Chapter 9
Pharmacoepidemiological Approaches in Health Care

Xiaojuan Li and Christine Y. Lu

Abstract Pharmacoepidemiology studies the utilization patterns of medicines—also known as drug utilization research—which is an important component of pharmacy practice research. Pharmacoepidemiology also studies the relationship between medicines or other medical treatments and outcomes in large populations under nonexperimental situations. Providing an introduction to pharmacoepidemiology, this chapter describes frequently used metrics to understand drug utilization and medication adherence. This chapter also covers the key concepts involved in studying the association between medical or surgical treatments and outcomes. These concepts include forming a research question, selecting sources of data, defining the study population, and defining drug exposures, covariates, and outcomes. The chapter also discusses a range of study designs used in pharmacoepidemiologic research, including, but not limited to, cohort studies, case-control studies, within-subject studies, cross-sectional studies, ecological studies, and quasi-experimental designs. Finally, the chapter draws on key challenges such as confounding bias as well as commonly used analytical techniques to overcome these challenges.

9.1 Pharmacoepidemiology and the Need for Pharmacoepidemiological Research

Pharmacologic treatments are a major component of modern medicine. Pharmacoepidemiology is a discipline that uses similar methods in epidemiologic studies to study pharmacologic treatments but focuses on the area of clinical pharmacology. The birth of pharmacoepidemiology may be dated to the early 1960s (Wettermark 2013). Initially pharmacoepidemiologic investigations focused on adverse drug reactions but in recent decades also include studies of the beneficial

X. Li · C. Y. Lu (✉)
Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA
e-mail: christine_lu@hphci.harvard.edu

Z. Babar (ed.), *Pharmacy Practice Research Methods*,
https://doi.org/10.1007/978-981-15-2993-1_9

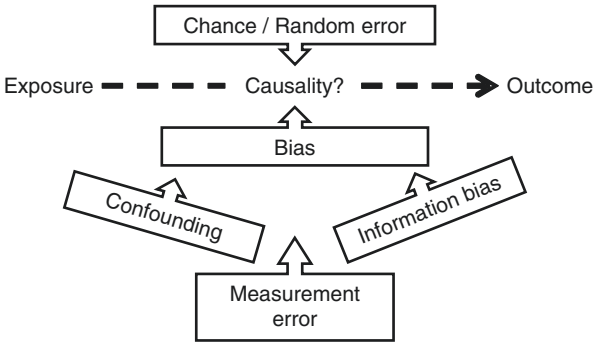

Fig. 9.1 Cause-effect relationship between an exposure and an outcome

effects of medicines. In general, pharmacoepidemiology studies the utilization patterns of medicines, and the relationship between medical treatment and outcomes (good and bad)—see Fig. 9.1—in large, often diverse populations under nonexperimental settings over time (Avorn 2004). The driving forces behind the development of pharmacoepidemiology are the increasing attention on the safety and effectiveness of medicines and the growing awareness that health outcomes of medicine use in the rigorous setting of randomized controlled trials (RCTs) are not necessarily the same as health outcomes of medicine use in real-world clinical practice.

Randomized trials are regarded as the gold standard for assessing the efficacy and safety of an intervention. Randomization is the most important feature of this study design in determining causality (see Fig. 9.1), which ensures that the groups formed are similar at time of randomization, except for chance difference, in all aspects. This method maximizes the internal validity by minimizing confounding biases at time of randomization. The internal validity of a study is the extent to which the observed difference in outcomes between the study groups can be attributed to the intervention rather than other factors. However, RCTs have several important limitations. They are resource intensive and focus on effects of an intervention among a small population of carefully selected patients, who are treated and followed up for a relatively short period of time under strictly controlled conditions. Trials typically have strict inclusion and exclusion criteria that underrepresent vulnerable patient groups (e.g., children, pregnant women, the elderly, individuals with multimorbidity). Because of these limitations, the external validity of RCTs is often limited. External validity, also known as generalizability or transportability, refers to whether the causal relationship holds beyond the individuals included in the study (e.g., other settings or populations). Because RCTs only provide results of average patients in a controlled environment, they do not provide a true reflection of how medication use will impact health outcomes in patients seen in the real-world setting. In addition, RCTs are not feasible to answer many questions of importance such as rare outcomes. Therefore, clinicians, patients, and policymakers must turn to pharmacoepidemiologic studies for best available evidence.

Pharmacoepidemiologic research has an important role in supporting the rational and cost-effective use of drugs in the real world, thereby improving health outcomes.

Specifically, pharmacoepidemiologic investigations can contribute in several ways (Avorn 2004; Lu 2009). We will discuss these main research questions that pharmacoepidemiologic research can help answer in the next section.

9.2 Major Research Questions in Pharmacoepidemiology

Pharmacoepidemiologic research can define medication needs by measuring the prevalence and burden of a particular clinical problem to identify the clinical place for the new therapeutic agent. Pharmacoepidemiologic research can assess utilization patterns of medicines (also referred to as drug utilization research) and issues such as medication adherence (sometimes noted as compliance). Importantly, pharmacoepidemiologic research can examine the safety and *effectiveness* of medicines in large, diverse populations; effectiveness describes how well a medication performs in the real-world setting, that is, when it is used by clinicians treating typical patients over a prolonged period of time and in comparison with other available therapeutic alternatives. Pharmacoepidemiologic research can be used for drug safety surveillance by quantifying the frequency and severity of adverse effects of a drug or drug class.

9.2.1 Drug Utilization Research

Drug utilization research is an essential part of pharmacoepidemiology and pharmacy practice as it describes the extent, nature, and determinants of drug exposure (Introduction to Drug Utilization Research 2019). Drug utilization research provides insights into the following aspects of drug prescribing and use. It can estimate the number of patients exposed to a drug or drug class within a given time period. We can estimate all drug users, regardless of when they started to use the drug (prevalence), or patients who started to use the drug within a given time period (incidence). Drug utilization research also describes the extent and profiles of medicines use at a certain time point and/or in a certain region (e.g., country, state, hospital) and trends and costs of medicines use over time. On the basis of epidemiologic data on a disease, drug utilization research also estimates the extent of appropriate use, overuse, or underuse of medicines. It describes the utilization pattern of a group of medicines and their relative market share for a certain disease. Examining utilization patterns by patient or prescriber characteristics (e.g., sociodemographic factors, provider specialty) can help identify the target population for educational interventions to improve medicines use. Drug utilization research also compares observed patterns of medicines use with clinical recommendations or guidelines for the treatment of a certain disease or local drug formularies. Such comparison can help generate hypotheses whether discrepancies represent less than optimal clinical practice, determine whether educational or other types of interventions are required,

or identify if the guidelines need to be reviewed in the light of actual practice. In addition, drug utilization research compares utilization patterns and costs of medicines between different regions and time periods. Such comparisons can generate hypotheses to further investigate reasons for and health implications of the differences found. Geographical variations and changes over time in medicines use may have medical, social, and/or economic implications both for the individual patient and for society and are thus important to identify, explain, and intervene, if necessary.

Drug utilization research often uses cross-sectional (see Sect. 9.4.3.5) or longitudinal study designs. Cross-sectional studies provide a snapshot of medicines use at a certain time (e.g., year 2019). Such studies may use similar data to compare medicines use between countries, different regions in a country, or different hospitals. Longitudinal data are often used to describe trends in medicines use (Vitry et al. 2011; Kelly et al. 2015; Chung et al. 2008; Lu et al. 2007a, b). Longitudinal data for drug utilization research can be obtained through administrative healthcare claims databases, based on a statistically valid sample of pharmacies or medical practices, or obtained from repeated cross-sectional surveys. Data collection in repeated cross-sectional surveys is continuous, but the patients or providers surveyed are continually changing. Thus, such data can reflect overall trends but cannot provide information about prescribing trends for individual practitioners or practices.

9.2.2 Drug Safety and Effectiveness Research

As indicated by its name, pharmacoepidemiology uses the study designs, methods, and techniques of epidemiology to study the uses and effects of medicines. In addition to characterize the use of medicines in drug utilization research, pharmacoepidemiology can study the effects of medicines in large numbers of people. One specific application is in the context of post-marketing drug surveillance, which has been broadened to include more areas in recent years, including effectiveness (Strom et al. 2012). As safety issues of medicines lead to major public concerns and both their effectiveness and safety affect evidence-based prescribing, studies discerning the effectiveness and safety of medicines have increasingly become a major emphasis.

Randomized trials are a great way to test the safety and efficacy of a new drug. The baseline randomization of the interventions, the careful collection and adjudication of the outcomes, and the execution of a rigorous, pre-specified protocol enable RCTs greater power to infer causal effects. However, RCTs are expensive, time- and resources-intensive, and sometimes unethical. RCTs also tend to have limited generalizability due to their strict eligibility criteria and other reasons mentioned in Sect. 9.1. In addition, some RCTs are powered for testing efficacy but are too small for studying adverse events (Evans 2012).

Pharmacoepidemiologic research, with large, diverse populations, can be used to examine the safety and effectiveness of medicines. In contrast to the well-"controlled" environment of RCTs, pharmacoepidemiologic studies can study the effects of medicines in the "real-world" setting where patients are treated in routine

clinical care. Due to the large number of individuals included in the studies, pharmacoepidemiologic studies can potentially detect the adverse effects, for which RCTs mainly targeted for efficacy are generally underpowered. In addition, pharmacoepidemiologic research can study long-term effects of treatments over a long period of time and in comparison with other available therapeutic alternatives, which would be too costly for RCTs.

While data from RCTs remain the cornerstone of regulatory decisions, there is growing interest in utilizing robust real-world evidence generated from high-quality pharmacoepidemiologic research with real-world data to support regulatory decision-making. Following the twenty-first Century Cures Act (Bonamici 2016), the US Food and Drug Administration (FDA) has determined real-world evidence one of the most important topics to be funded under the Prescription Drug User Fee Act VI (PDUFA VI 2019) and committed to facilitating the use of real-world evidence and considering its use in regulatory approval decisions.

