event associations are well-recognized (known positive associations) or highly unlikely ('negative controls') based on MEDLINE-indexed publications, drug product labels, spontaneous reports made to the World Health Organization's pharmacovigilance database, and expert opinion. Only drugs with adequate exposure in the EU-ADR database network to allow detection of an association were considered. Manual verification of positive associations and negative controls was independently performed by two experts proficient in clinical medicine, pharmacoepidemiology, and pharmacovigilance. A third expert adjudicated equivocal cases and arbitrated any disagreement between evaluators. 94 drug-event associations comprised the reference standard, which included 44 positive associations and 50 negative controls for the 10 events of interest: (1) bullous eruptions; (2) acute renal failure; (3) anaphylactic shock; (4) acute myocardial infarction; (5) rhabdomyolysis; (6) aplastic anemia/pancytopenia; (7) neutropenia/agranulocytosis; (8) cardiac valve fibrosis; (9) acute liver injury; and (10) upper gastrointestinal bleeding. For cardiac valve fibrosis, there was no drug with adequate exposure in the database network that satisfied the criteria for a positive association. We emphasize that this reference standard should be considered dynamic and as knowledge on drug safety evolves over time and new issues in drug safety arise, this reference standard must be re-assessed periodically.

To evaluate the relative performance of different statistical methods for detecting drug-adverse event associations in EHR data, we computed relative risk (RR) estimates for drug-event pairs using 10 different methods, including those developed for spontaneous reporting systems, cohort methods such as the Longitudinal Gamma Poisson Shrinker (LGPS), and case-based methods like case control. The method 'Longitudinal Evaluation of Observational Profiles of Adverse events Related to Drugs' (LEOPARD), developed in EU-ADR,[13] was used to remove associations likely due to protopathic bias. Data from the different databases was combined by pooling of data and by meta-analysis for random effects. We used the previously mentioned surrogate reference standard of known positive associations and 'negative controls' to evaluate method performance. We calculated the area under the receiver operating characteristic (AUC) curve (which is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one) for each method, both with and without LEOPARD filtering. The highest AUC (0.83) was achieved by the combination of either LGPS or case-control method with LEOPARD filtering, but the differences in performance among the different methods were marginal. LEOPARD increased overall performance, but flagged several known ADRs as caused by protopathic bias. Based on this preliminary evaluation, we concluded that combinations of methods demonstrate good performance in distinguishing known ADRs from negative controls, and that these methods could also be used to detect new drug safety signals.

Accuracy of outcome assessment is crucial to ensure validity when mining multiple EHR databases for drug safety signal detection. In twin studies, we evaluated the accuracy of various coding-based algorithms used to identify cases of upper gastrointestinal bleeding (UGIB) and acute myocardial infarction (AMI) in the databases in EU-ADR. We conducted two separate

validation studies in four databases of the EU-ADR network (only the first three were used for AMI validation): (1) IPCI (general practitioner (GP) records, Netherlands); (2) HSD (GP records, Italy); 3) Aarhus (administrative claims, Denmark); and (4) ARS (administrative claims, Tuscany region in Italy). We identified cases from GP and medical specialist diagnoses and from primary hospital discharge diagnoses and death registries using pre-defined coding algorithms in three disease terminology systems: (1) International Classification of Primary care (ICPC); (2) International Classification of Diseases-9th revision-clinical modification (ICD9-CM); and (3) ICD 10th revision. We also employed free text using key words consistent with the UMLS concepts for both AMI and UGIB. A random sample of 200 cases per database was obtained from all potential cases identified. Since free text search was known to be extensively used in IPCI, additional 200 potential cases identified by free text search was obtained in IPCI. Manual review of medical records and hospitalization charts was performed using a standardized questionnaire implemented as computerized data entry algorithm using custom-built software Chameleon,© locally installed in each database. Positive predictive values (PPV) were calculated overall and for each code and free text query. Overall case retrieval was good: 93.5% for AMI and 98.5% for UGIB. For UGIB, there were considerable differences in the PPV of the different algorithms, with ICD9-CM-based coding showing the highest PPV (77-90%), followed by ICD10 (66%). There was no difference in the PPV for UGIB between ICPC and free text search in IPCI (25% and 23%, respectively). For HSD, free text search had a 50% PPV for UGIB. For AMI, all ICD-10 codes used had 100% PPV. Overall the ICD9-CM codes had very good PPV, with the most frequently occurring code having a PPV of 96.5%. The ICPC coding-based algorithm had a PPV for AMI of 75%, while use of free text had a lower PPV: 60% in HSD and 19.7% in IPCI. No data regarding PPV of ICPC codes was available from the literature to compare our findings with, but the PPV estimates for ICD-9CM and ICD-10 for both AMI and UGIB were compatible with what has been cited in the literature. These two studies showed that EHR databases present a potentially good source of identifying patients with AMI and UGIB and that further studies are needed to optimize the value of free text search in the identification of events in EHR databases.

To further explore how EHR databases can augment safety surveillance, we responded to a request from the Dutch Medicines Evaluation Board to study the background epidemiology of progressive multifocal leukoencephalopathy (PML). PML is a rare central nervous system demyelinating disease caused by reactivation of the JC virus and occurs almost exclusively in immunosuppressed individuals such as patients with HIV/AIDS, leukemia, tumors, or those undergoing organ transplants. The disease has recently generated widespread concern after reports of PML allegedly developing after treatment with several new biologic agents. Because of its rarity, few data are available regarding incidence of PML in non-immunocompromised populations. Using demographic and clinical data from six databases in three countries within EU-ADR, we provided estimates of the incidence of PML in the general population. Cases of PML were identified from primary and secondary hospital discharge diagnoses as well as death registries (claims databases) and from general practitioner or specialist diagnoses (medical records databases) using diagnostic codes and free text. Case validation by medical chart/records review was performed in a

subset of cases. We identified a total of 1 150 cases of PML from data representing over 21 million individuals with 154 million PYs of follow-up. We concluded that PML is a rare disease among the general European population (overall incidence rate was 0.75 per 100 000 PYs, age-standardized rate across the databases ranged from 0.03 –2.29 per 100 000 PYs) and its incidence is largely associated with HIV disease, malignancy, and autoimmune conditions.

How to make sense of it all?

Like any signal detection system, there is a need to establish 'rules' how to trigger an alert, when to consider a signal likely enough to be true to warrant follow-up or even to require immediate health policy intervention. In the final paper of this thesis, we described the triage of drug-adverse event associations (i.e., potential signals) derived from EHR databases. Taking the event acute myocardial infarction (AMI) as an example, we proposed a strategy for combining evidence from different data sources in EU-ADR to prioritize signals that may represent genuine risk and, hence, necessitate further investigation and formal pharmacoepidemiologic studies. Association measures between drug use and AMI were generated by first applying various statistical methods on healthcare data from seven databases of the EU-ADR network. Association estimates were ranked based on the best performing method (Longitudinal Gamma Poisson Shrinker). Matched case-control and self-controlled case series methods were additionally employed to deal with temporality and confounding effects, while the LEOPARD method was applied to specifically detect protopathic bias. Consistency of the association among drugs of the same class and the number of excess cases attributable to the drug exposure were further assessed to prioritize the list of potential signals. Finally, signal filtering and signal substantiation were done using different bioinformatics workflows to determine the novelty and plausibility of the identified signals. Overall, 163 potential signals for AMI were identified based on statistical association. Of these, 69 signals were flagged by LEOPARD as likely due to protopathic bias. Further signal refinement to reduce possible confounding decreased the number of signals to 41. The following nine signals remained after applying the criteria for novelty and plausibility: (1) metoclopramide; (2) cisapride; (3) domperidone; (4) betamethasone; (5) erythromycin; (6) roxythromycin; (7) azithromycin; (8) fluconazole; and (9) megestrol. We further discussed the likelihood of these identified signals in the light of current evidence, including the possibility of an increasing, or previously unrecognized, role of Kounis syndrome (so-called 'allergic myocardial infarction') in drug-induced AMI. Because this is the first signal triage strategy developed for use – and tested – in electronic healthcare data, this strategy needs to be further tested using other EHR data sources and other adverse events. While the automated signal filtering and signal substantiation streamlined the triage and greatly reduced manual work, full automation of the signal prioritization process is still not possible at this time. Manual verification of the output produced by these workflows, in terms of both accuracy and completeness, remains a crucial step.

METHODOLOGICAL CONSIDERATIONS

While the motivation for merging disparate data sources primarily comes from the need to investigate drug safety in larger populations (there is strength in numbers), the enormity and multi-dimensionality of the data present various challenges both in the methodology and in the interpretation of results.

Ensuring patient confidentiality and data protection

The common data framework, implemented via the custom-built software Jerboa©, and the choice of a distributed data network were crucial steps in the development of the EU-ADR system. This distributed network takes advantage of multiple, routinely collected, aggregated healthcare data while minimizing sharing of confidential patient-level information. A similar concept of distributed processing of healthcare data has been employed by other research collaborations, albeit with different research objectives. This model has previously been described in bioterrorism and syndromic surveillance as well as vaccine safety surveillance.[25-27] The ongoing Sentinel Initiative of the FDA is also adopting a distributed data architecture for combining healthcare databases to improve drug safety monitoring. While the scale may be comparable, there are issues in combining data that are unique to Europe. Challenges stem from the fact that different countries have distinct natural languages, aside from having different drug and disease coding systems. The diversity of healthcare systems throughout Europe makes merging data from databases a more complex task that requires striking a balance between international cooperation and adequate protection of patient confidentiality.[28]

The problem of multidimensionality

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. It is, however, a monumental task to undertake safety monitoring of *all* drugs for *all* possible events (in EU-ADR, we have already limited the number of events) in *all* members of the population (including children and the elderly) under *all* actual circumstances of medical care. Signal detection in multiple EHR data requires striking a balance between customization of methodologies for particular outcomes (immediate reactions such as anaphylaxis vs. more insidious events such as upper gastrointestinal bleeding or aplastic anemia) and increasing applicability. Optimization of methods for signal detection, taking into account variations in exposure and outcome assessment, among others, is an ongoing work in EU-ADR, in collaboration with other international initiatives.

