*British Journal of Pharmaceutical Research*
*3(2): 161-175, 2013*

**SCIENCEDOMAIN** *international*
www.sciencedomain.org

# Methodological Challenges in Observational Research: A Pharmacoepidemiological Perspective

**Ayad K. Ali**[1*]

[1]*College of Pharmacy, University of Florida, and Eli Lilly and Company, USA.*

***Review Article***

***Received 19th December 2012***
***Accepted 22nd February 2013***
***Published 4th March 2013***

## ABSTRACT

Pharmacoepidemiology is the science of studying the effects of utilizing pharmaceutical and biological products in the population, and in most cases this science is conducted in observational epidemiological designs, including retrospective database analysis. Observational pharmacoepidemiology is associated with a myriad of methodological challenges that affect study conclusions and related causal inferences. However, if these challenges are addressed and effectively dealt with, observational studies can have important and impactful clinical, regulatory, and public health outcomes. This article examines common challenges in retrospective database analysis and serves as an introductory text to important methodological concepts in research involving medication use, including confounding by indication, time-dependent confounding, informative censoring, depletion of susceptibles, and immortal time bias.

*Keywords: Pharmacoepidemiology; Observational Research; Retrospective Database Analysis; Confounding; Bias; Drug Safety.*

## DEFINITIONS, ACRONYMS, ABBREVIATIONS

$E_t$: = *exposure at time t, which is a medicinal product;*

$E_{t-1}$ = *previous exposure.*

$E_{t+1}$ = *subsequent exposure.*

$O$ = *outcome, which is either a beneficial effect or an adverse event of the drug.*

$Z_t$ = *confounder at time t.*

**Corresponding author: Email: ayadali@ufl.edu;*

$Z_{t-1}$ = *the value/status of the confounder at previous time.*
$Z_{t+1}$ = *the value/status of the confounder at future time.*
*Ě* = *observed measure of the exposure*
*Ǒ* = *observed measure of the outcome*
*Ž* = *observed measure of the confounder*
*!e* = *measurement error in exposure*
*!o* = *measurement error in outcome*
*!z* = *measurement errors in confounder*

## 1. INTRODUCTION

Pharmacoepidemiology is the science of studying the effects of pharmaceutical and biological product utilization in the population, and in most cases this science is conducted in observational epidemiological designs. In observational pharmacoepidemiology, exposure to drugs is not at random and probabilities of exposures are not the same across different populations; certain people have higher tendencies to be exposed to a specific treatment than others. The non-experimental nature of observational designs presents challenges in accurately and reliably estimating the association between exposures and outcomes. Three sources of bias can distort this association in research involving medication use: confounding, selection bias, and measurement bias [1]. However, if these challenges are addressed and effectively dealt with, observational studies can have important and impactful clinical, regulatory, and public health outcomes. This paper serves as an introductory text to these concepts and can be used as a teaching tool in the syllabi of pharmacoepidemiology, research methods, and evaluation of drug literature courses within the curricula of professional and graduate programs in colleges of pharmacy.

## 2. CONFOUNDING

In clinical practice, medications are not prescribed (and therefore, not utilized) at random; prescribing by physicians take into considerations multiple factors, such as patient, disease diagnostic, and disease prognostic characteristics. Such factors play the role of exposure determinants, which lead to confounding by indication, and time-dependent confounding [2]. Furthermore, failure to account for all confounding characteristics in an observational study leads to residual confounding [3].

Confounding is Latin in origin (*confundere*), which means "*to mix together*". Epidemiologically, confounding is defined as the distortion of exposure-outcome relationship by some other variable, i.e. confounder [2]. The distortion can lead to overestimation, underestimation, or reversal of exposure effect [2]. A confounder $Z_t$ is associated with the exposure $E_t$ and the outcome $O$ (i.e. a common cause), accounts for some or all of the observed exposure effect, and not in the exposure-outcome causal pathway (Fig. 1A). Here, association means that the risk of the outcome is different among people with the confounder compared to those without, and the distribution of the confounder is different among people with the exposure compared to those without (or those with other comparator exposure).

## 2.1 Confounding by Indication

Confounding by indication is defined as a distortion in the association between the exposure and the outcome, which occurs when a drug (or a class of drugs) is preferentially prescribed to patients with preexisting comorbidity at baseline (channeling bias) or to those with specific baseline characteristics, such as patients with worse disease state at baseline (confounding by disease severity) [4]. In case of the latter, the observed exposure effect could be mixed with an effect of severer disease state [5]. Both types are common in chronic disease pharmacoepidemiology [6], e.g. asthma, diabetes, and rheumatoid arthritis.