9.2.3 Importance of a Well-Defined Research Question

In pharmacoepidemiology, a prior specification of the research question (and study population, study design, and data analysis plan) in the format of a study protocol is recommended to minimize the risk of "cherry-picking" interesting findings and a related issue of observing spurious findings because of multiple hypothesis testing (Austin et al. 2006). The rationale for the study should be explicitly stated, along with what a new study can add to existing knowledge. The research question should be concise and clearly articulate the exposure and outcome(s) of interest when the effects of medicines are of interest. The research question should be formulated considering the strengths and limitations of the available data.

9.3 Sources of Data in Pharmacoepidemiology

The research question should dictate the choice of data sources and whether the question can be appropriately addressed with a particular database. Knowing the relative strengths and limitations of the available data sources shall aim the selection of the appropriate data source for a particular research question.

9.3.1 Main Computer-Based Data Sources

Pharmacoepidemiology has grown rapidly as large-scale, computer-based databases have become increasingly available over the last two decades. There are three main types of large computer-based data sources frequently used for pharmacoepidemiologic research: administrative healthcare claims databases, electronic medical

records (EMR) databases, and patient registries. Administrative claims databases contain information about the delivery of services or a record of events, collected primarily for payment purposes. EMR data are recorded during the process of clinical care. While administrative claims and EMR databases are valuable resources, they are not designed for research (Motheral and Fairman 1997; Schneeweiss 2007). In contrast, patient registries, disease-based or drug-based, are established for the specific reporting of clinical information and management of certain diseases and procedures. A more comprehensive description of data sources used in pharmacoepidemiologic research, including the three main types, can be found elsewhere (Strom et al. 2012).

Administrative claims databases (Lu 2009) with millions of observations on the use of drugs, biologics, devices, and medical procedures along with health outcomes are valuable sources for drug safety and effectiveness studies (Gram et al. 2000). Rigorous longitudinal observational studies using large healthcare claims databases can complement results from RCTs by assessing treatment effectiveness in patients encountered in routine clinical practice. Comparisons of results from observational studies with RCTs have shown that these studies often produce similar results and that well-designed observational studies do not systematically overestimate the magnitude of treatment effects and do provide valid additional information (Benson and Hartz 2000; Concato et al. 2000). Furthermore, observational studies overcome the limitations found with current pharmacovigilance systems, many of which rely on voluntary reporting.

There has been an enormous growth in the use of large administrative healthcare claims databases for pharmacoepidemiology, including outcomes research, drug safety surveillance, and healthcare quality improvement programs. Table 9.1 lists a few examples of healthcare claims databases used in pharmacoepidemiology.

Table 9.1 Examples of large electronic healthcare databases

Country	Name	Website
United States	HMO Research Network	http://www.hmoresearchnetwork.org/
	Healthcare Cost and Utilization Project (HCUP)	http://www.hcup-us.ahrq.gov/databases.jsp
	SEER-Medicare Linked Database	http://appliedresearch.cancer.gov/seermedicare/
	Medicare and Medicaid Databases	https://www.resdac.org/
	Veterans Administration Databases	http://www.virec.research.va.gov/
Canada	Population Health Research Unit	http://metadata.phru.dal.ca/
	Population Data BC	https://www.popdata.bc.ca/researchers
United Kingdom	The Clinical Practice Research Datalink	http://www.cprd.com/intro.asp
The Netherlands	PHARMO Record Linkage System	http://www.pharmo.nl/
Australia	Medicare Benefits Scheme Data, Pharmaceutical Benefits Scheme Data	http://www.humanservices.gov.au/corporate/statistical-information-and-data/?utm_id=9

Administrative healthcare claims databases have several strengths (Lu 2009). There is a good level of compliance with reporting, and the accuracy of data submitted is usually high, because the data are collected for administrative purposes and often closely audited due to the importance of correct filling for reimbursement reason. These databases contain information on patient demography, some clinical diagnoses, use of medical services and drugs, and detailed information on charges. Data can be used to answer a variety of research questions at a low cost in a relatively short time span. In addition, routine healthcare data reflect drug effectiveness and safety in patients encountered in real-world practice. Moreover, large populations of patients can be followed over long time periods, making these databases a good source to identify clinically important, rare adverse events as compared with RCTs.

One concern about administrative healthcare claims databases is about the data incompleteness. The use of prescription medicines may not be captured in the claims in some situations; examples include when patients use medicines during hospital stay, use their partner's pharmacy benefit (Schneeweiss and Avorn 2005), use of free samples (Li et al. 2014), or pay out of pocket fully for prescription medicines (Choudhry and Shrank 2010). Therefore, caution must be exercised when determining the start date for drug exposure using the pharmacy-dispensing data from healthcare claims databases.

Electronic medical records (EMR) databases contain rich clinical information on patients that are often lacking in administrative databases (e.g., smoking status, body mass index, vital signs, laboratory data). EMR data can provide data for better confounding adjustment, particularly for studies that may be susceptible to confounding bias. However, while EMR data capture records of physician prescribing, they do not record all prescribed medications taken by patients and are generally not considered as a valid source for identifying drug exposure. Another major challenge is the variation in available data fields and data standards across EMR databases (Kush et al. 2008), which may limit data linkage and, subsequently, study sample sizes.

Patient registries are also valuable sources for tracking relevant clinical, economic, and humanistic (e.g., patient health-related quality of life, patient satisfaction) outcomes of therapeutic treatments, including medicines. Registries are prospective observational studies of patients with certain shared characteristics that collect ongoing and supporting data on well-defined outcomes of interest over time. Given patient registries are designed specifically for a purpose, they may not have data to answer a wide range of questions other than what has been pre-specified.

Merging administrative and EMR datasets or data from patient registries can provide the opportunity to leverage the strengths of each type of data. However, such practice must consider privacy issues, data quality and transferability, and feasibility of merging datasets. Data linkage is discussed in the next section. Ultimately, the choice of data sources depends on the research question and whether the question can be appropriately addressed with a particular database. It is important to note that databases do not have all the answers researchers seek in measuring drug exposure and outcomes. In selecting a data source, one must at least consider the

breadth and depth of the data in the database, quality of the database itself, the patient population that contributes data, and duration of information contained in the database.

For drug utilization research, household surveys are another data source to examine drug utilization and related issues such as adherence and access to medicines (Paniz et al. 2010; Bertoldi et al. 2008). Medicines available in households have been either prescribed or dispensed at health facilities or purchased at a pharmacy (with or without a prescription) or are over-the-counter medications. The medicines may be for the treatment of a current illness or leftover from a previous illness. Thus, dispensing data and utilization data are not necessarily equivalent because they have not been corrected for nonadherence, which is a common issue in real-world pharmacoepidemiologic studies. Drug utilization can be assessed by performing household surveys, counting leftover pills, or using special devices that allow electronic counting of the number of times a particular drug is administered.

9.3.2 Data Linkage

A pharmacoepidemiologic study may require data from more than one source either to enhance data available through linkage of disparate sources or to expand the size of the study population through combination of similar data sources. Person-level linkage of disparate databases can allow a more robust evaluation by providing a more complete picture of patient care and characteristics (Lu 2009). Such linkage can improve validity of a study (e.g., mitigating missing data, improving confounding control) or generalizability (e.g., increasing sample size).

Common linkages include the combination of inpatient, outpatient, and pharmacy data or linking cancer or death registries to medical records and may be within or across institutions. In the best scenario, each dataset will include several common relevant patient descriptors to allow a high-probability match (e.g., based on medical record number or other standardized person-level identifier, date of birth and residence); the more linkage variables are available the better. For common information across sources, rules for handling potentially duplicate information must also be specified (e.g., which record to be kept). In countries like the United States where no unified patient identifier is available, linking data from different sources typically require a probabilistic or deterministic linkage algorithm to account for ambiguity, for instance, slightly different spelling of names or addresses. The choice of linking method should be based on expertise in the approach used, previous linkage of the databases (if any), and the acceptable balance of false positives and false negatives, recognizing that some linkages will be incorrect and some will be missed. Furthermore, it is important to assess the overlap in populations because low linkage will affect sample size. Sensitivity analyses should be considered to evaluate potential linkage errors. Patient privacy is a concern when conducting linkages. Approaches have been developed for anonymous linkage (e.g., secure hashing algorithms) (Dusetzina et al. 2014), which are beyond

the scope of this chapter. In recent years, more electronic healthcare databases linking data from different sources are becoming available, offering a more complete picture of a patient's journey through time. The examples include SEER-Medicare Linked Database (National Cancer Institute Division of Cancer Control and Population Sciences 2019), linking data from the Surveillance, Epidemiology, and End Results (SEER) program and Medicare, and OptumLabs Data Warehouse (OptumLabs Health Care Collaboration and Innovation 2019) linking claims and EHR data for over 200 million individuals covered by a health plan.

Linkage of data sources containing similar information on different patients aims to expand the size of the study population (Brown et al. 2010). Many pharmacoepidemiologic studies require very large populations. Examples include research questions targeted on small population of interest (e.g., hypereosinophilic syndrome or chronic eosinophilic leukemia), uncommon exposures (e.g., safety surveillance of new treatments), and/or rare outcomes (e.g., rhabdomyolysis). Multiple sources, for instance, data from multiple health insurance plans, will be valuable and needed to identify an adequate size of study population when no single database is large enough to address such research questions in a timely and adequate way. Examples include FDA-funded Sentinel System (Sentinel Initiative 2019) and the Vaccine Safety Datalink Project (Vaccine Safety Datalink (VSD) 2019)

Assessment of comparability of data sources is needed before data linkage. Comparability of data sources refers to the way in which the data are captured and recorded so that the data can be reasonably combined with respect to data capture and terminology. Comparability should be assessed qualitatively through detailed understanding of the data source and quantitatively across all relevant variables to ensure that information from the different sources can be combined. For example, claims databases of different health insurers may be comparable, the data may be captured via a standardized reimbursement system, and the information is recorded using standardized coding schema. For multi-institutional studies through a distributed model (Brown et al. 2010, 2013), data partners maintain physical control of their data in adherence to their privacy and security rules instead of all data partners transferring data to a single site for analysis in a centralized model, thereby giving up control. Comprehensive analysis to characterize data should be conducted to evaluate variability across data partners with respect to overall cohort metrics (e.g., age and sex distribution) and study-specific metrics (e.g., exposure and outcome rates by age, sex, and year).

