

Opportunities and Challenges in Using Epidemiologic Methods to Monitor Drug Safety in the Era of Large Automated Health Databases

Elizabeth B. Andrews · Andrea V. Margulis ·
Patricia Tennis · Suzanne L. West

Published online: 12 October 2014
© Springer International Publishing AG 2014

Abstract Healthcare databases have been used in the past four decades to identify, refine, and evaluate potential safety signals of marketed medical products. Critics have challenged this research because data are from secondary sources and because some published studies have lacked robust methods for exposure and outcome definition and failed to adequately control for biases. We review the history of healthcare databases used in pharmacovigilance for quantifying adverse outcomes associated with therapeutics, methods to improve the quality of this research, and best practices for pharmacoepidemiologic studies. Drug and vaccine safety studies increasingly use information from multiple healthcare databases, with analyses that aim to keep patient-level identifying data with local research custodians. Analytic methods, including high-dimensional exposure propensity scores, use large numbers of variables to reduce confounding and further anonymize patient data. However, due to gaps in and complexities of the available databases, the value of the research depends on experts with knowledge about the clinical context (e.g., how products are prescribed and taken, how outcomes are diagnosed and recorded, what risk factors must be considered), understanding the nuances of individual databases and

the clinical practice patterns they represent, and utilizing study designs that minimize bias, particularly confounding by medication indication.

Keywords Pharmacoepidemiology · Drug safety · Signal detection · Electronic health records · Claims analysis · Propensity scores · Confounding factors · Bias epidemiologic · Pharmacovigilance · Product surveillance

Introduction

Existing healthcare databases, such as electronic medical records, insurance claims, and disease registries, have emerged over the last four decades as major sources of information on the safety of therapeutics such as medicines, biologics, vaccines, and devices following their marketing approval. Important adverse events can be identified in this general-use setting that might not be observed in preapproval clinical trials because the general-use setting includes more patients exposed to the therapeutics and greater diversity related to age, sex, comorbidities, co-medications, and treatment adherence.

Applications of these data for medical product safety evaluation fall into three general categories: (1) signal generation, that is data mining to identify new signals of possible but previously unknown exposure–outcome associations; (2) signal refinement, including routine and sequential monitoring for predefined exposure–outcome combinations to follow up on potential signals; and (3) signal evaluation, or protocol-driven studies for selected exposure–outcome combinations. The earliest uses of healthcare databases focused on signal evaluation, but all three applications are being used in some settings.

Cynics have long challenged the use of healthcare databases in the study of medical product safety. One of the earliest and most vociferous published criticisms came from Samuel Shapiro, who in 1989 evaluated ten published studies

E. B. Andrews (✉) · P. Tennis
RTI Health Solutions, 3040 Cornwallis Road, Post Office Box
12194, Research Triangle Park, NC 27709-2194, USA
e-mail: eandrews@rti.org

P. Tennis
e-mail: ptennis@rti.org

A. V. Margulis
RTI Health Solutions, Barcelona, Spain
e-mail: amargulis@rti.org

S. L. West
RTI International, 3040 Cornwallis Road, Post Office Box 12194,
Research Triangle Park, NC 27709-2194, USA
e-mail: swest@rti.org

against standard research validity criteria: (1) exposures and (2) outcomes should be appropriately defined; (3) exposure must precede the outcome in time; (4) bias and (5) confounding should be controlled; (6) findings should be internally and externally coherent; (7) findings should have statistical stability across logical strata such as medication dose; and (8) measures of association should have reasonable statistical precision, especially when increased risks are being excluded [1]. Of the ten early studies selected for review, Shapiro judged that six fulfilled none or only one of the eight criteria and only two fulfilled six or more. One may take exception to the specific criteria, their applicability to these studies, and the wisdom of publishing this critique; nevertheless, the paper reminded the research community that database studies deserve critical evaluation against scientific principles. In a commentary in 2010, David Grimes described research findings from database studies, including some conducted in the Danish national patient registries, as “garbage in, garbage out” [2]. Both papers sparked controversy and fueled a mostly healthy debate that led to improvements in database research. Over time, research methods, the quality and quantity of databases, and the sophistication of the research community evolved substantially. We have also seen the emergence of best practice guidance and better targeting of studies to databases that are ‘fit for purpose.’

Use of healthcare databases in pharmacoepidemiologic research has more recently been embraced with renewed enthusiasm, and even promoted by law in the USA [3]. The US Food and Drug Administration (FDA) Sentinel Initiative, the Observational Medical Outcomes Partnership (OMOP) [now part of the Innovation in Medical Evidence Development and Surveillance (IMEDS) program], and Exploring and Understanding Adverse Drug Reactions (EU-ADR; described further in this article) are projects that each employ a collective of health databases to address medical product safety issues, making data available on large numbers of individuals (up to hundreds of millions) but increasing the likelihood that numerically small exposure–outcome associations with very narrow confidence limits may lead readers to falsely accept the existence or absence of a true association, even if erroneous, biased, or biologically implausible.

Does the current enthusiasm for real-world data oversell the utility of these data sources for addressing all potentially important questions of medical product safety? This article traces the development of product safety research using healthcare databases, from single-database studies to research and monitoring programs using multiple databases. We review (1) the history of administrative claims, electronic health records, and multiple databases used for pharmacovigilance; (2) the evolution of methods to improve the quality of this research; and (3) best practice recommendations for meeting the challenges of conducting this research.

Healthcare Databases: From One to Many

When drugs receive regulatory approval, the *complete* safety profile is unknown [4], which is why many countries have established formal spontaneous adverse event detection and reporting activities [5]. However, the resulting signal information provides only qualitative information due to incomplete ascertainment of the count of adverse events (numerator of a measure) and the total number exposed to the product (denominator of a measure). More importantly, a noted signal does not imply a causal relation between the drug and the adverse event. In fact, an important next step is to consider whether the signal may be plausible, given the nature of the signal, mechanism of action of the product, and temporal and biological plausibility [6]. Once a safety signal has been identified and deemed a possibility, regulatory authorities typically request post-authorization safety studies that use appropriate study designs with comparison groups, rigorous methods, and data sources that can provide valid estimates of numerators and denominators. The need for this post-authorization research has influenced the design and analysis of pharmacoepidemiologic research studies using healthcare databases.