## 2.2 Time-Dependent Confounding

Time-dependent confounding is defined as an alteration in the association between the exposure and the outcome as a result of a variable that may vary over time and concurrently acts as a confounder between current exposure and outcome, and as an intermediary between previous exposure and current exposure (Fig. 1B) [7]. In another word, a time-dependent confounder is a variable that is predicted by previous exposure and a predictor of subsequent exposure. Conventional methods of confounding control inadequately accounts for time-dependent confounding because they cannot account for the confounder's effect between previous exposure and future exposure; on the other hand, methods like marginal structural models controls for that type of confounding under specific assumptions [8,9]; however, if previous exposure status does not affect present confounder's value (Fig. 1C), time-dependent confounding is not ensued and conventional regression modeling techniques can efficiently account for this type of confounding. In pharmacoepidemiology, drug effects are time-dependent, and are affected by time-dependent confounders that themselves are affected by previous drug exposure, and affect subsequent drug exposure and outcome. This phenomenon is common in usual-care "real-world" settings, and conclusions drawn from studies that fail to account for time-dependent confounding could be misleading.

## 2.3 Residual Confounding

Databases commonly used in pharmacoepidemiological research may contain an array of variables; however, even the most complete and detailed database fails to include all potentially important confounders. The presence of confounding despite adjustment is referred to as residual confounding. Failure to account for all confounders can rise from unmeasured variables or measurement errors in recorded variables [3,10]. Further, grouping a confounder that is on the continuous scale of measurement, e.g. age into few categories may result in inadequate confounding control in the observed exposure-outcome relationship [11], especially when some categories end with sparse number of observations. Unmeasured confounders are either measurable but unmeasured in the main study or immeasurable (unknown or difficult to measure) [2].

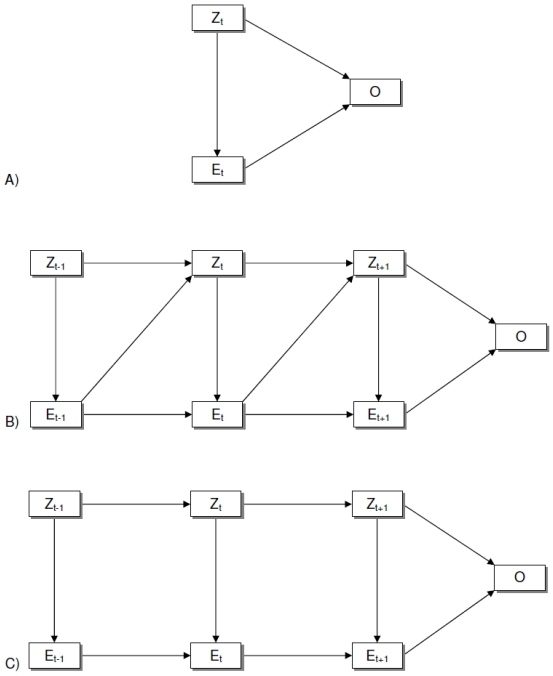


**Fig. 1. Illustration of confounding. A) classical time-independent confounding, B) time-dependent confounding and time-varying exposure, and C) time-varying exposure with no time-dependent confounding**

Examples of variables that are frequently unavailable (unmeasured) in databases that are used for pharmacoepidemiologic research include: behavioral information (smoking, alcohol drinking, nutrition and eating habits, exercise, substance abuse, and therapy adherence measures); laboratory information (blood tests, lung function tests, and other functionality tests); and exposure information (over-the-counter products, nutraceuticals and herbal remedies, and immunization history).

Observing exposure effect in a subset of patients defined by a particular characteristic is effect modification. In another word, the effect of the exposure on the outcome is different across levels of the effect modifying factor. Unlike confounding, effect modification is not a source of bias; rather it may be of public health importance, e.g. identifying exposure groups at risk. Both confounding and effect modification can exist together or separately. Stratified analysis is used to elucidate the association between exposure and outcome across levels of the factor. E.g. the association between obesity and asthma prevalence is different between men and women [12]. In some epidemiology references, the term "*effect-measure modification*" is used to signify the point that identifying the presence of effect modification is contingent upon the type of effect measure used, i.e. additive or multiplicative [13]. An effect modifier may change the direction or the magnitude of exposure effect. When the direction of exposure effect is the same in all subsets of the effect modifier but the magnitude is strengthened or weakened between subsets, quantitative effect modification occurs. When exposure effects are in opposite directions in subsets of the effect modifier (i.e. the exposure and the outcome are associated in the presence of the effect modifier but not associated in the absence of the effect modifier, and vice versa), qualitative effect modification occurs. In

case of qualitative effect modification, both additive and multiplicative effect modifications are present [14]. In the absence of qualitative effect modification, effect modification may be present in one scale of measurement than the other [13]. Therefore, the concept of effect modification in pharmacoepidemiology should always be examined in light of the chosen scale of effect measure, e.g. risk ratio (multiplicative) or risk difference (additive).

## 3. SELECTION BIAS

Selection bias is a distortion in the association between exposure and outcome that arises from systematic errors in the way the study sample was selected. In cohort studies, the outcome of interest can be over- or under-represented in the preferentially selected sample of patients who have higher or lower likelihood of presenting the outcome [4]; therefore, the sample will be unrepresentative of the target population to which study results are extrapolated (Fig. 2).

