METHODS REPORT

Use of Secondary Population-Based Databases to Evaluate the Safety of Medications

Patricia Tennis, Elizabeth Andrews, Lee Lanza, and Catherine Johannes

September 2008

RTI Press

RTI INTERNATIONAL

About the Author

Patricia Tennis, PhD, MPH, FISPE, is a senior director, Epidemiology, at RTI Health Solutions' office in Research Triangle Park, North Carolina.

Elizabeth Andrews, PhD, MPH, FISPE, is vice president of Pharmacoepidemiology and Risk Management at RTI Health Solutions in Research Triangle Park, North Carolina.

Lee Lanza, ScD, MPH, is director, Epidemiology, at RTI Health Solutions in Waltham, Massachusetts.

Catherine Johannes, PhD, MS, is also director, Epidemiology, at RTI Health Solutions in Waltham, Massachusetts.

RTI Press publication MR-0007-0809

Suggested Citation

Tennis, P., Andrews, E., Lanza, L., and Johannes, C. (2008). Use of Secondary Population-Based Databases to Evaluate the Safety of Medications. RTI Press publication No. MR-0007-0809. Research Triangle Park, NC: RTI International. Retrieved [date] from http://www.rti.org/rtipress.

This publication is part of the RTI Press Method Report series.

RTI International
3040 Cornwallis Road
PO Box 12194
Research Triangle Park, NC
27709-2194 USA

Tel: +1.919.541.6000
Fax: +1.919.541.5985
E-mail: rtipress@rti.org
Web site: www.rti.org

doi:10.3768/rtipress.2008.mr.0007.0809

www.rti.org/rtipress

Use of Secondary Population-Based Databases to Evaluate the Safety of Medications

Patricia Tennis, Elizabeth Andrews, Lee Lanza, and Catherine Johannes

Abstract

Public concern is increasing over the safety of medicines, particularly serious adverse events detected after extensive use of products in the general marketplace. This concern has led to the need for prompt evaluation of safety signals within large populations following drug approval. The most relevant and available data resources primarily include electronic health care claims and electronic medical records and can be used to identify new safety issues and to evaluate known or suspected signals. In this review paper, we (1) summarize the data resources available for detection and evaluation of safety signals and (2) critically describe these resources and methods used in drug safety research. For each type of data resource, we summarize the characteristics and describe the associated applications and appropriate methods. To place each data resource and method in perspective, we provide examples from disease areas with substantial public health impact. We conclude that in certain circumstances these data resources can be valuable for the relatively cost-effective evaluation of serious adverse events in users of specific medications. However, implementation of such research requires a thorough understanding of the strengths and weaknesses of the data sources and the pharmacoepidemiologic methods used for analysis.

Contents

RTI INTERNATIONAL

Executive Summary

Signal detection involves screening for previously unknown adverse events that might be caused by use of a specific medication, vaccine, or biologic; frequently no *a priori* hypotheses exist to guide the screening. Signal detection is most commonly conducted within databases consisting of adverse events spontaneously reported to regulatory agencies and pharmaceutical manufacturers. However, occurrence rates of adverse events among users cannot be calculated using such data. Consequently, the results of signal detection studies should not be used to compare adverse events across agents. Furthermore, signal detection can rarely establish that the agent caused the adverse event.

More systematic data collection efforts utilized for signal detection involve outcomes surveillance, such as the Drug Abuse Warning Network, and exposure cohorts, such as Prescription Event Monitoring in the United Kingdom. Once a signal is detected, however, more rigorous analysis approaches are needed for evaluating hypotheses about the relation of the adverse event to one or more specific medications.

Recent developments in technology allow for the construction of large electronic databases that represent the accrual of data on insurance claims for health care utilization of large numbers of people (administrative data) and data from electronic medical records (EMRs) in integrated health care systems. Such multipurpose electronic data are useful for the conduct of nonexperimental research to identify rare outcomes, such as adverse events related to medication use.

As the Medicare Modernization Act (MMA) is implemented, prescription drug use data on a large segment of the Medicare population will soon be available. Linking these drug use data with Medicare health care claims available through the Centers for Medicare and Medicaid Services (CMS) will provide a rich source of data for signal detection and drug safety research, as have existing selected private insurance claims databases and the Medicaid database. Prior experience with existing databases can be used to inform the issues that could arise with use of future Medicare data for such research.

Large multipurpose electronic databases can be used in different types of studies to evaluate safety signals for selected medications. Such study designs could include cohort, case-control, or case-crossover studies and could address the following questions:

- What is the risk or incidence of a specific adverse event in people taking a new drug?
- Are users of a specific medication at a higher risk for developing a certain type of adverse event than users of other medications or nonusers?
- What characteristics of the users (e.g., age, sex, severity of disease, presence of selected co-morbidities) or use of the drug of interest (e.g., dosage) are associated with the occurrence of a particular adverse event?
- What is the frequency of the event in the general population?
- What are risk factors for the event in the general population?
- How frequently does the natural history of the disease treated with a specific medication include the suspected adverse event?

Characteristics of electronic administrative databases and EMRs from integrated health care systems make them advantageous for evaluating safety signals for medications and include the following:

- A sufficient number of enrollees to meet the large study size requirement for evaluating most hypotheses about rare serious adverse events;
- Feasibility of constructing comparator group(s) to contrast with the group using the drug of interest by using information available from all covered health care services;
- Relatively efficient electronic retrieval by diagnoses, drugs, or procedures, which usually allows quicker analyses and reporting than prospective data collection;
- Relative efficiency of grouping individuals based on diagnosis or medication codes, limiting the misclassification that can occur from use of free text written by health care providers in medical records;
- Inclusion of health care services provided from all providers covered by the associated plan(s); and

- Coverage of real-world health care in general practice and specialty clinics.

Further advantages associated with administrative databases include the following:

- Data on prescription drug dispensing, whereas clinical records may frequently be limited to prescribing information;
- Availability of enrollment data, which allow enumeration of the population at risk for the condition(s) of interest; and
- Routine record quality and completeness audits conducted by payers for prevention of financial abuse.

Weaknesses of electronic claims-based databases and EMRs from integrated health care systems that should be considered when designing and interpreting studies of drug safety include the following:

- Lack of specificity in some diagnostic codes, which may not always be clarified in other parts of the EMRs;
- Lack of patient-reported information, such as quality of life (outcomes that require this type of information cannot be studied); and
- The fact that coverage of real-world health care includes restrictions imposed by health plans on prescribing of new drugs, which could lead to additional confounding if patients who receive new medications are different from those who receive older products.

Further weaknesses associated with administrative databases (but not with EMR databases) include the following:

- Lack of patient-reported information and/or detailed clinical information, such as smoking history and family history;
- Some codes that cover multiple conditions, meaning that one cannot differentiate among conditions within a code; for most diseases, key clinical details are not directly reflected in claims;
- Lack of data on laboratory tests performed and associated results (although some information is available in selected claims databases);
- Lack of information on medications administered in an inpatient setting;
- For most databases, lack of linkage to underlying medical records for review to validate electronic data;
- Diagnoses that reflect the reason for a billable service, not necessarily a confirmed condition;
- To date, minimal representation of the unemployed and people older than 64 (i.e., loss-to-follow-up at age 65 years as enrollees convert to Medicare);
- Follow-up that is limited to 2 or 3 years because of changes in insurance as employment changes;
- Sources of error that include misdiagnosis or false ranking of primary diagnosis; incomplete record keeping; miscoding of diagnoses or procedures; failure to submit claims; transaction errors; lag time between service, filing, and adjudication; and incorrect record linkage across files;
- Inability to document over-the-counter use of products and use of professional samples; and
- Inability to identify deaths occurring outside of the medical setting without linking to other data sources.

Because of these limitations, electronic claims databases are most useful in the following circumstances:

- Medications of interest are commonly used prescription drugs.
- Adverse events of interest are consistently coded as hospital discharge diagnoses or can be identified as procedures that are billed.
- Important alternative risk factors for the event of interest can be identified through available claims.

In conclusion, the use of large population-based claims and other electronic health care data permit the evaluation of the frequency of serious adverse events in users of specific medications in a relatively cost-effective manner. However, implementation of such studies requires a thorough understanding of the strengths and weaknesses of pharmacoepidemiologic methods and the available data.

Introduction

Background

Evaluation of drug safety has evolved significantly over the past 40 years. Until the 1960s, medication safety was evaluated primarily by using data from Phase III clinical trials and reports of individual cases of adverse events. After the discovery of birth defects associated with thalidomide use, the spontaneous events reporting system was developed. This system provided an important tool and enabled the creation of a new practice—pharmacovigilance, the process of identifying previously unrecognized potential hazards of marketed drugs. Over the past 20 years, this tool was supplemented with the use of more formal epidemiologic methods to identify and assess medication safety.

As information technology developed and administrative and medical record databases grew and became more prevalent, these databases were used as the basis for large population-based evaluation of safety signals related to medication use. In addition, the number of very large randomized clinical trials has been increasing; they are sponsored by pharmaceutical companies and government agencies and have evaluated both the effectiveness and the safety of medications.

The awareness of the possibility of studying adverse events in larger and more diverse populations than in the past has led to a demand for more safety information on marketed drugs and quicker action in response to safety questions. Two additional forces have added strength to that demand: (1) highly visible recalls of drugs from the market because of safety issues and (2) increasing public reimbursement for prescription drugs, especially under Medicare, and the attendant concerns over costs and benefits.

What is not always clear to the public and policy makers is why more work is not already being done, and why such work is not completed more quickly in response to potential safety signals. Policy makers are often unaware of what study approaches could be used, what the strengths and limitations of each approach are, and what types of conclusions can be appropriately drawn from each approach.

The public pressure for more work in the arena of drug safety and the expanding access to electronic patient-based data can lead to important development work to expand the utility of many data resources. However, it can also lead to inappropriate research if conducted without an understanding of the limitations and strengths associated with any body of research.

The purposes of this paper are to (1) briefly summarize the data resources available for detecting safety signals and (2) critically describe the data resources and methods used to follow up and evaluate signals about drug safety.

For each type of data resource, we summarize the characteristics and describe the associated applications and appropriate methods. To place each data resource and method into perspective, we provide examples from disease areas identified by the Centers for Medicare and Medicaid Services (CMS) and the Agency for Healthcare Research and Quality (AHRQ)[1] as high-priority for critical evaluation of therapies. When we started this report, these disease areas include the following:

- arthritis and nontraumatic joint disorders,
- cancer,
- chronic obstructive pulmonary disease and asthma,
- dementia, including Alzheimer's disease,
- depression and other mood disorders,
- diabetes mellitus,
- ischemic heart disease,
- peptic ulcer disease and dyspepsia,
- pneumonia, and
- stroke and hypertension.