9.4 Study Designs and Methods in Pharmacoepidemiology

As mentioned in Sect. 9.2.3, having a well-pre-specified study question is of great importance in a pharmacoepidemiologic study. A detailed study protocol should explicitly state the study question, the exposure and outcomes of interest, the study population, the measurement of study variables, the study design, and the analytic plans. In this section, we will describe the specific considerations for each of these elements.

9.4.1 Selecting the Study Population

The selection/creation of study populations in pharmacoepidemiologic studies is critically important because confounding bias is a particular concern in nonexperimental research. For pharmacoepidemiologic studies interested in assessing effects of medicines, study cohorts typically include a study group of patients who have had the drug exposure and a comparison group of patients who have *not* had the same drug exposure (but may be exposed to a comparison drug). To increase comparability of the study groups, study cohorts should be restricted to patients who are homogeneous regarding their indication for the study drug exposure, which will lead to more balance of patient characteristics that predict the outcome (Perrio et al. 2007; Schneeweiss et al. 2007). This approach will reduce but not completely eliminate confounding because it is likely that some factors that influenced prescribing decisions may not be available in the data.

There are two major exclusion criteria to consider in pharmacoepidemiologic studies to maximize internal validity by reducing confounding. First, if the objective is to examine the incidence, rather than reoccurrence, of an outcome of interest, make sure to exclude patients with a history of the outcome of interest; these patients may be at an increased baseline risk for the outcome and at the same time may be more likely to take a study medication. It is often better to exclude these patients in the cohort creation/design stage instead of adjustment in the later analysis stage, particularly if the condition is a strong risk factor for future events (and thus a confounder). Second, studies may restrict to incident users of the study medications. Incident users are those starting on a study medication without prior dispensings of study drugs (i.e., no drug exposure) during a predefined time interval (also known as washout period—see Fig. 9.2). An often-used washout period is 6 months. However, this period might not be long enough for some patients who might have taken the drug 9 months ago. Thus, a longer washout period can increase the certainty that patients are truly incident users. Unfortunately, using a longer washout period reduces the number of patients eligible for the study, thus reducing precision of the effect estimates (i.e., study results). Prevalent users are individuals who have been taking a study medication for some time. Prevalent users are likely to be those

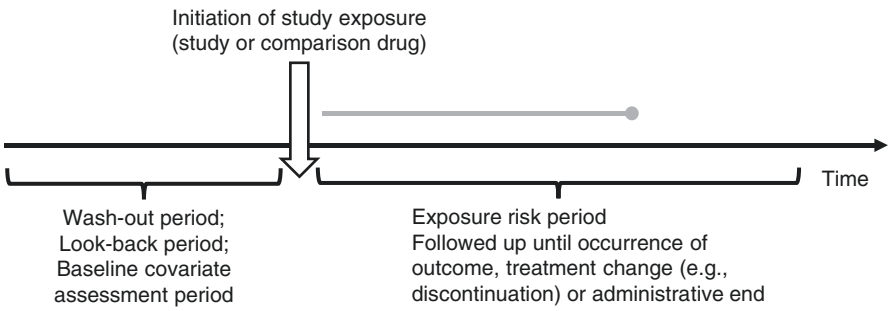

Fig. 9.2 Basic design of a pharmacoepidemiologic study

who tolerate the drug well, perceive some therapeutic benefits, and may lead to healthy user bias (Glynn et al. 2001). Restricting study cohorts to all patients in a defined population who start a course of treatment with the study medication ("new-user design") may reduce confounding (Johnson et al. 2013). The new-user design ensures the appropriate temporal ordering of baseline confounders, exposures, and outcomes, avoiding adjustment for intermediate variables that may be on the causal pathway between exposure and outcome. When used in combination with an active-comparator design (Schneeweiss et al. 2007) that compares the new users of the study drug to new users of a therapeutic alternative or comparator drug, the new-user design approach can help reduce the potential for immortal time bias (Suissa 2003) and also confounding by indication (Walker 1996) (see Sect. 9.6).

9.4.2 Defining Exposures and Outcomes

Drug exposures and outcomes in pharmacoepidemiologic studies must be operationally defined considering the formulated research question and the data source to be used. Because administrative claims data are recorded for billing purposes and not for research, both systematic and random errors can occur in the identification of exposure and outcome. Importantly, data are only captured for individuals who seek care and whose care is obtained through the insurance payment system. Claims for prescription drugs are generally considered a valid measure of drug exposure (Strom et al. 1991), although they may miss capturing some medication information, e.g., free drug samples, prescriptions paid in full by patients, etc. (see Sect. 9.3.1). Claims for medical procedures and services have been found to have a high level of specificity, but substantial variability in sensitivity exists across diagnoses when compared against the gold standard of medical records (Wilchesky et al. 2004)

Prescription claims data provide a wealth of information on drug exposure including dispensing date, pharmacy identifier, and drug information (generic and brand names, dose, duration in the format of days' supply). Drugs may be coded by established classification systems such as World Health Organization's Anatomical Therapeutic Chemical system. Using details like date of dispensing and days' supply, one can construct measures to assess medication adherence (discussed below). In comparison, while EMR data capture whether the physician prescribed medication for the patient, the dose, and intended regimen, they do not record whether the patient actually obtained the medication from the pharmacy. This nonadherence to initial treatment decision has been known as "primary nonadherence" or "primary noncompliance" and has been found to be substantial in real-world practices (Beardon et al. 1993; Fischer et al. 2010). This imperfect reflection of all dispensed medications taken by patients is a key limitation of the EMR data.

Medical claims data provide information on final end points such as fractures, stroke, myocardial infarction, or death but are limited for outcomes that involve intermediate biomarkers, self-reported symptom scales, or measures of patient functioning. Researchers may use a combination of diagnostic, procedures, and

facility codes to develop proxy measures of intermediate outcomes. For instance, a study that used diagnostic and inpatient hospital stays to classify severity of chronic obstructive pulmonary disease found moderate accuracy to medical charts (McKnight et al. 2005). Recent years have seen an increasing use of laboratory result data linked to administrative claims data, but these data are not available on a large scale across the globe.

To assess the occurrence of outcomes, study cohorts are typically observed (followed) for a certain period of time after the start of treatment—see Fig. 9.2. This is known as the exposure risk window (or period). The exposure risk window is the time period during which the medication puts individuals at risk for outcome(s) of interest. The choice of exposure risk period considers the duration of medicines use and the onset and persistence of drug toxicity. Typically, there is an extension after the drug is discontinued to account for the period when a drug is still biologically active in the body. The choice of exposure risk windows can influence the estimate of outcome risks. Risk windows should be carefully evaluated, or sensitivity analysis should be conducted on the varying length of exposure risk window.

9.4.3 Study Designs

Pharmacoepidemiologic research typically uses epidemiological study designs and methods. This section introduces a range of study designs often used in pharmacoepidemiologic studies; they are also summarized in Table 9.2. It is important to consider all potential study design options before choosing the most appropriate one for the study question of interest.

9.4.3.1 Cohort Studies

A cohort study typically follows a group of individuals in which some have had or continuing to have an exposure of interest in order to determine the occurrence of outcome(s). In pharmacoepidemiologic research, the exposure is typically a drug or a medical intervention. Usually a comparison group of individuals who have not been exposed to the same medication, unexposed or exposed to a comparator drug, is also included in the cohort study. The probability of developing the outcome in one group is compared with that in the other group; this is called the relative risk. Cohort design can be prospective or retrospective and has a number of applications, including the study of incidence, causes, and prognosis (Goldacre 2001; Gurwitz et al. 2005). In a prospective cohort study, individuals are enrolled into the study before none of them has developed outcomes of interest. In a retrospective cohort study, both the exposure and the outcome of interest have already occurred, but the investigators will go back in time and assemble a cohort at a point before the occurrence of outcome of interest. As a result, no matter whether a prospective or retrospective design is used, a cohort study enrolls individuals into the study based on their exposure status and measures subsequent outcome occurrence. In other words,

Table 9.2 Study designs for pharmacoepidemiology

- **Cohort studies** follow one group that is exposed to a drug or medical intervention and another group that is exposed to a comparison drug or unexposed to determine the occurrence of the outcome (estimating the relative risk). Cohort studies can examine multiple outcomes of a single exposure
- **Case-control studies** compare the proportion of cases with a specific exposure to the proportion of controls with the same exposure (estimating the odds ratio). Case-control studies can examine multiple factors that may be associated with the presence or absence of the outcome

Within-subject methods:
- *The self-controlled case series method* assesses the association between a transient exposure and an outcome by estimating the relative incidence of specified events in a defined time period after the exposure
- *Case-crossover design* estimates the odds of an outcome by comparing the probability of exposure between the at-risk and control periods
- *Case-time-control design* is case-crossover design with the addition of a traditional control group without occurrence of outcome

- **Cross-sectional studies** are used to determine prevalence, that is, the number of cases in a population at a certain time or time period and to examine the association between an exposure and an outcome
- **Ecological studies** focus on the comparison of groups. They can be used to identify associations by comparing aggregate data on risk factors and disease prevalence from different population groups

Quasi-experimental designs:
- *Interrupted time series design* involves a time series (repeated observations of a particular outcome collected before and after the implementation of an intervention to evaluation its effects). It can be conducted without or with a time series from a comparison group (interrupted time series with comparison series)
- *Pre-post with/without comparison group design* involves one measurement of a particular outcome before and another measurement after the implementation of an intervention to evaluate its effects. Intervention effect is estimated by a difference-in-differences approach when there are also pre-post measurements from a comparison group
- *Post-only with/without comparison group design* involves only measurements of a particular outcome after the implementation of an intervention to evaluate its effects

cohort studies measure exposure and outcome in temporal sequence, thereby avoiding the debate as to which comes first and making it possible to demonstrate causal relationships. Another advantage of the cohort design compared with the case-control approach, discussed in the next section, is that one can examine a wide range of possible outcomes from the same exposure in one cohort study. A cohort study is usually cheaper and easier than a RCT.