Single Databases

In the USA, use of single databases for pharmacoepidemiologic research began in 1979 when Jick and colleagues evaluated the association between postmenopausal estrogens and endometrial cancer using the Group Health Cooperative (GHC) of Puget Sound database [7]. GHC is a managed care organization that was initiated in Seattle, Washington, in 1947. GHC covered outpatient and inpatient care and prescriptions for approximately 250,000 members at the time the Jick et al. study was completed. Since then, US health insurance databases from commercial payers (e.g., UnitedHealthcare and WellPoint) and federal payers (e.g., Medicaid and Medicare) have been used for pharmacoepidemiology studies [8]. These claims databases typically have information on outpatient and inpatient services experienced, outpatient drugs dispensed, emergency care, mental health care, and laboratory and radiographic tests [8, 9]. They do not usually have clinical information such as the results of laboratory tests or vital signs such as blood pressure.

Similar database research has been conducted in Canada using administrative claims data from Saskatchewan province [10]. In the late 1980s, general practitioners in the UK established Value Added Information Medical Products (VAMP) to facilitate management of medical record data and build an information database [11, 12]. The VAMP database later became the General Practice Research Database (GPRD) and is now the Clinical Practice Research Datalink (CPRD). In The Netherlands, two databases were initiated, PHARMO [13]

and the Integrated Primary Care Information database [14]. Some websites maintain a catalog of databases that can be used for research, along with contact information [15–17].

Multiple Databases

Because an individual database may not be large enough to evaluate rare outcomes that may occur as a result of exposure to biologics or medications, initiatives such as the Vaccine Safety Datalink (VSD) and the HMO Research Network (HMORN) include multiple data sources. Sponsored by the US Centers for Disease Control and Prevention, the VSD was begun in 1990 to monitor the safety of vaccines using data from Kaiser Permanente Northwest, Northern California, and Southern California, and from GHC of Puget Sound [18]. It now uses data from an additional five healthcare organizations: HealthPartners, Minneapolis, Minnesota; Harvard Pilgrim Health Plan, Boston, Massachusetts; Kaiser Permanente, Colorado; Kaiser Permanente, Georgia; and Marshfield Clinic, Marshfield, Wisconsin [19]. Only Kaiser Permanente Northwest, Northern California, Southern California, and Colorado, GHC, and the Marshfield Clinic provide data routinely, with the remaining three participating in select studies (DeStefano F. In: West SL, editor, 2014, personal communication). Since its inception, VSD researchers have published almost 150 peer-reviewed articles on investigations related to influenza; diphtheria, tetanus, pertussis; rotavirus; human papillomavirus; pneumonia; measles, mumps, rubella; zoster; hepatitis; and meningococcal vaccines. They have also conducted studies describing the design and analysis of studies for signal detection of adverse events that might be associated with specific vaccines [20–24], vaccine coverage [19, 24, 25], and algorithms for identifying outcomes [26, 27], and they have published articles on using electronic healthcare databases for assessing vaccine safety [28–35].

HMORN was initiated in 1995 when several healthcare research networks decided to pool data to “increase sample size and population diversity” [36]. HMORN includes the same members as the VSD but has expanded to include nine others [37]. In the 19 years since its inception, HMORN researchers have published more than 2,000 articles, many of which were protocol-driven single-exposure outcome studies.

The success of HMORN’s Virtual Data Warehouse and distributed data model led to the FDA’s Mini-Sentinel pilot program, which was established to perform active surveillance [38] using 16 data sources, including HMORN and additional commercial insurance claims data sources (Table 1). In 2011, Canada launched the Canadian Network for Observational Drug Effect Studies (CNODES), which uses claims data from seven provinces: British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Quebec, and Nova Scotia [40].

Researchers in Europe have also developed an ongoing data network, the EU-ADR, consisting of eight databases

(one is a pediatric database) from The Netherlands, Denmark, the UK, and Italy, facilitating both surveillance [39, 41] and protocol-driven studies to evaluate signals [39] (Table 1).

Currently, healthcare data from more than 100 million individuals are available from the Mini-Sentinel project, 21 million from EU-ADR, and possibly another 40 million from CNODES, with other networks in development [41]. Cross-continent collaborations are ongoing, and the potential to pool data from across multiple continents exists.

Evolution of Methods Applied to Healthcare Databases

In this section, we briefly describe methods and processes to handle the increasingly vast amounts of information contained in automated health databases to conduct product safety research. Sound design and analysis are prerequisites for valid results, but mistakes in the design phase are often impossible to recoup later and may have a major influence on results. In contrast, analytic mistakes are often revealed because multiple analyses are performed, are straightforward to address.

Study Design Aspects

In the 1990s and early 2000s, a number of papers noted that jazz musicians did not live shorter lives despite a life of excesses [42], that Oscar winning actors and actresses lived longer than non-winning candidates [43], and Popes lived longer than artists [44]. Letters [45] and re-analyses of the original data [46–48] pointed to immortal-time bias, which had been described previously in textbooks [49] but not widely recognized. Two methods publications described this bias, caused by person-time not at risk of the outcome that is retrospectively assigned an incorrect exposure status or incorrectly excluded from the study population experience [50, 51]. These publications highlighted the methodological problems of using future information (accessible in database studies) to characterize exposure in observational cohort studies.

Disagreements in results from interventional and observational research on the coronary safety of hormonal replacement therapy led to development of the currently considered state-of-the-art new-user design [52]. With this design, only new users of the study exposures are eligible for inclusion, thereby reducing the risk of adjusting for factors that may be on the causal pathway and that may have been affected by treatment before study entry as well as ensuring that events occurring prior to study entry have been ascertained [52]. The best study-specific definition of new or incident drug use will address by a trade-off between internal validity and applicability, especially in the context of comparative effectiveness or safety research [53], in which the treatment of interest is compared with alternative treatments. Comparative

Table 1 Populations included in the US Food and Drug Administration's Mini-Sentinel Pilot and the Exploring and Understanding Adverse Drug Reactions (EU-ADR) programs