Signal Detection vs. Signal Evaluation

According to the World Health Organization (WHO), a drug safety signal consists of "reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented previously."[2] An adverse event is an unanticipated health problem that results in harm to the individual. To guide the discussion, we differentiate between signal detection and signal evaluation.

Signal Detection

Signal detection involves screening for previously unknown adverse events that might be caused by use of a specific medication. Sometimes there are known hypotheses about events owing to the nature or class of the drug of interest. Frequently, however, no *a priori* hypotheses exist to guide the screening, because some adverse events, such as a serious allergic reaction (e.g., anaphylactic shock) are idiosyncratic and unpredictable. Signals can arise from clinical trials or spontaneous reports after a drug is marketed. A signal indicates that a formal analysis of a possible causal relationship between a drug and a particular adverse event may be needed. Further evaluation of the signal is needed to assess causality by a candidate medication or determine the extent of the problem among users of the medication, particularly in the case of serious adverse events, such as heart attack or liver failure.

Signal Evaluation

Although drug safety signals may arise through individual case reports, signal evaluation calls for studies that require a more systematic approach to data gathering and analysis. Such studies may be clinical trials or nonexperimental epidemiologic studies. These studies generally require certain data elements for each person within the study population. These elements include timing and dose of the medications taken and presence of, timing, and clinical details of the adverse events. In addition, a comparison group of people who did not use the drug of interest is generally required to assess whether the people who used the drug of interest are experiencing the event at a greater frequency than people who did not use the drug. Only through quantifying and characterizing the occurrence of the events can signals be confirmed and placed into the appropriate context.

Examples of typical signal evaluation research questions include the following:

- Does use of COX-2 inhibitor drugs increase the risk of myocardial infarction?
- Is high dosing of statins, used to control elevated cholesterol, associated with increased risk of rhabdomyolysis (muscle weakness caused by the breakdown of muscle fibers)?

In the next two sections, we describe approaches to signal detection and signal evaluation separately. Within each section are descriptions of data sources used and methods applied to the evaluation of the data sources.

Signal Detection

Signal detection for adverse events is routinely conducted by drug manufacturers and regulatory agencies such as the US Food and Drug Administration (FDA), using accumulated data on spontaneously reported adverse events. Investigators also conduct signal detection using systematic data collection, such as surveillance registries, as their data sources. Because of the big differences between the two types of data sources and the approaches to using them, we address them separately in the sections below.

Spontaneously Reported Adverse Events

Systematic collection and evaluation of spontaneously reported adverse events is one of the fundamental functions of pharmacovigilance. A spontaneous report is a clinical observation that is not part of a formal study; it is a form of passive surveillance.[3] A physician, pharmacist, or patient reports a particular occurrence of an adverse event to a manufacturer or regulatory agency, often in the context of a discussion about whether such events have been previously reported, or in conversations between physicians and company sales representatives.

Reporting of such events requires that a clinician or patient perceives that the event occurred in association with use of one or more medications. Whether the medication(s) being used caused the event (i.e., whether it is a true adverse drug reaction [ADR]) is, however, frequently unknown.

Collection of Data

Events collected in the spontaneous reporting system are those that occur in patients treated in routine clinical practice, unlike the events collected in the restricted setting of a clinical trial. The FDA relies on voluntary reporting of adverse events by health care providers and consumers through the MedWatch program and on mandatory reporting by

manufacturers.[4] Requirements for the reporting of post-marketing adverse events have been harmonized across the United States, the European Union, and Japan through guidance issued by the International Conference on Harmonisation.[5]

In the US, all adverse events reported to the FDA are entered into the Adverse Events Reporting System (AERS), a client server, Oracle-based relational database maintained by FDA's Office of Drug Safety.[6] A similar US database, the Vaccine Adverse Event Report System, captures adverse events related to vaccines. The WHO, through the Uppsala Monitoring Center (UMC), maintains an international adverse event database, Vigibase.[2]

Evaluation of Data

Analysis of spontaneous adverse event reports is useful for generating hypotheses about ADRs that are unusual, unexpected, and infrequent and for events that occur close in time to the start of treatment or a change in dose. Patterns observed in the analysis of spontaneously reported adverse events result in the identification of safety signals (i.e., the identification of a possible ADR). Signal detection must usually be followed by further investigation to support or refute a causal relationship between the drug and suspected reaction, estimate the extent of the problem, determine the mechanisms, and identify special risk groups.[3] The evaluation methods for generating hypotheses regarding adverse events include careful review of individual case reports and systematic review and analysis of aggregated reports. These methods are described in the following sections.

Case Review

Potential safety signals are identified by ongoing manual review of individual reports by safety evaluators at regulatory agencies and at pharmaceutical companies. In the United States, reports of serious adverse events and important medical events, such as liver failure, cardiac arrhythmias, renal failure, and rhabdomyolysis, are electronically transmitted to the FDA postmarketing safety evaluators, who review them daily.

When the evaluators identify a potential signal, they conduct further research to investigate the signal. This approach involves collecting all similar cases in the database, searching for cases from international databases, searching the medical literature for case reports, and reviewing the series of cases for common characteristics. This research seeks out further evidence that could indicate a possible causal association between the drug and event, such as timing of exposure and onset of event, whether a dose-response relationship exists, or whether the effect makes biological sense.[6]

Review of Aggregated Data

Beyond individual case reviews, patterns of reporting can be assessed to help identify new signals. For example, spontaneous reporting rates may be generated by dividing the number of reported cases by a measure of the suspect drug's utilization, often the number of prescriptions dispensed or sold.

Identifying appropriate data sources and methods to estimate numbers of people who have used a specific medication nationally or worldwide has been a significant challenge. The reporting rates can be compared over time to identify trends. These reporting rates should not be considered estimates of the true rate of event occurrence in users of the drug, as the available denominator is usually not an accurate estimate of the number of persons in the population exposed to the drug, and the numerator, the number of reported cases, is typically a serious underestimate of the actual number of cases of the event in the population.

The FDA and the WHO are increasingly employing formal data mining techniques to augment the work of safety evaluators because of the vast amount of data collected. The goal of data mining is to identify combinations of drugs and adverse events that would warrant a more in-depth investigation. Data mining is a systematic, unbiased approach used to analyze large amounts of data for signal detection, but it does not replace the clinical judgment of safety reviewers.[7] Potential signals identified by data mining must still be confirmed by additional investigation.

Data mining techniques search for drug-event patterns that stand out from the background experience in the database. The expected number of a specific adverse event reported for any given drug is compared with the observed number. The expected number may be derived from the

frequencies observed in the past or may be derived based on the frequency of the adverse event reported for all other medications in the database.[8] Mining of spontaneously reported adverse events data (i.e., searching for signals) largely involves numerator-based methods such as those listed below and described briefly thereafter:

- short memory schemes,[9]
- reporting odds ratios,[10]
- proportional reporting ratios,[11] and
- Bayesian data mining:
  - empirical Bayes screening[12] and
  - Bayesian Confidence Propagation Neural Network.[12]

Short memory schemes compare the number of adverse events reported in the current time period with the number of adverse events during a prior reference period.[13]

Reporting ratios—such as the reporting odds ratios and proportional reporting ratios—quantify how many times more frequently the combination of drug *i* and event *j* is reported than would be expected to occur if reports involving drug *i* and event *j* were statistically independent.[8,14-17]

Reporting ratios are easily interpreted but are subject to large sampling variability, especially when the expected and observed frequencies are small. In such cases, the signal-to-noise ratio becomes small. To improve the signal-to-noise ratio in estimating the reporting ratio, analysts can use the Bayesian Confidence Propagation Neural Network and Empirical Bayes Screening.[14]

Examples of the application of proportional reporting ratios to multiple spontaneous reporting systems (e.g., ADROIT Yellow Card database in the UK, the WHO database, and the FDA AERS database) show that data mining algorithms not only detect signals earlier than in the past in most cases, but also detect signals that were missed by traditional signal detection.[11,12,18] However, no algorithm has emerged as the method of choice, no validated performance criteria exist, and using such algorithms can generate signals that require careful review by clinical experts.[14] Furthermore, the analysis may be very sensitive to the choice of a comparator agent or agents, particularly when the agent of interest is utilized by an age group not well represented in the spontaneous reports database or is commonly used by healthy people. Finally, the value of data mining is a function of the relative value of detecting true ADRs with these methods versus the cost of finding false positives. The value of detecting true ADRs is that new risks can be identified and consequent actions taken to reduce future risks. By contrast, identification of false positives can waste limited resources used to verify or refute each signal and may have adverse clinical implications if they lead to inappropriate warnings.

Limitations of Spontaneous Adverse Event Reporting Systems

Spontaneous reporting systems have important drawbacks. A spontaneous report typically stems from a caretaker, usually a physician, making a judgment about potential causality of the event that occurred. Reported events may not necessarily be caused by the medication used. Also, reporting may be influenced by recent publicity. Adverse event reports are less useful for identifying adverse events with a gradual onset, those requiring prolonged exposure to a drug, or those with a high background rate, because such events are less likely to be perceived as drug-related.[19] Other limitations of spontaneous reporting systems include the following:

- Incompleteness and poor quality of data (e.g., frequently, important clinical information is not reported);
- Underreporting, a problem inherent in all voluntary reporting systems (it has been long understood that only a fraction of relevant events are reported);
- Underreporting of certain types of events, such as events with long latency (e.g., cancer) and events with high background rates in the population;
- Absence of denominator data to place the number of reported events into context (one cannot calculate the rate of adverse events because the number of people who have used the medication is not known and frequently surrogates are used to estimate denominators); and
- Biases that influence reporting, such as time since drug release, publicity about possible events, the extent of use of the drug, and the type and severity of the event.[20]

These issues together or individually can result in biased comparisons of reporting rates among different therapies. For example, Niu and colleagues[21] showed that differences in reporting rates of serious adverse events associated with two different vaccines were not confirmed when a more rigorous epidemiologic cohort approach was applied to data from a large health maintenance organization.

Strengths of Spontaneous Adverse Event Reporting

Despite the limitations of spontaneous adverse event reporting systems, they are the only practical way to monitor all drugs throughout their complete life cycle in large populations. Experience over many decades has shown them to be useful in alerting the medical community and the public to serious drug safety issues, and adverse event systems have stimulated many regulatory actions.