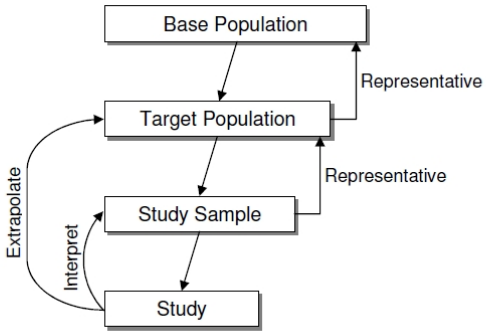


**Fig. 2. The relationship between study sample and target population in pharmacoepidemiologic studies**

There are many forms of selection bias, including: informative censoring; depletion of susceptibles (survivor bias); admission (referral) bias; diagnostic bias; volunteer (self-selection) bias; and healthy-user effect—the latter is considered a type of confounding more than of selection bias [15]. In retrospective database analyses, informative censoring and depletion of susceptibles are probably the most important types of selection bias.

### 3.1 Informative Censoring

In the context of survival analysis, censoring refers to the termination of observations when they are not followed long enough to observe the outcome of interest [4]; the conditions for such termination are usually defined by study investigator. When censoring occurs due to reasons that are not under the control of the investigator, random censoring is said to happen [16]. Within this framework, informative censoring is likely to occur (Fig. 3) [16]. If the censored individuals are biased subsample of the uncensored individuals who have similar covariate distribution (i.e. censored individuals have systematically higher or lower risks of observing the outcome than the uncensored counterparts), the censoring mechanism could be due to the exposure or the outcome, i.e. informative. Informative censoring can lead to biased estimates of the association between exposure and outcome, which is likely in longitudinal studies with time-dependent confounding [17].

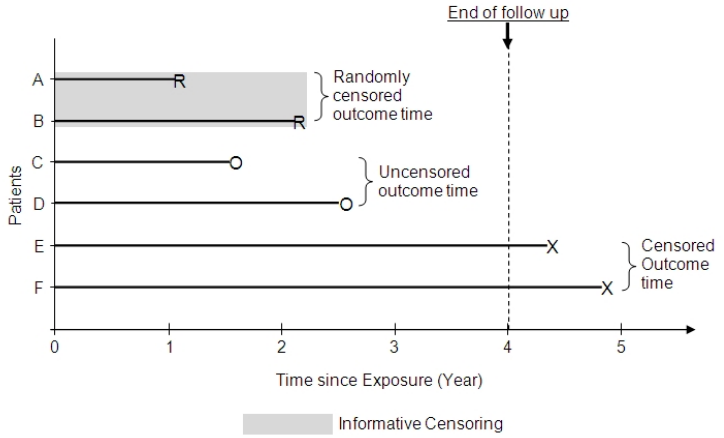


**Fig. 3. Informative censoring in pharmacoepidemiology**

*O = observed outcome of interest, i.e. uncensored;*
*R = randomly censored outcome other than the outcome of interest and was not defined by the investigator as a censoring criterion; X = censored outcome occurred after study termination.*

Selection bias due to random censoring can be prevented by including the variables that affect censoring and event times in the multivariate regression model, e.g. study starting time (index date) in case of random study entry and termination times [16]. In case of censoring due to competing outcomes, e.g. asthma hospitalization and asthma death, if the rate of competing outcome is higher in the exposed, the exposure effect, e.g. asthma death will be overestimated. Sensitivity analyses on the other hand can be employed to see how sensitive the estimates are to informative censoring [16]. For example, the investigator may assume that censored individuals are at high risk of observing the outcome (outcome occurs immediately after censoring), or the reverse (censored individuals have longer time-to-event than any other individual in the sample). When association estimates from sensitivity analyses bracket the estimates from original analysis (i.e. act as confidence limits), conclusions are not affected by treating censoring reason (e.g. other death when asthma-related death is the outcome of interest) as non-informative.

## 3.2 Depletion of Susceptibles

There are many terms for this type of selection bias, including “survivor bias” [3], “crossing of hazards” [18], and “survivor cohort effect” [19]. Depletion of susceptibles is the most commonly used term in pharmacoepidemiology. Depletion of susceptibles is defined as the gradual exclusion of a subgroup of patients who are susceptible to the outcome of interest from one exposure group, which leads to subsequent selection of another subgroup of patients who are less susceptible to the outcome from another exposure group (provided that exposure groups are comparable with regard to other factors other than exposure categories, e.g. confounders). This phenomenon is common in studies using prevalent users instead of incident users [20].