Cohort design is inefficient for studying the incidence of a latent or rare outcome (e.g., cancer) because individuals would need to be followed for a long time. The major challenges include (1) selection bias caused by potential systematic differences between the study groups in factors related to the outcome, (2) the inability to control for *all* extraneous factors (confounders) that might be associated with the outcome and might differ between the study groups, and (3) bias caused by differential loss to follow up due to migration, death, or dropouts (Gurwitz et al. 2005). Bias and confounding are discussed later in the chapter.

9.4.3.2 Case-Control Studies

In comparison to cohort studies, case-control studies enroll individuals into study based on their status of outcome and then ascertain their prior exposure status. Thus, case-control studies are usually retrospective. One group would include individuals who have the outcome of interest (i.e., cases), and they are matched with a control group who do not (i.e., controls or non-cases). Same information on prior exposure is collected from both groups (Breslow 1982). The key measure of association in case-control study report is the odds ratio, comparing the proportion of cases with a specific exposure to the proportion of controls with the same exposure, that determines the relative importance of the exposure with respect to the presence or absence of the outcome. Due to the lack of the denominators for the two exposure groups, case-control studies cannot directly report the incidence rates or incidence ratios of the outcomes. In cases of rare diseases, the odds ratio approximates the relative risk.

As some of the individuals have been deliberately chosen because they have the outcome, case-control studies are more cost-efficient than cohort studies—that is, a smaller sample size is sufficient to generate adequate information because of a higher percentage of cases per study. Further, a large number of variables can be examined at one time while the outcome being studied is limited (i.e., presence or absence of the outcome). Case-control studies are commonly used for initial, inexpensive evaluation of risk factors and are particularly useful when there is a long time period between an exposure and the occurrence of the outcome or when the outcome is rare. The main problems with the case-control design are confounding, selection bias, and recall bias because people with the outcome are more likely to remember certain antecedents or exaggerate or minimize what they consider to be risk factors.

9.4.3.3 Nested Case-Control Studies

A nested case-control study is comprised of individuals sampled from a well-defined cohort study. The case-control study is thus "nested" inside the cohort study (Etminan 2004). Analytical methods appropriate for case-control studies are applicable to nested case-control studies with computation of an odds ratio. The nested case-control design is flexible in that it allows examination of an exposure not planned in advance if records of a specific exposure of the cases and a subset of non-cases are available. This design also reduces selection bias because case and controls are sampled from the same source population. In some settings, a nested case-control design may involve less complex analysis compared to a standard cohort design because confounding is controlled for through matching and thus avoids sophisticated statistical techniques such as propensity scores (Etminan 2004).

Traditionally, case-control and nested case-control designs are favored due to its improved efficiency relative to the cohort design in that they reduce the costs and burden of data collection. In contemporary era of pharmacoepidemiology where electronic healthcare databases are the main data sources, all exposure, covariate, and outcome data for the entire cohort are already available, so the cost of data collection

for a single study approximates zero. Another feature that previously makes nested case-control studies attractive was its computational efficiency in the setting of time-dependent exposures (Essebag et al. 2005). Recent advances in computational sciences and technology make this advantage less relevant. Increasingly, researchers argue that these designs should no longer be used in secondary databases where data are already available (Schuemie et al. 2019) and suitable for a cohort design.

9.4.3.4 Within-Subject Methods (Case-Only Designs)

Cohort and case-control studies are useful for examining cumulative effects of chronic exposures. In situations where suitable comparison groups or controls are difficult to identify, within-subject methods that use self-controls offer a good alternative. The within-subject methods, also referred to as case-only designs, have the advantages that they don't require a separate comparison group and that all fixed confounders, unmeasured or unmeasured, are well controlled for (Petersen et al. 2016). These methods include self-controlled case series method, case-crossover design, and case-time-control design (Maclure et al. 2012).

In contrast to the case-crossover design discussed below, the self-controlled case series design derives from the cohort (fixed exposure, random event) rather than case-control (fixed event, random exposure) logic (Farrington 2004). The self-controlled case series method was originally published by Farrington et al (1995) to investigate the association between vaccination and acute potential adverse events and has also been used to examine effects of chronic exposures such as antidepressants (Hubbard et al. 2003). Using data on cases only, it is an alternative to cohort or case-control methods for assessing the association between a transient exposure and an outcome by estimating the relative incidence of specified events in a defined period after the exposure. This design retrieves the entire exposure history inside a given time window. Time within the observation period is classified as at-risk period or as control period in relation to the exposure. The key advantages are that the design controls for individual-level confounders (measured and unmeasured) that are stable over time and allows for changes in exposure with time (i.e., exposure trends) (Whitaker et al. 2006). Therefore, it provides valid inference about the incidence of events in at-risk periods relative to the control period and is suitable for studying recurrent outcomes.

Case-crossover studies can also eliminate within-person confounding that is stable over time because the exposure history of each case is used as his/her own control thus (Maclure 1991). They are useful for examining effects of transient exposures (e.g., use of benzodiazepine) on acute events (e.g., car accidents) and the time relationship of immediate effects to the exposure. It estimates the odds of an outcome by comparing the probability of exposure between the at-risk and control periods. However, the underlying probability of exposure must be constant (i.e., no exposure trends) so that the at-risk and control periods are comparable. Therefore, changes in prescribing over time or within-person confounding, including transient indication or changes in disease severity, may be problematic because

they can influence the probability of exposure, that is, the case-crossover design may have time trend bias (Schneeweiss et al. 1997).

Case-time-control design is an elaboration of the case-crossover design (Suissa 1995). This design uses data from a traditional control group (without occurrence of outcomes) to estimate and adjust for time trend bias and control-time selection bias (Schneeweiss et al. 1997). The trend-adjusted measure of association is obtained by dividing the observed odds ratio in cases by the observed odds ratio in controls.

9.4.3.5 Cross-Sectional Studies

Cross-sectional studies are primarily used to determine prevalence, that is, the number of cases in a population at a certain time or time period. This method is also used to examine the association between an exposure and an outcome, rather than establishing causation. The subjects are assessed at one point in time to determine whether they are exposed to a medication and whether they have the outcome. A difference between cross-sectional studies and cohort and case-control designs is that some of the individuals in the study sample will not have been exposed nor have the outcome of interest. The major advantage of cross-sectional studies is that they are generally quick to conduct and inexpensive because there is no follow up. However, this method cannot differentiate between cause and effect due to the inability to discern the sequence of events and is inefficient when the outcome is rare.

9.4.3.6 Ecological Studies

Ecological or correlational studies focus on the comparison of groups rather than individuals and are typically based on aggregate secondary data. The unit of analysis in an ecological study is an aggregate of individuals, and variables are often aggregate measures collected on this group. One can use ecological studies to identify associations by comparing aggregate data on risk factors and disease prevalence from different population groups. Because all data are aggregate at the group level, relationships between exposure and outcome at the individual level cannot be empirically determined. An error of reasoning—"ecological fallacy"—occurs when conclusions are drawn about individuals on the basis of group-level data, as relationships between variables observed for groups may not necessarily hold for individuals (Wilchesky et al. 2004). Ecological studies provide relatively cheap and efficient source for generating or testing the plausibility of hypotheses for further investigation by other study designs (e.g., case-control, cohort, or experimental studies) to test whether the observations made on populations as a whole can be confirmed in individuals. Despite these practical advantages, there are major methodological problems that limit causal inference, including ecologic and cross-level bias, problems of confounder control, within-group misclassification, temporal ambiguity, collinearity, and migration across groups (Morgenstern 1995). Therefore, ecological studies should only be conducted when individual-level data are unavailable.

9.4.3.7 Quasi-Experimental Study Designs

Similar to RCTs, quasi-experimental studies aim to estimate causal effect of an intervention on an outcome, but quasi-experimental studies do not use randomization. For such studies, interventions of interest are often educational interventions, quality improvement initiatives, and health policies, rather than drug exposure in typical pharmacoepidemiologic studies. The intervention often cannot be randomized; reasons include (1) ethical considerations, (2) infeasibility to randomize patients, (3) infeasibility to randomize locations, and (4) a need to intervene quickly.

An interrupted time series design is a strong quasi-experimental design that evaluates the longitudinal effects of interventions through regression modelling (Wagner et al. 2002). It consists of repeated measures of an outcome taken at regular intervals of time (e.g., monthly or quarterly) both before and after an intervention that occurs at a defined point in time. For example, studies may aim to assess the impact of a policy or regulatory actions on drug utilization and immediate outcomes (Lu et al. 2010, 2011, 2012, 2014; Adams et al. 2009). This method can control for most threats to internal validity (e.g., secular changes in prescribing, aging of the population) because it adjusts for baseline trends in study outcomes that are unrelated to the intervention. In an interrupted time series study, the post-intervention outcomes that might have occurred in the absence of the intervention are predicted based on patterns of historical data before the intervention of interest, so it is possible to get more valid and accurate measures of intervention effects. A challenge for interrupted time series design is the typical need for relatively large effect sizes.

In an interrupted time series study, it might be challenging to conclude the observed effect was not due to co-intervention or some other events occurring around the time of intervention of interest. One useful design to minimize such confounding is the interrupted time series with comparison series design that includes a comparison time series from another region or group of providers or patients.

Pre-post with non-randomized comparison group design is another commonly used quasi-experimental study design. This design examines a single measurement before and a single measurement after an intervention in the intervention group as well as in a comparison group. The inclusion of an observation before the intervention provides some information about what rates might have been had the intervention not occurred. In most cases, if the intervention achieves its expected impacts, the differences in effects observed between the groups should come from changes in the study group. It is therefore important to show that the intervention and comparison groups were similar on a variety of factors before the intervention takes place. Statistical methods (e.g., propensity scores) are sometimes used to adjust for differences in baseline characteristics between the groups. However, studies that depend on statistical adjustment alone without strong study designs provide less convincing results.

Quasi-experimental studies can also use “pre-post without comparison group” or “post-only” designs. Pre-post without comparison group designs examine a single measurement before and a single measurement after an intervention in a single

group. In contrast, post-only designs examine only measurements collected after an intervention has occurred. A pre-post study is a weak design; we cannot be confident that observed changes would have occurred anyway without the intervention due to previous trends or to external changes. A post-only study is also a weak design because of the lack of knowledge of previous levels and trends of the measured effect; thus, we cannot be certain that observed effects are due to the intervention and not to some other factors. Even if the study includes a comparison group ("post-only with comparison group"), there is no way to know whether observed effects in study and comparison groups would have been different anyway without the intervention.