The perennial threat of false alarms

The advantages of automated surveillance and quantitative signal detection are obvious, but there is a legitimate concern that such data mining on a grand scale may generate more signals than can be followed up effectively with currently available resources. The unpleasant implications

of spurious and unsubstantiated signals for public health, as well as for regulatory agencies and the pharmaceutical industry, cannot be underestimated and is an important consideration.[29-32] We have tried to set some limits to unconstrained data mining by initially monitoring only the events considered to be most relevant from a public health and pharmacovigilance perspective. We employed signal detection methods that allow shrinkage of the risk estimates so that those supported by large data are given more weight compared to those supported by sparse data. In the estimation of how much drug exposure data would be necessary for signal detection, we employed a 5% significance level, which 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, such requirement may not be appropriate in the context of exploratory signal detection. We investigated how correction for multiple comparisons would influence the number of drugs that can be monitored by applying the Bonferroni correction and 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 in both instances. However, with multiple testing already factored in the exposure calculations, there is a very small gain (with the use of 10% significance level) in the 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. This was true for the current database network as well as for the simulated expanded database platform.

Caveat

Secondary use of data for purposes other than they have originally been intended for brings with it many opportunities as well as challenges. The literature is replete with discussions on the merits and challenges of recycling routinely-collected EHR data, including how the type of database influences the structure and content of the data.[33-35] Potential associations are inferred outside the actual patient-physician encounter that leads to suspicion of an ADR – something that is inherent in SRS. Data in medical records databases, recorded in the course of routine clinical care, provide a different perspective from data in databases that document claims for utilization of healthcare services. We have shown that with use of multiple databases from different countries and different healthcare systems, the issue of accuracy of information vis-à-vis heterogeneity is even more formidable. There are limitations with respect to validity of coding systems and potential misclassification of both drug exposure and outcome and, like any other observational data, control for bias and confounding may not always be easy and is often inadequate. EHR data derived from reimbursement claims, for example, are affected by a lack of incentive to record sufficient data to allow proper case classification. Drug utilization patterns derived from population-based EHR data reflect 'real-world' conditions, but are thereby also influenced by changes in clinical practice, including changes brought about by preferential prescribing and disease management guidelines, and may lead to underestimation of risks.

FUTURE PERSPECTIVES

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.[35] 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.[36] 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. We have shown in simulated analyses that even with expansion of the EU-ADR network to 10 times its current size of over 21 million individuals and ~154 million person-years of follow-up, there would still be unmet needs in signal detection and that further collaboration is in order. 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. Extending the EU-ADR 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.[37] Further collaboration with other ongoing safety surveillance initiatives across the continent such as FDA Sentinel and OMOP is essential, and is already underway.

An EHR-based signal detection system, like any other surveillance system, provides no definitive answers. There are still a lot of spots to cover before rational decisions can be made, including weighing the implications of potential risks against the benefits of therapy, as well as risk communication and its many complexities. There are further opportunities to evaluate comparative effectiveness and risk management activities in EHR databases and such opportunities are worthwhile to explore.

Finally, the dynamic nature of medicine, health care, and drug development renders drug safety a moving target. There is the overriding need not only to streamline the process of signal detection, evaluation, and confirmation, but also to continually update existing data and explore other sources of safety information.

REFERENCES

1. World Health Organization. Safety of medicines. A guide to detecting and reporting adverse drug reactions 2002; http://whqlibdoc.who.int/hq/2002/WHO_EDM_QSM_2002.2.pdf Accessed 10 July 2011.
2. Hauben M, Aronson JK. Defining 'signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions. *Drug Saf.* 2009;32(2):99-110.
3. Institute of Medicine. The Future of Drug Safety - Promoting and Protecting the Health of thePublic. 2006. http://www.iom.edu/Reports/2006/The-Future-of-Drug-Safety-Promoting-and-Protecting-the-Health-of-the-Public.aspx. Accessed Last accessed 20 October 2011.
4. Blum MD, Graham DJ, McCloskey CA. Temafloxacin syndrome: review of 95 cases. *Clin Infect Dis.* Jun 1994;18(6):946-950.
5. Darpo B. Detection and reporting of drug-induced proarrhythmias: room for improvement. *Europace.* Sep 2007;9 Suppl 4:iv23-36.
6. Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. *Arch Intern Med.* Jun 27 2005;165(12):1363-1369.
7. US Food and Drug Administration. FDA Announces Withdrawal Fenfluramine and Dexfenfluramine http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm179871.htm. Accessed 18 October 2011.
8. Desmond P. Flucloxacillin hepatitis - an Australian epidemic. *Aust NZ J Med* 1995;25(3):195-196.
9. Centers for Disease Control and Prevnetion (CDC). Intussusception among recipients of rotavirus vaccine--United States, 1998-1999. *MMWR Morb Mortal Wkly Rep.* 1999 Jul 16;48(27):577-581.
10. Coulter DM, Edwards IR. Cough associated with captopril and enalapril. *Br Med J (Clin Res Ed).* Jun 13 1987;294(6586):1521-1523.
11. Coulter DM, Edwards IR. Mianserin and agranulocytosis in New Zealand. *Lancet.* Sep 29 1990;336(8718):785-787.
12. Bate A, Edwards I, Edwards J, Swahn E, Noren GN, Lindquist M. Knowledge finding in IMS disease analyser Mediplus UK database – effective data mining in longitudinal patient safety data. Paper presented at: ISoP Annual Conference: Pharmacovigilance - Current and Future Challenges; 2004; Dublin, Ireland.
13. Schuemie MJ. Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD. *Pharmacoepidemiol Drug Saf.* Mar 2011;20(3):292-299.
14. Curtis JR, Cheng H, Delzell E, et al. Adaptation of Bayesian data mining algorithms to longitudinal claims data: coxib safety as an example. *Med Care.* Sep 2008;46(9):969-975.
15. Svanstrom H, Callreus T, Hviid A. Temporal data mining for adverse events following immunization in nationwide Danish healthcare databases. *Drug Saf.* Nov 1 2010;33(11):1015-1025.
16. 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.* Aug 13 2009;361(7):645-647.
17. 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.* Nov 2 2010;153(9):600-606.
18. Hernandez-Diaz S, Rodriguez LA. Association between nonsteroidal anti-inflammatory drugs and upper gastrointestinal tract bleeding/perforation: an overview of epidemiologic studies published in the 1990s. *Arch Intern Med.* Jul 24 2000;160(14):2093-2099.
19. Hammond IW, Gibbs TG, Seifert HA, Rich DS. Database size and power to detect safety signals in pharmacovigilance. *Expert Opin Drug Saf.* Nov 2007;6(6):713-721.
20. Le J, Nguyen T, Law AV, Hodding J. Adverse drug reactions among children over a 10-year period. *Pediatrics.* Aug 2006;118(2):555-562.
21. Johnson TN. The development of drug metabolising enzymes and their influence on the susceptibility to adverse drug reactions in children. *Toxicology.* Oct 1 2003;192(1):37-48.
22. Carleton B, Primmett D. Recognizing idiosyncratic adverse drug reactions in children: A practice imperative. *Paediatr Child Health.* Apr 2001;6(4):187-189.

23. Rashed AN, Wong IC, Cranswick N, Tomlin S, Rascher W, Neubert A. Risk factors associated with adverse drug reactions in hospitalised children: international multicentre study. *Eur J Clin Pharmacol.* Dec 14 2011.
24. Mason J, Pirmohamed M, Nunn T. Off-label and unlicensed medicine use and adverse drug reactions in children: a narrative review of the literature. *Eur J Clin Pharmacol.* Jan 2012;68(1):21-28.
25. Brown JS, Kulldorff M, Chan KA, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. *Pharmacoepidemiol Drug Saf.* Dec 2007;16(12):1275-1284.
26. Lazarus R, Yih K, Platt R. Distributed data processing for public health surveillance. *BMC Public Health.* 2006;6:235.
27. Lieu TA, Kulldorff M, Davis RL, et al. Real-time vaccine safety surveillance for the early detection of adverse events. *Med Care.* Oct 2007;45(10 Supl 2):S89-95.
28. Rynning E. Public trust and privacy in shared electronic health records. *Eur J Health Law.* Jul 2007;14(2):105-112.
29. Hauben M, Reich L, Van Puijenbroek EP, Gerrits CM, Patadia VK. Data mining in pharmacovigilance: lessons from phantom ships. *Eur J Clin Pharmacol.* Nov 2006;62(11):967-970.
30. Smyth P. Data mining: data analysis on a grand scale? *Stat Methods Med Res.* Aug 2000;9(4):309-327.
31. Waller P. Dealing with uncertainty in drug safety: lessons for the future from sertindole. *Pharmacoepidemiol Drug Saf.* Jun 2003;12(4):283-287; discussion 289-290.
32. Avorn J, Schneeweiss S. Managing drug-risk information--what to do with all those new numbers. *N Engl J Med.* Aug 13 2009;361(7):647-649.
33. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol.* Apr 2005;58(4):323-337.
34. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects--advantages and disadvantages. *Nat Clin Pract Rheumatol.* Dec 2007;3(12):725-732.
35. Schneeweiss S. Understanding secondary databases: a commentary on "Sources of bias for health state characteristics in secondary databases". *J Clin Epidemiol.* Jul 2007;60(7):648-650.
36. McClellan M. Drug safety reform at the FDA--pendulum swing or systematic improvement? *N Engl J Med.* Apr 26 2007;356(17):1700-1702.
37. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. http://www.encepp.eu. Accessed [Last accessed 01 May 2011].

Summary

The Imperative of Change

Neither randomized clinical trials, considered the gold standard in the evaluation of therapeutic interventions, nor the drug approval process of the most powerful drug regulatory agencies in the world, can guarantee safety of all drugs that enter the market. A life-cycle approach to drug safety monitoring has been advocated because once drugs are marketed, they are used in a more diverse group of people, often for much longer periods, and sometimes with a wider range of therapeutic indications. There has been a growing clamor for improving the current passive-reactive paradigm of drug safety monitoring so that the thalidomide tragedy and the rofecoxib and practolol disasters do not have to happen again.

