A Textbook of Pharmacoepidemiology and Pharmacovigilance for Pharmacy Students

Lesson 3:

Confounding and bias

Fernando Fernandez-Llimos

ISBN-13: 978-84-124195-4-2

Porto, Portugal.

Published by CIPF, S.L.

Suggested citation: Fernandez-Llimos F. Lesson 3: Confounding and bias. In: A Textbook of Pharmacoepidemiology and Pharmacovigilance for Pharmacy Students. Porto: CIPF; 2025.

Available at http://www.cipf-es.org

Lesson 3: Confounding and bias

Contents

3.1. Introduction to Systematic Errors

Pharmacoepidemiologic research bridges the gap between controlled clinical trials and real-world drug use, but its observational nature renders it vulnerable to systematic distortions that compromise **causal inference**. Unlike random errors, which reduce precision but average out across studies, **systematic errors** like **bias** and **confounding** skew results in consistent, directional ways, potentially leading to erroneous conclusions about drug safety and efficacy. These distortions carry profound implications: overestimation of a drug's risk may prompt unnecessary regulatory restrictions, while underestimation could endanger public health.

At the heart of this challenge lies the interplay between prescribing patterns and patient characteristics. For instance, **confounding by indication** arises when the clinical rationale for prescribing a drug (e.g., antipsychotics for dementia-related agitation) becomes entangled with the baseline risk of the outcome (e.g., pneumonia in frail elderly patients). Similarly, **selection**

bias may artificially inflate risk estimates if study populations disproportionately include high-risk individuals, as seen in early studies linking acid-suppressive drugs to pneumonia, where hospitalized patients — already predisposed to infections — are overrepresented.

Systematic error and random error

Systematic error, often referred to as bias, is a consistent deviation from the true value that occurs in a predictable direction. This type of error arises from flaws inherent in the design, conduct, or analysis of a study, such as using a faulty measurement instrument, applying incorrect data collection procedures, or selecting an unrepresentative sample. Because systematic error consistently skews results in one direction, it leads to inaccuracy, or a lack of validity, in the findings. Importantly, **increasing the sample size does not reduce systematic error**; the bias remains regardless of how much data is collected. Common examples of systematic error include selection bias, where certain groups are more likely to be included or excluded from a study, and information bias, where there are systematic differences in how data is obtained or recorded.

In contrast, **random error** refers to the unpredictable fluctuations that occur due to chance or inherent variability in the measurement process. Unlike systematic error, random error does not consistently push results in one direction; instead, it causes results to scatter around the true value, sometimes overestimating and sometimes underestimating it. Random error reflects the imprecision of a measurement or sampling process, and is expected in any study due to biological variability, sampling fluctuations, or minor inconsistencies in measurement techniques. One key feature of **random error is that it can be minimized by increasing the sample size** or by improving measurement precision, as repeated observations tend to average out these random fluctuations. However, random error cannot be completely eliminated.

To illustrate the difference, consider a study measuring blood pressure using a sphygmomanometer. If the device is miscalibrated and always reads 5 mmHg higher than the true value, every measurement will be systematically too high — this is systematic error. If, however, small variations in readings occur due to slight changes in the position of the cuff or observer technique, these represent random error, as they may be higher or lower than the true value and tend to cancel out over many measurements.

The consequences extend beyond academic debate. Regulatory decisions, such as the withdrawal of rofecoxib due to cardiovascular risks, depend on distinguishing true drug effects from confounding factors, such as the severity of underlying arthritis. This chapter explores the mechanisms by which **bias** and **confounding** distort pharmacoepidemiologic evidence, equipping students with frameworks to critically appraise studies and design robust real-world analyses. Subsequent sections will dissect specific error types, mitigation strategies, and contemporary case studies, culminating in methodologies to navigate these challenges in an era of expanding real-world data.

Understanding Causality and Association

Distinguishing between **causality** and **association** is fundamental to interpreting research findings and making informed clinical decisions. An **association** refers to **a statistical relationship between two variables**, such as a drug exposure and a health outcome, where these events occur together more or less often than would be expected by chance alone. However, **the presence of an association does not automatically imply that one variable causes the other**; it simply indicates that they are linked in some way, which could be due to chance, bias, confounding factors, or a true causal relationship.