9.5 Common Measures for Medication Use

This section introduces frequently used metrics to understand drug utilization and medication adherence, key study outcomes in pharmacoepidemiologic and pharmacy practice research.

9.5.1 Drug Utilization Metrics

The World Health Organization has recommended a number of quality indicators of medicines use (WHO 2018) that can be constructed from prescription or dispensing data. These include but are not limited to:

- Average number of drugs per prescription (per encounter or per patient).
- Percentage of drugs prescribed by generic name.
- Percentage of encounters with an antibiotic prescribed.
- Percentage of encounters with an injection prescribed.
- Percentage of drugs prescribed from essential drugs list or formulary.
- Proportion of treatment according to standard treatment guidelines.
- Average drug cost per encounter.

Data on drug costs are important for policy design and development to manage drug supply, pricing, and use. Costs may be determined at government, health facility, hospital, health insurance plan, or other levels within the health sector. Costs are often broken down according to drug group or therapeutic area to determine, for example, the reason for an increase in drug costs. For instance, the introduction of new, expensive oncology therapies may be found to be driving the increases in drug costs in a hospital. Changes in drug costs can result from changes in prescription volumes, quantity per prescription, or the average cost per prescription. Common cost metrics include total drug costs; cost per prescription; cost per treatment day, month, or year; cost as a proportion of total health costs; and cost as a proportion of average income (Introduction to Drug Utilization Research 2019).

A commonly used measure of drug utilization is defined daily doses (DDDs) per 1000 inhabitants per day, the standard unit recommended by the World Health Organization (WHOCC 2019). This measure allows comparisons of medication use independent of the country's population, the pack size, and dosage of the medication dispensed. The DDD is the assumed average maintenance dose per day for a drug used for its main indication in adults. Based on available information about doses (e.g., sales, prescription, or dispensing data), DDD/1000 inhabitants/day provides a crude estimate of the proportion of the study population that may be treated daily with certain medicines. For example, 10 DDDs/1000 inhabitants/day indicates that 1% of the population on average might get a certain drug or group of drugs every day. This estimate is most useful for chronically used drugs when there is good agreement between the average prescribed daily dose and the DDD. This method facilitates comparisons between drugs in the same therapeutic class and between different settings or geographic areas.

The DDD should be interpreted with caution. First, this metric is a technical unit of comparison and not a recommended dose and so does not reflect the actual prescribed dose. Second, the DDD describes the medication use in adults and needs to be adjusted first if pediatric use needs to be included. Finally, the DDD method does not consider variations in medication adherence.

9.5.2 Medication Adherence Metrics

Medication adherence generally refers to whether a patient takes a medication as prescribed, while persistence generally indicates how long a patient continues with the therapy regimen. The definitions and methods to determine adherence and persistence differ substantially in the published literature. Studies of medication adherence and persistence in large populations are important to understand factors related to low adherence (which will allow development of necessary interventions to improve adherence) and to assess clinical and economic outcomes related to low adherence and persistence. Medication adherence can be assessed by biochemical measures (e.g., levels of the drug or its metabolites in the blood or urine), patient interviews, medication diaries, pill counts, electronic drug monitors, and clinician assessments. However, these approaches are generally not practical to perform on large populations.

Administrative pharmacy claims databases are valuable sources for assessing medication adherence and persistence efficiently. One major limitation worth noting is that actual utilization is likely to differ from observed utilization, and based on utilization data only, we cannot determine if the patient actually consumed the dispensed medication. Here we discuss some common measures of medication adherence using the pharmacy claims data (Andrade et al. 2006).

Two most common methods are medication possession ratio (MPR), which estimates the proportion (or percentage) of days medication was supplied during a specified time period, and proportion of days covered (PDC), which estimates the number of days covered over a time interval. Other related measures of medication

availability include adherence ratio, refill adherence, compliance rate, continuous multiple-refill-interval measure of medication availability, adherence index, compliance ratio, or total number of days' supply dispensed during a specified time interval. The adherence measure is often dichotomized or categorized so that patients are considered adherent if a specified threshold was attained. A value of 80% or higher is generally considered adherent (Michael Ho et al. 2009).

In measurement of medication adherence, switching between drugs within a therapeutic class is defined as the dispensing of a different drug within the same class at some point during the study period (following the dispensing of the initial drug). Medication gap-related measures (e.g., continuous measure of medication gaps, cumulative gap ratio) are based on the number of days a patient is without medication. They can be determined for each refill interval using days' supply information in claims and the duration between refills. This allows calculation of proportion of days without medication during a specified time interval.

Metrics including discontinuation and continuation rates, often known as persistence, or the frequency of patients discontinuing/continuing medications are indicators of the acceptability of that medication. Discontinuation is generally defined by gaps between one dispensing of a drug and a subsequent dispensing, with continuous use based on the days' supply of medication dispensed or a specified time period after each dispensing (e.g., days' supply dispensed plus a grace period in days).

9.6 Challenges of Pharmacoepidemiologic Studies

It is critical to minimize the effects of chance, confounding, and other biases in pharmacoepidemiologic studies in order to provide results that are credible and convincing. Chance, confounding, and other biases are major threats to internal validity of a study and should always be considered as alternative explanations when interpreting the relationship between an exposure and the outcome. This section introduces major challenges in pharmacoepidemiologic research: misclassification, selection bias, and confounding, which are also summarized in Table 9.3.

Table 9.3 Major challenges in pharmacoepidemiologic studies

- **Selection bias**: Systematic error in creating comparison groups, such that they differ with respect to prognosis. That is, the groups differ in measured or unmeasured baseline characteristics because of the way participants were selected or assigned. This also used to mean that the participants are not representative of the population of all possible participants
- **Confounding**: A situation in which the seeming association or lack of association is due to another factor that determines the occurrence of the outcome of interest but that is also associated with the exposure, such as baseline characteristics, prognostic factors, or concomitant medications. For a factor to be a confounder, it must differ between the comparison groups and predict the outcome of interest
- **Information bias**: This occurs when systematic differences in the completeness or the accuracy of data lead to differential misclassification of individuals regarding exposures or outcomes

Definitions derived from the STROBE statement (Vandenbroucke et al. 2007)

9.6.1 Misclassification

A major challenge using claims data for defining exposure, covariates, and outcome is misclassification (information bias) (Vandenbroucke et al. 2007), that is, subjects may be classified as being exposed to a drug when they are not or as being unexposed when they are, similarly for the classification of covariate and outcome events. The likelihood of misclassification may differ between the exposure and nonexposed groups, often noted as differential misclassification. In general, the exposed group may have a lower likelihood of outcome misclassification because they have encounters with the healthcare system, which increases the likelihood of recording a diagnosis for the outcome event. In contrast, the nonexposed group is more likely to be misclassified as not having the outcome, which is an artifact of not entering the healthcare system.

With respect to drug exposure, the main data sources are prescribing or dispensing data. Research using these data sources need to be aware that drugs that are prescribed are not necessarily dispensed (primary nonadherence) and drugs that are dispensed are not necessarily taken (secondary nonadherence) (Beardon et al. 1993; Fischer et al. 2010), contributing to exposure misclassification. Misclassification can also occur when subjects receive their medications outside of the reimbursement system through multiple channels, including medication samples, patient assistance programs, paying out of pocket, taking medications belonging to someone else, secondary insurance coverage, and low-cost generic programs offered by retail pharmacies. Misclassification of drug exposure can impact outcome measurement because the risk of outcome is assessed during the time window when patients are considered "exposed" (exposure risk window—see Fig. 9.2). Misclassification of drug exposure can also affect the interpretation of the study results (Li et al. 2018).

With respect to outcomes, they are normally identified using a list of diagnostic or procedure codes. Misclassification of diagnostic or procedure codes can occur due to payment arrangements. For instance, clinicians are less incentivized to submit claims documenting care under capitated payment systems. Coding practices also vary under fee-for-service systems (e.g., upcoding—billings deliberately exaggerated to obtain higher payments, or undercoding—to avoid penalty). Ideally, researchers should consult clinicians who are familiar with the coding practice within the field under study or use definitions that have been validated against medical chart reviews in a similar setting. When several approaches are available to define the outcome, sensitivity analysis should be conducted to understand the implications of the various definitions on the results.

Correct classification of covariates is also essential for the validity of a research study. As patient characteristics and covariate status may vary with time, the assessment window of covariates is important. A common approach is to assess covariates in a fixed time window prior to start of exposure (i.e., fixed look-back period). Another approach is doing the assessment using all available historical data (Brunelli et al. 2013) and has shown to result in estimates with less bias but that requires more data assessment.

Misclassification of follow-up time in the exposure risk window can result in time-related bias, including immortal time bias (Suissa 2008), and produce delusive results. They can generally be avoided using appropriate study design and correct classification of follow-up time and exposure status in the analyses.

9.6.2 Selection Bias

Selection bias is a systematic error due to design and execution errors in sampling, selection, or classification methods that cause a distortion in the measure of association such that it does not accurately reflect the target population (Gurwitz et al. 2005). Selection bias will occur in cohort studies if the rates of enrollment into the study or the rates of loss to follow up differ by both the exposure and outcome status. Selection bias can also occur in case-control studies when the controls are not truly representative of the source population that produced the cases. One example is Berkson's bias (Berkson 1946), also known as hospital patient bias, that may occur when hospital controls are used in a case-control study.

In pharmacoepidemiologic studies, efforts should be made to avoid biased selection of study groups. Careful selection and clear identification of study population at the design stage are an important first step. Study groups need to be selected without knowing the outcome. Analytical methods including inverse probability of censoring weights can adjust for selection bias arising from the follow-up stage, such as differential loss to follow up (Robins and Finkelstein 2000).

9.6.3 Confounding

Confounding occurs when the study groups differ with respect to other factors that influence the outcome (Mamdani et al. 2005). For a variable to confound an association, it must be associated with both the exposure and the outcome, and its relation to the outcome should be independent of its association with the exposure. Confounding can cause over- or underestimation of the true exposure-outcome relationship and may even change the direction of the observed effect. Left unadjusted, results of nonexperimental studies may lead to invalid inference regarding the effects of the exposure.