Our overview of traditional pharmacovigilance systems (spontaneous reporting systems and prescription event monitoring) and ongoing initiatives exploring EHR data (**Chapter 2**) showed that the surveillance of drugs post-licensure has become both a science and a crusade. The search goes on for viable strategies and data sources in pharmacovigilance that can facilitate the systematic and timely detection, assessment, and understanding of potential safety signals. Electronic healthcare record (EHR) databases have been heralded as an important resource for active surveillance that can augment existing pharmacovigilance systems. The aim of this thesis was to demonstrate whether mining of multiple EHR databases can give a comprehensive and comprehensible view of the many complex interdependent factors that play a role in the development of drug-related adverse events in a real-world setting.

Finding Opportunity in Diversity

We have developed and tested a series of methodologies that enable combining data from EHR databases of different countries and origins (medical records, administrative registries, record-linkage databases) to scale up safety signal detection. We first determined what events would be considered important to monitor from a pharmacovigilance and public health perspective (**Chapter 3.1**). The databases in the EU-ADR network have well-developed safeguard mechanisms ensuring compliance with the European directives and national regulations as well as database governance rules. Since no new data are collected, other than those made available by the participating databases, the cornerstone of ensuring proper ethical and legal conduct is set down in the rules and regulations that govern each database. Rather than imposing a one-size-fits-all approach and compel people to change their data, we leveraged on the diversity of the databases to use local expertise and maximize extraction of relevant information (**Chapter 3.2**). Within this distributed network system, databases retain ownership of their respective data, extraction being done locally, and only the aggregated, non-identifiable data are shared with the rest of the network (**Chapter 3.3**).

Challenge of Surveillance

Secondary use of data for purposes other than they have originally been intended for brings with it many opportunities as well as challenges. Data in medical records databases, recorded

in the course of routine clinical care, provide a different perspective from data in databases that document claims for utilization of healthcare services. We have shown that with use of multiple databases from different countries with different healthcare systems, the issue of accuracy of information vis-à-vis heterogeneity is even more formidable (**Chapter 3.2, Chapter 4**). There are limitations with respect to validity of coding systems and potential misclassification of both drug exposure and outcome and, like any other observational data, control for bias and confounding may not always be easy and is often inadequate.

While the motivation for merging disparate data sources primarily comes from the need to investigate drug safety in larger populations (there is strength in numbers), the enormity and multi-dimensionality of the data present various challenges both in the methodology and in the interpretation of results. In **Chapter 4**, we gave ourselves a sanity check on what these healthcare data-based surveillance systems can do. We demonstrated how much leverage EHR databases can provide for monitoring the safety of medicines in both adults and children (**Chapters 4.1 and 4.2**). We provided estimates of the number and types of drugs that can be monitored in these systems as a function of actual drug use, minimal detectable relative risk, and incidence rates of outcomes of interest. Use of EHR databases for safety surveillance will be most powerful for more frequently used classes of drugs and for outcomes with a high background incidence in the general population. Simulated analyses showed that even with expansion of the EU-ADR network to 10 times its current size of over 21 million individuals and ~154 million person-years of follow-up, there would still be unmet needs in signal detection and that further collaboration with other initiatives (such as the OMOP and FDA Sentinel) is in order. Signal detection in multiple EHR data also involves striking a balance between customization of methodologies for particular outcomes and increasing applicability. Optimization of methods for signal detection, taking into account variations in exposure and outcome assessment, among others, is an ongoing work in EU-ADR, also in collaboration with other international initiatives.

The lack of a true 'gold standard' (i.e., list of all known adverse reactions and which drugs can cause them) hinders methods development for signal detection. We have proposed a strategy for the construction of a surrogate reference standard to evaluate drug safety signal detection methodologies using EHR (**Chapter 4.3**). This reference standard should be considered dynamic and as knowledge on drug safety evolves over time and new issues in drug safety arise, this reference standard should be periodically re-assessed.

The advantages of automated surveillance and quantitative signal detection are obvious, but there is a legitimate concern that such data mining on a grand scale may generate more signals than can be followed up effectively with currently available resources (**Chapter 5**). The unpleasant implications of spurious and unsubstantiated signals for public health, as well as for regulatory agencies and the pharmaceutical industry cannot be underestimated and is an important consideration.

From Statistics to Decision-Making

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.

As had been emphasized in the introductory chapter of this thesis, there is no drug – or any intervention for that matter – that is without risk. Even placebo has been incriminated for causing various adverse effects. Nevertheless, there is no excuse for putting patients' lives in unnecessary danger by inappropriate prescribing and irrational drug use. Especially in this era of greater longevity, chronic diseases and polypharmacy, rational drug therapy plays an even more vital role in pharmacovigilance. We have shown that EHR data may provide a snapshot of how drugs are actually being prescribed and being used (pharmacy dispensings as proxy) in 'real-world' circumstances and that there are differences among databases, countries, and sub-groups of the population. We have demonstrated the feasibility of studying various outcomes and events within multiple EHR databases - outcomes that span a wide range of incidence and affected populations (**Chapters 4.5, 4.6, 4.7**). We have also shown that, using a variety of statistical methods, mining of multiple EHR databases can corroborate drug-adverse event associations previously described in the literature and can identify associations that signal potential increased risk (**Chapter 4.4**).

Despite the limitations and caveats, there is no denying the important contribution of multi-country EHR databases, not only with respect to augmenting drug safety surveillance but also to a better understanding of the role of healthcare systems in disease. The work described in this thesis was an attempt to provide a glimpse of what can be done so that medicines can continue to be valuable, comforting, and life-saving to patients. The signal triage strategy that we have described (**Chapter 5**) is an open-ended question that demands further scrutiny and testing by various stakeholders such as regulators, drug manufacturers, pharmacoepidemiologists and pharmacovigilance experts, public health practitioners, and physician prescribers. An EHR-based signal detection system, like any other surveillance system, provides no definitive answers. There are still a lot of spots to cover before rational decisions can be made, including weighing the implications of potential risks against the benefits of therapy, as well as risk communication and its many complexities. There are further opportunities to evaluate comparative effectiveness and risk management activities in EHR databases and such opportunities are worthwhile to explore.

Finally, the dynamic nature of medicine, health care, and drug development renders drug safety a moving target. There is the overriding need not only to streamline the process of signal detection, evaluation, and confirmation, but also to continually update existing data and explore other sources of safety information.

Samenvatting

De noodzaak van verandering

Gerandomiseerde klinische studies worden gezien als de gouden standaard voor het beoordelen van therapieen, echter noch deze studies noch het registratieproces van nieuwe geneesmiddelen door de regulatoire agentschappen in de wereld kunnen garanderen dat alle geneesmiddelen die worden toegelaten veilig zijn. Na toelating op de markt worden geneesmiddelen vaak gebruikt door groepen mensen die meer divers zijn dan in de registratiestudies, vaak voor langere periodes en soms ook voor een grotere verscheidenheid aan indicaties. Daarom wordt er gepleit voor een "life-cycle" benadering voor geneesmiddelenbewaking. Meer en meer wordt aangedrongen om het huidige passief-reactieve model voor het bewaken van de risico's van geneesmiddelgebruik te verbeteren om zo te voorkomen dat een tragedie als thalidomide of de catastrofes met rofecoxib en practolol nogmaals kunnen optreden.

Toezicht op geneesmiddelen na registratie is zowel een wetenschap als een kruistocht geworden. We hebben dit laten zien met het overzicht van de klassieke farmacovigilantie systemen (spontane meldingen systemen en 'prescription event monitoring') en lopende initiatieven die data uit "Electronic healthcare record" (EHR) databanken, oftewel elektronische gezondheidszorgdossiers databanken verkennen (**Hoofdstuk 2**). Binnen de farmacovigilantie gaat de zoektocht naar haalbare strategieën en data bronnen die stelselmatige en tijdige detectie, beoordeling en begrip van potentiële veiligheidssignalen kunnen faciliteren door. EHR databanken worden beschouwd als een belangrijk bron voor proactieve geneesmiddelenbewaking in aanvulling op bestaande farmacovigilantie systemen. Het doel van dit proefschrift was om te onderzoeken of het samenbrengen van meerdere EHR databanken, met als doel bijwerkingen te onderzoeken, kan bijdragen aan het verkrijgen van een volledig en begrijpelijk beeld van de vele complexe en van elkaar afhankelijke factoren die een rol kunnen spelen in de ontwikkeling van bijwerkingen van geneesmiddelen onder levensechte omstandigheden.

Verscheidenheid biedt kansen

Wij hebben een reeks aan werkwijzen ontwikkeld en getest die ons de gelegenheid hebben gegeven om gegevens uit meerdere EHR databanken uit verschillende landen en van verschillende oorsprong (patiëntendossiers, verzekeringsclaims en databanken met verschillende registraties) te combineren om zo de schaal van het detecteren van veiligheidssignalen van geneesmiddelen (signaaldetectie) te vergroten. Ten eerste hebben we bepaald voor welke bijwerkingen prospectieve geneesmiddelenbewaking van belang is uit het oogpunt van zowel farmacovigilantie en volksgezondheid (**Hoofdstuk 3.1**). De EHR databanken die onderdeel van het EU-ADR netwerk uitmaken, hebben voorzorgsmaatregelen ontwikkeld om naleving van de Europese en nationale wetgeving, en van de toezichthouders op de databanken te waarborgen. Binnen de studies werd geen nieuwe data verzameld, maar werd er slechts gebruik gemaakt van gegevens die reeds in de databanken binnen het netwerk aanwezig waren. Het ethisch gebruik van de data binnen de bestaande wetgeving wordt gewaarborgd door regels en maatregelen die zijn vastgelegd door iedere databank. In plaats van databankhouders te vragen om hun gegevens op eenzelfde manier

aan te leveren, waardoor wellicht wijzigingen van de data noodzakelijk zijn, hebben we juist gebruik gemaakt van de diversiteit binnen de verschillende databanken. Dit door met behulp van de locale expertise de relevante data te extraheren (**Hoofdstuk 3.2**). Binnen het gedistribueerde systeem van netwerken wordt de data lokaal geëxtraheerd en wordt alleen de geaggregeerde, niet identificeerbare data gedeeld met de rest van het netwerk (**Hoofdstuk 3.3**).