Causality, on the other hand, implies **a direct cause-and-effect relationship**, where exposure to a particular factor (such as a medication) is responsible for producing a specific outcome (such as an adverse event). Establishing causality requires more rigorous evidence and often involves applying criteria such as those proposed by **Bradford Hill**, including the strength and consistency of the association, the temporal sequence (exposure precedes the outcome), biological plausibility, and the presence of a dose-response relationship. Only when these criteria are met can we confidently infer that the exposure is likely to have caused the observed effect.

For example, observational studies may find that patients taking a certain antihypertensive drug have a lower risk of developing stroke compared to those not taking the drug. This observed association could be due to the drug's effect, but it could also result from confounding factors, such as healthier lifestyle choices among those prescribed the medication. Only through well-designed randomized controlled trials (or interrupted time series analyses) and careful application of causal inference criteria can we determine whether the drug truly reduces stroke risk (causality), rather than merely being associated with it.

Reference:

Hill AB. The environment and disease: association or causation? Proc R Soc Med. 1965;58(5):295-300. doi:10.1177/003591576505800503

3.2. Bias vs. Confounding

Bias and **confounding** represent distinct mechanisms of **systematic error**, each requiring specific methodological approaches to mitigate their distorting effects. While both threaten the validity of causal inferences, their origins and pathways differ fundamentally.

Bias arises from flaws in study design, data collection, or analysis that systematically skew observations away from the true association. It operates through non-random distortions in how exposure or outcome information is gathered or interpreted. For example, **recall bias** in **case-control studies** occurs when patients with adverse events (cases) retrospectively report medication use more thoroughly than controls, artificially inflating risk estimates. Similarly, **immortal time bias** in cohort studies emerges when unexposed follow-up periods are misclassified, as seen in early analyses of statins and sepsis risk where pre-treatment intervals were improperly excluded.

The Bradford Hill Criteria of Causality

The Bradford Hill criteria, introduced by Sir Austin Bradford Hill in 1965, provide a framework for assessing causality in epidemiological research. These nine viewpoints — not rigid rules — guide the evaluation of whether an observed association may reflect a causal relationship. Below is a summary of each criterion:

- Strength of Association: Stronger associations (e.g., large effect sizes) are more suggestive of causality, though weak associations may still be causal.
- Consistency: Repeated observations of the association across different studies, populations, and settings bolster causal inference.
- Specificity: A cause leading to a single specific outcome supports causality, but this is not essential (e.g., smoking causes multiple diseases).
- Temporality: The cause must precede the effect; this is the only indispensable criterion.
- Biological Gradient (Dose-Response): Increasing exposure intensity should correlate with higher outcome risk (e.g., smoking more cigarettes → greater lung cancer risk).
- Plausibility: A biologically plausible mechanism strengthens causality, but this depends on current scientific knowledge.
- Coherence: The causal hypothesis should align with existing understanding of disease biology and epidemiology.
- Experiment: Evidence from interventions (e.g., reduced exposure leading to fewer outcomes) supports causation.
- Analogy: Similar cause-effect relationships in analogous scenarios may support causality (e.g., known teratogens suggesting risks for new drugs).

Key Considerations

- These criteria are viewpoints, not a checklist; no single criterion is mandatory (except temporality).
- They complement modern causal frameworks (e.g., Directed Acyclic Graphs) but remain foundational in pharmacoepidemiology for evaluating drug-outcome relationships.
- Limitations: Critics note that confounding, bias, or chance can mimic these criteria, necessitating careful integration with study design and statistical methods.

Reference:

Hill AB. The environment and disease: association or causation? Proc R Soc Med. 1965;58(5):295-300. doi:10.1177/003591576505800503

Confounding, by contrast, stems from **pre-existing differences** between exposed and unexposed groups that independently influence both treatment allocation and outcome risk. A confounder must:

1. Be associated with the exposure.
2. Predict the outcome independent of the exposure.
3. Not lie on the causal pathway between exposure and outcome.