Confounding by indication (also known as channeling bias) occurs when treatments are preferentially prescribed to groups of patients based on their underlying risk profile (Psaty et al. 1999). Patients with more severe disease are more likely to be treated (with higher doses) but also have higher risk of adverse outcomes. This confounding tends to make the study drug look worse when compared with nonexposed individuals. Confounding by indication is one of the most important, frequent problems encountered in pharmacoepidemiologic studies due to the natural presence of incomparability of prognosis between subjects receiving the drug and those who do not.

Confounding by frailty occurs when individuals close to death (or frailer) are more likely to receive certain drug classes or palliative treatments and are less likely to receive preventive treatments due to more focus on the main medical problem (Glynn et al. 2001; Redelmeier et al. 1998). This confounding tends to make the study drug look better when compared with nonexposed individuals.

The next section will introduce a few appropriate study designs and analytic methods that can help mitigate the potential confounding and selection bias.

9.7 Common Design Options and Analytical Techniques

Pharmacoepidemiologic studies typically use data that were originally collected for other purposes, so not all the relevant information may have been available for analysis, resulting in unknown and/or unmeasured potential confounders. This section discusses approaches that have been developed and adopted to improve the comparability between groups while limiting confounding and selection bias. Table 9.4 summarizes these strategies.

Table 9.4 Strategies to reduce confounding

Design phase
- *New-user design* restricts the study sample to individuals who are new users of a drug and follows them from the initiation of the treatment
- *Active comparator, new-user design* compares a cohort of new users for the study drug of interest to a cohort of new users of a therapeutic alternative or comparator drug, rather than a nonuser group
- *Restriction*: inclusion to the study is restricted to a certain category of a confounder (e.g., male)
- *Matching* of controls to cases (in case-control studies) to enhance equal representation of subjects with certain confounders among study groups

Analytical phase
- *Stratification*: the sample is divided into subgroups or strata on the basis of characteristics that are potentially confounding the analysis (e.g., age)
- *Statistical adjustment* estimates the association of each independent variable with the dependent variable (the outcome) after adjusting for the effects of other variables

Confounder summary scores

- **Propensity score**: the conditional probability of exposure to an intervention given a set of observed variables that may influence the likelihood of exposure

- **Disease risk score**: the conditional probability or hazard of having the study outcome conditional on their baseline characteristics
- Both scores can be used to control for confounding via matching, stratification, weighting (except for disease risk score), and regression

G-methods, including parametric g-formula, inverse probability weighting of marginal structural models, and g-estimation, have been developed to adjust for time-varying confounding affected by past treatment

Instrumental variables: a pseudo-randomization method that divides patients according to levels of a covariate that is associated with the exposure but not directly associated with the outcome unless through exposure

9.7.1 Study Design Options

The new-user design (Ray 2003), widely used in pharmacoepidemiology, restricts the study sample to individuals who are new users of a drug and follows them from the initiation of the treatment. This design avoids biases associated with prevalent users and adjusts for covariates at study entry that have been impacted by the drug already, also known as mediators.

The active comparator, new-user design is one option of new-user design that compares a cohort of new users for the study drug of interest to a cohort of new users of a therapeutic alternative or comparator drug, rather than an unexposed group. Coupled with an active-comparator design, the new-user design can help mitigate many of the biases discussed in the last section. This study design is regarded as the standard for comparative research in pharmacoepidemiology (Johnson et al. 2013).

There are additional methods to control for confounding in the design phase. First, restriction–inclusion to the study is restricted to a certain category of a confounder (e.g., male). However, strict inclusion criteria can limit generalizability of results to other segments of the population. In addition, in a case-control study, researchers can match controls to cases on certain confounders via frequency matching or one-to-one matching. However, the effect of the variable used for restriction or matching cannot be assessed and is a disadvantage of these approaches.

9.7.2 Analytic Options

In the analysis phase, stratifying the study sample into subgroups or strata on the basis of characteristics that are potentially confounders (e.g., age) can reduce confounding. The effects of the treatment are measured within each subgroup and can be summarized using the Mantel-Haenszel method (Mantel and Haenszel 1959). This approach may result in reduced power to detect effects because the number of participants in each stratum is smaller than the total study population. Subgroups may not be balanced with respect to other characteristics after stratification. It might not be appropriate to summarize the stratum-specific effects. Significant heterogeneity between stratum-specific effects suggests the presence of treatment effect modification, which is a characteristic of the effect under study rather than a source of bias that needs to be eliminated. In this case, stratum-specific estimates should be reported rather than a summarized estimate.

Statistical adjustment for dissimilarities in characteristics between study groups by including them in the regression model is a commonly used method to control for confounding (Normand et al. 2005). Regression analyses estimate the association of each independent variable (i.e., the treatment and certain characteristics of interest) with the dependent variable (the outcome) after adjusting for the effects of all the other variables.

Regardless of the approach used to control for confounding, the first important step is to capture and assess all potential confounders for the exposure-outcome relationship under study. A thorough literature review should be conducted to

identify variables that can influence treatment selection or the risk of outcome. Complete adjustment for confounding would require detailed information on these variables that sometimes include clinical parameters and lifestyle changes, which are not well captured in electronic healthcare databases. Residual confounding bias due to unmeasured confounders would occur. The impact of residual confounding should be systematically evaluated in sensitivity analyses or mitigated via external adjustment if such data are available (Schneeweiss 2006).

9.7.3 Confounder Summary Scores

The rich information contained in electronic healthcare databases enable the study to control for an extensive list of potential confounders, but its sheer volume can pose challenges for statistical analyses. To adjust for the large number of confounders, confounder summary scores—the propensity score and the disease risk score—can condense the information contained in individual confounders into a single variable.

Propensity score, proposed by Rubin and Rosenbaum (1984), is the conditional probability of having the drug exposure given patients' characteristics that may influence the likelihood of exposure. Disease risk score is the conditional probability or hazard of having the study outcome conditional on their baseline characteristics (Arbogast and Ray 2011). The propensity score can be estimated from a multivariable logistic regression model, while the disease risk score can be estimated using a logistic or Cox regression model. The most critical issue of the confounder summary score techniques is the appropriate selection of covariates to include in the model to generate the score. For propensity scores, all factors that are related to the treatment selection and/or outcome should be carefully considered for inclusion (Brookhart et al. 2006). Instrumental variables, discussed below, should be excluded from the propensity score model.

Both confounder summary scores can be incorporated in the analysis via matching, stratification, weighting (except for disease risk score), and regression. When correctly estimated, matching, stratifying, or weighting treated and comparison individuals on estimated scores tend to balance the observed characteristics across groups (McWilliams et al. 2007). However, balance between unmeasured variables cannot be assumed across groups when these scores are used for confounding control.

9.7.4 Instrumental Variable

In recent years, the instrumental variable method, a technique that originates from the field of econometrics, has been used more commonly in pharmacoepidemiologic studies to overcome the potential lack of balance on unobserved prognostic factors (e.g., health behavior) (Greenland 2018). In brief, this pseudo-randomization method divides patients according to levels of a covariate that is associated with the

exposure but not directly associated with the outcome unless through the exposure. The method may lead to equal distribution of characteristics in both exposed and nonexposed people and thus reduce potential confounding. For example, Brookhart et al (2006) used the prescribing physician's preference to cyclooxygenase-2 inhibitors or nonselective, nonsteroidal anti-inflammatory drugs as an instrumental variable to compare the risk of gastrointestinal complications associated with the use of these medicines. However, finding good instrumental variables has demonstrated to be remarkably difficult. Researchers should focus efforts on reducing the sources of bias (e.g., measurement error, omitted variables) instead of wishing for a "magic bullet" from instrumental variables.

9.7.5 Time-Varying Confounding

In real-world clinical practice, the treatment for a condition, in particular chronic conditions, often changes across time. To estimate the effect of the treatment, a study needs to appropriately control for time-varying confounding in the regression model. In situations where time-varying confounders are themselves affected by past treatment, standard regression methods for confounding control will be biased even when all relevant confounders are included and correctly specified in the regression model. An example is myocardial infarction in the estimation of effect of aspirin on the risk of cardiac death (Cook et al. 2002). Prior myocardial infarction affects subsequent aspirin use and the risk of subsequent cardiac death; it itself is also affected by previous aspirin use. Prior myocardial infarction is thus a time-dependent confounder between aspirin and cardiac death that is also affected by previous treatment.

Several approaches have been proposed to estimate effects of treatment in the presence of time-varying confounding affected by past treatment. Collectively referred to as "g-methods," these approaches include the parametric g-formula (Robins 1986), inverse probability weighting of marginal structural models (Robins et al. 2000; Hernán et al. 2000), and g-estimation. Few applications of these approaches exist in pharmacoepidemiology due to lack of sufficient information on time-varying confounders in administrative healthcare databases and limited availability of or familiarity with analytical tools to implement the relatively complex algorithms (Li et al. 2017). Fortunately, the increasing availability of data sources that contain more complete longitudinal information and better understanding of the g-methods begins to facilitate the use of these methods to estimate effects of complex time-varying treatment and treatment strategies.

9.8 The Future of Pharmacoepidemiology

We have been fortunate to live in an era where large amounts of data are available for research, including genetic information. Genetic information in the field of medical care includes a person's genetic predisposition to disease (e.g., results of specific

genetic tests), diagnosis of heritable medical conditions, or family history of disease with a known pattern of inheritance. Genetic testing may help identify DNA variants that predict an individual's response to a drug or course of therapy, resulting in identifying groups that may benefit most in terms of treatment effectiveness while avoiding adverse effects.

The public, patients, and consumers have a lot of concerns about confidentiality and the inappropriate use of the sensitive genetic information that may affect employment or health insurance rights. Higher privacy standards may be required than those for other medical information. To address these concerns, the International Declaration on Human Genetic Data was adopted in October 2003; this and the Universal Declaration on the Human Genome and Human Rights are the only international points of reference in the field of bioethics (International Declaration on Human Genetic Data: UNESCO 2019). Furthermore, in May 2007, member countries of Organization for Economic Cooperation and Development (OECD) adopted the Guidelines for Quality Assurance in Molecular Genetic Testing, which provides principles and best practice for the quality assurance of molecular genetic testing (OECD Guidelines for Quality Assurance in Genetic Testing—OECD 2019). Based on OECD Privacy Guidelines, the protection of patient privacy has generally been safeguarded by laws in some countries including OECD member countries. Pharmacoepidemiology has begun to see increasing research questions involving genetic information and will see more in the future. Researchers in the field should pay attention to legislations, policies, and guidelines for use of genetic information for research.