Bewaking is een uitdaging

Het gebruik van gegevens die oorspronkelijk voor een ander doel zijn verzameld biedt veel mogelijkheden, maar kent ook uitdagingen. Gegevens die zijn verzameld tijdens dagelijkse klinische zorg in patiëntendossiers bieden een ander perspectief dan gegevens verzameld op basis van verzekeringsaanspraken. We hebben laten zien dat bij het gebruik van meerdere databanken uit verschillende landen met elk eigen gezondheidszorgsystemen het afwegen van de betrouwbaarheid van de informatie ten opzichte van de verscheidenheid van de informatie van groot belang is. (**Hoofdstuk 3.2, Hoofdstuk 4**). Er zijn tekortkomingen op het gebied van validiteit van coderingssystemen en potentiële misclassificatie van zowel blootstellingen aan geneesmiddelen als de gevolgen hiervan. Zoals bij alle observationele gegevens is het corrigeren voor bias en confounding niet altijd gemakkelijk en vaak onvoldoende.

Het belangrijkste motief voor het samenvoegen van verschillende bronnen is de behoefte om risico's van geneesmiddelengebruik te bestuderen in grote populaties (de kracht van grote getallen). Echter de omvang en veelzijdigheid van de gegevens brengt verschillende uitdagingen met zich mee, zowel in de methodologie als in de interpretatie van de resultaten. In **Hoofdstuk 4** zijn we nagegaan wat mogelijk is met monitorsystemen op basis van EHR databanken. We hebben gedemonstreerd hoeveel EHR databanken kunnen bijdragen aan de geneesmiddelenbewaking van zowel volwassenen als kinderen (**Hoofdstuk 4.1 en 4.2**). Op basis van het werkelijk geneesmiddelgebruik, een minimaal detecteerbaar relatief risico en incidentiecijfers van uitkomsten hebben we geschat hoeveel geneesmiddelen bewaakt kunnen worden. De kracht van het gebruik van EHR databanken voor geneesmiddelenbewaking ligt in bewaking van geneesmiddelen die frequent gebruikt worden en aandoeningen met een hoge incidentie in de algemene populatie. Middels een simulatie hebben we aangetoond dat zelfs bij een tienvoudige uitbreiding van het EU-ADR netwerk, dat nu 21 miljoen personen and ~150 miljoen persoonsjaren aan follow-up omvat, er nog altijd niet aan alle behoeften voor signaaldetectie voldaan kan worden en dat verdere samenwerking met bestaande initiatieven (zoals OMOP en FDA Sentinel) noodzakelijk lijkt. Het slaan van de balans tussen het aanpassen van de methoden voor specifieke uitkomsten en toenemende toepasbaarheid is onderdeel van signaaldetectie in meerdere EHR databanken. Onder andere het optimaliseren van methoden voor signaaldetectie, rekening houdend met verschillen in de bepaling van expositie of uitkomst, is werk dat wordt voortgezet door EU-ADR ondermeer in samenwerking met andere internationale initiatieven.

Het ontbreken van een echte "gouden standaard" (bijvoorbeeld een lijst van alle bekende bijwerkingen en de geneesmiddelen die deze kunnen veroorzaken) hindert het ontwikkelen van methoden voor signaaldetectie. Wij hebben een strategie voorgelegd voor het samenstellen van een

surogaat referentieset om zo methoden voor signaaldetectie binnen EHR te evalueren (**Hoofdstuk 4.3**). Dit is een dynamische referentieset die, naarmate de kennis over geneesmiddelenveiligheid toeneemt over de tijd en naarmate zich nieuwe kwesties voordoen, van tijd tot tijd geherevalueerd zal moeten worden.

De voordelen van geautomatiseerde bewaking en kwantitatieve signaal detectie zijn onmiskenbaar, maar de zorg dat het op grote schaal exploiteren van data meer signalen genereert dan dat er bevestigd kunnen worden is terecht (**Hoofdstuk 5**). De nadelige gevolgen van onterechte en ongefundeerde signalen op de volksgezondheid, voor geneesmiddelproducten en regulatoire agentschappen moeten niet worden onderschat en zullen altijd overwogen moeten worden.

Van statistiek tot beslisvorming

We hebben benadrukt dat voldoende statistische capaciteit van belang is voor de detectie van veiligheidssginalen. Het is echter niet zo dat de capicteit van een EHR databank netwerk alleen uitgedrukt wordt in de hoeveelheid data die het kan genereren. Vooral voor longitudinale EHR databanken kan worden gesteld dat er een veel grotere waarde schuilt in de mogelijkheid van het schatten van de incidentie van bijwerkingen gedurende het gebruik van geneesmiddelen in verschillende subgroepen van de populatie. Onafhankelijk van voldoende capaciteit om een statistisch significant verschil aan te tonen met géén gebruik van geneesmiddelen zijn deze gegevens op zich zelf belangrijk voor zowel de farmacovigilanctie als de volksgezondheid.

Zoals benadrukt in het inleidende hoofdstuk van dit proefschrift is er geen geneesmiddel of interventie zonder risico's. Ook bij gebruik van placebo's worden bijwerkingen gemeld. Het leven van patiënten moet niet nodeloos in gevaar gebracht worden door ongepast en irrationeel gebruik van geneesmiddelen. Naarmate mensen ouder worden en de prevalentie van chronische ziekten en polyfarmacie toeneemt, speelt rationeel gebruik van geneesmiddelen een nog belangrijkere rol in de farmacovigilantie. We hebben aangetoond dat EHR data een inzicht kunnen geven in hoe geneesmiddelen daadwerkelijk worden voorgeschreven en worden gebruikt (geschat op basis van apotheek uitgiften) en dat dit gebruik verschilt tussen de databanken, tussen landen en tussen verschillende subgroepen van de populatie. We hebben laten zien dat het mogelijk is om binnen meerdere EHR databanken diverse uitkomsten en bijwerkingen te bestuderen. Deze uitkomsten hebben een groot bereik aan incidenties en aangedane populaties (**Hoofdstukken 4.5, 4.6, 4.7**). Ook hebben we door het gebruik aan een verscheidenheid aan statistische methoden binnen meerdere EHR databanken, bijwerkingen die eerder zijn beschreven in de literatuur bevestigd. Daarnaast is het mogelijk om associaties met een mogelijk verhoogd risico aan het licht te brengen (**Hoofdstuk 4.4**).

Ongeacht de beperkingen en valkuilen van het gebruik van meerdere EHR databanken leveren deze een belangrijke bijdrage, niet alleen als aanvulling op de geneesmiddelenbewaking maar ook in het beter begrijpen van de rol van deze systemen in ziekte. Het werk zoals beschreven in dit proefschrift geeft een kijk op wat gedaan kan worden om te zorgen dat geneesmiddelen waardevol en levensreddend voor patiënten kunnen blijven. De strategie die we hebben beschreven voor triage van signalen (**Hoofdstuk 5**) is een vraag met een open einde die verder

nauwkeuriger onderzocht en getest zal moeten worden door verschillende belanghebbenden zoals toezichthouders, geneesmiddelproducten, farmacoepidemiologen, experts in farmacovigilantie, mensen werkzaam in de volksgezondheid en voorschrijvende artsen. Een signaaldetectie systeem gebruik makend van EHR databanken zal, zoals elk ander bewakingssysteem, geen definitieve antwoorden geven. Veel andere aspecten zullen ook moeten worden bekeken, voordat er rationele beslissingen gemaakt kunnen worden. Voorbeelden hiervan zijn het afwegen van de gevolgen van potentiële risico's en voordelen van behandeling, maar ook het communiceren van risico's. Er zijn nog verdere mogelijkheden om effectiviteit van middelen met elkaar te vergelijken en het risico management activiteiten te evalueren –verdere verkenning hiervan is sterk aan te raden.

Ten slotte, de dynamische aard van de geneeskunde, de gezondheidszorg en ontwikkeling van geneesmiddelen maken geneesmiddelenbewaking een 'moving target.' Het is van doorslaggevend belang om niet alleen het proces van signaal detectie, evaluatie en bevestiging te stroomlijnen, maar ook om de bestaande gegevens continue up-to-date te houden en verdere bronnen van informative over geneesmiddelveiligheid te verkennen.

ACKNOWLEDGMENTS

I have the good fortune of being raised in a family where excellence is anticipated and versatility encouraged. To my Dad, Atty. Herman Coloma - my ideal of the renaissance man - who inculcated in me a displeasure with mediocrity and an understanding of the value of hard work – this thesis is both a tribute and an atonement. The only let-down in this whole enterprise is you not being around to witness the ending, but I hope I make you proud nevertheless. To my Mom, Dr. Teresita Coloma – whose own experiences at the School of Management in Cranfield and at the Kennedy School of Government in Harvard made me open to the prospect of expanding my intellectual horizons abroad – and to whom I owe my mastery of, and respect for, the English language – thank you for everything! To my siblings Eloisa, JJ and Sette, Dulce and Fred, Alma and Dominic, Mak and Calen, and Sushi and Boyet, and their families (nephews and nieces Lander, Cedrick, Ethan, Gabriel, Nicole, Bambam, and most especially to Katniss, who did the design and layout of the thesis cover) - who are my biggest 'fans' as well as my biggest critics – thank you! Despite my tireless wandering, your love and support remain unconditional. To the rest of my family all over the world, especially to Perri and Carri, Ray and Laura, Lora and David, Simonie and Fe, Malou and Abet, Grima and Ric, Lola Darling, the Anchetas, Sally [who's like family] – there are just too many of you to list here but that does not make my overwhelming gratitude any less.

To the Dutch Duncs – Derek, Geertje, Benjamin, Naomi, Abigail, and Joella, and the Bosmans – thank you for making me a part of your extended family and for introducing me to all the wonders and peculiarities of the Dutch culture. I probably would not have made it this far in this country without the knowledge that family is just close by.