Consider studies investigating NSAIDs and gastrointestinal bleeding. Physicians often avoid prescribing NSAIDs to high-risk patients (e.g., those with prior ulcers), creating **confounding by contraindication** where the unexposed group inherently carries lower bleeding risk. This differs fundamentally from bias, as the distortion originates from real-world prescribing patterns rather than measurement or design flaws.

Mechanistic distinctions

- **Directionality**: Bias typically introduces unidirectional error (e.g., overestimation of risk through preferential outcome ascertainment), while confounding can create spurious positive or negative associations depending on the confounder's distribution.
- **Temporal relationships**: Confounders precede both exposure and outcome, whereas bias often originates during study execution. **Protopathic bias** — a hybrid phenomenon — occurs when early symptoms of the outcome drive exposure, as seen when acid-suppressive drugs are prescribed for undiagnosed gastric cancer precursors.
- **Adjustability**: Measured confounders can be addressed through stratification or multivariate models, but many biases (e.g., selection bias from non-random loss to follow-up) resist statistical correction.

The interplay between these forces is exemplified in research on antidepressants and suicide risk. **Confounding by severity** occurs because severely depressed patients — already at higher suicide risk — receive antidepressants more frequently. Simultaneously, **surveillance bias** may exaggerate risk estimates if medicated patients undergo closer psychiatric monitoring, increasing outcome detection. Disentangling these requires both analytic adjustments for depression severity and sensitivity analyses quantifying potential surveillance effects.

Recognizing these distinctions is pivotal for evidence interpretation. **While confounding reflects real-world clinical complexity, bias often signals methodological limitations.** Advanced designs like active comparator new-user cohorts simultaneously address both by anchoring analyses to treatment initiation times and selecting comparators with similar indication profiles. Subsequent sections will explore specific manifestations of these errors and their solutions in pharmacotherapeutic research.

3.3. Types of Bias

3.3.1. Types of Bias: Selection Bias

Channeling Bias

Channeling bias occurs when prescribers systematically direct specific medications to patients based on unmeasured clinical characteristics or risk profiles, creating **non-comparable exposure groups**. This bias is particularly pervasive when comparing newer drugs to established therapies, as prescribers often **reserve novel agents for patients perceived to have higher clinical complexity or treatment resistance**. For example, in studies comparing sodium-glucose cotransporter-2 (SGLT-2) inhibitors to older antidiabetic drugs like sulfonylureas, SGLT-2 inhibitors are preferentially prescribed to patients with established cardiovascular disease due to their known cardioprotective benefits. This creates an exposed cohort with inherently higher baseline cardiovascular risk, leading to inflated estimates of adverse renal outcomes if analyses fail to adjust for this channeling effect. Similarly, early observational studies linking proton pump inhibitors (PPIs) to dementia risk were confounded by channeling bias: patients with gastroesophageal reflux disease (GERD) often have vascular risk factors independently associated with cognitive decline, creating a spurious association between PPI use and neurodegeneration.

Mitigating channeling bias requires designs that balance indication profiles across exposure groups. **Active comparator new-user cohorts, which restrict analyses to patients initiating therapy without prior use of either drug**, help align baseline risks. Additionally, negative control outcome analyses — testing associations with biologically implausible endpoints (e.g., influenza vaccination rates) — can reveal residual channeling effects.

Prevalent User Bias

Prevalent user bias distorts risk estimates when studies include long-term medication users rather than new initiators, **disproportionately capturing "survivors" who tolerate the drug well**. This bias is magnified in comparisons between newly marketed drugs and established therapies. Consider a study comparing the novel oral antidiabetic drug SameAsMetformin[1] to metformin, where the cohort includes both **incident and prevalent users**. Metformin, having been available for decades, has a pool of long-term users who discontinued therapy due to early adverse effects (e.g., gastrointestinal intolerance), leaving a residual population with greater drug tolerance. In contrast, SameAsMetformin's cohort consists primarily of new users, including both tolerant and susceptible individuals. If the study finds higher lactic acidosis risk with SameAsMetformin, this may reflect differential depletion of susceptible metformin users rather than true pharmacological risk. **Restricting the analysis to incident users of both drugs often resolves this bias by ensuring comparable susceptibility profiles**.