The Drug Safety Research Unit (DSRU), an independent registered medical nonprofit organization, recently completed a review of publicly available information concerning 11 drug product withdrawals in the UK and US from 1999 to 2001. Of the 11 products, evidence from spontaneous reports supported the withdrawal of eight, four of which were withdrawn on the basis of spontaneous reports alone.[22]

Recently, a sudden increase in spontaneous reports of pure red-cell aplasia associated with epoetin use prompted a worldwide investigation that identified risk factors that could have contributed to the adverse event. After corrective interventions were established, there was an 83 percent reduction in reported cases.[19,23] This event was an example of a very rare and unusual outcome generating a high level of suspicion through the existing spontaneous reporting systems.

Systematic Data Collection

In contrast to spontaneous reporting of data, in which no actual rates of events can be calculated or compared across drugs of interest, a variety of sources of data are systematically collected in a manner that enables evaluation of patterns and rates of events. Some data sources are not necessarily designed for monitoring for safety signals related to medication use and do not systematically collect data on medication use. These sources include surveillance registries of outcomes such as birth defects and cancer.

Spontaneous Adverse Event Reporting
Definition
• Reporting of individual cases of possible adverse events
Methods for Analysis
• Review of individual cases
• Review of patterns of cases and reporting rates
• Data mining methods to find patterns of drug-event combinations that deviate from expected patterns if no association with a specific drug is found
Best Uses of Spontaneous Reporting Data
• Identify serious events that are too rare to be observed otherwise
• Identify serious events close in time to the exposure
Limitations
• Can rarely establish that the drug causes the adverse outcome
• Rates of events cannot be calculated
• Direct comparisons of adverse events frequencies across drugs are not valid

Other systems and approaches are designed around monitoring for drug safety signals. These approaches include surveillance programs that do include collection of information on medication exposure and large health-based databases that can be used to conduct drug safety signal detection, such as administrative health databases and electronic medical records (EMRs).

Outcomes Surveillance and Outcomes Registries

Public health surveillance systems that systematically collect data on outcomes within a large population can be used for identifying signals, although they do not routinely collect information on medication use. These systems include vital records, such as birth and death records, as well as specific registries (e.g., birth defects registries, cancer registries). These databases are used to examine patterns in outcomes over time, by geographic area, and by specific types of outcomes (e.g., patterns of bone cancer by age and country).

Unfortunately, many surveillance systems do not simultaneously collect data on medication use. For example, cancer registries do not collect data

on past medication exposures or pharmaceutical treatments. These types of surveillance systems can be used for "ecologic" studies, in which one observes patterns in numbers of events at a population level. For example, the introduction of fenoterol, a beta-agonist for asthma, in some countries was followed by an increased rate of asthma mortality, leading to a debate about whether the increase was caused by fenoterol.[24] Subsequent work evaluated the impact of corticosteroids in reducing asthma mortality.[25,26] Other examples include neonatal mortality, which was found to decrease after the introduction of neonatal lung surfactant therapy.[27]

Studies of time trends in population-based rates of exposures and outcomes may be useful in signal detection in a manner similar to trends in reporting adverse experiences, but time trends cannot be used to evaluate signals because of the "ecologic fallacy." The exposures and outcomes in such studies are observed at a population level, not at a patient level, which can lead to spurious conclusions. A decrease in suicide rates in the years following introduction of antidepressants, for example, cannot refute a potential increased risk of suicide in subpopulations of antidepressant users.

As a counterexample of a surveillance system that collects data on medication use, the Drug Abuse Warning Network (DAWN) collects data on acute medication exposures preceding emergency department visits.[28] Within this network, hundreds of hospital emergency departments report the numbers of cases associated with abuse of and/or poisoning from specific medications.[29] To place these reports into perspective, investigators calculate denominators of use from the Automation of Reports and Consolidated Orders System (ARCOS) database, to which manufacturers are required to report the total weight of medication distributed to medical institutions and pharmacies. By evaluating temporal trends for individual medications and comparing rates of abuse relative to the amount used across medications, researchers can monitor for signals of abuse.

Other outcomes, such as poisoning from and accidental ingestion of prescription medications that are not generally abused, may be reported through poison control centers, which are sometimes used for medication safety surveillance. However, the low frequency of events may hinder the evaluation of trends.

Using the DAWN data source, signals of abuse may be monitored via patterns in reporting over time and comparing reporting rates among similar medications. When signals arise, researchers/analysts can follow up to understand and remedy the problem. For example, a sudden increase in abuse cases of a particular medication within a small geographic area may be the result of inappropriate prescribing by a single physician, who can be targeted for intervention.

Another example of an outcomes registry that systematically collects data on medication use at the time of the event is the Acute Liver Failure Study Group, which collects data on acute liver failure at 22 tertiary care centers in the United States. Recent results from this registry confirmed that, for the United States, acetaminophen poisoning is the leading cause of acute liver failure and unintentional overdose is the leading form of acetaminophen toxicity. Recognition of this drug-related, preventable cause of acute liver failure may lead to measures to reduce its incidence in this country.[30]

Outcomes surveillance data are useful because they can be used for signal detection years after data collection. However, use of such data can often result in false signals, requiring linkage to other data sources to validate signals or to address specific questions about specific medications.

Case-Control Surveillance

A more formal approach to identifying adverse events and screening for medications that may cause such events has been described by the Slone Epidemiology Center.[31] The Center identifies all birth defects that occur within participating centers and interviews the mothers about medication use during pregnancy. Using cases of a specific birth defect and controls without the birth defect, the Slone Epidemiology Center identifies potential signals by conducting case-control surveillance to monitor for drug exposures in cases that are more frequent than in controls (see Case-Control Study, page 18, for more information on these studies).

On a regular basis the Center calculates odds ratios, which are approximations of risk ratios or prevalence ratios, to evaluate the associations for all combinations of birth defects and drugs (individually or in drug classes), accounting for age, sex, and geographic region. The investigators establish a priori rules for identifying potential signals. The screening analyses are considered to be the first step. More rigorous case-control analysis is performed for verification and for determining whether other factors related to the drug exposure and the adverse event could have been the source of the apparent relation between the drug exposure and the event.

The case-control surveillance approach is an efficient way to detect signals of relatively rare outcomes. If appropriate data on potential confounding factors are collected, the signals can then be evaluated relatively quickly. However, if a particular drug exposure is rare, identifying enough relevant cases for meaningful analyses may not be possible

Signal Detection in Cohorts

Prospective Data Collection

Another approach to screening for safety signals is to use data systematically collected on large groups of people who have received a medication of interest. When data are collected for this purpose, the organized effort is sometimes referred to as an exposure registry. The frequency of adverse events in the users of the medication of interest is quantified. However, because there is no simultaneous collection of data on an unexposed comparator group, it may not be clear whether the observed frequency in the exposed group is greater than what would be experienced by a comparable group unexposed to the medication. In some drug and vaccine exposure registries, historical experience from published literature, clinical trials, and population surveys have been used for comparison.

The Prescription Event Monitoring (PEM) system in the UK is similar in design to an exposure registry. PEM is a form of postmarketing drug surveillance that uses a noninterventional, nonexperimental cohort technique.[3] Unlike spontaneous reporting systems, PEM takes an active approach to surveillance. The program was initiated in the UK, but similar programs are in use in other countries, including New Zealand and Japan.[32,33] PEM is most suited for the monitoring of new drugs with anticipated widespread use but is also used for established products.[22]

The DSRU selects new or established drug products for study by PEM and notifies the Prescription Pricing Authority, the agency that reimburses all dispensed prescriptions in the UK. Under a confidential arrangement, the Prescription Pricing Authority provides the DSRU with electronic copies of all prescriptions for the drugs being monitored. Three to 12 months after the date of the first prescription for an individual drug in an individual patient, the DRSU sends a "green form" follow-up questionnaire to the prescribing general practitioner. This form collects information on events that may have occurred for each patient while taking the drug or just after stopping the drug. Recently, questions have been included to capture information on prior and concurrent illnesses and concomitant medications to aid in assessing causality and possible biases. To date, 90 completed PEM studies are listed on the DSRU Web site, with an overall general practitioner response rate of 56 percent and a mean cohort size per study of about 10,000 patients.

The usefulness of PEM in generating safety signals was shown in a DSRU study of 10,033 users of the antiepileptic drug vigabatrin conducted between 1991 and 1994.[34] Four cases of bilateral, persistent visual field defect were identified from the initial cohort; this signal prompted investigators to conduct a follow-up study that identified at least 30 more vigabatrin-associated cases, confirming the safety signal.

The strengths of PEM include the following:

- Provision of a numerator (number of events) and denominator (number of exposed patients) for a known period of observation.[35] Thus, the incidence of adverse events occurring after exposure to drugs can be calculated, and the exposure data may be more reliable than prescription data because they are obtained from dispensed prescriptions and do not count prescriptions that do not get filled.[3]
- Monitoring of patients in real-world clinical practice. Thus, this information is more representative than clinical trial data.

- Generation of more complete adverse event reporting compared with the spontaneous reporting system in the UK.[15]

  In a comparison of events detected by PEM with spontaneous reports in the UK, only 9 percent of all adverse events detected by PEM were reported via the spontaneous adverse event reporting system; however, the proportion was higher (53 percent) for serious adverse events.[15] Thus, spontaneous systems suffer from selective reporting of more serious events in addition to underreporting of events overall.

Limitations of PEM include the following:

- Underreporting, as demonstrated by the relatively low proportion of completed forms returned by general practitioners (from 50 percent to 70 percent).[35] This factor could result in an underestimate of the true incidence of adverse events and possibly distorted information about risk if the patients whose forms are returned differ substantially from those whose forms are not returned.
- Incomplete reporting of events not observed by general practitioners.
- Lack of information about drugs used in hospitals.
- Possible inadequate size of the study cohorts to detect very rare adverse events.[3]
- Limited follow-up that does not allow for the detection of longer-term outcomes or events with long latency periods, such as many forms of cancer.