I have always been an impressionable student and the reason why I have taken interest in a wide variety of subjects is that I never lacked for great teachers:

To Prof. Miriam Sturkenboom – dear Miriam, thank you for taking me in [from seemingly out of nowhere] and entrusting me with the work in EU-ADR [and now Eu2P]. It has been an interesting journey of trial and error, hide and seek, agreement and disagreement, but on the whole it has been a remarkable learning experience. None of these would have happened were it not for you taking a chance on me.

To Prof. Johan van der Lei – dear Johan, I remember you making the candid remark, 'Prove it' after I said, when meeting you for the first time, that I felt honored for the opportunity to work in the Department. Three-and-a-half years later, I feel even more honored to be working with you. Though usually impervious to the inconvenience of proving myself [doing good work is primary, proving myself is subsidiary], I hope I have managed to gain your confidence. Many thanks for all your support and guidance.

To my co-promotors Drs. Gianluca Trifirò and Martijn Schuemie, who have taken very good care of me [including sharing Work Package 9 photos with my paranymphs!] – thank you! Dear Gianluca and Martijn, it was a challenge being supervised by a clinical pharmacologist on the one hand and a computer scientist on the other hand, but we did [and will continue to do] great work together.

I have had the amazing opportunity to listen to [and in few instances, be listened to by] some of the luminaries in the field of epidemiology and pharmacovigilance, including Prof. Bruno Stricker, Prof. Kenneth Rothman, Prof. Sander Greenland, Prof. Albert Hofman, Prof. Moyses Szklo, Dr. Sebastian Schneeweiss, Dr. Stella Blackburn, Prof. Jeffrey Aronson, Prof. Robin Ferner, and Prof. Stephen Evans. My encounters with them ranged from two hours to three years [half of the time I was just among a crowd of avid listeners] and most of them will probably not remember me, but that does not diminish their influence on how I view the work that I do.

To Dr. Nelia Cortes-Maramba, pediatrician/clinical pharmacologist/toxicologist - who made me fall in love with pharmacology at first sight, and whose teachings made me understand the importance of rational drug therapy – thank you!

To Dr. Jeane Azarcon, plastic and burn surgeon and mentor - who welcomes me back with love whenever I am home, although she cannot imagine me being anywhere else other than the O.R. [or perhaps the Burn Unit] – thank you! Your invaluable support has made this journey more enlightening.

To Dr. Raffy Consunji, my favorite surgical intensivist and who remains to be one of my most influential mentors – thank you! For showing me how important it is to not only 'talk the talk,' but also 'walk the walk.' And for reassuring me that [of course!] I can have my cake and eat it too!

To the collaborators and management staff of the EU-ADR Consortium – thank you for all your support. The list of people I am thankful to would take up at least two pages, but special mention goes to Rosa Gini, Giampiero Mazzaglia, Paul Avillach, Lars Pedersen, Gino Picelli, Ron Herings, Lorenza Scotti, Julia Hippisley-Cox, Mariam Molokhia, Annie Fourrier-Reglat, Antoine Pariente, Francesco Salvo, Erik van Mulligen, and Jan Kors.

To my paranymphs Sandra de Bie [and, by extension, Jelmer Alsma], Leonoor Wijnans, and Anne Chain [my twin sister and shadow nymph] – who celebrated great parties with me, who got amused by [and worried about the lack of] my cycling skills, who taught me the 'must know' in the Dutch language, who tolerated my temper tantrums, who commiserated with me for the many sleepless nights, who provided level-headed discussions during times of distress, and whose love and support made difficulties bearable and everything else enjoyable – you all know I wouldn't have made it without you.

To Seppe and Eva Koopman, who took the time to get to know me and assist me [with practically everything] when I was still new and uninteresting to most people, who helped me find and put together a place I could call home, and who drank wine and broke bread with me both in good times and in bad times – thank you!

Thank you to all dear colleagues and friends in IPCI and Pharmacoepi, especially to Vera who has been a great collaborator both in research and in the more important thing in life called

'fun,' to amaz-Inge who always brought me sunshine, to Dr. Katia Verhamme and Dr. Sabine Straus for all the advice and support, to Silvana, Alessandro, Eline, Rikje, Marjolein, Gwen, René, Gina, Fedde, Daniel, Ann V and former colleagues Fatma, Emine [who gave me the essential refresher course in cycling], Ria, Jeanne, Ana, and Roelof. Thank you [and perhaps apologies] to Carmen, Osemeke, and Nico for being an unwitting audience to the pendulum of utter silence and frantic activity during the crucial period[s] of my PhD. To Ingrid and Peter R, who bore the brunt of the work in Eu2P while I was busy writing my thesis. To Mees, Kris, and Marcel - thank you for all the technical assistance.

To the wonderful people of BIGR, especially to my Dinner Club – Renske, Marcel K, Esben, and Michiel – thank you! For letting me trespass into geekdom and into all the fun that comes with it [including learning how to cycle, safely, under the influence].

To colleagues in BIOS and to other colleagues in Medical Informatics, and to the Secretariat: Desiree; Tineke; Carmen; Petra; and also Sander – thank you for your support. Were it not for you, the wheels taking me here wouldn't have gotten in motion.

Staying in the Netherlands would not have been as pleasant and smooth-sailing were it not for the support and assistance of Andreas Engelhard and Solange Degenaar. It is never easy for a *buitenlander,* even for an international institution like Erasmus MC, but you've made it as close to hassle-free as possible.

To Dr. Janice Campos-Caoili and to Dr. Jong de Castro, who gave me the encouragement and moral support to pursue research abroad and to explore further opportunities – thank you!

To Atty. Leo Domingo, my legal adviser, sanity check, and unconditional friend for life – thank you! For cheering me on and letting me be.

To Dr. Jody Junia, who gives me my wake-up call [more often than what he believes should be enough], who will always be my lifeline, and who, despite [or perhaps because of] knowing me too well, still looks forward to my day of enlightenment – thank you, my dear.

And to the perennially inspiring Dr. Rodney Dofitas, who doesn't mind being my accidental life coach and bomb squad – thank you! For always reminding me never to take things for granted. And for the unrelenting belief that I will always be a good doctor no matter how I decide to do it.

LIST OF PUBLICATIONS

Publications related to this thesis

- Trifirò G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salamé G, Catania MA, Salvo F, David A, Moore N, Caputi AP, Sturkenboom M, Molokhia M, Hippisley-Cox J, Diaz Acedo C, van der Lei J, Fourrier-Reglat A, on behalf of the EU-ADR Project.
  Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor?
  Pharmacoepidemiol Drug Saf. 2009 Dec; 18(12):1176-84.

- Coloma PM, Schuemie MJ, Trifirò G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom M, 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 Jan; 20(1):1-11. doi: 10.1002/pds.2053. Epub 2010 Nov 8.

- Trifirò 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 Group.
  EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection.
  Stud Health Technol Inform. 2011; 166:25-30.

- Coloma PM, Trifirò 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 EU-ADR consortium.
  Electronic healthcare databases for active drug safety surveillance:
  is there enough leverage?
  Pharmacoepidemiol Drug Saf. 2012 Feb 8. doi: 10.1002/pds.3197.[Epub ahead of print]

- Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourrier-Réglat A, Molokhia M, Patadia V, van der Lei J, Sturkenboom M, and Trifirò G, on behalf of the EU-ADR consortium.
  A Reference Standard for Evaluation of Methods for Drug Safety Signal
  Detection using Electronic Healthcare Record (EHR) Databases.
  Drug Safety 2012 [In press]

Other publications

- Coloma PM et al., **The Activity of *Centella asiatica (L. Urb)* on mice with Induced Depression.** *Acta Medica Philippina* 1997; *33 (1 & 2): 1-25*

- Coloma PM. *Building an Ethic Against Violence Directed at Children*. in Children: A Cause for Peace. 1998. Medical Students for Social Responsibility (MSSR), UP College of Medicine and International Physicians for the Prevention of Nuclear War (IPPNW).

- Technical Working Group (TWG) for the Pain Management Standards and Clinical Practice Guidelines on the Management of Acute Pain. **For the Philippine College of Surgeons-Philippine Society of Anesthesiologists: Evidence-based clinical practice guidelines for the management of acute postoperative pain**. Philipp J Surg Spec 2007; 62(2): 79-106.

PHD PORTFOLIO

Name	Preciosa M. Coloma
Erasmus MC Department	Medical Informatics
Research School	Netherlands Institute for Health Sciences
PhD period	August 2008 to May 2008
Promotor	Prof.dr.M.C.J.M. Sturkenboom

1 PHD TRAINING

Research Skills

Data Management, Statistics, and Methodology

2008-2010 Master of Science in Clinical Epidemiology
Netherlands Institute for Health Sciences, Rotterdam
30 ECTS

Oral Presentations

2010 *Opportunities for large-scale drug safety monitoring using multiple healthcare databases.*
16th IUPHAR World Congress of Basic and Clinical Pharmacology (WorldPharma 2010) July 17-23, 2010, Copenhagen, Denmark

2010 *Psychotropic Drug Use in the Elderly across Europe: A Perspective from the EU-ADR Project.*
26th International Conference on Pharmacoepidemiology and Therapeutic Risk Management. August 19-22, 2010, Brighton, United Kingdom

2011 *Methods for drug safety signal detection using EHR databases – the added value of longitudinal patient information.*
27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management. August 14-27, 2011, Chicago, USA.

2011 *A Reference Standard for Evaluating Methods for Drug Safety Signal Detection using Electronic Healthcare Record (EHR) Databases.*
11th Annual Meeting of the International Society of Pharmacovigilance (ISOP). October 26-28, 2011, Istanbul, Turkey.

Poster Presentations

2009 *Opportunity in Diversity: Merging Healthcare Databases in Europe for Safety Signal Detection in Pharmacovigilance.*
2nd place, Best Poster Presentation
25th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE). August 16-19, 2009, Providence, Rhode Island, USA.

2010 *The Third Way in Pharmacovigilance: Drug Safety Signal Detection using Multiple Electronic Healthcare Record (EHR) Databases.*
46th Annual Meeting of the Drug Information Association (DIA)
June 13-17, 2010, Washington, DC, USA.