This bias also affects chronic disease research. In studies of bisphosphonates and osteonecrosis of the jaw, prevalent users — who persisted with therapy despite early side effects — may underrepresent the true risk profile observed in new initiators. Advanced methods like clone-censoring analyses, which account for treatment discontinuation patterns, can partially adjust for this depletion of susceptible ones.

Immortal Time Bias

Immortal time bias arises from misclassifying unexposed follow-up periods as exposed, typically due to **flawed cohort entry definitions**. In a landmark study examining β-blockers and mortality post-myocardial infarction (MI), researchers compared patients who received β-blockers during hospitalization to those who did not. However, β-blocker initiation often occurred days after admission, creating an "immortal" period during which **only survivors could receive the drug**. By excluding early deaths that occurred before β-blocker administration, the analysis falsely attributed survival benefits to the medication. Similarly, early studies linking statins to reduced sepsis risk erroneously defined exposure as any statin prescription within a follow-up window, ignoring pre-treatment intervals when outcomes could not occur.

Modern approaches to mitigate this bias include:

1. **Time-varying exposure analysis**: Treating drug exposure as a time-dependent variable to account for initiation timing.
2. **Lag periods**: Exposing follow-up time immediately after cohort entry to avoid misclassifying pre-treatment events.

---

[1] Fake name created for educational purposes

3. **Clone-censoring**: Creating matched clones for each patient at risk of exposure switch, ensuring comparable follow-up periods

3.3.2. Types of Bias: Information Bias

Information bias arises from systematic errors in measuring exposures, outcomes, or confounders, distorting the validity of pharmacoepidemiologic studies. Unlike random errors that reduce precision, information bias skews effect estimates in predictable directions, potentially leading to false conclusions about drug safety or efficacy. This section examines three critical subtypes: misclassification, recall bias, and detection bias, with examples illustrating their mechanisms and impacts.

Misclassification

Misclassification occurs when **exposure, outcome, or confounder variables are inaccurately measured or categorized**.

Non-differential misclassification introduces random errors affecting all study groups equally, typically biasing results toward the null. For instance, using prescription records to define statin adherence may misclassify non-adherent patients as adherent across both cases and controls, diluting true associations with myopathy risk. In studies of antidepressant use and suicide risk, non-differential misclassification of suicide attempts (e.g., relying on incomplete hospital coding) could obscure genuine pharmacological effects.

Differential misclassification, by contrast, introduces systematic errors that disproportionately affect one group. A case-control study linking antidepressants to birth defects may suffer from this if mothers of affected infants (cases) meticulously recall prenatal drug exposures, while controls underreport due to less salient memories. This differential recall inflates odds ratios, creating illusory associations. Similarly, electronic health records (EHRs) may undercapture lifestyle factors in schizophrenia patients — who often receive fragmented care — leading to overestimates of antipsychotic-related metabolic risks compared to healthier populations.

Modern pharmacoepidemiologic datasets amplify these challenges. For example, studies comparing antidiabetic drugs like SGLT-2 inhibitors with metformin may misclassify exposure if electronic health records (EHRs) fail to capture over-the-counter metformin use, artificially inflating the perceived superiority of newer agents. **Validation studies using pharmacy dispensing data or patient surveys can mitigate such errors**, though residual misclassification often persists.

Recall Bias

Recall bias, a subtype of **differential misclassification**, emerges when participants' ability to report exposures differs based on outcome status. This bias predominantly afflicts **case-control** studies, where cases — aware of their condition — may scrutinize past exposures more thoroughly than controls. An example is the association between maternal antidepressant use and congenital heart defects: mothers of affected infants often retrospectively report medication use with greater detail than controls, exaggerating risk estimates. Conversely, controls may underreport exposures perceived as socially stigmatizing (e.g., opioid use), further distorting associations.