Data partner	Population size (approximate)	Source
Mini-Sentinel		
Aetna	45 million	http://www.aetna.com/about-us/aetna-facts-and-subsidiaries/aetna-facts.html
HealthCore, Inc.	43 million	http://healthcore.com/home/hc_home.php
HealthPartners	12 million	https://www.healthpartners.com
Group Health Cooperative	600,000	https://www1.ghc.org/html/public/about/index.html
Harvard Pilgrim Health Care Institute	1.2 million	https://www.harvardpilgrim.org/pls/portal/docs/PAGE/MEMBERS/ABOUT/FINANCIALS/FINANCIAL_FILES/Q1-2014-FINANCIALS.PDF
HealthPartners Institute for Education and Research	1.5 million	https://www.healthpartners.com/public/about/quick-facts/
Henry Ford Health System (Health Alliance Plan of Michigan)	> 675,000	http://www.henryford.com/body.cfm?id=38768 https://www.hap.org/
Marshfield Clinic Research Foundation	400,000	https://www.marshfieldclinic.org/about-us
Meyers Primary Care Institute	500,000	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3421389/
Humana's Comprehensive Health Insights, Inc.	11.3 million	http://www.competitive-health-analytics.com/research/data_assets.asp
Kaiser Permanente	9.3 million	http://share.kaiserpermanente.org/article/fast-facts-about-kaiser-permanente/
UnitedHealthcare	70 million	https://www.unitedhealthcareonline.com/b2c/CmaAction.do?channelId=6518c7958f5fa010VgnVCM100000c520720a_____
Vanderbilt University Medical Center	1.2 million	http://www.tn.gov/tenncare/
EU-ADR		
Italy		
Pedinet	160,000 children	
HSD	1.5 million	
Lombardy Regional	9 million	
Tuscany Regional	3.5 million	
Netherlands		
IPCI	1.5 million	
PHARMO	3 million	
UK		
QRESEARCH		
Denmark	1.8 million	
Aarhus		

HSD Health Search/CSD LPD, *IPCI* Integrated Primary Care Information

effectiveness/safety research combined with the new-user design allows researchers to minimize confounding by indication at the design stage by comparing subjects with similar baseline risks related to the indication [54], while providing clinicians with evidence directly applicable to their practice. Selection bias related to time on treatment and created by the healthy user effect is also eliminated at baseline [55].

Analytical Aspects

Starting in the late 19th century, techniques related to what we currently know as correlation and regression relied on stratification of data and tabular analyses [56] and were based on a single or a limited number of predictors. Now, we know of

numerous patient and physician characteristics that influence disease risk, prescribing, and diagnosing and that may act as confounding or effect-modifying factors in pharmacoepidemiologic research. Also, an individuals' healthcare services utilization affects prescribing and eventual diagnosis and is often treated as a confounding factor [57].

Healthcare databases containing administrative data or electronic medical records can contain a large number of variables. Although less so than early stratification methods, regression models may nonetheless be limited in the number of covariates they can accommodate [58, 59] depending on the frequency of the outcomes. Methods to limit the number of variables to include in statistical models while retaining the ability to control for confounding summarize, in one score, the

effect of many measured variables (exposure propensity scores [60, 61] and disease risk scores [62]) or rely on a proxy for measured and unmeasured confounders (instrumental variables [63]). Exposure propensity scores can take advantage of the availability of large numbers of variables for the study of common exposures and rare outcomes [64]. A semi-automated version of propensity scores called high-dimensional propensity score combines subject matter knowledge and epidemiologically appropriate automated variable selection from large pools of variables in healthcare data [65]. Variables derived from healthcare data, such as prescribing preference at the physician or hospital level, have been used as instrumental variables [66, 67].

Methods Appropriate for Multiple Data Sources

Meta-analytical techniques allow analytical results or individual-level data from different studies to be combined. While most commonly used to combine results from published literature, meta-analytical techniques can also be used to pool results from prospectively planned research. Thus, in prospectively designed multinational studies, which may use retrospective data, a parent protocol is adapted to the local data specifications to decrease study design heterogeneity across sites [68–70].

Newer approaches standardize automated healthcare data to a common data model to create very large analytic datasets not limited to specific exposure–outcome associations, including initiatives such as Mini-Sentinel [71••] and the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) [72]; these initiatives typically put in place processes for running common programming code as well. Mini-Sentinel data partners extract data, transform them into the Mini-Sentinel common data model, and load their source data, with the resulting data stored as tables within a relational database. The Mini-Sentinel Operations Center sends an executable program to the secure Mini-Sentinel Secure Portal to query the Mini-Sentinel Distributed Database. Each data partner runs the program on its own data transformed to the common data model behind its data security firewalls. Only aggregated, rather than patient-level, results are uploaded to the secure portal for retrieval by the Mini-Sentinel Operations Center. Mini-Sentinel researchers have also investigated a privacy-maintaining method that allows pooled analyses on individual-level data with full confounder adjustment [73•].

A different approach to data sharing has been taken in the signal-evaluation observational component of programs such as SOS (Safety Of non-Steroidal anti-inflammatory drugs) [70] and SAFEGUARD [74] aimed at studying specific groups of exposure–outcome associations. Research partners extract and elaborate data following criteria agreed among the research partners and create aggregated tables in a standardized format for submission to custom-built Jerboa software.

Jerboa combines the tables and runs the statistical analyses at the central level [39]. Appropriate software and hardware infrastructure ensures central data storage and remote secure data access by geographically dispersed research partners. The EU-ADR Alliance also relies on Jerboa software to create an ongoing platform that maintains the ability to study a wide scope of associations [41]. Methods developments within the EU-ADR include the harmonization of event definition and validation across databases [75–77].

Methods for sequential safety analysis have been in use for some time in VSD and are being implemented in Mini-Sentinel, including a variation of the log likelihood ratio test, the creation of propensity score-matched cohorts (sequential cohort designs), and implementation of generalized estimating equations [23, 31, 78, 79•, 80].

Improving the Methodology

OMOP, now part of IMEDS [81•, 82, 83], was a 5-year public–private partnership begun in the USA in 2008 that focused on identifying good methods for medical product safety research in healthcare data and establishing a shared resource for scientific collaboration [68, 84]. OMOP maintains a publicly available methods library including methods for sequential safety monitoring [85]. OMOP's research and tools have been instrumental in evaluating the utility of identical methods on results from a variety of US data sources [86] and the effect of applying a variety of study designs to address a single question [87]. OMOP methods have also been replicated in six European databases that contribute to EU-ADR [88].

Best Practices

As the availability of large clinical and claims databases increased, their use in non-interventional research to evaluate the effect of pharmaceuticals on health outcomes has increased. With this, it has been essential that principles of collaboration, patient privacy, and methodologic rigor be developed and followed. Table 2 lists available guidelines on a variety of pharmacoepidemiologic topics, stratified by scope of guideline.