Fig. 4 illustrates the depletion of susceptibles concept in a retrospective cohort design assessing the association of a drug with an adverse drug reaction. At earlier time of the study (Period 1), patients at risk of experiencing the outcome are expected to develop the outcome and thereby excluded from the follow up. This will eventually leave a subgroup of patients who have low risk of experiencing the outcome at later times (Periods 2-3). If the

drug of interest causes the outcome of interest (i.e. effective in comparative effectiveness research—CER, or harmful in a drug safety study), susceptible patients will be differentially excluded from the drug exposure group (Drug A, gray shaded blocks) than the other comparator group (Drug B, not shaded area). The overtime depletion of susceptibles from Drug A group leads to the selection of susceptibles (who were less susceptible in prior periods) from Drug B group. Overtime (Periods 3-6), Drug A will appear inferior to Drug B (less effective in CER, or protective/safe in a drug safety study) when it is not in earlier periods. This phenomenon leads to the crossing of survival curves at a point of time due to the depletion of susceptibles and differential selection of less susceptibles overtime (hazard ratio>1 in Periods 1-3; hazard ratio<1 in Periods 4-6). Therefore, caution should be exercised in interpreting "time-varying period-specific hazard ratios" [18], where they are prone to selection bias due to this phenomenon.

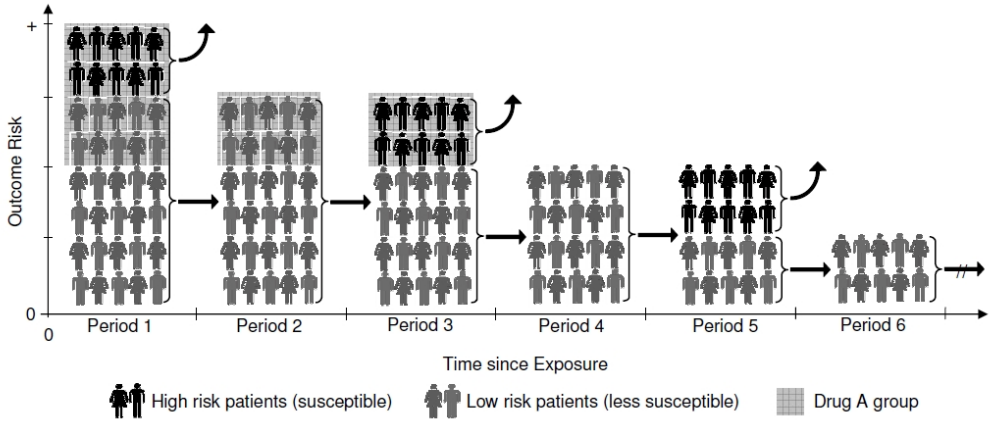


**Fig. 4. Depletion of susceptibles in a cohort study examining the association of a drug with an adverse reaction**

Furthermore, prevalent users invoke selection bias in two mechanisms: depletion of susceptibles, and adjusting for confounders that are affected by past exposures [21]. Prevention of selection bias induced by prevalent users is possible at the design stage of the study, where follow up is restricted to those patients who were naïve to the exposure and first time users, or those who were not exposed to the drug of interest for a period of time, then became exposed after the end of the specific period, when they will be considered exposure initiators after being unexposed [21]. Patient's past experience with a drug affects current and future utilization and risks of adverse events associated with current or future drug use. Prevalent users (those with past experience) may not have the same risk of an adverse event as incident users (first-time users); those who tolerate the drug continue using it, and those who do not (susceptible to the adverse event) stop using it. Table 1 depicts an example of attenuating the risk of gastrointestinal bleeding among 100 prevalent users of a non-steroidal anti-inflammatory drug (NSAID) in contrast to 100 incident users of NSAID.

**Table 1. Depletion of susceptibles among prevalent users of a NSAID**

| Drug | Incident users | | | | Prevalent users | | |
|---|---|---|---|---|---|---|---|
| | N | GI bleeding | IR | RR | GI bleeding | IR | RR |
| NSAID | 100 | 40 | 0.4 | 2.0 | 30 | 0.3 | 1.5 |
| COX II inhibitors | 100 | 20 | 0.2 | | 20 | 0.2 | |

*COX = cyclo-oxygenase; GI = gastrointestinal; IR = incidence rate; N = total sample size; RR = relative risk*

## 4. MEASUREMENT BIAS

Measurement (information) bias is a distortion in the association between exposure and outcome that arises from systematic errors in the way variables of interest are measured for the comparison groups [4]. In retrospective database analysis, information ("unobserved construct"; E, O, Z) regarding true exposures, outcomes, and other variables are recorded in form of indicators ("observed measures"; Ĕ, Ŏ, Ž) to mimic these true experiences [22]; thus for example, drug exposure ascertainment in a database is not the reflection of true drug utilization, rather a reflection of a measured drug utilization. When these observed measures do not accurately reflect unobserved constructs, measurement error—or misclassification (!e, !o, !z) occurs [23].

Measurement bias in pharmacoepidemiology is classified into four types (Fig. 5): independent nondifferential misclassification; independent differential misclassification; dependent nondifferential misclassification; and dependent differential misclassification [22]. Independent nondifferential misclassification denotes to the independence between the measurement errors in the exposure and that in the outcome; however, exposure misclassification is not affected by true outcome value (i.e. exposure misclassification is nondifferential across levels of the outcome), and likewise in outcome misclassification (Fig. 5A). Independent differential misclassification refers to the independence between the measurement errors in the exposure and that in the outcome; however, the true value of the outcome affects exposure classification (i.e. exposure misclassification is differential across levels of the outcome) (Fig. 5B), and similarly in outcome misclassification (Fig. 5C).