In both PEM and exposure registries, analysts can calculate actual rates of adverse events by counting the number of events that occur in users of the drugs or in the person-time contributed by the users. Because exposure registries usually do not include an unexposed group for comparison with the exposed group for estimation of excess risk, creative analysis methods may be used to identify possible signals. For example, within a single PEM cohort, the frequency of events occurring during the first month of use is compared with the frequency of events during the second through sixth months of use.[36] This approach is based on the assumption that most adverse events associated with use of the new medication are most likely to occur during the first month and that the events occurring during the second through sixth months most likely represent events of other causes and serve as a "background" rate. This approach is useful for identifying acute reactions to the drug (e.g., allergic reactions), but it is not useful for assessing the long-term safety of a medication.

PEM researchers have created comparator cohorts, either (1) by using previously collected cohorts of individuals who were initiating use of other medications[37] or (2) by pooling all data from all past cohorts for any medication accumulated in the PEM database.[38]

By comparing the adverse event frequencies occurring in the current cohort of interest with the adverse event frequency in the cumulative PEM database, the investigators can calculate relative risks for each event type, and they then identify potential signals from the highest relative risks.

Use of Administrative Data and Electronic Medical Records

Approaches have been used to systematically screen for drug safety signals in administrative data[7] and in EMRs. Safety signals are generated using data mining techniques on prescription and outcomes records to identify diagnoses that occur at increased frequency within the exposed group after drug initiation compared with other drugs. Similar analytic methods are used as described above for data mining. These analyses can be updated over time, as the treatment patterns may change as more people take new drugs and physicians change their prescribing patterns.

The main strengths of this signal detection method are the systematic ways in which adverse events are detected, in contrast to voluntary reporting systems. However, this method is limited by the nature of the data, which often do not contain direct clinical measures of underlying disease severity, results of diagnostic tests or procedures, or verified evidence that patients actually used the drugs. In addition, follow-up time is often limited in some databases, so the detection of events occurring after long periods of exposure is difficult. In signal detection approaches that use administrative claims data, data mining is likely to generate large numbers of false-positive signals of irrelevant events because the outcomes used are based on ICD-9 (International Classification

of Disease, Ninth Revision) codes. Individual codes can represent serious events, such as myocardial infarction, which could be drug-related. Because all medical services are screened, many other irrelevant codes—including delivery of a newborn, broken bones, and others that are unlikely to be related to medication exposure—could arise as potential signals. Because this approach is intended for signal detection, it would not replace the need for more in-depth confirmatory research. As with any data mining activity, the value of identifying true signals must be weighed against the costs of generating false signals, including resources applied to verifying and refuting them and clinical consequences of warning based on a false signal.

Drug safety signaling techniques used in the postmarketing setting can be applied to several data sources, including spontaneous reports, prospectively collected cohorts and case-control samples, and existing large multipurpose electronic health care databases. Yet the availability of standards is quite variable across these data sources. Spontaneous reports, which are consistently required by regulatory agencies for all drugs, are the subject of the most extensive regulation and standardization. Table 1 summarizes the approaches to signal detection, and Table 2 summarizes selected strengths and limitations of these approaches.

Table 1: Summary of approaches to signal detection

Type of Approach	Definition	Types of Events	Information on Medication Use	Methods of Signal Detection	Examples
Spontaneous reports	Collection of individual events and associated medication use reported by clinicians and patients	Any reported medical event is included	Provided by reporter for individual cases	• Review of single cases • Review of characteristics of similar cases • Time trends in number reported • Data mining techniques	• AERS • WHO
Outcomes surveillance and outcomes registries	Systematic collection of data on all cases of a specific outcome that occur within a population	Single type of event (see examples)	None collected	Ecologic studies that compare total number of cases with total sales of medications	• Birth defects registries • Cancer registries
Outcomes surveillance for known drug-related outcomes	Systematic collection of data on all cases of a specific outcome; specifically designed to monitor safety of medicines	Single type or group of events (see examples)	Medication use is systematically collected	Descriptive analyses (e.g., of all events that were reported, proportion that occurred in association with a specific medication)	• Drug Abuse Warning Network • Acute Liver Failure Study Group Registry
Case-control surveillance	Systematic collection of data on medication use for a specific outcome or group of related outcomes	Specific event or group of events	Systematically collected for all	Case-control analysis—compare frequency of medication use in one type of case with frequency in other types of cases	Birth defects case-control surveillance
Exposure cohorts or registries	Systematic collection of data on outcomes occurring in a group of people using a specific medication	All incident events of any type	Inclusion criteria based on medication use	Incidence of specific event (1) compared with incidence in another population or (2) during first 2 months of treatment compared with incidence during later periods	• Pregnancy exposure registries • Prescription event monitoring
Administrative data cohort surveillance	Systematic review of incidence of all events within a cohort of people dispensed a specific medication and an appropriate comparator group identified by administrative claims database	Any event identified by a diagnosis code associated with specified types of medical service	Claim for a specific medication = basis for inclusion criteria	For each medical code associated with a pre-specified type of service, compare incidence in user group with incidence in comparator group	Prescription drug screening for subsequent carcinogenicity[39]

AERS = adverse events reporting system; WHO = World Health Organization.

Table 2. Strengths and limitations of various approaches to signal detection

Type of Approach	Definition	Strengths	Limitations
Spontaneous reports traditional analysis	• Review of single cases • Review of characteristics of similar cases • Analysis of temporal trends in reporting rates	Astute clinician observation and subsequent reporting can result in very early identification of a signal	• Small proportion of events are reported • Number of users is unknown and rates of events cannot be evaluated
Spontaneous reports data mining	Comparison of observed vs. expected numbers of event/medication combination	Efficient approach to screening large numbers of reports	• Potential large number of false signals • Potential for biased reporting can influence results
Outcomes surveillance and outcomes registries	• Systematic data collection on all cases of a specific outcome that occur within a population • Registries usually involve follow-up	• A resource for identifying outcomes of subjects identified from other data sources if linkage can be made • Useful for ecologic analyses where total numbers of cases are correlated with total use of medication	No information on medication use is collected; therefore, signals regarding medication use cannot be detected unless linkage to other data about exposures is feasible
Outcomes surveillance for known drug-related outcomes	Systematic collection of data on all cases of a specific outcome or event in a defined population and specifically designed to monitor safety of medicines	Useful for generating hypotheses about causes for the events	Without data on medication use in the population giving rise to the reported outcomes, false signals based on numbers of cases can easily arise
Case-control surveillance	Use of the case-control study design to identify risk factors, including medications, for a specific outcome or group of related outcomes	More methodologically rigorous than spontaneous reporting and outcomes surveillance	• Rates of events cannot be estimated • Limited by number of cases, so signal detection is limited to commonly used medications • Requires *a priori* ideas about possible risk factors to ensure that appropriate data are collected
Exposure cohorts or registries	Systematic collection of data on incidence of events occurring in a group of people using a specific medication	• Prior hypotheses about events of interest are not required • Rates of events can be estimated	• Limited by numbers of people exposed • Can be resource-intensive • No comparison group for evaluating potential causality • Potential for bias if there is selective loss to follow-up • Not likely to detect events involving long latency
Cohort surveillance based on administrative data or electronic medical record databases	Systematic review of incidence of all events within a cohort of people utilizing a specific medication and within an appropriate comparator group identified by administrative claims database	• Based on readily available data • Allows calculation of rates and controls for some confounding	• Potential to generate false signals • Can be a challenge to distinguish a real signal from multiple false signals

Signal Evaluation

General Approach

As noted previously, signals of a safety concern may arise from spontaneously reported adverse events, systematic surveillance, published case reports, and clinical trials. Once a signal has been detected, the next step frequently involves conducting one or more nonexperimental studies to test hypotheses about the signal.

Of course, in some circumstances, large randomized clinical trials may be the only way to definitively answer a comparative safety question, and large simple trials are being used increasingly by government agencies and pharmaceutical companies. Randomization of patients to the medication or vaccine of interest and a comparator or placebo may be the only way to fully control for confounding by indication at the initiation of treatment. However, a discussion of the use of large trials for safety evaluation is not the focus of our paper.

Nonexperimental studies are usually focused on assessing whether the signal represents a real safety concern in the setting of actual clinical practice, unlike the more restrictive clinical trial setting. Nonexperimental studies may be the only practical way to study rare adverse events for a number of reasons. Once a drug is on the market, conducting clinical trials may sometimes be considered unethical, especially if effective treatment would be denied to some patients assigned to either placebo or an inferior treatment. In addition, a trial of the size or duration of follow-up necessary to detect rare outcomes is extremely expensive and sometimes logistically not feasible.

Typical hypotheses and associated operational research questions to evaluate safety signals include the following:

- What is the risk or incidence of a specific adverse event in people taking a new drug?
- Are users of a specific medication at a higher risk for developing a certain type of adverse event than users of other medications or nonusers?
- What characteristics of the users (e.g., age, sex, severity of disease, presence of selected comorbidities) or use of the drug of interest (e.g., duration of use, daily dosage) are associated with the occurrence of a particular adverse event?
- What is the frequency of the event in the general population?
- What are risk factors for the event in the general population?
- How frequently does the natural history of the disease treated with a specific medication include the suspected adverse event?

In the last decades of the twentieth century, the increasing availability of electronic record keeping made some drug safety studies feasible, when earlier they would have been too costly or time-consuming. In the arena of government-sponsored reimbursement programs, much work has been conducted in electronic records to assess quality of care, access, and other policy issues.[40] However, little work has been conducted comparing the safety or effectiveness of medication. Such work is the focus of an initiative established by the Agency for Healthcare Research and Quality (AHRQ) in 2005, the DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) Network of research centers.[1]

This section reviews some of the major sources of automated data for drug safety research and describes their strengths and weaknesses and examples of their use.

Methods and Study Designs

Investigators use a variety of study designs to test hypotheses about drug safety signals (see Table 3). The choice of the study design and analysis methods depends upon the research question, the types of data available, and resources available to conduct the study. In the following sections, we first describe confounding by indication, a central issue that must be addressed within nonexperimental comparisons of treatments. We then describe the most commonly used study designs for evaluating drug safety signals, followed by a description of the types of data sources and associated issues in using these data for evaluating drug safety signals.