The International Society for Pharmacoepidemiology guidance [91] recommends that a protocol be written before and followed during the conduct of any pharmacoepidemiologic study and that the staff be qualified for such research; the European Medicines Agency (2014) provides details on the organization of such a protocol. More recently, the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) methodologic guidance [90••] catalogs study designs and analytical methods that are

Table 2 Published sources that provide guidance on conducting pharmacoepidemiologic studies or studies with electronic healthcare data

Scope	Guidance and source
Guidelines on conducting pharmacoepidemiologic studies	<p><i>Guideline on Good Pharmacovigilance Practices</i>, European Medicines Agency [89]</p> <p><i>Guide on Methodological Standards in Pharmacoepidemiology</i> (Revision 3), European Network of Centres for Pharmacoepidemiology and Pharmacovigilance [90••]</p> <p><i>Guidelines for Good Pharmacoepidemiology Practices (GPP)</i>, International Society for Pharmacoepidemiology [91]</p> <p><i>Guideline on Conduct of Pharmacovigilance for Medicines Used by the Paediatric Population</i>, European Medicines Agency [92]</p>
Guidelines on conducting non-interventional studies using a large multipurpose healthcare database	<p><i>Guidance for Industry and FDA Staff. Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data</i>, Food and Drug Administration [93•]</p> <p><i>Guidelines for Good Database Selection and Use in Pharmacoepidemiology Research</i>, Hall et al. [94]</p> <p><i>GPS – Good Practice in Secondary Data Analysis: Revision After Fundamental Reworking</i>, Working Group for the Survey and Utilization of Secondary Data (AGENS) [95]</p>
Guidance on collaborations across multiple investigators and use of multiple data sources to increase the study size for a single study	<p><i>Mini-Sentinel Statement of Principles and Policies</i>, Mini-Sentinel Coordinating Center [96]</p> <p><i>Vaccine Safety Surveillance Using Large Linked Databases: Opportunities, Hazards and Proposed Guidelines</i>, Verstraeten et al. [97]</p> <p><i>Combining Electronic Healthcare Databases in Europe to Allow for Large-Scale Drug Safety Monitoring: the EU-ADR Project</i>, Coloma et al. [39]</p>

commonly used in non-interventional research of medications and describes the importance of implementing and documenting quality control and quality assurance procedures in non-interventional and randomized studies.

Guidance on the Use of Electronic Healthcare Databases

A number of guidances specific to use of electronic healthcare data sources in pharmacoepidemiologic research have been developed. To understand and select appropriate data sources for pharmacoepidemiologic research, one should consider data lags (i.e., time between occurrence of a medical service or prescription and its mention in the data source), sources of variable values (e.g., dispensed prescriptions vs. prescribed medications), sources of bias (e.g., lack of insurance coverage for selected services), population covered, within-patient linkage coverage (e.g., sufficient linkage between prescription files and hospital files), similarity of data sources if multiple sources are used to increase numbers of included patients [94], and reasons why patients ‘leave’ the data source (e.g., in commercial claims in the USA, reaching age 65 years and qualifying for Medicare). The FDA guidance also recommends that an assessment of inappropriate data or missing data be implemented and that the researcher understand the

limitations of the available data for exposure assessment [93•]. For example, exposure assessment often involves outpatient prescription data; however, actual consumption of a medication is rarely documented.

Outcome assessment in electronic healthcare data sources has similar challenges. Because outcomes are frequently ascertained via hospital claims diagnoses or clinical recording in electronic medical records, it is important to validate the outcome of interest [93•, 97]. Claims diagnoses are generally coded into categories that may not be specific enough, and claims diagnoses and free text in electronic medical records may occur because the diagnosis is being ruled out, not because the condition has been diagnosed. In addition, sensitivity analyses should be used to assess the effect of various definitions [93•].

Guidances for Collaborations

The growing use of multiple electronic health data sources requires cross-center rules for addressing technical issues and policy issues. In general, each collaboration has developed its own set of rules, and several have been published and serve as good models (Table 2).

Verstraeten and colleagues [97] recommend the following steps for data sources used for hypothesis generating and hypothesis testing: (1) evaluate the data quality; (2) provide a detailed description of data source and data linkage methods; (3) for screening, consider multiple comparisons; (4) use positive controls (e.g., seizures after pertussis vaccination); (5) split datasets into subsets and conduct multiple analyses to evaluate the consistency of results; (6) adhere to predefined statistical criteria; and (7) for vaccination studies, perform minimal matching or stratification on age, time, and socioeconomic status, if feasible.

The US Mini-Sentinel investigators have developed extensive policies and procedures [96] to address, among other things, agreements among collaborators on how to work and publish results, ensuring transparency, understanding public health practice versus research, maintaining patient privacy protections, ensuring use of minimum required patient-specific data, agreements on safety communications, ensuring protocol-based assessments, and conflicts of interest. Along similar lines, the ENCePP code of conduct has developed guidance to promote scientific independence and transparency in the implementation of pharmacoepidemiologic studies [98].

Other Recommendations

Many of the recommendations have been focused on signal evaluation, with much less attention to signal refinement/routine monitoring or signal generation using existing databases. The routine monitoring process generally involves multiple exposures and/or multiple outcomes and is addressed through standardized computer programs or modules that are not as focused on controlling for confounding and bias as signal evaluation studies. Signal evaluation studies can tailor the design, selection of covariates, and follow-up time to the specific exposure–outcome pair. This granularity is not yet possible in routine monitoring, and therefore results of such monitoring generally need additional confirmation through signal evaluation studies. A challenge faced in the era of frequent use of multi-database studies is that the sources used for routine monitoring may be the only good sources for signal evaluation studies. A group of North American and European pharmacoepidemiologists, convened at the request of the FDA, recently published recommendations regarding the implications of using signal refinement modular programs on the ability to perform signal evaluation studies in the same data sources [99].

Other recommendations not necessarily specific to pharmacoepidemiologic research include understanding sources of unidentified confounding, providing sufficient patient identity protections and data security [94], conducting quality checks at every step of analysis, starting with data extraction (e.g., subsetting of study patients from the original data source), and documenting all quality checks [94].

Complexities of Using Multiple Databases for Research

As described earlier, the use of multiple databases to evaluate drug safety signals has been ongoing since 1995 [36]. Country- and database-specific differences in diagnosis and treatment patterns have to be considered when pooling the results across the databases. Source data are extracted, transformed, and loaded into a common data model that uses universal coding schemes for consistency across databases. New methods were developed to deal with data that had to be kept behind local research center firewalls to maintain patient confidentiality and adhere to institutional ethical requirements. The efforts put forth and the progress made by those conducting research using distributed databases have moved the pharmacoepidemiology discipline forward by giant leaps. Some of the challenges these innovators faced are described below.

Because the data maintained by the data partners are often coded using differing coding nomenclatures and data structures, those working with multiple databases use a common data model to create a consistent layout for their research studies. The process of deriving a common data model requires extensive discussions to flesh out nuances specific to each partner's data (e.g., the meaning of the ordering and maximum number of the diagnosis codes on the outpatient and inpatient claims files, as well as whether to maintain the detail of the original data versus bringing along the added complexity of derived variables). Along with determining a common data model, multiple database projects promote consistency across databases by using centrally developed analytic code that is distributed to the data partners for execution on their data [41, 100].