Dependent misclassification happens when measurement errors in the exposure and that in the outcome are dependent on each other, e.g. because they share same mechanism of abstracting information such as recall bias in pregnancy outcomes. Like independent misclassification, dependent misclassification can be nondifferential (Fig. 5D) and differential (exposure misclassification, Fig. 5E and outcome misclassification, Fig. 5F). Confounder misclassification is depicted in Fig. 6, which is similar to the framework of residual confounding. This type of misclassification may incline investigators to erroneously conclude the presence of effect modification by the confounder, when none actually exists [22].

Nondifferential misclassification drives estimates of association between exposure and outcome towards the null hypothesis; differential misclassification on the other hand drives estimates either towards or away from the null values [1]. In addition to measurement errors in abstracting the information, grouping categorical or continuous variables into fewer categories can revert nondifferential misclassification to differential misclassification [1]; and association estimates can be driven away from the null estimates even when the misclassification is nondifferential. This may occur in multi-category exposures and when exposure or outcome misclassification depends on misclassification in other variables in the dataset, e.g. confounders [1].

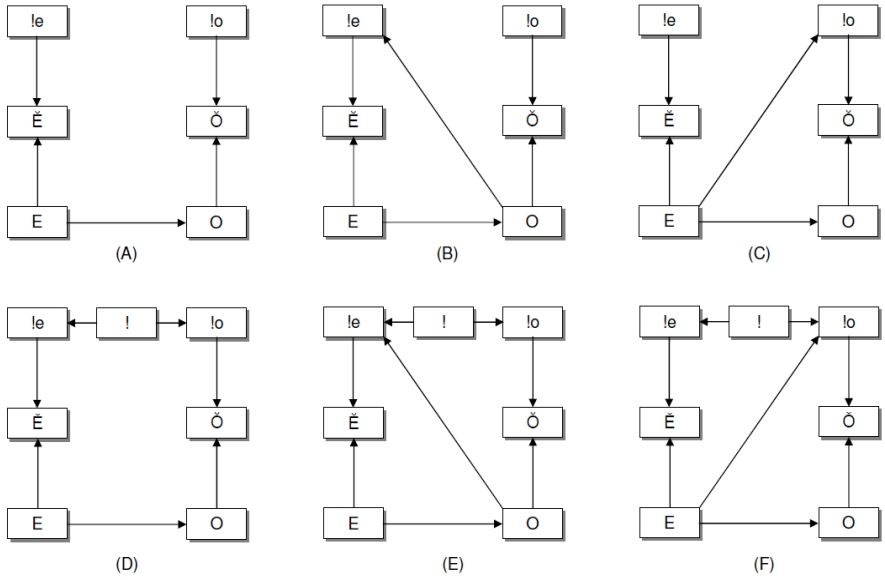


**Fig. 5. Measurement bias in pharmacoepidemiology. A) independent nondifferential, B) independent differential exposure, C) independent differential outcome, D) dependent nondifferential, E) dependent differential exposure, and F) dependent differential outcome misclassification**

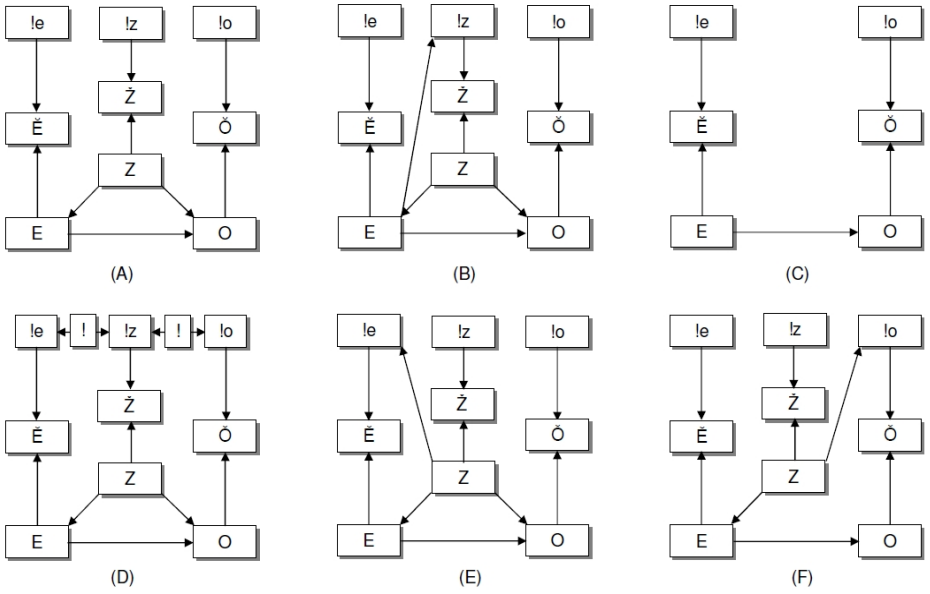


**Fig. 6. Confounder misclassification in pharmacoepidemiology. A) independent nondifferential, B) independent differential, C) independent differential depends on outcome, D) dependent nondifferential, E) exposure misclassification depends on confounder, and F) outcome misclassification depends on confounder**