The availability of electronic healthcare databases and advances in pharmacoepidemiologic methods enable researchers to identify products in which effectiveness in the real world does not match efficacy shown in the trials. This will challenge the actions of all concerned—industry, regulators, payers, healthcare providers, and patients. In recent years, the European Medicines Agency and the US FDA have required risk management plans or risk evaluation and mitigation strategies as part of the drug approval process to help ensure that the benefits of a particular medicine outweigh its risks in the real-world setting. Observational studies are also increasingly requested by payers and other agencies to assess the value of medicines. Patients may also demand better systems to monitor effectiveness and safety of medicines. In fact, it is best practice to establish a systematic, comprehensive approach to monitor all marketed drugs postlaunch, and abundant electronic healthcare databases present a unique opportunity. Such monitoring may range from descriptive utilization statistics to sophisticated comparative effectiveness research, depending on the budget impact and level of uncertainty about the risk-benefit of the medicine at the time of marketing.

The data explosion in modern society will surely continue. As presented in this chapter, the nature of drug monitoring activities will be determined by the availability of data, advances in research methods and biostatistics, and competent pharmacoepidemiologists. Pharmacoepidemiology will also continue to be an area for collaboration between multiple stakeholders, including physicians, regulators, payers, manufacturers, patients, and the general public. Given the important contribution of pharmacoepidemiologic studies, collaboration should also involve

decision-makers for drug formularies, health economists, and health policy researchers. Pharmacoepidemiology will likely continue to be one of the most dynamic and challenging research areas for the coming decades.

Acknowledgement We thank Caitlin Lupton, M.Sc., of Harvard Pilgrim Health Care Institute for research assistance and administrative support.

References

Adams AS, Zhang F, LeCates RF, et al. Prior authorization for antidepressants in Medicaid: effects among disabled dual enrollees. Arch Intern Med. 2009;169(8):750–6. https://doi.org/10.1001/archinternmed.2009.39.

Andrade SE, Kahler KH, Frech F, Chan KA. Methods for evaluation of medication adherence and persistence using automated databases. Pharmacoepidemiol Drug Saf. 2006;15(8):565–74. https://doi.org/10.1002/pds.1230.

Arbogast PG, Ray WA. Performance of disease risk scores, propensity scores, and traditional multivariable outcome regression in the presence of multiple confounders. Am J Epidemiol. 2011;174(5):613–20. https://doi.org/10.1093/aje/kwr143.

Austin PC, Mamdani MM, Juurlink DN, Hux JE. Testing multiple statistical hypotheses resulted in spurious associations: a study of astrological signs and health. J Clin Epidemiol. 2006;59(9):964–9. https://doi.org/10.1016/j.jclinepi.2006.01.012.

Avorn J. The role of pharmacoepidemiology and pharmacoeconomics in promoting access and stimulating innovation. Pharmacoeconomics. 2004;22(2):81–6. https://doi.org/10.2165/00019053-200422002-00009.

Beardon PH, McGilchrist MM, McKendrick AD, McDevitt DG, MacDonald TM. Primary non-compliance with prescribed medication in primary care. BMJ. 1993;307(6908):846–8.

Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med. 2000;342:1878–86.

Berkson J. Limitations of the application of fourfold table analysis to hospital data. Biometrics. 1946;2(3):47–53.

Bertoldi AD, Barros AJ, Wagner A, Ross-Degnan D, Hallal PC. A descriptive review of the methodologies used in household surveys on medicine utilization. BMC Health Serv Res. 2008;8(1):222. https://doi.org/10.1186/1472-6963-8-222.

Bonamici S. Text—H.R.34—114th congress (2015–2016): 21st century cures act. https://www.congress.gov/bill/114th-congress/house-bill/34/text. Published 13 Dec 2016. Accessed 15 Oct 2019.

Breslow N. Design and analysis of case-control studies. Annu Rev Public Health. 1982;3(1):29–54. https://doi.org/10.1146/annurev.pu.03.050182.000333.

Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. Am J Epidemiol. 2006;163(12):1149–56. https://doi.org/10.1093/aje/kwj149.

Brookhart MA, Wang P, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiology. 2006;17(3):268–75. https://doi.org/10.1097/01.ede.0000193606.58671.c5.

Brown J, Holmes J, Shah K, Hall K, Lazarus R, Platt R. Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. Med Care. 2010;48(6):S45. https://doi.org/10.1097/MLR.0b013e3181d9919f.

Brown J, Kahn M, Toh S. Data quality assessment for comparative effectiveness research in distributed data networks. Med Care. 2013;51:S22. https://doi.org/10.1097/MLR.0b013e31829b1e2c.

Brunelli SM, Gagne JJ, Huybrechts KF, et al. Estimation using all available covariate information versus a fixed look-back window for dichotomous covariates. Pharmacoepidemiol Drug Saf. 2013;22(5):542–50. https://doi.org/10.1002/pds.3434.

Choudhry NK, Shrank WH. Four-dollar generics—increased accessibility, impaired quality assurance. N Engl J Med. 2010;363(20):1885–7. https://doi.org/10.1056/NEJMp1006189.

Chung Y, Lu CY, Graham GG, Mant A, Day RO. Utilization of allopurinol in the Australian community. Intern Med J. 2008;38(6a):388–95. https://doi.org/10.1111/j.1445-5994.2008.01641.x.

Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 2000;342(25):1887–92.

Cook NR, Cole SR, Hennekens CH. Use of a marginal structural model to determine the effect of aspirin on cardiovascular mortality in the physicians' health study. Am J Epidemiol. 2002;155(11):1045–53. https://doi.org/10.1093/aje/155.11.1045.

Dusetzina SB, Tyree S, Meyer A-M, Meyer A, Green L, Carpenter WR. Linking data for health services research: a framework and instructional guide. Rockville, MD: Agency for Healthcare Research and Quality (US); 2014. http://www.ncbi.nlm.nih.gov/books/NBK253313/. Accessed 15 Oct 2019.

Essebag V, Platt RW, Abrahamowicz M, Pilote L. Comparison of nested case-control and survival analysis methodologies for analysis of time-dependent exposure. BMC Med Res Methodol. 2005;5(1):5. https://doi.org/10.1186/1471-2288-5-5.

Etminan M. Pharmacoepidemiology II: The nested case-control study—a novel approach in pharmacoepidemiologic research. Pharmacotherapy. 2004;24(9):1105–9. https://doi.org/10.1592/phco.24.13.1105.38083.

Evans SJW. An agenda for UK clinical pharmacology: pharmacoepidemiology. Br J Clin Pharmacol. 2012;73(6):973–8. https://doi.org/10.1111/j.1365-2125.2012.04248.x.

Farrington CP. Re: "Risk analysis of aseptic meningitis after measles-mumps-rubella vaccination in Korean children by using a case-crossover design". Am J Epidemiol. 2004;159(7):717–8. https://doi.org/10.1093/aje/kwh093.

Farrington P, Pugh S, Colville A, et al. A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines. Lancet. 1995;345(8949):567–9.

Fischer MA, Stedman MR, Lii J, et al. Primary medication non-adherence: analysis of 195,930 electronic prescriptions. J Gen Intern Med. 2010;25(4):284–90. https://doi.org/10.1007/s11606-010-1253-9.

Glynn R, Knight E, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. Epidemiology. 2001;12(6):682–9.

Goldacre M. The role of cohort studies in medical research. Pharmacoepidemiol Drug Saf. 2001;10(1):5–11. https://doi.org/10.1002/pds.562.

Gram LF, Hallas J, Andersen M. Pharmacovigilance based on prescription databases. Pharmacol Toxicol. 2000;86(s1):13–5. https://doi.org/10.1034/j.1600-0773.2000.d01-4.x.

Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol. 2018;47(1):358. https://doi.org/10.1093/ije/dyx275.

Gurwitz JH, Sykora K, Mamdani M, et al. Reader's guide to critical appraisal of cohort studies: 1. Role and design. BMJ. 2005;330(7496):895–7.

Hernán M, Brumback B, Robins J. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology. 2000;11(5):561–70.

Ho PM, Bryson CL, Rumsfeld JS. Medication adherence: its importance in cardiovascular outcomes. Circulation. 2009;119(23):3028–35. https://doi.org/10.1161/CIRCULATIONAHA.108.768986.

Hubbard R, Farrington P, Smith C, Smeeth L, Tattersfield A. Exposure to tricyclic and selective serotonin reuptake inhibitor antidepressants and the risk of hip fracture. Am J Epidemiol. 2003;158(1):77–84. https://doi.org/10.1093/aje/kwg114.

International Declaration on Human Genetic Data: UNESCO. http://portal.unesco.org/en/ev.php-URL_ID=17720&URL_DO=DO_TOPIC&URL_SECTION=201.html. Accessed 15 Oct 2019.

Introduction to Drug Utilization Research. https://apps.who.int/medicinedocs/en/d/Js4876e/. Accessed 15 Oct 15 2019.

Johnson ES, Bartman BA, Briesacher BA, et al. The incident user design in comparative effectiveness research. Pharmacoepidemiol Drug Saf. 2013;22(1):1–6. https://doi.org/10.1002/pds.3334.

Kelly E, Lu CY, Albertini S, Vitry A. Longitudinal trends in utilization of endocrine therapies for breast cancer: an international comparison. J Clin Pharm Ther. 2015;40(1):76–82. https://doi.org/10.1111/jcpt.12227.

Kush RD, Helton E, Rockhold FW, Hardison CD. Electronic health records, medical research, and the tower of babel. N Engl J Med. 2008;358(16):1738–40. https://doi.org/10.1056/NEJMsb0800209.