Besides the need for procedures promoting consistency among the databases as described above, researchers using multiple databases for pharmacoepidemiologic research have also needed to develop new analytic approaches such as high-dimensional propensity scores [79•] for controlling bias and confounding to enhance the validity of the results [39]. Linkage across files within a single healthcare system is feasible in the EU-ADR because each of the four countries has national health identifiers. In the USA, additional methodologic research is needed to determine appropriate linkage algorithms so that individuals can be linked across databases (e.g., from different insurance companies or linking electronic health record data to claims) in order to construct longitudinal histories, exclude duplicate histories, or identify certain outcomes (e.g., cancer or death) [100].

Much has been accomplished by researchers using multiple databases for drug safety research, but more needs to be done. For example, a recent methodologic study conducted by OMOP researchers evaluated 53 drug–outcome pairs in ten different databases using two study designs: a new-user cohort design with propensity score adjustment and a self-controlled

case series [86]. The proportion of test cases that yielded estimates of the association that were consistent in the direction of the association across databases was 23 of 53 (43 %) for the cohort method and 18 of 53 (34 %) for the self-controlled case series method. This heterogeneity across data sources needs further exploration because it may affect the utility of multiple database studies.

Conclusion

Healthcare databases can be very useful for protocol-driven studies and for signal refinement, specifically monitoring of predefined exposure–outcome pairs for specific outcomes that have been validated in the individual database. Much progress has been made since the early studies. Databases have improved in number and quality, allowing us to study more products in diverse settings and to evaluate some outcomes that are extremely rare. Our knowledge of the strengths and limitations of the databases has improved as more validation studies have been conducted to compare secondary data to source data. Some databases, such as those reflecting electronic medical records, are the source data. The use of new methods such as the new user design, greater use of active comparators and propensity scores, and recognition of immortal time bias have improved the ability to control for confounding. Use of common data models and efficient methods for conducting analyses in a distributed fashion while pooling aggregated results have facilitated multicenter research in settings where patient-level identifiable data must remain behind the firewall of the individual data center, allowing for greater research collaboration and larger studies while protecting the privacy of patients' personal health information. Moreover, the evolution of policies and common expectations for collaborative research is creating a multinational research community in which multicenter database studies are becoming common.

As a result of these advances, most safety studies published in recent years fully address the research criteria enumerated by Shapiro [1], particularly those relating to exposure and outcome definition, bias and confounding, and coherence of evidence. When researchers cannot fully address these methodological questions through the research methods, they describe the limitations of methods and potential impact of the limitations, for example by conducting sensitivity analyses.

Despite these advances, the monitoring of medical product safety cannot yet be delegated to smart algorithms applied to healthcare databases. There is substantial heterogeneity across databases in content, coding systems and practices, duration of available medical history and follow-up time, and quality of outcome information (e.g., linkage to cancer registries and mortality details in selected databases), and databases reflect differences in clinical practice patterns. This heterogeneity

may be important for some exposure–outcome relationships but not for others. Understanding the sources and implications of heterogeneity is the subject of an active IMEDS research program [101].

Moreover, there are data gaps owing to the secondary nature of the data. Some of the data problems noted by Grimes [2] have been overcome by using appropriate exposure and outcome definitions, such as prescription records of diabetes medications to help identify patients with diabetes, rather than using only inpatient diagnosis data. However, many data sources do not include information on potential risk factors that affect health outcomes, such as use of illicit substances, use of over-the-counter medicines, smoking, and actual adherence to the medication. Some long-term outcomes, such as cancer, cannot be studied easily because of the relatively short-term follow-up contained in many databases. This limits the characterization of exposure dynamics and generally truncates follow-up before a reasonable latency period has elapsed for cancer development and diagnosis. Reasons for a physician prescribing one drug over another can be measured only by proxy indicators, meaning that confounding by indication remains a relevant topic of concern for many exposure–outcome pairs except when an active comparator drug is used interchangeably for the same indication.

Even if the data gaps were filled and methods and understanding of the data perfected, research is still limited by the number of persons with relevant exposures for some important exposures and populations. Thus, the spontaneous adverse experience reporting systems are still needed to generate signals for infrequent exposures and rare outcomes. Despite having 59,594,132 person-years of follow-up in the EU-ADR, it was estimated that it would be possible to detect relative risks for two outcomes of interest for commonly used medications: a relative risk of 2 for only 23 % of available medications for outcomes as frequent as myocardial infarction (an association that would not be detectable from spontaneous reports alone) and for only 1 % for events as rare as rhabdomyolysis [102]. Similar challenges have been observed for studies on non-steroidal anti-inflammatory drugs in pediatric populations in a multi-database project in Europe and studies of asthma mortality in users of long-acting β -agonists in chronic asthma in nine US databases [77, 103].

Use of healthcare databases for 'high-throughput' signal refinement and wide-scale signal generation activities may not be advisable until we can better target these applications to the right data sources appropriate for specific exposures and outcomes. Currently, such targeting is dependent on experts who are knowledgeable about the clinical context (e.g., how products are prescribed and taken, how outcomes are diagnosed and recorded, what risk factors must be considered). In addition, experts need to understand the nuances of individual databases and the clinical practice patterns they represent, and must be facile with methods for minimizing bias,

particularly confounding by medication indication. The increasing amounts of potentially linkable healthcare and non-healthcare records from multiple ‘big data’ sources add new challenges to such targeting. Without appropriate care, we risk finding an overabundance of false signals and false assurance from the absence of signals, with associated consequences for patients and the healthcare system.

Acknowledgments The authors wish to acknowledge the significant editorial contributions of Adele C. Monroe, and the review and suggestions by Kenneth J. Rothman.