2011 *For how many and what types of drugs can longitudinal healthcare databases detect safety signals? A view from the EU-ADR Project*.
27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management. August 14-27, 2011, Chicago, USA

2011 *Electronic Healthcare Record (EHR) Databases for Drug Safety Signal Detection: What Can We Expect?*
11th Annual Meeting of the International Society of Pharmacovigilance, October 26-28, 2011, Istanbul, Turkey.

2011 *Drug Use and Acute Liver Injury in Children: Signal Detection Using Multiple Healthcare Databases.*
11th Annual Meeting of the International Society of Pharmacovigilance, October 26-28, 2011, Istanbul, Turkey.

2011 *Comparison of Signal Detection Using Healthcare Database Network versus Spontaneous Reporting System Database: the EU-ADR Experience.*
11th Annual Meeting of the International Society of Pharmacovigilance, October 26-28, 2011, Istanbul, Turkey.

International Conferences

2010 46th Annual Meeting of the Drug Information Association (DIA)
June 13-17, 2010, Washington, DC, USA.

2010 26th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE). August 19-22, 2010, Brighton, United Kingdom

2010 16th IUPHAR World Congress of Basic and Clinical Pharmacology (WorldPharma 2010) July 17-23, 2010, Copenhagen, Denmark

2011 27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management. August 14-27, 2011, Chicago, USA

2011 11th Annual Meeting of the International Society of Pharmacovigilance, October 26-28, 2011, Istanbul, Turkey.

Other Presentations, Seminars, and Workshops

(both as presenter and as collaborator)

2009, 2010	Medical Informatics Days
2008-2012	Research Seminars in Pharmacoepidemiology Research Colloquia in Medical Informatics
2010	Workshop*: Design, development and validation of a computerised system that exploits data from electronic health records and biomedical information for the early detection of adverse drug reactions. The EU-ADR project: Preliminary Results*. 13th International Congress on Medical Informatics (Medinfo 2010) September 12-15, 2010, Cape Town, South Africa.
2010	*Harmonising definitions of adverse events among 8 European healthcare databases participating in the EU-ADR Project*. EUROEPI 2010 Epidemiology and Public Health in an Evolving Europe November 6-9, 2010, Florence, Italy.

Others

2010, 2011	Reviewer for the *Annals of Internal Medicine*
2012	Reviewer for *Drug Safety*
2011	Contributor to the GRACE Checklist validation The Good ReseArch for Comparative Effectiveness (GRACE) Initiative
2011	Contributor, OMOP Epidemiology Design Decision Inventory and Evaluation (EDDIE) Survey
2011-present	Member, Extended Consortium The Observational Medical Outcomes Partnership (OMOP)

2 TEACHING ACTIVITIES

2011	Lecturer, Module on *Principles of identifying and recognizing adverse events and safety signal*s Eu2P European Programme in Pharmacoepidemiology and Pharmacovigilance
2011-present	Course Coordinator for the Module on *Substantiation and Quantification of Risks* and Assistant Director for the Course Domain *Medicines Risk Identification and Quantification* Eu2P European Programme in Pharmacoepidemiology and Pharmacovigilance (http://www.eu2p.org/)

ABOUT THE AUTHOR

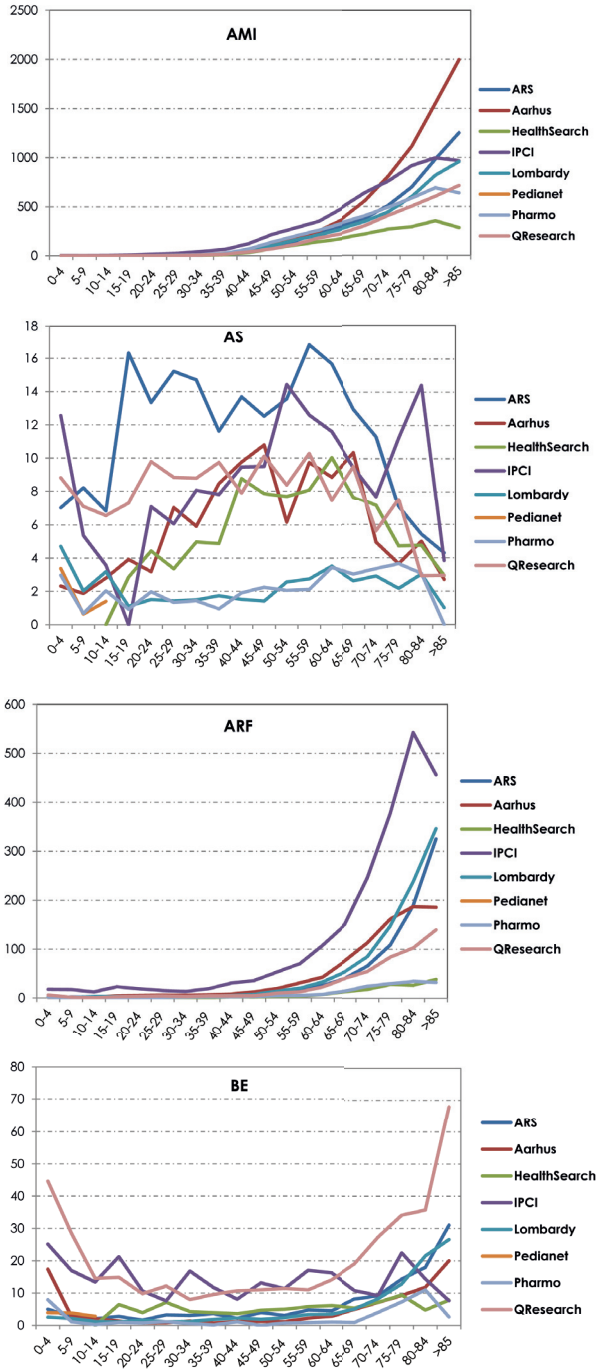
Preciosa Coloma was born in Laoag City in the Philippines and is a graduate of The Philippine Science High School. She obtained her Bachelor of Science degree in Industrial Pharmacy, *cum laude* and college valedictorian and overall most outstanding student, from the University of the Philippines in Manila. She finished her graduate degree in Medicine from the same University and did her rotating clinical internship at the 102-year old Philippine General Hospital, University teaching hospital and national tertiary referral center. She is licensed by both the Philippine Board of Pharmacy and the Board of Medicine. She is certified by the United States Educational Commission for Foreign Medical Graduates [since 2006] and has successfully passed the US Medical Licensure Examinations [USMLE] Steps 1 to 3.

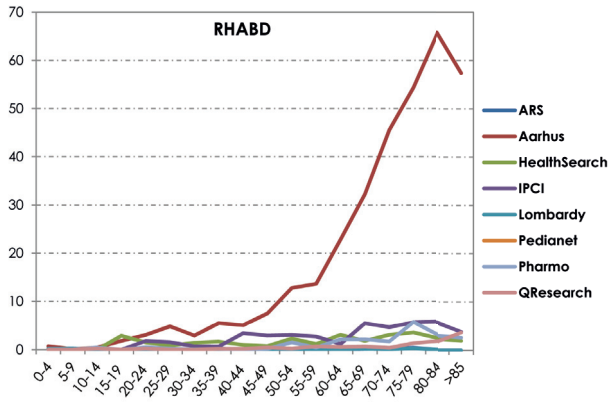
Although Internal Medicine was her first love, finishing at the top of her class in Integrated Clinical Clerkship in Internal Medicine, she started her residency in Plastic and Reconstructive Surgery and Burns at the Philippine General Hospital after internship. She completed her training in General Surgery and Trauma and was Most Outstanding Resident in Surgical Intensive Care in her senior year. After alternating between surgical house officer and researcher, community-based general practitioner and TB awareness advocate, copy editor of modern Shakespearean plays and house-sitter, she decided to go on yet another detour to take up translational medicine and international health and ended up doing drug safety research in the Netherlands [a serendipity of sorts]. She started working on her PhD in pharmacopidemiology and medical informatics at the Erasmus University Medical Center in Rotterdam in 2008 and finished her Master of Science in Clinical Epidemiology at the Netherlands Institute for Health Sciences in 2010.

Aside from being involved in the EU-ADR Project and its offshoot, the EU-ADR Alliance, she is currently assistant Domain director and Module coordinator for the Eu2P European Programme in Pharmacovigilance and Pharmacoepidemiology and will continue as post-doctoral researcher at the Department of Medical Informatics at Erasmus MC. She still considers herself a surgeon, which probably explains her veneration for mentorship, her rigidly severe [severely rigid] work ethics, and her terrible sense of humor. Her blog [incidentally called '*a-fish-outtawater*'] contains a biased selection of her anecdotes and musings on wondering, wandering, the praxis of living, and the practice of medicine.

Colour figures

Appendix 3 | Age-specific incidence rates of the different events across the databases (per 100 000 person-years) ***(see page 81 and 82)*****.**





Appendix 3 (Continued)

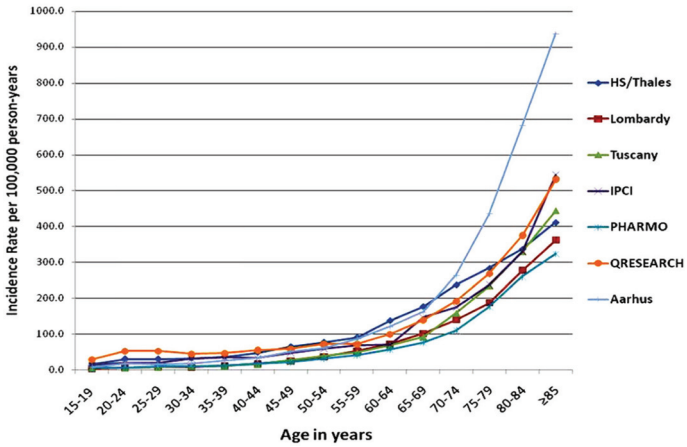


Figure 3.3.4 | Incidence rates of UG1B across databases, age 15 years and above *(see page 96)*.