Table 3. Strengths and limitations of various epidemiologic study designs for signal evaluation

Study Design	Description	Strengths	Limitations
Cohort study	Follow exposed (and/or unexposed) groups forward in time to measure AE incidence	• Can assess frequency (incidence) of AE • Can study more than one type of AE • Can identify patients lost to follow-up and potential related bias	• Requires large numbers of people to test hypotheses about rare AEs • Unmeasured or poorly measured confounders can complicate the analysis
Case-control study	Identify cases (people with AE) and controls (people without AE) and assess exposures and other risk factors backward in time	• Can estimate the ratio of rates of AEs • Can test multiple hypotheses about factors associated with the AE (case status) • Number of people required is substantially smaller than for cohort study	• Cannot measure incidence • Selection of appropriate controls to minimize bias can be complicated • Medication exposures are often too rare to evaluate
Case-crossover study	In cases, compare frequency of exposure immediately before AE to frequency of exposure at another time period	• No need for external controls (an individual serves as his or her own control) • Useful to control for some types of unmeasured confounding	• Cannot measure incidence • May not control for confounding by indication if disease severity changes over time within patients

AE = adverse event.

Confounding by Indication

It is critical to design all studies so that other factors associated with high risk of the event of interest do not bias the results of the study. When other known factors can cause the adverse event of interest, and when those factors occur more frequently in one group of the groups being compared, the result is referred to as confounding. For example, when comparing the risk of suicide among people using specific antidepressants, identifying and adjusting for differences in severity of depression among the users of the antidepressants being compared is essential. In this case, if one does not control for the severity of depression when comparing drug exposures, the risk may seemingly be "caused by" the antidepressant being used to treat more severe depression.

This type of confounding, in which the drug of interest is used to treat a specific subset of patients who may be at higher risk of complications in general, is called confounding by indication. It is one of the most common types of confounding encountered in nonexperimental studies that are aimed at evaluating safety signals for medications. Several strategies are available to reduce confounding by indication. These include multivariable regression and propensity score methods. These methods, however, typically can address only part of the bias that stems from confounding by indication.[41] To the extent that confounders remain unmeasured (e.g., smoking in a claims database study of respiratory events), most analytic methods cannot remove their impact.

Multivariable Regression

Multivariable (linear or nonlinear) regression can be used to estimate treatment effects in nonexperimental data by regressing the outcome on the covariates and including an indicator variable for treatment status and interactions between the treatment variable and each of the covariates. A large coefficient for a term in the equation that represents treatment indicates a treatment effect.

A disadvantage of multivariable regression is the fact that the regression model depends on the correct specification of the functional form of the relationship (e.g., linearity or log linearity) between the outcome and the covariates. Such specific assumptions may not be a problem when the treatment groups have similar covariate distributions. When, however, the covariate distributions in the two groups are very different, linear regression models depend on the specific form of the model to extrapolate estimates of treatment differences.

An advantage of multivariable regression is that a regression model may indicate a difference between the outcome frequency of one treatment group and the outcome frequency of another group due to an interaction with other covariates, such as age, sex, or geographic region. In addition to estimating any treatment effects, the regression model also describes the effects of other measured covariates. Of course, the fact that some important covariates may not be measured must be considered in the interpretation of results.

Propensity Scores

Propensity scoring is a strategy available to researchers using observational data to balance an entire collection of background characteristics between treatment groups; this enables them to make valid statistical inferences about the differences in treatment patterns or effectiveness.[42] The propensity score provides a scalar summary of the covariate information and is a measure of the likelihood that a person would have been treated with a specific treatment using only their covariate scores.[43] The propensity score can be estimated using discriminant analysis or logistic (or probit) regression. Both of these techniques lead to estimates that rank the probability of receiving treatment as a function of observed variables. However, the observed variables are assumed to have a multivariate normal distribution (conditional on treatment assignment) when discriminant analysis is used, whereas this assumption is not needed for logistic regression.[44]

Propensity scores can be used to adjust treatment effects by matching,[43] by stratification,[43] by regression adjustment,[43] or by weighting individual observations by the inverse of their propensity scores.[45] These methods were recently compared by Kurth and colleagues in evaluating the effect of tissue plasminogen activator on mortality among ischemic stroke patients.[46] By definition, subjects in the treated and control groups with equal (or nearly equal) propensity scores will tend to have the same (or nearly the same) distributions on their background covariates. Exact adjustments made using the propensity score will, on average, remove all of the bias in the background covariates.[47] Therefore, bias-removing adjustments can be made using the propensity scores rather than all of the background covariates individually.

Propensity score methods provide a flexible and convenient way to adjust for preexisting between-group differences. The choice of factors should be based, in part, on prior research that has identified factors as shaping service use or health outcomes.[48] Furthermore, the efficiency of the methods depends on the proportion of individuals in each treatment group and the adequacy of the sample size in addition to the degree of overlap of the propensity scores between treatment groups. This latter condition is described in more detail in Rosenbaum and Rubin.[47] As with other methods, confounding cannot be reduced if information on important confounders has not been gathered for the analysis.

Cohort Study

The cohort approach is used to accomplish the following:

- Measure the frequency at which an adverse event occurs within a group of people who use a specific medication or who have a particular disease;
- Evaluate whether one group of people using a particular medication is experiencing a particular adverse event more frequently than another group not using the medication; and
- Identify whether other measured factors are related to increased frequency of the adverse event. Identification of these factors can be accomplished only if these factors are measured in the cohorts of people under study.

In a cohort study, a group of people who meet specified inclusion criteria (e.g., fall within a specified age range during a prespecified calendar period) and who have used or have been dispensed the medication of interest (exposed group) and a group of unexposed persons (comparator group) are followed forward in time to quantify their experience and to assess the frequency of newly developed adverse events. By comparing rates of the events (number of events that occur during the follow-up time) in the exposed and unexposed persons and analytically adjusting for confounding factors, the

researcher can estimate the relative rate of the event among those exposed to the drug of interest relative to those who are not exposed.

In general, to evaluate the period of highest risk for many events, analysts will select a population receiving a new medication for the first time (i.e., an inception cohort). The alternative approach, namely, including all users of a product, will generally underestimate a risk because that cohort will include people who have already demonstrated tolerance to the product.

Persons who use the drug of interest may be compared with persons who do not use the medication and/or to those who use a comparable medication given for a similar indication. Researchers can match persons receiving the comparator medication to those receiving the medication being evaluated according to many factors that may have led to the decision to use each medication, or they can use alternative methods that work as effectively as matching. Control of confounding can address the problem that comes from recipients of a new drug (exposed group) being sicker and at higher risk of a certain adverse event before being treated than those who receive older therapies. The adverse event of interest may be a particular disease that is chronic in nature or has long latency (e.g., rheumatoid arthritis or cancer), or it may be an acute event with sudden onset (e.g., myocardial infarction).

If the adverse event of interest is rare, a cohort study requires a very large population for meaningful analysis. In addition, some events of interest require a long period of time for follow-up because they may be associated with a long induction time between exposure and diagnosis of the adverse event. The cost and logistics for completing such a study could be prohibitive if subjects were recruited from the population at large. For this reason, researchers often use existing large electronic databases that are collected for administrative purposes, such as Medicare or Medicaid, or for insurance billing claims.

For example, one cohort study conducted using Tennessee Medicaid data evaluated whether increased frequency of acute myocardial infarction (AMI) occurred in more than 181,000 people with a prescription for one or more nonaspirin, nonsteroidal anti-inflammatory (NSAID) drugs.[49] The frequency of AMI in this exposed group was compared with the frequency of AMI in a comparable number of people who did not have a prescription for an NSAID who were selected to match to the exposed group by age and sex. Ray and colleagues evaluated whether users of specific NSAIDs or specific COX-2 inhibitors experienced increased incidence of hospitalization for AMI or death from coronary heart disease.[50]

In these studies, the research team identified people who received a prescription of the relevant medication, estimated person-time of exposure from the number of prescriptions, and assessed outcomes occurring within each exposure category by identifying diagnoses associated with hospitalizations. For each study subject, they classified person-time into the exposed group following a prescription and into the unexposed group once the prescription was estimated to be finished.

In the study of COX-2 inhibitors,[50] study subjects with prescriptions for COX-2 inhibitors were compared with subjects who did not have a prescription for an NSAID or COX-2 inhibitor within 365 days of Medicaid enrollment. The investigators simultaneously identified and controlled for a cardiovascular risk score developed from relevant cardiovascular disease prescriptions and medical services, age, sex, ethnic origin, calendar year, and other factors.

Frequently, other factors of interest are measured at the time of study start (i.e., at the time when the first relevant prescription is dispensed). These other characteristics are taken into account during the analysis to separate the possible impact of the medication(s) being studied from the effects of other measured factors. However, some analysis techniques allow adjustment for the impact of factors that change throughout the time that patients are being followed when these changes are measurable.

To conduct such cohort studies within administrative databases, a unique set of each database should meet a number of research requirements, such as the following:

- Adequate numbers of patients being dispensed the drug of interest.
- Patients with adequate duration of history before receiving the drug of interest to identify other risk factors.
- Patients with adequate duration of history after receiving the drug of interest to identify the adverse event of interest.
- Collection of claims for services required to identify the adverse event of interest. For example, if researchers wish to measure the frequency of AMI, then implementation of the study is limited to databases that systematically receive and record information on hospital discharge diagnoses.
- Complete data on the exposure of interest. For example, data from plans with large deductible limits and no corresponding record of patient payment for prescription drugs may be missing key exposure information and would therefore not be useful for studying adverse reactions to prescription drugs. In addition, many plans do not collect data on in-hospital medications. Exposures to within-hospital or over-the-counter drugs are currently neither identified within useful research-oriented claims databases nor studie.
- Availability of an enrollment file to identify patients who are still in the plan but not receiving medical care. These data are not always essential but are very useful for some research questions requiring a denominator for deriving a population-based rate.

Many important research questions to evaluate signals with a cohort study can be addressed by administrative databases, but research topics regarding the safety of medicines are generally limited to the following:

- Exposure to prescription drugs (and not within-hospital drugs or over-the-counter drugs) that are used continuously following dispensing. Although for most patients dispensing is related to the start of medication use, for patients using medications on an as-needed basis, timing of use and of exposure may not be continuous throughout subsequent time periods, which may pose a challenge to researchers evaluating the safety of these medications. No information will be available on over-the-counter medications or on medications taken but not reimbursed.
- Adverse events that are consistently coded as hospital discharge diagnoses or can be identified as procedures that are billed.
- Commonly used medications.
- Research questions that require control for confounding only by information identified in claims. That is, conditions not coded for in claims (e.g., body mass index, a function of body weight and height) are not important confounders.
- Events that occur within 2 to 3 years of a prescription.

Information on use of over-the-counter medications or medications dispensed but not taken is not required.