Conflict of Interest E.B. Andrews, A.V. Margulis, P. Tennis, and S.L. West all declare no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Shapiro S. The role of automated record linkage in the postmarketing surveillance of drug safety: a critique. *Clin Pharmacol Ther.* 1989;46(4):371–86.
2. Grimes DA. Epidemiologic research using administrative databases: garbage in, garbage out. *Obstet Gynecol.* 2010;116(5):1018–9.
3. Food and Drug Administration. Public Law 110–85. Section 905. Active postmarket risk identification and analysis. September 27, 2007. <http://www.gpo.gov/fdsys/pkg/PLAW-110publ85/pdf/PLAW-110publ85.pdf>. Accessed 11 Aug 2014.
4. Honig PK. Advancing the science of pharmacovigilance. *Clin Pharmacol Ther.* 2013;93(6):474–5.
5. Jones JK, Kingery E. History of pharmacovigilance. In: Andrews EB, Moore N, editors. *Mann’s pharmacovigilance*. 3rd ed. Chichester: Wiley; 2014. p. 11–24.
6. Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965;58:295–300.
7. Jick H, Watkins RN, Hunter JR, Dinan BJ, Madsen S, Rothman KJ, et al. Replacement estrogens and endometrial cancer. *N Engl J Med.* 1979;300(5):218–22.
8. Strom BL, Schinnar R, Kinman JL. Overview of North American databases. In: Andrews EB, Moore N, editors. *Mann’s pharmacovigilance*. 3rd ed. Chichester: Wiley; 2014. p. 409–19.
9. West SL, Ritchey ME, Poole C. Validity of pharmacoepidemiology drug and diagnosis data. In: Strom BL, Hennessy S, Kimmel SE, editors. *Pharmacoepidemiology*. 5th ed. Chichester: Wiley; 2012. p. 757–94.
10. West R, Sherman GJ, Downey W. A record linkage study of valproate and malformations in Saskatchewan. *Can J Public Health.* 1985;76(4):226–8.
11. Jick H, Hall GC, Dean AD, Jick SS, Derby LE. A comparison of the risk of hypoglycemia between users of human and animal insulins. 1. Experience in the United Kingdom. *Pharmacotherapy.* 1990;10(6):395–7.
12. van Staa TP, Dyson L, McCann G, Padmanabhan S, Belatri R, Goldacre B, et al. The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. *Health Technol Assess.* 2014;18(43):1–146.
13. Herings RM, Stricker BH, Leufkens HG, Bakker A, Sturmans F, Urquhart J. Public health problems and the rapid estimation of the size of the population at risk. Torsades de pointes and the use of terfenadine and astemizole in The Netherlands. *Pharm World Sci.* 1993;15(5):212–8.
14. Vlug AE, van der Lei J, Mosseveld BM, van Wijk MA, van der Linden PD, Sturkenboom MC, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf Med.* 1999;38(4–5):339–44.
15. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Search database. 2013. <http://www.encepp.eu/encepp/search.htm>. Accessed 4 Aug 2014.
16. DGI, LLC. B.R.I.D.G.E. TO DATA®. 2014. <http://www.bridgetodata.org/Database-ProfileListing>. Accessed 4 Aug 2014.
17. International Society for Pharmacoeconomics and Outcomes Research. ISPOR international digest of databases. 2014. <http://www.ispor.org/DigestOfIntDB/CountryList.aspx>. Accessed 4 Aug 2014.
18. Wassilak SG, Glasser JW, Chen RT, Hadler SC. Utility of large-linked databases in vaccine safety, particularly in distinguishing independent and synergistic effects. The Vaccine Safety Datalink Investigators. *Ann N Y Acad Sci.* 1995;754:377–82.
19. McCarthy NL, Irving S, Donahue JG, Weintraub E, Gee J, Belongia E, et al. Vaccination coverage levels among children enrolled in the Vaccine Safety Datalink. *Vaccine.* 2013;31(49):5822–6.
20. Xu S, Newcomer S, Nelson J, Qian L, McClure D, Pan Y, et al. Signal detection of adverse events with imperfect confirmation rates in vaccine safety studies using self-controlled case series design. *Biom J.* 2014;56(3):513–25.
21. Jones SG, Coulter S, Conner W. Using administrative medical claims data to supplement state disease registry systems for reporting zoonotic infections. *J Am Med Inform Assoc.* 2013;20(1):193–8.
22. McClure DL, Xu S, Weintraub E, Glanz JM. An efficient statistical algorithm for a temporal scan statistic applied to vaccine safety analyses. *Vaccine.* 2012;30(27):3986–91.
23. Nelson JC, Cook AJ, Yu O, Dominguez C, Zhao S, Greene SK, et al. Challenges in the design and analysis of sequentially monitored postmarket safety surveillance evaluations using electronic observational health care data. *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 1:62–71.
24. Glanz JM, McClure DL, Xu S, Hambidge SJ, Lee M, Kolczak MS, et al. Four different study designs to evaluate vaccine safety were equally validated with contrasting limitations. *J Clin Epidemiol.* 2006;59(8):808–18.
25. Lin ND, Kleinman K, Chan KA, Yu XJ, France EK, Wei F, et al. Variation in hepatitis B immunization coverage rates associated with provider practices after the temporary suspension of the birth dose. *BMC Pediatr.* 2006;6:31.
26. Naleway AL, Gold R, Kurosky S, Riedlinger K, Henninger ML, Nordin JD, et al. Identifying pregnancy episodes, outcomes, and mother-infant pairs in the Vaccine Safety Datalink. *Vaccine.* 2013;31(27):2898–903.
27. Shui IM, Shi P, Dutta-Linn MM, Weintraub ES, Hambidge SJ, Nordin JD, et al. Predictive value of seizure ICD-9 codes for vaccine safety research. *Vaccine.* 2009;27(39):5307–12.
28. Baggs J, Gee J, Lewis E, Fowler G, Benson P, Lieu T, et al. The Vaccine Safety Datalink: a model for monitoring immunization safety. *Pediatrics.* 2011;127 Suppl 1:S45–53.