Protopathic bias, a related phenomenon, occurs when early symptoms of an undiagnosed outcome influence exposure measurement. In studies linking acid-suppressive drugs to gastric cancer, early dyspepsia (a precursor to cancer) may prompt PPI prescriptions, creating the illusion that PPIs cause cancer. Similarly, insulin initiation in diabetic patients may coincide with undiagnosed pancreatic cancer — a condition causing glucose dysregulation — leading to spurious associations between insulin and cancer risk. **Mitigation strategies include lag periods (excluding exposure immediately before the outcome diagnosis) and validation through biomarkers or procedural records**.

Detection Bias

Detection bias arises when **surveillance intensity differs between exposure groups**, leading to unequal outcome ascertainment. Patients on chronic therapies often undergo more frequent medical monitoring, increasing incidental diagnosis of conditions unrelated to the drug. For example, anticoagulant users receive regular INR testing, elevating detection of asymptomatic liver enzyme abnormalities compared to non-users. Early studies misinterpreted these findings as drug-induced hepatotoxicity, overlooking the role of surveillance patterns.

This bias also distorts comparative effectiveness research. In analyses of biologics versus conventional DMARDs for rheumatoid arthritis, biologics' higher cost and perceived risk may prompt more frequent imaging, increasing detection of subclinical cardiovascular events. Conversely, conventional DMARD users might undergo less rigorous monitoring, underdiagnosing similar outcomes. **Active comparator designs — pairing drugs with similar monitoring protocols — help equalize surveillance, while negative control outcomes (e.g., fractures in antidepressant studies) test for residual detection bias**.

3.3.3. Types of Bias: Confounding by indication

Confounding by indication represents one of the most pervasive and methodologically challenging biases in pharmacoepidemiologic research. It arises when **the clinical rationale for prescribing a drug, the indication, becomes entangled with the baseline risk of the outcome under study**, creating spurious associations that obscure true drug effects. Unlike conventional confounding, where extraneous variables distort results, this bias originates from the intrinsic link between therapeutic decision-making and patient prognosis.

At its core, confounding by indication occurs when:

1. The indication for treatment (e.g., disease severity, comorbidities) is a risk factor for the outcome.
2. Treatment allocation correlates with this indication.
3. The indication is not adequately measured or adjusted for in analyses.

For example, early observational studies suggested that hormone replacement therapy (HRT) reduced cardiovascular risk in postmenopausal women. However, HRT was preferentially prescribed to healthier women with fewer cardiovascular risk factors — a phenomenon termed **healthy user bias**. The apparent cardioprotective effect vanished when **adjusting for baseline health status**, revealing that the indication (healthier profile) rather than the drug drove the observed association.

Text Box: Why confounding by indication is a bias, not just confounding

Confounding by indication represents a specialized form of bias within pharmacoepidemiology, distinct from general confounding due to its origin in clinical decision-making. While **confounding** broadly refers to any distortion in exposure-outcome associations caused by a third variable influencing both, **confounding by indication** arises specifically when the clinical rationale for treatment (e.g., disease severity or comorbidities) acts as that confounder. This creates systematic bias because treatment allocation is intrinsically linked to prognosis: sicker patients receive specific drugs, making the drug appear causally associated with worse outcomes.

The critical distinction lies in **non-random channeling**: unlike generic confounders (e.g., age or smoking), the "indication" directly determines treatment assignment while also influencing outcomes. For example, anticoagulants are prescribed to high-risk thrombosis patients, so any observed harm from the drug may reflect the underlying risk profile rather than the drug's effect. This violates the principle of exchangeability between treated and untreated groups.

Why It constitutes a bias:

1. **Structural entanglement**: The indication (e.g., disease severity) is inseparable from treatment decisions, creating a self-reinforcing distortion that standard adjustment cannot fully resolve.
2. **Measurement intractability**: Clinical indications are rarely captured comprehensively in real-world data (e.g., claims databases), leaving residual confounding even after statistical adjustments.
3. **Directional distortion**: It systematically skews results toward paradoxical conclusions (e.g., beneficial drugs appearing harmful when prescribed to high-risk groups).