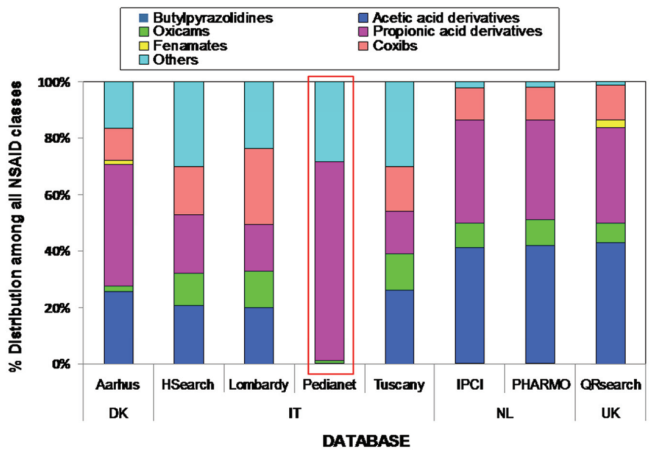


Figure 3.3.5 | Comparison of use of specific NSAID classes across databases*(see page 96)*.

CHAPTER 4

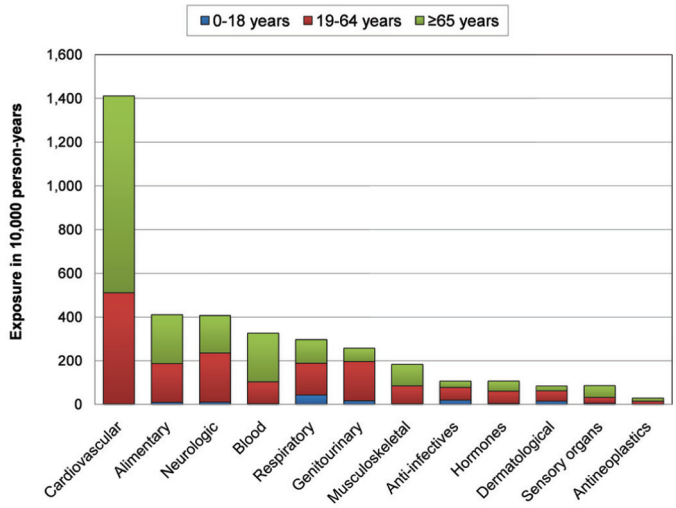


Figure 4.1.1 | Overview of exposure to all drugs (grouped according to ATC first level) by age categories ***(see page 111).***

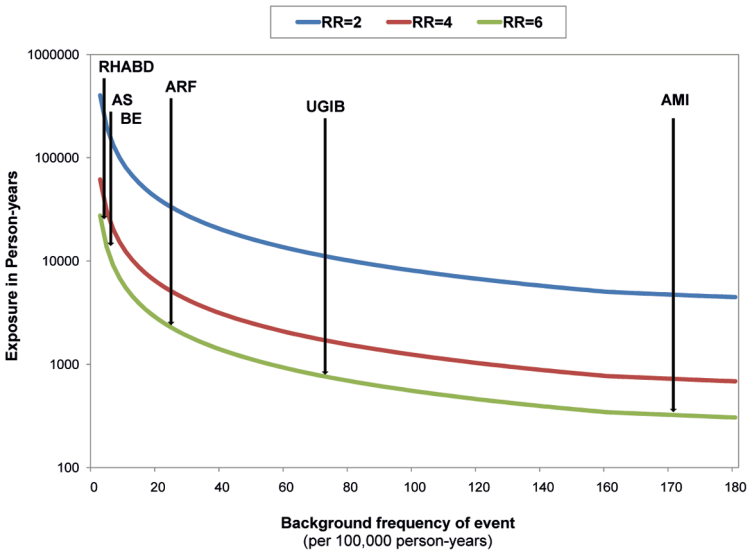
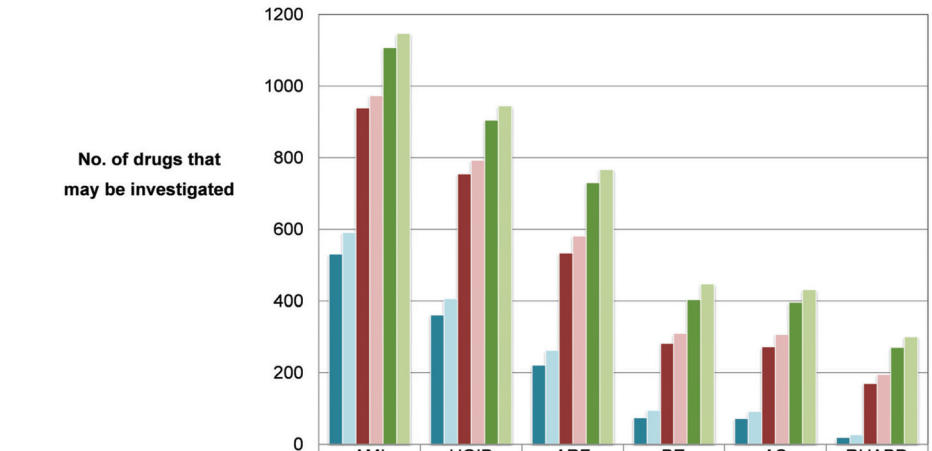


Figure 4.1.2 | Relationship between background incidence rate and amount of drug exposure required to identify a potential safety signal ***(see page 112).***



	AMI	UGIB	ARF	BE	AS	RHABD
Weak associations at α=0.05	531	361	221	74	72	19
Weak associations at α=0.10	591	407	262	95	92	27
Moderate associations at α=0.05	939	755	534	282	272	170
Moderate associations at α=0.10	973	793	582	310	307	195
Strong associations at α=0.05	1107	905	730	404	396	270
Strong associations at α=0.10	1147	945	767	448	432	300

Figure 4.1.3 | Change in the number of drugs that may be investigated for signal detection when the acceptable probability of error is increased from α= 0.05 to α= 0.10 *(see page 117)*.

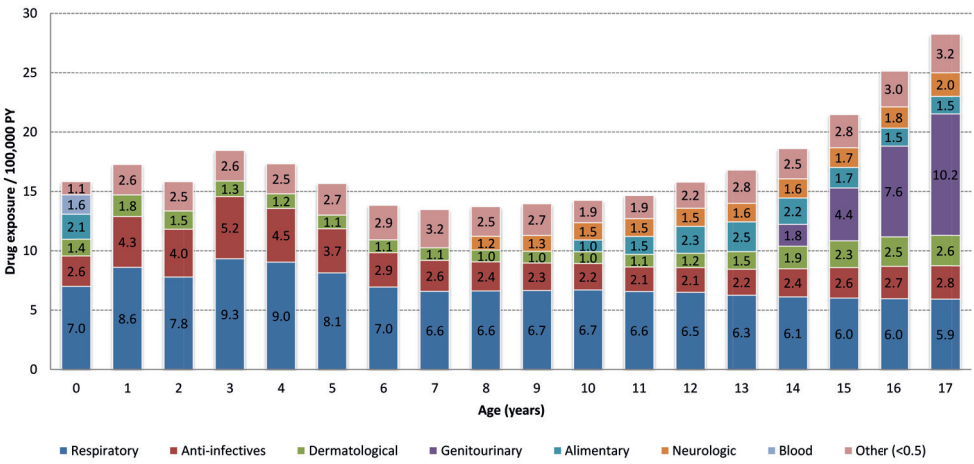
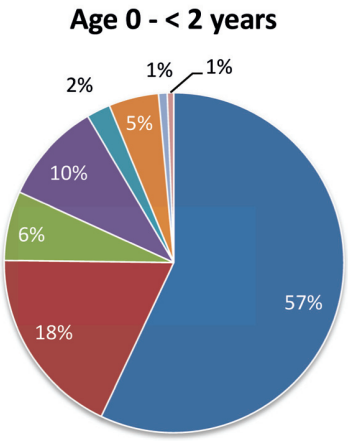
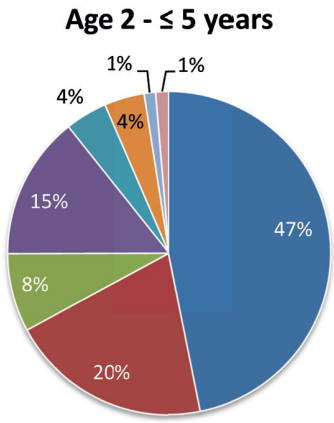
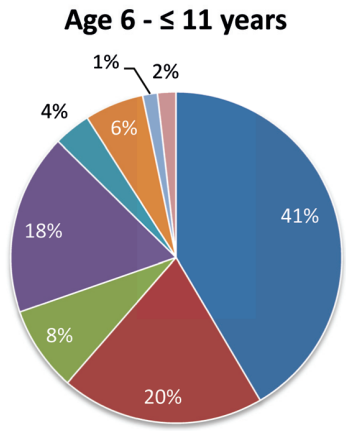
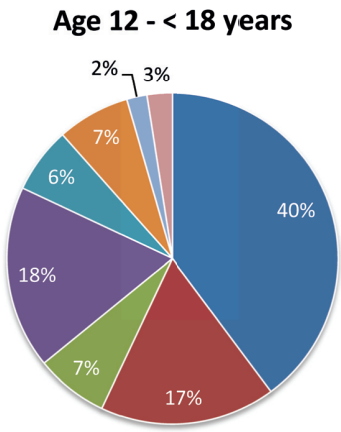


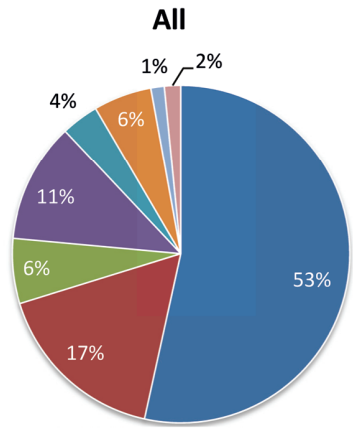
Figure 4.2.1 | Drug exposure in person-years/ 100 000 person-years by age *(see page 133)*.