29. Hazlehurst B, Naleway A, Mullooly J. Detecting possible vaccine adverse events in clinical notes of the electronic medical record. *Vaccine*. 2009;27(14):2077–83.
30. Mullooly JP, Donahue JG, DeStefano F, Baggs J, Eriksen E. VSD Data Quality Working Group. Predictive value of ICD-9-CM codes used in vaccine safety research. *Methods Inf Med*. 2008;47(4):328–35.
31. Lieu TA, Kulldorff M, Davis RL, Lewis EM, Weintraub E, Yih K, et al. Real-time vaccine safety surveillance for the early detection of adverse events. *Med Care*. 2007;45(10 Suppl 2):S89–95.
32. Mullooly J, Drew L, DeStefano F, Maher J, Bohlke K, Immanuel V, et al. Quality assessments of HMO diagnosis databases used to monitor childhood vaccine safety. *Methods Inf Med*. 2004;43(2):163–70.
33. Chen RT, DeStefano F, Davis RL, Jackson LA, Thompson RS, Mullooly JP, et al. The Vaccine Safety Datalink: immunization research in health maintenance organizations in the USA. *Bull World Health Organ*. 2000;78(2):186–94.
34. Mullooly J, Drew L, DeStefano F, Chen R, Okoro K, Swint E, et al. Quality of HMO vaccination databases used to monitor childhood vaccine safety. Vaccine Safety DataLink Team. *Am J Epidemiol*. 1999;149(2):186–94.
35. Davis RL, Black S, Vadheim C, Shinefield H, Baker B, Pearson D, et al. Immunization tracking systems: experience of the CDC Vaccine Safety Datalink sites. *HMO Pract*. 1997;11(1):13–7.
36. Selby JV. Linking automated databases for research in managed care settings. *Ann Intern Med*. 1997;127(8 Pt 2):719–24.
37. Platt R, Davis R, Finkelstein J, Go AS, Gurwitz JH, Roblin D, 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(5):373–7.
38. Platt R, Camahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 1:1–8.
39. Coloma PM, Schuemie MJ, Trifiro G, Gini R, Herings R, Hippisley-Cox J, 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):1–11.
40. Suissa S, Henry D, Caetano P, Dormuth CR, Ernst P, Hemmelgam B, et al. CNODES: the Canadian Network for Observational Drug Effect Studies. *Open Med*. 2012;6(4):e134–40.
41. Trifiro G, Coloma PM, Rijnbeek PR, Romio S, Mosseveld B, Weibel D, et al. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? *J Intern Med*. 2014;275(6):551–61.
42. Spencer FJ. Premature death in jazz musicians: fact or fiction? *Am J Public Health*. 1991;81(6):804–5.
43. Redelmeier DA, Singh SM. Survival in Academy Award-winning actors and actresses. *Ann Intern Med*. 2001;134(10):955–62.
44. Carrieri MP, Serraino D. Longevity of popes and artists between the 13th and the 19th century. *Int J Epidemiol*. 2005;34(6):1435–6.
45. Rothman KJ. Longevity of jazz musicians: flawed analysis. *Am J Public Health*. 1992;82(5):761.
46. Hanley JA, Carrieri MP, Serraino D. Statistical fallibility and the longevity of popes: William Farr meets Wilhelm Lexis. *Int J Epidemiol*. 2006;35(3):802–5.
47. Redelmeier DA, Singh SM. Reanalysis of Survival of Oscar Winners. *Ann Intern Med*. 2006;145(5):392.
48. Sylvestre MP, Huszti E, Hanley JA. Do OSCAR winners live longer than less successful peers? A reanalysis of the evidence. *Ann Intern Med*. 2006;145(5):361–3. *discussion* 92.
49. Walker AM. Chapter 4: Time. In: Observation and inference: an introduction to the methods of epidemiology. Chestnut Hill: Epidemiology Resources Inc.; 1991.
50. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf*. 2007;16(3):241–9.
51. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167(4):492–9.
52. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915–20.
53. Johnson ES, Bartman BA, Briesacher BA, Fleming NS, Gerhard T, Komegay CJ, et al. The incident user design in comparative effectiveness research. *Pharmacoepidemiol Drug Saf*. 2013;22(1):1–6.
54. Brookhart MA, Patrick AR, Dormuth C, Avorn J, Shrank W, Cadarette SM, et al. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. *Am J Epidemiol*. 2007;166(3):348–54.
55. Sturmer T, Jonsson Funk M, Poole C, Brookhart MA. Nonexperimental comparative effectiveness research using linked healthcare databases. *Epidemiology*. 2011;22(3):298–301.
56. Stigler SM. Regression towards the mean, historically considered. *Stat Methods Med Res*. 1997;6(2):103–14.
57. Brookhart MA, Sturmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. *Med Care*. 2010;48(6 Suppl):S114–20.
58. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373–9.
59. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol*. 2007;165(6):710–8.
60. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127(8 Pt 2):757–63.
61. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163(12):1149–56.
62. Glynn RJ, Gagne JJ, Schneeweiss S. Role of disease risk scores in comparative effectiveness research with emerging therapies. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 2:138–47.
63. Heman MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology*. 2006;17(4):360–72.
64. Arbogast PG, Seeger JD, DECIDE Methods Center Summary Variable Working Group. Summary variables in observational research: propensity scores and disease risk scores. Effective Health Care Program Research Report No. 33. Agency for Healthcare Research and Quality; 2012. http://www.effectivehealthcare.ahrq.gov/ehc/products/441/1084/Decide33_SummaryVariables_FinalReport_20120515.pdf. Accessed 1 Aug 2014.
65. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology*. 2009;20(4):512–22.
66. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006;17(3):268–75.
67. Ionescu-Ittu R, Abrahamowicz M, Pilote L. Treatment effect estimates varied depending on the definition of the provider prescribing preference-based instrumental variables. *J Clin Epidemiol*. 2012;65(2):155–62.
68. Fireman B, Toh S, Butler MG, Go AS, Joffe HV, Graham DJ, et al. A protocol for active surveillance of acute myocardial infarction in association with the use of a new antidiabetic pharmaceutical agent. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 1:282–90.
69. Sturkenboom MC, Fourier-Re'glat A, Perez-Gutthaus S, Garbe E, Nicotra F, Romio S, et al. How to conduct collaborative studies in the EU? Lessons and results from the completed safety of NSAIDs (SOS) project [abstract]. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 3:S38.