Confounding by indication bias manifests in multiple forms:

- **Disease severity gradients**: In studies of antipsychotics and pneumonia risk in dementia patients, advanced dementia itself predisposes individuals to aspiration due to dysphagia. Since antipsychotics are prescribed more frequently in severe dementia, the underlying disease — not the drug — explains the elevated pneumonia risk.
- **Therapeutic contraindications**: Physicians often avoid NSAIDs in high-risk patients with prior gastrointestinal bleeding, creating an unexposed group with inherently lower bleeding risk. This **confounding by contraindication** artificially inflates the perceived safety of alternative analgesics
- **Protopathic bias**: Early symptoms of an undiagnosed outcome may prompt drug initiation. Studies linking acid-suppressive drugs to gastric cancer initially failed to account for dyspepsia — a precursor symptom of cancer that triggered PPI prescriptions

Confounding by indication is particularly **resistant to conventional adjustment methods** because clinical decision-making often incorporates unmeasured or poorly captured factors. Electronic health records (EHRs) rarely document nuanced therapeutic rationale, such as prescribers' subjective assessments of disease severity. Even when measurable, indication variables may exhibit structural confounding — being too tightly correlated with treatment choices to permit statistical disentanglement. For instance, tumor necrosis factor (TNF) inhibitors are reserved for severe rheumatoid arthritis cases, making it impossible to separate drug effects from the inflammatory burden driving both treatment selection and cardiovascular outcomes.

3.4. Sources of confounding in drug outcome studies

Disease severity gradients and treatment allocation patterns

Disease severity gradients represent one of the most recalcitrant sources of confounding in pharmacoepidemiologic research, arising when treatment decisions correlate with unmeasured or incompletely adjusted disease progression markers. Clinicians naturally escalate therapies in patients with worsening symptoms or biomarker profiles, creating exposure groups with intrinsically divergent outcome risks. This phenomenon, termed **confounding by severity**, distorts causal inferences by conflating pharmacological effects with the natural history of untreated disease.

A paradigmatic example emerges from studies of aldosterone antagonists in heart failure. As demonstrated, physicians preferentially prescribe these drugs to patients with advanced systolic dysfunction and elevated natriuretic peptide levels — markers of poor prognosis. If analyses fail to adequately **adjust for ejection fraction** or functional class, the drug appears harmful due to its concentration in high-risk patients, despite randomized trials showing mortality benefits. This "severity skew" explains why early observational studies of β-blockers in heart failure paradoxically associated treatment with increased mortality — a finding later reversed through **propensity score matching** on disease severity parameters.

This challenge is intensified in chronic diseases with heterogenous phenotypes. In multiple sclerosis, disease-modifying therapies (DMTs) like natalizumab are reserved for aggressive relapsing-remitting subtypes. Comparisons with interferon-β users — a cohort dominated by milder cases — artificially inflate DMT efficacy estimates unless stratified by magnetic resonance imaging activity and relapse frequency. Source highlights how structural confounding arises when treatment indications become inseparable from disease subtypes, necessitating active comparator designs that match therapies within indication niches.

Comorbidity-driven prescribing behaviors

Comorbidity-associated confounding arises when coexisting conditions influence both treatment selection and outcome risk through pathways independent of the drug's mechanism. This bias operates through two primary mechanisms: **therapeutic channeling**, where comorbidities dictate drug choice, and **risk amplification**, where comorbidities independently predispose to adverse outcomes.

A classic example involves NSAIDs and renal failure. Physicians avoid NSAIDs in patients with chronic kidney disease (CKD), creating an unexposed control group enriched with low-risk individuals without renal impairment. This prescribing pattern artificially inflates NSAIDs' apparent nephrotoxicity unless analyses adjust for baseline estimated glomerular filtration rate (eGFR) and proteinuria. Conversely, preferential NSAID use in osteoarthritis patients — a population with lower cardiovascular risk than rheumatoid arthritis patients receiving biologics — may mask true cardiorenal risks.

In psychopharmacology, the association between atypical antipsychotics and metabolic syndrome illustrates comorbidity entanglement. These drugs are preferentially prescribed to schizophrenia patients with preexisting obesity and insulin resistance — factors independently linked to diabetes and dyslipidemia.