Range PY	Incid /1,000 RR=2	Incid /1,000 RR=4	Incid /1,000 RR=6
<10	<765	<12	<5.2
>10 - ≤50	<160	<2.4	<1.1
>50 - ≤100	<80	<1.2	<0.55
>100 - ≤500	<16	<0.25	<0.11
>500 - ≤1,000	<8.0	<0.12	<0.055
>1,000 - ≤5,000	<1.6	<0.02	<0.011
>5,000 - ≤10,000	<0.8	<0.01	<0.0055
>10,000	≥0.8	≥0.01	≥0.0055

Figure 4.2.2 | Distribution of exposure in PY by age-groups (5th ATC level, chemical subgroup) ***(see page 137)*****.**

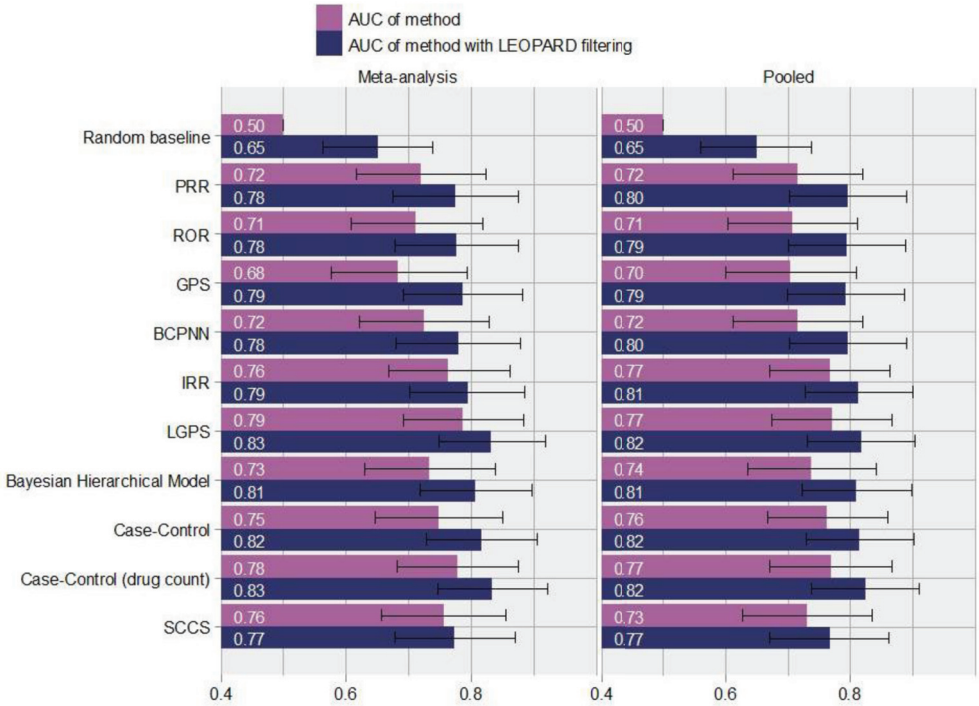


Figure 4.4.1 | Area Under the ROC Curve (AUC) for all methods, with and without LEOPARD filtering. Combination across databases was performed using meta-analysis for random effects (left panel) and pooling (right panel). Error bars indicate the 95% confidence interval. See Appendix for area under ROC curves ***(see page 182)*****.**

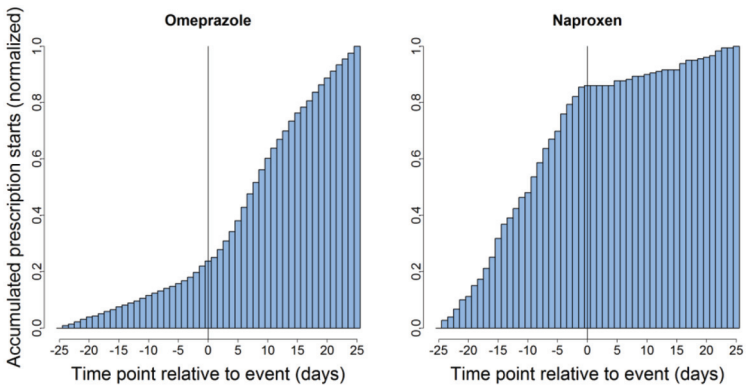


Figure 4.4.A1 | Empirical cumulative distribution functions of prescription starts in a window around upper gastrointestinal bleeding occurrences for Omeprazole and Naproxen ***(see page 187)*****.**

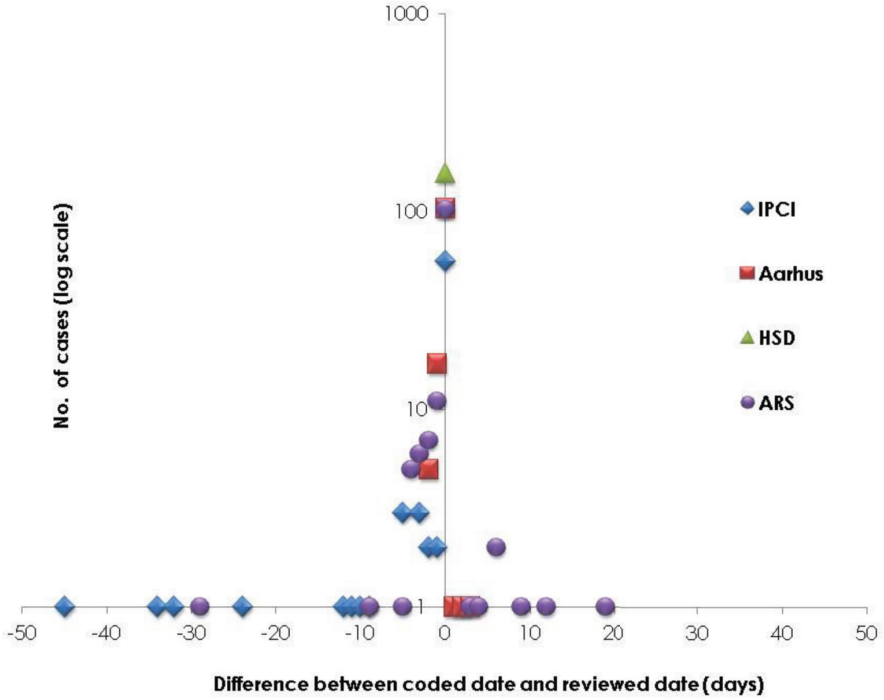


Figure 4.5.1 | Distribution of difference in days between coded date and reviewed date of event in confirmed UGIB cases only ***(see page 203)*****.**

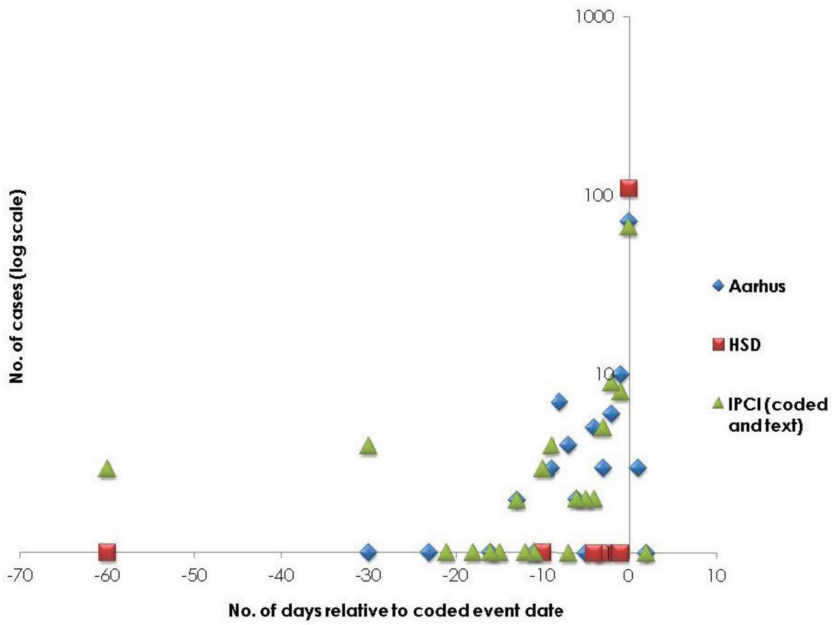


Figure 4.6.2 | Differences in the coded event date and date of onset of symptoms across the three databases ***(see page 221)*****.**

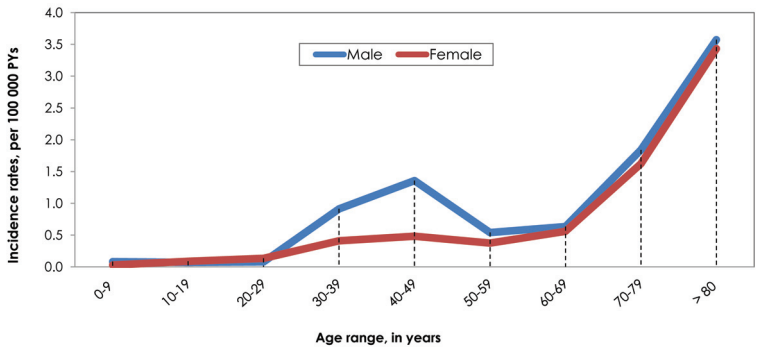


Figure 4.7.1 | Incidence rates of PML across age groups and sex (all databases) ***(see page 236).***

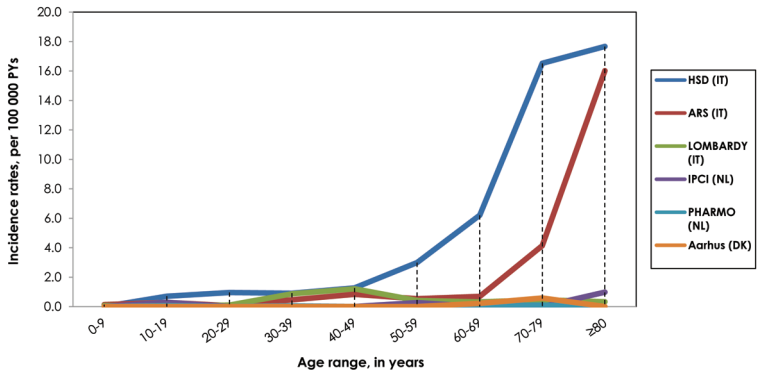


Figure 4.7.2 | Age-specific incidence rates of PML across the six databases from four countries ***(see page 236).***