## 4.1 Exposure Misclassification

Characterization of exposure in observational designs is a challenging task. There is a myriad of reasons for imposing measurement bias in exposure characterization when utilizing retrospective database analyses. In pharmacoepidemiology, measures of drug exposure are not necessarily accurate reflections of true drug exposure. In databases, prescribing a drug to a patient doesn't necessarily mean that the patient has actually received the medication from the pharmacy and has actually taken it with adherence to prescriber's instructions. Patients may take or receive too little (e.g. incorrect inhaler technique) or too much of the correct drug (e.g. too high dose for the indication, or interactions), or they may not take or receive the drug prescribed (e.g. product is not affordable to the patient or not reimbursed by the healthcare insurance system; product is inconvenient to use, e.g. inhaler coordination; instructions not understood or remembered or even agreed upon by the patient; experiencing undesirable effects; failure to experience perceived benefits; product unavailability; or health beliefs and cultural factors). Thus, such patients may be classified as exposed to the drug of interest, when they may be not exposed or exposed to a variable amount of the drug.

Medication adherence measures are developed to measure the extent of drug utilization in pharmacy administrative claims databases that record prescription filling and medication dispensing information. However, these measures have limited applications in prescribing databases, e.g. the Clinical Practice Research Datalink, where dispensing information is not recorded.

### 4.1.1 Immortal time bias

Immortal time bias is a form of exposure misclassification that is increasingly found in pharmacoepidemiologic studies [24]; it arises from cohort definitions in which patients must meet certain survival criterion of follow up from index date, in which patients should survive for a specific duration of time since exposure (Fig. 7) [25]. Failure to account for immortal person-time results in biased estimates of the association between exposure and outcome, which is an underestimation of the true association in the absence of this bias (If survival rate is higher in the exposed group, the hazard ratio will be reduced), and the magnitude of bias is proportional to the duration of immortal person-time [24,26].

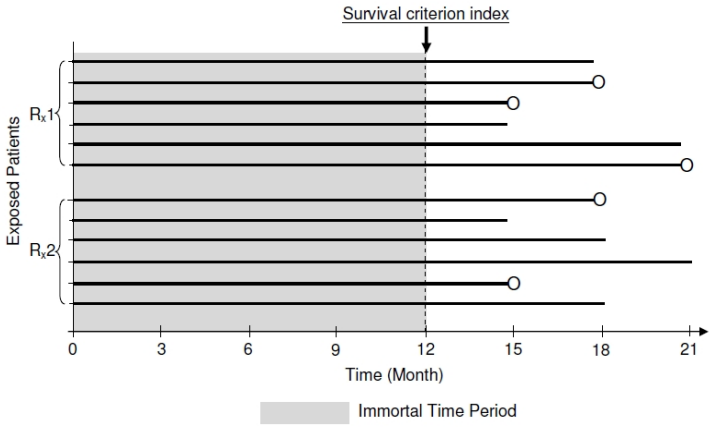


**Fig. 7. Immortal time bias in pharmacoepidemiology**
*O = outcome of interest; Rx1 = first drug group; Rx2 = second drug group.*

Immortal time bias can be prevented at the study design and analysis stages. At the design stage, patient follow up should start after the immortal person-time period; while at the analysis stage, this period can be excluded from the analysis of the denominator of person-time calculations for the risk estimates [24,25].

## 4.2 Diagnostic Misclassification

Diagnostic (disease) misclassification refers to the differential classification of the clinical condition among patients. Diagnostic misclassification is a form of measurement bias and should be distinguished from diagnostic bias, which is a type of selection bias in which patients are differentially diagnosed depending on exposure to specific risk factors [4]. Specifying the clinical condition of interest is usually the first step in pharmacoepidemiologic research design. The validity of disease diagnoses in pharmacoepidemiologic databases varies by the type and the source of the database. Some databases record diagnoses utilizing disease classification codes, e.g. International Classification of Diseases (ICD) and Read clinical terms; others use clinical terminologies, e.g. Medical Dictionary for Regulatory Activities (MedDRA) and Systematic Nomenclature of Medicine (SNOMED). Incorrect coding may arise at the general practice, hospital, health insurance system, or the database administrator. Patients may be misclassified as having the disease of interest, when in reality they are not and vice versa. In databases, the accuracy of the diagnosed condition is contingent upon the validity of the diagnosis process (at the practitioner level), and the extent of association of the classification code with the documented diagnosis (at the database level), where the code serves as a surrogate measure for the true diagnosis [27].