Temporal confounding in longitudinal datasets

Temporal confounding introduces unique challenges in longitudinal studies, where time-varying factors influence both treatment trajectories and outcome risks. This manifests through two principal mechanisms: immortal time bias, time-dependent confounding, and lagged exposure effects, each requiring distinct analytic approaches.

Time-dependent confounding occurs when time-varying factors (e.g., hemoglobin A1c, blood pressure) influence both treatment intensification and outcomes. Declining hemoglobin triggers erythropoietin-stimulating agents (ESA) dose escalation while independently predicting mortality. Conventional time-dependent Cox models fail here because hemoglobin lies on the causal pathway between prior ESA use and survival. Marginal structural models with inverse probability weighting resolve this by creating pseudo-populations where time-varying confounders are balanced across exposure histories.

Lagged exposure effects emerge when drugs have delayed therapeutic or adverse effects. Studies of bisphosphonates and atypical femoral fractures must account for cumulative exposure over decades, as transient users (<3 years) have fundamentally different risk profiles than persistent users.

3.5. Impact on Study Validity

False-Positive and False-Negative Associations in Drug Safety Signals

False-positive signals frequently arise from unmeasured confounding and selection biases. For example, early observational studies linking proton pump inhibitors (PPIs) to dementia risk conflated the underlying vascular comorbidities of GERD patients with drug effects, creating illusory associations. Similarly, protopathic bias — where early symptoms of an undiagnosed outcome prompt drug initiation — inflated perceived risks of NSAIDs in myocardial infarction studies, as chest pain preceding infarction drove NSAID prescriptions.

False-negative associations often stem from non-differential misclassification or channeling bias. Non-differential exposure misclassification (e.g., using prescription fills as a proxy for adherence) dilutes true effect sizes toward the null. Channeling bias obscures risks when high-risk patients avoid certain drugs; for instance, physicians withhold NSAIDs from renal-impaired patients, making these drugs appear safer in healthier cohorts. Such distortions carry grave implications: a meta-analysis of corticosteroid trials underestimated gastrointestinal bleeding risks in ambulatory patients by excluding high-risk hospitalized populations.

Magnification of Effects in Rare Outcome Studies

Rare adverse events amplify the impact of biases due to low baseline incidence. Even minor misclassification or residual confounding disproportionately inflates odds ratios (ORs) in case-control studies. For example, early investigations associating aspirin with Reye's syndrome reported ORs >40, but later studies adjusting for viral illness severity reduced estimates to OR 12–18, demonstrating how unmeasured confounders magnify effects in rare outcomes.

This magnification is exacerbated in agnostic medication-wide analyses. A Swedish registry study found that published drug-cancer associations had smaller effect sizes than agnostic analyses, as hypothesis-driven studies selectively adjusted for confounders, while agnostic approaches captured inflated signals from rare, unadjusted associations. Rare outcomes also increase vulnerability to **competing biases**: immortal time bias in statin-sepsis studies and detection bias in anticoagulant-liver injury research both produced order-of-magnitude overestimates by misaligning exposure windows and surveillance intensity.

Compromised External Validity and Generalizability

Pharmacoepidemiologic findings often fail to generalize beyond study populations due to **selection biases** and **indication creep**. RCTs of antidepressants systematically exclude patients with suicidality or comorbidities, limiting applicability to real-world psychiatric populations. Similarly, electronic health record (EHR) studies overrepresent insured, tech-literate patients, undercapturing marginalized groups.

Therapeutic heterogeneity further erodes generalizability. Drugs like SGLT-2 inhibitors show cardioprotective effects in diabetes trials but increased hospitalization risks when used off-label for heart failure in older adults with multimorbidity. Temporal confounding compounds this: COVID-19 vaccine safety signals initially appeared stronger in early-adopter cohorts (healthcare workers with high SARS-CoV-2 exposure) than in later-vaccinated general populations.