Alternatively, EMRs can be extremely useful for addressing research questions that require information on laboratory test results or clinical data such as body height and weight. However, because such records in the United States frequently do not cover all medical facilities that patients visit, use of EMR databases to evaluate safety signals is usually not feasible in the current database environment. For example, if a rheumatologist prescribes a medication that results in an arrhythmia for which the patient is hospitalized, this event may not be recorded in the rheumatologist's record if no one identifies the causative agent or communicates to the rheumatologist that the patient was hospitalized. However, such information will be available in integrated health care systems that use EMRs throughout the provider network.

Case-Control Study

A case-control study is used to identify risk factors for the event of interest, including medication use. *Cases* are people who have experienced the event, and *controls* are selected from among those in the source population who also meet most case criteria but have not yet experienced the event. For very rare events or events that take a long time to develop, such as some

types of cancers, a case-control study can be a much more efficient and cost-effective option than a cohort study, as a cohort study would require keeping a very large population under surveillance until an adequate number of events occur for analysis.

In case-control studies, information on medication use in the past is obtained after cases and controls are selected for the study. If the frequency of use of the medication of interest is sufficiently higher in cases than controls after adjusting for the influence of other factors that may be differently distributed among cases and controls, then use of the medication of interest is considered to be associated with the adverse event that was used for case definition.

Solomon and colleagues conducted a case-control study to evaluate whether people with prescriptions for NSAIDs are at increased risk of AMI.[51] In this example, they identified cases that had a hospitalization for AMI. Subsequently, among these cases, those with prescriptions for NSAIDs during the 6 months before the AMI were identified to assess exposures of interest. Other factors that were identified and controlled for in the analysis included history of hypertension, history of diabetes, number of prescribed medications, and number of comorbidities.

As with cohort studies, research questions conducted within claims data to evaluate signals with a case-control study are limited to the following:

- Exposure to prescription drugs (not hospital-dispensed drugs or over-the-counter drugs);
- Adverse events that are consistently coded as hospital discharge diagnoses or that can be identified as procedures that are billed;
- Commonly used medications. Most medications are used by a small proportion within the general population. Because case-control studies of adverse events associated with medication use usually require hundreds of cases when use of the medication of interest is between 4 percent and 10 percent, or even thousands of cases when medication use is less than 4 percent in the population, most case-control studies are limited to evaluating signals about commonly used medications; and
- Research questions that do not require information on variables not mentioned within claims, such as body weight and height.

One limitation of the case-control study is that it usually cannot be used to estimate the absolute rate of an event within the total number of patients taking the medication. However, this limitation is overcome by having access to the medication utilization data in the administrative database, which provides a denominator. In this situation, cohorts who can be used to generate rate estimates can be used to identify cases and controls for a nested case-control study.

Case-Crossover Study

The case-crossover study design is used to evaluate whether specific medications may cause adverse events of sudden onset that follow shortly after exposure starts. This approach involves selecting cases and using the cases themselves to provide comparison information. One compares the frequency of medication use preceding the event with the frequency of medication use at a different time (e.g., at 1 year before the event).

Case-crossover studies are useful for situations when measuring all important factors associated with the adverse event within the case sample is impractical. By making internal comparisons within cases (e.g., comparing exposure at one time with exposure at another time) the differences across individuals do not affect the comparison. The case-crossover study is a type of case-control study, and like most case-control studies, it does not generate estimates of the absolute frequency with which an event occurs, but such estimates can be derived from the full administrative database.

This approach provides a way to control for confounding by indication or when confounding involves factors that cannot be easily measured. For example, genetic risk factors for disease, which can vary between people and thus can confound comparisons that contrast different groups of people, are automatically controlled for in case-crossover studies, because comparisons are always within individuals. As with other study designs, if the appropriate data are available, these methods can be

quite valuable for evaluating medications as acute risk factors for some events. For example, Mittelman and colleagues used this approach to assess whether sildenafil treatment in men was a risk factor for triggering a myocardial infarction.[52]

Types of Data Sources

Nonexperimental studies can be performed by using data that already exist in medical records or other electronic sources or by prospectively and systematically collecting data from health care providers and/or patients who complete data collection forms. The remainder of this paper will focus on issues that relate to utilizing data within existing records. The database can be divided into two general categories: EMRs, which include records maintained for the management of patient's clinical care; and administrative databases, which include transactions performed primarily to achieve administrative purposes, such as making claims for reimbursement from insurance carriers.

Electronic Medical Records

Clinical records, whether kept on paper or in computer files, are maintained primarily for documenting the patient's condition and treatments. In most locations, health care providers have legal responsibilities for the completeness and accuracy of information they enter into a patient's record.

Strengths of the original patient medical record are its finely detailed nature, its longitudinal aspect (often having months or years of consecutive episodes of care), and the variety of data it contains. Physicians maintain records of all visits, diagnoses, prescriptions written, laboratory tests ordered and results, and clinical recommendations. Hospitals maintain a folder (paper or electronic) per patient with one or more new pages generated from each episode of care, including physicians' notes, nurses' notes, medication records, vital signs, operative reports, pathology reports, and results of various types of tests. A limitation of medical records is that certain information, such as smoking status, may not be reliably ascertained or recorded.

In the United States, one limitation of basing nonexperimental studies solely on patient charts is that the chart usually contains records for only one institution or provider. Documentation of treatment from other medical institutions is included only if reported by the patient or sent from other providers. For paper records, deciphering handwriting can be a challenge. Conducting large safety studies with cohorts of thousands of patients using paper charts is extremely labor-intensive, prohibitively expensive, and not generally feasible. Paper records can be useful, however, for verification of adverse events identified through another source, such as an administrative claims database.

The advent of EMR systems has remedied some of these problems. Large amounts of readable or coded clinical data are available on many patients and can be aggregated across many providers. However, unless records are available across an integrated health care system, including physicians, hospitals, and other clinical settings, the EMR may not form a comprehensive record of the patient's medical history.

Administrative Databases

Databases constructed from administrative and financial billing records are created when providers of health care submit claims for health care services to insurers. Limited information is required for billing, including the type of insurance coverage, dates of medical service, associated diagnoses, tests performed, and prescriptions dispensed by outpatient pharmacies. Insurance databases include data only on prescriptions and health care services covered by the insurance plan. Administrative data may be used for reimbursement and are also used for tracking needs for services.

Some insurance-based and administrative databases offer anonymized electronic records for research. In these records, linkage to other sources of data (e.g., death certificates or cancer registries) and linkage to the original patient medical charts are either not available or can be accomplished only through a trusted third party (owing to data privacy concerns).

Some examples of administrative databases and their characteristics (as of 2006) are provided in Table 4.

Table 4. Summary of selected data sources, 2006 (✔ = present; ◗ = partial)

Data Source	Country	Cumulative Number of People in Data Source	Available Data Elements						
			Rx	Outpatient Diagnoses	Hospital Diagnoses	Outpatient Procedures	Hospital Procedures	Outpatient Lab Tests	Medical Record Review
Administrative									
Commercial/Private									
Integrated Health Care Information Services[a]	US	> 25 million	✔	✔	✔	✔	✔	✔	
PharMetrics[b]	US	55 million	✔	✔	✔	✔	✔	✔	
i3DrugSafety and LabRx[c]	US	> 10 million	✔	✔	✔	✔	✔	✔	✔
MedStat[d]	US	> 10 million	✔	✔	✔	✔	✔		
Public									
Medicaid[e]	US	41 million	✔	✔	✔	✔	✔		
Saskatchewan Health[f]	Canada	1 million	✔	◗	✔	✔	✔		✔
Regié de l'Assurance Maladie du Québec[g]	Canada	> 2.3 million	✔	✔	✔	✔	✔		
Electronic Medical Records									
North America									
Kaiser Permanente[h]	US	8.2 million	✔	✔	✔	✔	✔	✔	✔
HMO Research Network[i]	US	> 10 million	✔	✔	✔	✔	✔	✔	✔
Europe									
General Practice Research Database[j]	England & Wales	> 6 million	✔	✔	◗	✔	◗	✔	✔
Medicines Monitoring Unit, Tayside[k]	Scotland	> 400,000	✔	◗	✔	◗	✔	✔	✔
IMS Disease Analyzer[l]	Germany, France, UK, Austria	10.5 million	✔	✔	✔	✔	✔	◗	
PHARMO Institute[m]	Netherlands	> 950,000	✔		✔				

[a] Large database of health care insurance claims; lab results on 2 million members.[53]

[b] Insurance claims on people in the US with health plan or commercial health insurance; approximately 1 million with lab result data.[54]

[c] Large database of commercial health care insurance claims with lab results on approximately 3 million people.[55,56]

[d] Large claims database including retired employees; can be linked with other data sets of Medicare, Medicaid, and managed care data; lab results to be added in 2006.[57]

[e] US national insurance program for low-income families.[58]

[f] Provincial administrative health insurance database for all of Saskatchewan.[59]

[g] Database of the Régie de l'Assurance Maladie du Québec with data on services covered by the Health Insurance Plan and Public Prescription Drug Insurance Plan.[60]

[h] Electronic medical record databases from large health maintenance organizations with inpatient and outpatient care, prescriptions, and lab results.[61]

[i] Consortium of 14 health plans in the US with health care databases; types of records and data elements vary among the plans.[62]

[j] Longitudinal patient medical records database from over 600 general practices in England and Wales, with more than 35 million patient-years of experience.[63]

[k] Database of linked health care records in Scotland including pharmacy, outpatient care, emergency care, inpatient care, and laboratory tests.[64]

[l] Longitudinal patient database with diagnoses, treatment, demographics, and costs from 2,500 physicians in Germany, Austria, France, and the United Kingdom (personal communication, IMS Health).

[m] Record linkage database from 25 cities in the Netherlands with hospital discharge records and pharmacy data.[65]

Discussion

Administrative databases and EMRs from integrated health care systems offer many advantages for implementing nonexperimental studies to evaluate the safety of medications in a real-world setting. These advantages include the following:

- Large number of enrollees, which frequently meets the large study size requirement for testing hypotheses about rare serious adverse events;
- Feasibility of constructing comparator group(s) to be comparable with the group using the drug of interest by using information available from all covered health care services;
- Relatively efficient electronic retrieval by diagnoses, drugs, or procedures, which usually allows quicker analyses and reporting than prospective data collection;
- Relative efficiency of grouping individuals based on diagnosis or medication codes, limiting the misclassification that can occur from use of free text written by health care providers in medical records;
- Inclusion of health care data from all providers covered by the associated plan(s); and
- Coverage of real-world health care in general practice and specialty clinics.