70. SOS. WP6: conduct of additional observational studies. 2014. <http://www.sos-nsaids-project.org/?q=content/wp6-conduct-additional-observational-studies>. Accessed 4 Aug 2014.
72. Platt R, Carnahan RM, editors. Supplement: The U.S. Food and Drug Administration's Mini-Sentinel Program. *Pharmacoepidemiol Drug Saf.* 2012;21(S1):1–303. *This supplement is devoted to the structure, governance, methods, and findings of the Mini-Sentinel program.*
72. Andrade SE, Davis RL, Cheetham TC, Cooper WO, Li DK, Amini T, et al. Medication exposure in pregnancy risk evaluation program. *Matern Child Health J.* 2012;16(7):1349–54.
74. Rassen JA, Avorn J, Schneeweiss S. Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple health care utilization databases. *Pharmacoepidemiol Drug Saf.* 2010;19(8):848–57. *This paper presents a new method of sharing a small number of variables at the patient level, controlling for many other variables using propensity scores, while protecting privacy.*
74. SAFEGUARD. WP4: observational studies. 2011. <http://www.safeguard-diabetes.org/?q=content/wp4-observational-studies>. Accessed 4 Aug 2014.
75. Coloma PM, Valkhoff VE, Mazzaglia G, Nielsson MS, Pedersen L, Molokhia M, et al. Identification of acute myocardial infarction from electronic healthcare records using different disease coding systems: a validation study in three European countries. *BMJ Open.* 2013;3(6):e002862. doi:10.1136/bmjopen-2013-002862.
76. Avillach P, Coloma PM, Gini R, Schuemie M, Mouglin F, Dufour JC, et al. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. *J Am Med Inform Assoc.* 2013;20(1):184–92.
77. Valkhoff VE, Coloma PM, Masclee GM, Gini R, Innocenti F, Lapi F, et al. Validation study in four health-care databases: upper gastrointestinal bleeding misclassification affects precision but not magnitude of drug-related upper gastrointestinal bleeding risk. *J Clin Epidemiol.* 2014;67(8):921–31.
78. Cook AJ, Tiwari RC, Wellman RD, Heckbert SR, Li L, Heagerty P, et al. Statistical approaches to group sequential monitoring of postmarket safety surveillance data: current state of the art for use in the Mini-Sentinel pilot. *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 1:72–81.
80. Rassen JA, Schneeweiss S. Using high-dimensional propensity scores to automate confounding control in a distributed medical product safety surveillance system. *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 1:41–9. *This paper presents an automated method for variable selection using high-dimensional propensity scores.*
80. Coloma PM, de Bie S. Data mining methods to detect sentinel associations and their application to drug safety surveillance. *Curr Epidemiol Rep.* 2014;1(3):110–9.
82. Evans SJW. editor. Supplement: Studying the science of observational research: empirical findings from the Observational Medical Outcomes Partnership. *Drug Saf.* 2013;36(S1):1–204. *This supplement is dedicated to the reporting OMOP methods and results.*
82. Observational Medical Outcomes Partnership (OMOP) website. OMOP; 2014. <http://omop.org/>. Accessed 11 Aug 2014.
83. Reagan-Udall Foundation. Innovation in Medical Evidence Development and Surveillance (IMEDS) Program: IMEDS program overview. 2014. <http://www.reaganudall.org/our-work/safety-and-better-evidence/imeds-program/>. Accessed 11 Aug 2014.
84. Evans SJ. Moving along the yellow brick (card) road? *Drug Saf.* 2013;36 Suppl 1:S3–4.
85. Observational Medical Outcomes Partnership (OMOP). Methods library. OMOP; 2014. <http://omop.org/MethodsLibrary>. Accessed 4 Aug 2014.
86. Madigan D, Ryan PB, Schuemie M, Stang PE, Overhage JM, Hartzema AG, et al. Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol.* 2013;178(4):645–51.
87. Ryan PB, Stang PE, Overhage JM, Suchard MA, Hartzema AG, DuMouchel W, et al. A comparison of the empirical performance of methods for a risk identification system. *Drug Saf.* 2013;36 Suppl 1:S143–58.
88. Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RM, Pedersen L, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. *Drug Saf.* 2013;36 Suppl 1:S159–69.
89. European Medicines Agency. Guideline on good pharmacovigilance practices [all modules]. 2014. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000345.jsp&mid=WC0b01ac05804fdb1. Accessed 11 Aug 2014.
91. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Guide on methodological standards in pharmacoepidemiology (revision 3). EMA/95098/2010 Rev.3. 2014. http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml. Accessed 11 Aug 2014. *Of major importance, particularly for safety studies required by the European Medicines Agency.*
91. International Society for Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practices (GPP). *Pharmacoepidemiol Drug Saf.* 2008;17(2):200–8.
92. European Medicines Agency. Guideline on conduct of pharmacovigilance for medicines used by the paediatric population. Committee for Medicinal Products For Human Use (CHMP); January 25, 2007. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003764.pdf. Accessed 11 Aug 2014.
94. Food and Drug Administration. Guidance for industry and FDA staff. Best practices for conducting and reporting pharmacoepidemiologic safety studies using electronic healthcare data. US Department of Health and Human Services; May 2013. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM243537.pdf>. Accessed 4 Aug 2014. *Of importance, particularly for safety studies required by the FDA.*
94. Hall GC, Sauer B, Bourke A, Brown JS, Reynolds MW, LoCasale R. Guidelines for good database selection and use in pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf.* 2012;21(1):1–10.
95. Working Group for the Survey and Utilization of Secondary Data (AGENS), with representatives from the German Society for Social Medicine and Prevention (DGSMP) and the German Society for Epidemiology (DGEpi), and Working Group for Epidemiological Methods, with representatives from the German Society for Epidemiology (DGEpi), the German Society for Medical Informatics, Biometry and Epidemiology (GMDS) and the German Society for Social Medicine and Prevention (DGSMP). GPS – good practice in secondary data analysis: revision after fundamental reworking. January 2008. http://dgepi.de/fileadmin/pdf/leitlinien/gps-version2-final_ENG.pdf. Accessed 4 Aug 2014.
96. Mini-Sentinel Coordinating Center. Principles & policies. 2014. http://www.mini-sentinel.org/about_us/principles_and_policies.aspx. Accessed 4 Aug 2014.
97. Verstraeten T, DeStefano F, Chen RT, Miller E. Vaccine safety surveillance using large linked databases: opportunities, hazards and proposed guidelines. *Expert Rev Vaccines.* 2003;2(1):21–9.
98. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. The ENCEPP code of conduct for scientific independence and transparency in the conduct of pharmacoepidemiological and pharmacovigilance studies. Revision 3. February 21, 2014. http://www.encepp.eu/code_of_conduct/

- documents/ENCePPCodeofConduct_Rev3.pdf. Accessed 11 Aug 2014.
99. Toh S, Avorn J, D'Agostino Sr RB, Gurwitz JH, Psaty BM, Rothman KJ, et al. Re-using Mini-Sentinel data following rapid assessments of potential safety signals via modular analytic programs. *Pharmacoepidemiol Drug Saf.* 2013;22(10):1036–45.
 100. Curtis LH, Weiner MG, Boudreau DM, Cooper WO, Daniel GW, Nair VP, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 1:23–31.
 101. Reagan-Udall Roundation. Innovation in Medical Evidence Development and Surveillance (IMEDS): IMEDS-methods research agenda. 2014. <http://www.reaganudall.org/wp-content/uploads/2013/04/IMEDS-Methods-Research-Agenda.pdf>. Accessed 11 Aug 2014.
 102. Coloma PM, Trifiro G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? *Pharmacoepidemiol Drug Saf.* 2012;21(6):611–21.
 103. Tennis PS, Johannes CB, Camargo C, Sturmer T, Lanes S, Brown J, et al. Feasibility of ruling out small treatment-associated increase in asthma mortality risk [abstract]. *Pharmacoepidemiol Drug Saf.* 2013;22 Suppl 1:S34–5.