3.6. Methodological strategies for bias mitigation

3.6.1. Study Design Approaches

Active comparator new-user (ACNU) designs

The ACNU design emulates randomized controlled trial (RCT) principles by comparing new initiators of a target drug to those starting a therapeutically similar alternative, ensuring both groups share comparable indications and baseline risks. This approach minimizes confounding by indication — a pervasive issue when comparing treated and untreated populations. For example, studies comparing SGLT-2 inhibitors to GLP-1 agonists in type 2 diabetes restrict analyses to patients initiating either drug within 30 days of diagnosis, balancing cardiovascular risk profiles that might otherwise skew results. By excluding prevalent users, the design avoids "healthy survivor" bias, where long-term users represent a selectively tolerant subgroup. ACNU frameworks also align follow-up periods, eliminating immortal time bias that plagued early statin studies misclassifying pre-treatment intervals as exposed time.

Nested case-control studies

Nested within well-defined cohorts, this design efficiently studies rare outcomes while preserving temporality. Cases are identified during cohort follow-up, with controls sampled from risk sets matched on time and covariates. For instance, in assessing COVID-19 vaccine-associated thrombosis, cases were matched to controls vaccinated in the same week, adjusting for age and comorbidities. This method leverages cohort data richness (e.g., EHRs with longitudinal biomarkers) while reducing costs: a study of antipsychotics and diabetes used nested sampling to analyze stored blood samples from 1,200 cases/controls instead of testing all 50,000 cohort members. Time-varying exposures are naturally handled, as in research linking cumulative bisphosphonate exposure to atypical fractures, where exposure windows were anchored to each participant's unique follow-up timeline.

3.6.2. Analytic Techniques

Propensity Score Methods

Propensity scores (PS) — the probability of treatment assignment conditional on covariates — enable adjustment for measured confounders through matching, stratification, or inverse probability weighting. In a study of NSAID-related gastrointestinal bleeding, PS matching paired each NSAID user with a non-user having similar age, renal function, and ulcer history, reducing bias from contraindication-driven prescribing. However, PS effectiveness fails in capturing all relevant confounders. **High-dimensional propensity scoring (hdPS)** addresses this by algorithmically selecting hundreds of covariates from large datasets, including proxies for unmeasured factors. For example, hdPS analyses of antidepressant safety automatically identified "number of psychiatrist visits" and "prior sleep disorder diagnoses" as surrogates for depression severity, improving control over confounding by indication.

3.6.3. Sensitivity Analyses

Quantitative Bias Analysis (QBA)

QBA quantifies how residual biases might alter results under specified assumptions. For unmeasured confounding, **E-value estimation** calculates the minimum strength of association an unmeasured confounder would need with both exposure and outcome to negate the observed effect. A study linking opioids to depression (RR=1.8) reported an E-value of 2.9, meaning any unmeasured confounder (e.g., chronic pain severity) must triple depression risk and opioid use to explain the association. For measurement error, **probabilistic bias analysis** simulates misclassification rates: a misclassification-adjusted analysis of statin-diabetes risk changed the hazard ratio from 1.25 to 1.12 when assuming 80% sensitivity/90% specificity in exposure ascertainment.

Text box: the E-value

The E-value is a sensitivity analysis metric used in observational studies to quantify the minimum strength of association that an unmeasured confounder must have with both the exposure and outcome to fully explain away an observed association. Developed to address unmeasured confounding, it provides a standardized way to assess the robustness of causal inferences.

For an observed risk ratio (RR) > 1:

$$\text{E-value} = RR + \sqrt{RR \times (RR - 1)}$$

For RR < 1, first take the inverse (1/RR1/RR) before applying the formula.

- A **large E-value** (e.g., >3) implies that substantial unmeasured confounding would be needed to nullify the association, supporting result robustness.
- A **small E-value** (e.g., <2) suggests weak unmeasured confounding could explain the association, undermining causal claims.

Triangulation with negative controls

Testing associations with implausible outcomes (e.g., antihypertensives and fracture risk) detects residual confounding. A study of ACE inhibitors and lung cancer found similar elevated risks for negative control outcomes (e.g., skin cancer), prompting reanalysis that attributed the signal to smoking-related surveillance bias.