Li X, Cole SR, Westreich D, Brookhart MA. Primary non-adherence and the new-user design. Pharmacoepidemiol Drug Saf. 2018;27(4):361–4. https://doi.org/10.1002/pds.4403.

Li X, Stürmer T, Brookhart MA. Evidence of sample use among new users of statins: implications for pharmacoepidemiology. Med Care. 2014;52(9):773–80. https://doi.org/10.1097/MLR.0000000000000174.

Li X, Young JG, Toh S. Estimating effects of dynamic treatment strategies in pharmacoepidemiologic studies with time-varying confounding: a primer. Curr Epidemiol Rep. 2017;4(4):288–97. https://doi.org/10.1007/s40471-017-0124-x.

Lu CY. Pharmacoepidemiologic research in Australia: challenges and opportunities for monitoring patients with rheumatic diseases. Clin Rheumatol. 2009;28(4):371–7. https://doi.org/10.1007/s10067-009-1102-6.

Lu CY, Law MR, Soumerai SB, et al. Impact of prior authorization on the use and costs of lipid-lowering medications among Michigan and Indiana dual enrollees in Medicaid and Medicare: results of a longitudinal, population-based study. Clin Ther. 2011;33(1):135–44. https://doi.org/10.1016/j.clinthera.2011.01.012.

Lu CY, Soumerai SB, Ross-Degnan D, Zhang F, Adams AS. Unintended impacts of a medicaid prior authorization policy on access to medications for bipolar illness. Med Care. 2010;48(1):4–9. https://doi.org/10.1097/MLR.0b013e3181bd4c10.

Lu CY, Srasuebkul P, Drew AK, Ward RL, Pearson S-A. Positive spillover effects of prescribing requirements: increased cardiac testing in patients treated with trastuzumab for HER2+ metastatic breast cancer. Intern Med J. 2012;42(11):1229–35. https://doi.org/10.1111/j.1445-5994.2011.02604.x.

Lu CY, Williams KM, Day RO. Has the use of disease-modifying anti-rheumatic drugs changed as a consequence of controlled access to high-cost biological agents through the Pharmaceutical Benefits Scheme? Intern Med J. 2007a;37(9):601–6. https://doi.org/10.1111/j.1445-5994.2007.01396.x.

Lu CY, Williams KM, Day RO. The funding and use of high-cost medicines in Australia: the example of anti-rheumatic biological medicines. Aust New Zealand Health Policy. 2007b;4(1):2. https://doi.org/10.1186/1743-8462-4-2.

Lu CY, Zhang F, Lakoma MD, et al. Changes in antidepressant use by young people and suicidal behavior after FDA warnings and media coverage: quasi-experimental study. BMJ. 2014;348:g3596. https://doi.org/10.1136/bmj.g3596.

Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. Am J Epidemiol. 1991;133(2):144–53. https://doi.org/10.1093/oxfordjournals.aje.a115853.

Maclure M, Fireman B, Nelson JC, et al. When should case-only designs be used for safety monitoring of medical products? Pharmacoepidemiol Drug Saf. 2012;21(S1):50–61. https://doi.org/10.1002/pds.2330.

Mamdani M, Sykora K, Li P, et al. Reader's guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. BMJ. 2005;330(7497):960–2. https://doi.org/10.1136/bmj.330.7497.960.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719–48. https://doi.org/10.1093/jnci/22.4.719.

McKnight J, Scott A, Menzies D, Bourbeau J, Blais L, Lemière C. A cohort study showed that health insurance databases were accurate to distinguish chronic obstructive pulmonary disease from asthma and classify disease severity. J Clin Epidemiol. 2005;58(2):206–8. https://doi.org/10.1016/j.jclinepi.2004.08.006.

McWilliams JM, Meara E, Zaslavsky AM, Ayanian JZ. Use of health services by previously uninsured Medicare beneficiaries. N Engl J Med. 2007;357(2):143–53. https://doi.org/10.1056/NEJMsa067712.

Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health. 1995;16(1):61–81. https://doi.org/10.1146/annurev.pu.16.050195.000425.

Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. Clin Ther. 1997;19(2):346–66. https://doi.org/10.1016/S0149-2918(97)80122-1.

National Cancer Institute Division of Cancer Control & Population Sciences. SEER-medicare linked database. https://healthcaredelivery.cancer.gov/seermedicare/. Published 3 Oct 2019. Accessed 15 Oct 2019.

Normand S-LT, Sykora K, Li P, Mamdani M, Rochon PA, Anderson GM. Readers guide to critical appraisal of cohort studies: 3. Analytical strategies to reduce confounding. BMJ. 2005;330(7498):1021–3. https://doi.org/10.1136/bmj.330.7498.1021.

OECD Guidelines for Quality Assurance in Genetic Testing—OECD. http://www.oecd.org/sti/emerging-tech/oecdguidelinesforqualityassuranceingenetictesting.htm. Accessed 16 Oct 2019.

OptumLabs Health Care Collaboration & Innovation. https://www.optumlabs.com/. Accessed 15 Oct 2019.

Paniz VMV, Fassa AG, Maia M de FS, Domingues MR, Bertoldi AD. Measuring access to medicines: a review of quantitative methods used in household surveys. BMC Health Serv Res. 2010;10:146. https://doi.org/10.1186/1472-6963-10-146.

PDUFA VI: Fiscal years 2018–2022. FDA. June 2019. http://www.fda.gov/industry/prescription-drug-user-fee-amendments/pdufa-vi-fiscal-years-2018-2022. Accessed 15 Oct 2019.

Perrio M, Waller PC, Shakir SAW. An analysis of the exclusion criteria used in observational pharmacoepidemiological studies. Pharmacoepidemiol Drug Saf. 2007;16(3):329–36. https://doi.org/10.1002/pds.1262.

Petersen I, Douglas I, Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. BMJ. 2016;354:i4515. https://doi.org/10.1136/bmj.i4515.

Psaty BM, Koepsell TD, Lin D, et al. Assessment and control for confounding by indication in observational studies. J Am Geriatr Soc. 1999;47(6):749–54. https://doi.org/10.1111/j.1532-5415.1999.tb01603.x.

Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol. 2003;158(9):915–20. https://doi.org/10.1093/aje/kwg231.

Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. N Engl J Med. 1998;338:1516–20.

Robins J. A new approach to causal inference in mortality studies with a sustained exposure period–application to control of the healthy worker survivor effect. Math Model. 1986;7(9–12):1395–512.

Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics. 2000;56(3):779–88. https://doi.org/10.1111/j.0006-341X.2000.00779.x.

Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. [Editorial]. Epidemiology. 2000;11(5):550–60.

Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc. 1984;79(387):516–24. https://doi.org/10.2307/2288398.

Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf. 2006;15(5):291–303. https://doi.org/10.1002/pds.1200.

Schneeweiss S. Developments in post-marketing comparative effectiveness research. Clin Pharmacol Ther. 2007;82(2):143–56. https://doi.org/10.1038/sj.clpt.6100249.

Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005;58(4):323–37. https://doi.org/10.1016/j.jclinepi.2004.10.012.

Schneeweiss S, Patrick AR, Stürmer T, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. Med Care. 2007;45(10 Suppl):S131–42. https://doi.org/10.1097/MLR.0b013e318070c08e.

Schneeweiss S, Stürmer T, Maclure M. Case-crossover and case-time-control designs as alternatives in pharmacoepidemiologic research. Pharmacoepidemiol Drug Saf. 1997;6(Suppl 3):S51–9. https://doi.org/10.1002/(SICI)1099-1557(199710)6:3+<S51::AID-PDS301>3.0.CO;2-S.

Schuemie MJ, Ryan PB, Man KKC, Wong ICK, Suchard MA, Hripcsak G. A plea to stop using the case-control design in retrospective database studies. Stat Med. 2019;38(22):4199–208. https://doi.org/10.1002/sim.8215.

Sentinel Initiative. https://www.sentinelinitiative.org/. Accessed 16 Oct 2019.

Strom BL, Carson JL, Halpern AC, et al. Using a claims database to investigate drug-induced Stevens-Johnson syndrome. Stat Med. 1991;10(4):565–76. https://doi.org/10.1002/sim.4780100408.

Strom BL, Kimmel SE, Hennessy S. Pharmacoepidemiology. 5th ed. Chichester: John Wiley & Sons Ltd; 2012.

Suissa S. The case-time-control design. Epidemiology. 1995;6(3):248–53. https://doi.org/10.1097/00001648-199505000-00010.

Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. Am J Respir Crit Care Med. 2003;168(1):49–53. https://doi.org/10.1164/rccm.200210-1231OC.

Suissa S. Immortal time bias in pharmacoepidemiology. Am J Epidemiol. 2008;167(4):492–9. https://doi.org/10.1093/aje/kwm324.

Vaccine Safety Datalink (VSD) | VSD | Monitoring | Ensuring Safety | Vaccine Safety | CDC. https://www.cdc.gov/vaccinesafety/ensuringsafety/monitoring/vsd/index.html. Published 17 June 2019. Accessed 16 Oct 2019.

Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. Ann Intern Med. 2007;147(8):W-163–94.

Vitry AI, Thai LP, Lu CY. Time and geographical variations in utilization of endocrine therapy for breast cancer in Australia. Intern Med J. 2011;41(2):162–6. https://doi.org/10.1111/j.1445-5994.2010.02304.x.

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther. 2002;27(4):299–309. https://doi.org/10.1046/j.1365-2710.2002.00430.x.

Walker AM. Confounding by indication. Epidemiology. 1996;7(4):335–6.

Wettermark B. The intriguing future of pharmacoepidemiology. Eur J Clin Pharmacol. 2013;69(1):43–51. https://doi.org/10.1007/s00228-013-1496-6.

Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Stat Med. 2006;25(10):1768–97. https://doi.org/10.1002/sim.2302.

WHO. How to investigate drug use in health facilities: selected drug use indicators—EDM Research Series No. 007. https://apps.who.int/medicinedocs/en/d/Js2289e/. Published 29 Oct 2018. Accessed 15 Oct 2019.

WHOCC. Definition and general considerations. https://www.whocc.no/ddd/definition_and_general_considera/. Accessed 15 October 2019.

Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. J Clin Epidemiol. 2004;57(2):131–41. https://doi.org/10.1016/S0895-4356(03)00246-4.