## 4.3 Outcome Misclassification

In retrospective database research, patient outcomes of interests are measured by a set of definitions that serve as a proxy to the true outcomes, e.g. asthma morbidity is sometimes measured by patient's referral to emergency departments or hospitals due to asthma exacerbations, or prescribing and utilization rates of rescue inhaler medications and oral corticosteroids [28]. Like exposures, outcomes are prone to misclassification in pharmacoepidemiologic research. When outcome misclassification is nondifferential (Figs. 5A & 5D), the association will be driven towards the null; and it will be driven away from the null if the misclassification is differential (Figs. 5C & 5F). For example, if exposure to a specific asthma drug increases (or decreases) the probability of misclassifying the patient as having hospitalized for asthma exacerbations, the association between the exposure and the outcome will be biased away from the null estimate.

Developing algorithms to identify outcomes that include a combination of clinical, referral, and prescription information can improve the validity of outcome measurement [26,29]. For example, if the outcome of interest is asthma exacerbations, the investigator can identify diagnosis codes for asthma exacerbations in addition to referral information to an emergency department and relevant medication history. Likewise, utilizing linked databases can help in the identification of relevant outcomes.

## 5. CONCLUSION

In randomized, controlled clinical trials, randomization equally (or near equally) balances confounders across exposure groups. Pharmacoepidemiologists aim at evaluating the causality of associations between drugs and outcomes; lack of randomization in

observational studies casts additional challenges in evaluating causal associations. Moreover, depending on the therapeutic effectiveness of a medication in an individual patient, subsequent medication use may change; patients who tolerate the unintended side-effects or experience the intended therapeutic effects of a medication are more likely to continue using it in the future compared to those who experience side-effects or those who do not perceive the beneficial effects of a medication.

Fig. 8 outlines different approaches to minimize confounding and bias in pharmacoepidemiology at the design and analysis stages of the research. In observational studies, measured confounders can be accounted for at the design stage through restriction, matching, or implementation of a new-user design; and at the analysis stage through stratification, standardization, or regression including propensity scores techniques. Likewise, unmeasured confounders can be addressed at the design stage through the implementation of crossover active comparator, or validation designs; and at the analysis stage through instrumental variable techniques, sensitivity analyses, and utilizing proxy measures [30].

In addition to selection and measurement bias, confounding is one of the most important sources of bias in observational studies, and any observed association in such studies is confounded to some degree [31]. However, well designed and well conducted rigorous observational studies have increasingly important clinical, regulatory, and public health impacts.

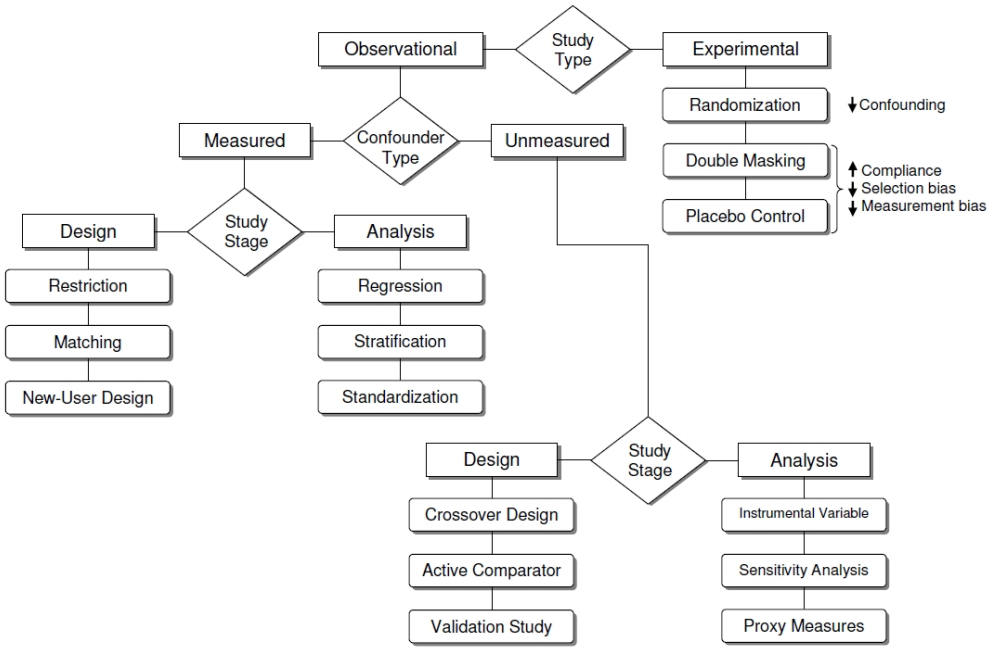


**Fig. 8. Pharmacoepidemiologic approaches to account for confounding and bias**

## CONSENT

Not applicable.

## ETHICAL APPROVAL

Not applicable.

## ACKNOWLEDGEMENTS

This article was part of author's doctoral dissertation. Author thanks the following individuals from University of Florida for their advisory role: Abraham Hartzema, Richard Segal, and Almut Winterstein from department of pharmaceutical outcomes and policy, college of pharmacy; Leslie Hendeles from department of pharmacotherapy and translational research, college of pharmacy; and Xiaomin Lu from department of biostatistics, college of public health and health professions and college of medicine.