Further advantages associated with administrative databases include the following:

- Data on prescription drug dispensings, whereas clinical records may frequently be limited to prescribing information;
- Availability of enrollment data, which are frequently required to enumerate the population at risk for the condition(s) of interest; and
- Routine record quality and completeness audits conducted by payers for prevention of financial abuse.

Weaknesses of administrative databases and EMR from integrated health care systems include the following:

- Lack of specificity in some diagnostic codes, which other parts of the EMRs may not always clarify;
- Lack of patient-reported information, such as quality of life (so outcomes that require this type of information cannot be studied); and
- Potential restrictions imposed by health plans on prescribing of new drugs, which could lead to additional confounding if patients who receive new medications are different from those who receive older products.

Further weaknesses associated with administrative databases (but not with EMR databases) include the following:

- Lack of specificity in some diagnostic codes (Some codes cover multiple conditions, limiting differentiation among conditions within a code. For most diseases, key clinical details are not directly reflected in claims.);
- Lack of linkage to underlying medical records for review to validate electronic data;
- Lack of patient-reported information and/or detailed clinical information, such as smoking history and family history;
- Diagnoses that reflect reason for service, not necessarily a confirmed condition (therefore, researchers develop more complicated rules requiring multiple diagnoses to identify individuals with a specific diagnosis);
- Lack of data on laboratory tests performed and results (although some information is available through some claims databases);
- Lack of information on medications administered in an inpatient setting;
- To date, minimal representation of the unemployed or people older than 64 and loss to follow-up at age 65 years as enrollees convert to Medicare;
- Follow-up that is limited to 2 or 3 years because of changes in insurance as employment changes;

- Sources of error such as misdiagnosis or false ranking of primary diagnosis; incomplete record keeping; miscoding of diagnoses or procedures; failure to submit claims; transaction errors; lag time between service, filing, and adjudication; and incorrect record linkage across files[66];
- Inability to document over-the-counter use of products and use of professional samples; and
- Inability to identify deaths occurring outside of the medical setting without linking to other data sources.

Typical study designs to evaluate safety signals require study subjects to be available in the records for some time before the exposure of interest and after the exposure of interest to identify the occurrence of adverse events. Using databases that compile consistent and complete medical coverage over several years is essential. Long-term follow-up within such databases is extremely important because measurement of risk factors, such as concomitant conditions, usually requires evaluation of data covering periods before the exposure of interest. In Medicare data, loss to follow-up is reduced if individuals' data are linked as they switch from plan to plan. Examples of databases for which longer and more complete follow-up are available include Saskatchewan Health (public payer data), which has been assembling administrative data for more than 30 years, and the General Practice Research Database (containing EMRs from general practitioner practices in the UK), which has been assembling EMR data for almost 20 years.

Once a safety issue has been confirmed and risk factors identified, approaches to reducing or minimizing the risk in clinical practice can be developed and implemented. Multipurpose large health care databases are being used more and more to monitor for inappropriate use of selected medications. This is a growing area of drug safety surveillance, and new approaches are being developed.

Existing EMRs and administrative databases can be powerful tools for conducting signal evaluation studies in the right circumstances. The following list of questions may guide decisions about using one or more studies in these databases versus selecting a different approach:

Nonexperimental studies conducted within large existing databases can be advantageous when

- A comparison group is needed
- The outcome of interest is routinely ascertained and correctly diagnosed in an administrative database
- The number of exposures can be accrued within health plans within the needed timeline
- All variables of interest can be identified in administrative data or electronic medical record data
- If, in administrative data, there are limited restrictions on payment through relevant health plans.

- Is the adverse event of interest routinely ascertained and consistently diagnosed in real-world clinical practice? If yes, a database study may be useful.
- Are the background rate of and risk factors for the event of interest well understood? If no, then replicating studies with several different data sources and/or approaches is preferable.
- Is the event usually caused by medications? If yes, then design may be relatively simple because no background rate is relevant other than that related to specific medications.
- Is the use of the medication generally sporadic or chronic? If sporadic, use of databases may result in substantial exposure misclassification and may reduce likelihood of detecting a true association between medication and event of interest. If chronic, the databases may be more useful.
- For new drugs, is the exposure common enough so that an appropriate study size can be quickly accrued in an administrative database used for epidemiologic research? If not, an alternative study would be desirable if answers are needed quickly.
- Is the needed follow-up time short enough to be feasible? Generally, outcomes that occur after long latency periods may not be studied well because of lack of adequate enrollment in the system and corresponding follow-up time.

Table 5 provides a few examples of research questions relevant to a Medicare population, along with comments about strengths and limitations of these databases.

Table 5. Examples of potential research questions and the strengths and limitations of administrative databases to address them

Research Question	Advantages of Administrative Databases for the Question	Limitations of Administrative Databases for the Question	Conduct the study?
Are certain medications used for joint pain likely to increase the risk of heart attack compared with no pain medication?	• Data on medications of interest will be available. • Large numbers of patients with these medications will be available. • Events of interest can be identified because patients seek medical care for heart attacks, and there is good ICD-9 coding for heart attack.	• Will not include over-the-counter aspirin use, known to decrease risk of heart attack, or smoking status.	Yes, recognizing that there may still be some unmeasured confounding.
Are certain medications used for joint pain more likely than others to increase the risk of heart attack?	• Data on medications of interest will be available. • Large numbers of patients with these medications will be available. • Events of interest can be identified because patients seek medical care for heart attacks, and there is good ICD-9 coding for heart attack.	• Data may not be sufficient to determine differences between patients taking different drugs that might confound the relationship with heart attack. For example, if smokers are more likely to receive drug *a* than drug *b*, drug *a* will appear to be associated with heart attacks. • No information on OTC aspirin use or smoking status; these important risk factors cannot be measured or adjusted for.	Yes, with reservations.
Will medications used to protect against osteoporosis lead to an increased risk of cancer?	• Data on medications of interest will be available. • Future Medicare data will include large numbers of women over age 65 at risk of or with osteoporosis.	• Current administrative databases have limited numbers of individuals over age 65. • Follow-up duration in most administrative databases may not be long enough to account for latency between drug exposure and tumor diagnosis. • Cancer diagnoses in administrative databases are not sufficiently detailed to distinguish between primary tumor in a site (e.g., bone) and a metastasis from another primary tumor to the site.	No
Will a chemotherapeutic agent used to treat breast cancer lead to increases in other forms of cancer, such as lymphoma?	• Some Medicare data already contain specific chemotherapeutic agents. • Limited information on cancer diagnosis (e.g., stage at diagnosis) will be available.	• May not include chemotherapy administered in a physician's office. • Follow-up duration in current databases may not be long enough to account for latency between use of agent and development of new tumor. • Duration of follow-up could increase with new Medicare coverage.	Possibly, if • Duration is sufficient. • Tumor type is unambiguous.
Will the new medications available for asthma reduce the risk of bone loss compared with corticosteroids that are generally recommended?	• Data on medications of interest will be available. • Large numbers of patients with asthma medications will be available.	• Bone loss measures will not be available because - Measurement is not made routinely. - Results of bone mineral density testing are not captured in administrative data.	No
What treatments for type II diabetes are safest for patients who already have heart failure (e.g., NYHA class III or IV)?	• Data on medications of interest will be available. • Many measures of serious cardiovascular outcomes will be available.	• Data may not be sufficient to determine differences between patients taking different drugs that might confound the relationship between the different medications and cardiovascular outcomes. • Data on severity of heart failure are not available from records.	No

ICD-9 = International Classification of Diseases, Ninth Revision; NYHA = New York Heart Association; OTC = over the counter

No ready guide exists to determine exactly what circumstances do and do not lend themselves to signal evaluation studies in these types of databases. Several factors are important to make appropriate use of these data sources:

1. Researchers need to understand the strengths and limitations of the databases as they relate to specific study questions.
2. Such research requires a multidisciplinary research team, with knowledge of epidemiology methods, understanding of clinical practices in the disease area of interest, and understanding of the claims recording and coding practices, as well as expertise in database programming and statistical analysis.
3. Before undertaking any study, the investigators should write a protocol.

The International Society of Pharmacoepidemiology has developed guidelines on developing protocols for such studies.[67] These guidelines recommend that before conducting such studies, potential investigators should develop a protocol that describes the researchers' qualifications, the data and methods to be used, how subjects' privacy will be protected, and how results will be communicated to the public. These guidelines were developed to encourage pharmacoepidemiologists to use appropriate methods and safeguards and to make such research as transparent as possible.

Conclusions

Large electronic databases with relatively complete information on health care services (both EMRs from integrated health care services and administrative databases) on large populations of people have made evaluating the safety of medications in a real-world setting relatively efficient and more feasible than such studies in the past. The future availability of prescription drug data from Medicare could be an invaluable resource on patients ages 65 and older. In addition, the study of safety can be extended to the study of the impact of policy changes, such as risk management interventions to reduce medication risks, using these same databases. It will be tempting for policy makers and the public to assume that all important medication safety questions will be able to be addressed by these databases.

However, as we have attempted to describe, significant methodologic challenges arise in using these databases, and they cannot be viewed as the single resource for all research questions. Failure to understand these challenges can lead to selection of inappropriate study methods and hinder interpretation of results. Other research methods, including studies that collect information directly from clinicians and patients, as appropriate, will be important to address many of the emerging drug safety questions that cannot be answered with these tools. Nevertheless, administrative databases and EMR databases will be extremely valuable tools to help advance our understanding of medication safety in broader populations in the future.

References

1. Agency for Healthcare Research and Quality. Effective health care. 2005. [cited 2006 Apr 17]. Available from: http://effectivehealthcare.ahrq.gov/
2. WHO Collaborating Centre for International Drug Monitoring. The Uppsala Monitoring Center. The importance of pharmacovigilance [Internet]. p. 42. Geneva (Switzerland): World Health Organization, 2002 [cited 2008 Sept 13]. Available from: http://whqlibdoc.who.int/hq/2002/a75646.pdf
3. Arlett P, Moseley J, Seligman PJ. A view from regulatory agencies. In: Strom BL, editor. Pharmacoepidemiology, 4th ed. Chichester, West Sussex, England: John Wiley & Sons; 2005, p.116-7.
4. Ahmad SR, Goetsch RA, Marks NS. Spontaneous reporting in the United States. In: Strom BL, editor. Pharmacoepidemiology, 4th ed. Chichester, West Sussex, England: John Wiley & Sons; 2005, p. 135-59.

5. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Maintenance of the ICH guideline on clinical safety data management: data elements for transmission of individual case safety reports E2B(R2) [Internet]. 2001[cited 2008 Sept 13]. Available from: http://www.ich.org/LOB/media/MEDIA2217.pdf
6, Ahmad SR. Adverse drug event monitoring at the Food and Drug Administration. J Gen Intern Med 2003;18:57-60.
7. Chan KA, Hauben M. Signal detection in pharmacovigilance: empirical evaluation of data mining tools. Pharmacoepidemiol Drug Saf 2005;14:597-9.
8. Gould AL. Practical pharmacovigilance analysis strategies. Pharmacoepidemiol Drug Saf 2003;12(7):559-74.
9. Shore DL, Quade D. A surveillance system based on a short memory scheme. Stat Med 1989;8:311-22.
10. van der Heijden PGM, van Puijenbroek EP, van Buuren S, van der Hofstede JW. On the assessment of adverse drug reactions from spontaneous reporting systems: the influence of underreporting on odds ratios. Stat Med 2002;21:2027-4.
11. 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 Oct-Nov;10(6):483-6.
12. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. Am Stat 1999;53(3):177-90.
13. Bate A, Lindquist M, Edwards IR, Orre R, Lansner A, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998;54:315-21.
14. Hauben M, Patadia V, Gerrits C, et al. Data mining in pharmacovigilance: the need for a balanced perspective. Drug Saf 2005;28(10):835-42.
15. Heeley E, Lynda VW, Shakir SA. Automated signal generation in prescription-event monitoring. Drug Saf 2002;25(6):423-32.
16. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. Drug Saf 2003;26(3):159-86.
17. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol Drug Saf 2004;13:519-23.
18. Lindquist M, Stahl M, Bate A, Edwards IR, Meyboom RH. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO database. Drug Saf 2000;23 (6):533-42
19. Trontell A. Expecting the unexpected—drug safety, pharmacovigilance, and the prepared mind. N Engl J Med 2004;351:1385-7.
20. Edwards R, Faich G, Tilson H. Points to consider: the roles of surveillance and epidemiology in advancing drug safety. Pharmacoepidemiol Drug Saf 2005;14:665-7.
21. Niu MT, Rhodes P, Salive M, Lively T, Davis DM, Black S, et al. Comparative safety of two recombinant hepatitis B vaccines in children: data from the Vaccine Adverse Event Reporting System (VAERS) and Vaccine Safety Datalink (VSD). J Clin Epidemiol 1998; Jun;51(6):503-10.
22. Drug Safety Research Unit (DSRU) of the United Kingdom, University of Portsmouth. Press Release [cited 2006 Mar 9]. Available from: http://www.dsru.org
23. Bennett CL, Luminari S, Nissenson AR, Tallman MS, Klinge SA, McWilliams N, et al. Pure red-cell aplasia and epoetin therapy. New Engl J Med 2004;351:1403-8.
24. Pearce N, Beasley R, Crane, J, Burgess C, Jackson R. End of the New Zealand asthma mortality epidemic. Lancet 1995;345:41-4.
25. Suissa S, Ernst P. Optical illusions from visual data analysis: example of the New Zealand asthma mortality epidemic. J Clin Epidemiol 1997;50(10):1079-88.

26. Suissa S, Ernst P. Use of anti-inflammatory therapy and asthma mortality in Japan. Eur Respir J 2003;21(1):101-4.

27. Horbar JD, Wright EC, Onstad L. Decreasing mortality associated with the introduction of surfactant therapy: an nonexperimental study of neonates weighing 601 to 1300 grams at birth. The Members of the National Institute of Child Health and Human Development Neonatal Research Network. Pediatrics 1993 Aug;92(2):191-6.

28. DAWN: Interim National Estimates of Drug-related Emergency Department Visits. DAWN Series D-26. Rockville (MD): Department of Health and Human Services. 2004 Apr. DHHS Publication No. (SMA) 04-3972.

29. Novak S, Nemeth WC, Lawson KA. Trends of medicinal use and abuse of sustained-release opioid analgesics: a revisit. Pain Medicine 2001;5:59-65.

30. Larson AM, Polson J, Fontana RJ, Davern TJ, Lalani E, Hynan LS, et al; for Acute Liver Failure Study Group. Acetaminophen-induced acute liver failure: results of a United States multi-center prospective study. Hepatology 2005;6:1364-72.

31. Kaufman DW, L Rosenberg, AA Mitchell, Signal generation and clarification: use of case-control data. Pharmacoepidemiol Drug Saf 2001;10:187-203.

32. Zhou L, Harrison-Woolrych M, Coulter DM. Use of the New Zealand Intensive Medicines Monitoring Programme to study the levonorgestrel-releasing intrauterine device (Mirena). Pharmacoepidemiol Drug Saf 2003;12:371-7.

33. Tanaka K, Morita Y, Kawabe E, Kubota K. Drug use investigation (DUI) and prescription-event monitoring in Japan (J-PEM). Pharmacoepidemiol Drug Saf 2001;10:653-8.

34. Wilton LV, Stephens MDB, Mann RD. Visual field defect associated with vigabatrin: observational cohort study. BMJ 1999;319:1165-6.

35. Shakir SAW. PEM in the UK. In: Mann R, Andrews E, editors. Pharmacovigilance, 2nd ed. Chichester, West Sussex, England: John Wiley & Sons, Inc. 2007, p. 308-9.

36. Barrera BA, Wilton L, Harris S, Shakir SA. Prescription-event monitoring study on 13,164 patients prescribed risedronate in primary care in England. Osteoporos Int 2005 Dec;16(12):1989-98. Epub 2005 Aug 31.

37. Layton D, Pearce GL, Shakir SA. Safety profile of tolterodine as used in general practice in England: results of prescription-event monitoring. Drug Saf 2001;24(9):703-13.

38. Layton D, Heeley E, Shakir SA. Identification and evaluation of a possible signal of exacerbation of colitis during rofecoxib treatment, using prescription-event monitoring data. J Clin Pharm Ther 2004 Apr;29(2):171-81.

39. van den Eeden SK, Friedman, GD. Prescription drug screening for subsequent carcinogenicity. Pharmacoepidemiol Drug Saf 1995;4(5):275-87.

40. Centers for Medicare & Medicaid Services (CMS). [cited 2006 Apr 17] Available from: http://www.cms.hhs.gov/Reports/Reports/list.asp#TopOfPage

41. Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol 2006;59:437-47.

42. Normand ST. Propensity Scores for Causal Inference, ISPOR News 2001;7(4):12-13.

43. D'Agostino, RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17:2265-81.

44. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: John Wiley and Sons; 1989.

45. Lunceford J, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med 2004;23:2937-60.

46. Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. Am J Epidemiol 2006;163:262-70.

47. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41-55.

48. Foster E. Propensity score matching: an illustrative analysis of dose response. Medical Care 2003;41:1183-92.

49. Ray W, Stein C, Hall K, Daugherty JR, Griffin MR. Nonsteroidal anti-inflammatory drugs and risk of serious coronary heart disease an observational study. Lancet 2002a;359:118-23.

50. Ray W, Stein C, Daugherty J, Hall K, Arbogast PG, Griffin MR. COX-2 nonsteroidal anti-inflammatory drugs and risk of serious coronary heart disease, Lancet 2002b;360:1070-3.

51. Solomon D, Glynn R, Levin R, Avorn J. Nonsteroidal anti-inflammatory drug use and acute myocardial infarction. Arch Intern Med 2002;162:1099-104.

52. Mittleman MA, Maclure M, Glasser DB. Evaluation of acute risk for myocardial infarction in men treated with sildenafil citrate. Am J Cardiol 2005; Aug 1;96(3):443.

53. ingenix.com [homepage on the Internet]. Ingenix (formerly IHCIS) [cited 2006 Mar 15]. Available from: www.ingenix.com

54. PharMetrics database [Internet]. IMS Health Inc. [cited 2006 Mar 6]. Available from: www.pharmetrics.com

55. i3global.com/Businesses/i3DrugSafety/ [homepage on the Internet]. i3 Drug Safety [cited 2006 Mar 17]. Available from: www.i3drugsafety.com

56. i3global.com/Solutions/PharmaInformatics/i3Aperio/ [homepage on the Internet]. i3 Pharma Informatics [cited 2006 Mar 17]. Available from: www.i3aperio.com

57. Thomson MedStat database [Internet]. Thomson Reuters [2008 Sept 13]. Available from: http://research.thomsonhealthcare.com/Products/view/?id=71

58. Centers for Medicare & Medicaid Services (CMS) [homepage on the Internet]. Baltimore (MD): US Department of Health and Human Services [cited 2006 Mar 20]. Available from: www.cms.hhs.gov/home/medicaid.asp

59. Saskatchewan Health database [Internet]. Government of Saskatchewan [cited 2006 Mar 21]. Available from: http://www.health.gov.sk.ca/

60. La Régie de l'assurance maladie du Québec [Internet]. Government of Québec [cited 2006 Mar 20]. Available from: http://www.ramq.gouv.qc.ca

61. dor.kaiser.org [homepage on the Internet]. Oakland (CA): Kaiser Permanente Division of Research [cited 2006 Mar 6]. Available from: www.dor.kaiser.org

62. HMO Research Network [Internet]. Seattle (WA): Group Health Center for Health Studies[cited March 29, 2006]. Available from: http://hmoresearchnetwork.org/home.htm

63. General Practice Research Database [Internet]. London: General Practice Research Database Group, The Medicines and Health Care Products Regulatory Agency. [cited 2006 Mar 21]. Available from: www.gprd.com

64. The Medicines Monitoring Unit, Pharmacoepidemiology Research website [Internet]. Dundee (Scotland): University of Dundee [cited 2006 Mar 17]. Available from: http://www.dundee.ac.uk/memo/

65. PHARMO database [Internet]. Utrecht (Netherlands): The PHARMO Institute [cited 2006 Mar 6]. Available from: www.pharmo.nl

66. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323-37

67. International Society for Pharmacoepidemiology (ISPE). Guidelines for good pharmacoepidemiology practices (revision 2) [Internet]. Bethesda (MD): ISPE; 2007. Available from: http://www.pharmacoepi.org/resources/guidelines_08027.cfm

Acknowledgments

The development of this paper was funded by the Pharmaceutical Research and Manufacturers of America (PhRMA).

www.rti.org/rtipress **RTI Press publication MR-0007-0809**