## COMPETING INTERESTS

Author has declared that no competing interests exist.

## REFERENCES

1. Rothman KJ, Greenland S, Lash TL. Validity in Epidemiologic Studies. In Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins. 2008;128-147.
2. Schneeweiss S. Confounding. In: Hartzema AG, Tilson HH, Chan KA, editors. Pharmacoepidemiology and Therapeutic Risk Management. Cincinnati, OH: Harvey Whitney Books Company. 2008;273-299.
3. Glymour MM, Greenland S. Causal Diagrams. In: Rothman KJ, Greenland S, Lash TL, editors. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008;183-209.
4. Bégaud B. Dictionary of Pharmacoepidemiology. West Sussex, UK: John Wiley & Sons, Ltd; 2000.
5. Hudson M, Suissa S. Avoiding common pitfalls in the analysis of observational studies of new treatments for rheumatoid arthritis. Arthritis Care & Research. 2010;62(6):805-810.
6. Psaty BM, Siscovick DS. Minimizing bias due to confounding by indication in comparative effectiveness research: the importance of restriction. Journal of the American Medical Association. 2010;304(8):897-898.
7. Kattan MW, Cowen ME. Causal Inference in Medical Decision Making. In: Encyclopedia of Medical Decision Making, vol. 1. Thousands Oaks, CA: SAGE Publications, Inc; 2009;117-121.
8. Robins JM, Hernán MA, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. Epidemiology. 2000;11(5):550-560.
9. Fewell Z, Hernán MA, Wolfe F, Tilling K, Choi H, Sterne JAC. Controlling for Time-Dependent Confounding Using Marginal Structural Models. The Stata Journal. 2004;4(4):402-420.
10. Stram DO, Huberman M, Wu AH. Is residual confounding a reasonable explanation for the apparent protective effects of beta-carotene found in epidemiologic studies of lung cancer in smokers? American Journal of Epidemiology. 2002;155(7):622-628.

11. Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. Epidemiology. 1997;8(4):429-434.
12. Loerbroks A, Apfelbacher CJ, Amelang M, Strümer T. Obesity and adult asthma: potential effect modification by gender, but not by hay fever. Annals of Epidemiology. 2008;18(4):283-289.
13. Greenland S, Rothman KJ, Lash TL. Measures of Effect and Measures of Association. In: Rothman KJ, Greenland S, Lash TL, editors. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008;51-70.
14. Thompson WD. Effect modification and the limits of biological inference from epidemiologic data. Journal of Clinical Immunology. 1991;44(3): 221-232.
15. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15(5):615-625.
16. Allison PD. Survival Analysis Using SAS®: A Practical Guide. 2nd ed. Cary, NC: SAS Institute Inc; 2010.
17. van der Laan MJ, Robins JM. Unified Methods for Censored Longitudinal Data and Causality. New York, NY: Springer-Verlag; 2003.
18. Hernán MA. The hazards of hazard ratios. Epidemiology. 2010;21(1):13-15.
19. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. Epidemiology. 1994;5(2):189-196.
20. Szklo M, Nieto FJ. Understanding lack of validity: Bias. In Epidemiology, Beyond the Basics. Sudbury, MA: Jones and Bartlett Publishers, Inc.; 2004;125-176.
21. Danaei G, Tavakkoli M, Hernán MA. Bias in observational studies of prevalent users: Lessons for comparative effectiveness research from a meta-analysis of statins. American Journal of Epidemiology. 2012;175(4):250-262.
22. Hernán MA, Cole SR. Causal diagrams and measurement bias. American Journal of Epidemiology. 2009;170(8):959-962; discussion 963-964.
23. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002;359(9302):248-252.
24. Suissa S. Immortal time bias in pharmaco-epidemiology. American Journal of Epidemiology 2008;167(4):492-499.
25. Rothman KJ, Greenland S. Cohort Studies. In: Rothman KJ, Greenland S, Lash TL, editors. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008;100-110.
26. Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. American Journal of Respiratory and Critical Care Medicine. 2003;168(1):49-53.
27. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. Journal of Clinical Epidemiology. 2012;65(2):126-131.
28. National Heart, Lung and Blood Institute (NHLBI). Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma; 2007. Accessed 15 October 2012. Available: http://www.nhlbi.nih.gov/guidelines/asthma/asthgdln.pdf.
29. Toh S, Hernán MA. Causal inference from longitudinal studies with baseline randomization. International Journal of Biostatistics. 2008;4(1):Article 22.

30. Hartzema AG, Tilson HH, Chan KA, editors. Pharmacoepidemiology and therapeutic Risk Management. Cincinnati, OH: Harvey Whitney Books Co.; 2008.
31. Jepsen P, Johnson SP, Gillman MW, Sørensen HT. Interpretation of observational studies. Heart. 2004;90(8):956-960.



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*http://www.sciencedomain.org/review-history.php?iid=197&id=14&aid=1032*