A Textbook of Pharmacoepidemiology and Pharmacovigilance for Pharmacy Students

Lesson 1:

Foundations of Pharmacoepidemiology

Fernando Fernandez-Llimos

ISBN-13: 978-84-124195-2-8

Porto, Portugal.

Published by CIPF, S.L.

Suggested citation: Fernandez-Llimos F. Lesson 1: Foundations of Pharmacoepidemiology. In: A Textbook of Pharmacoepidemiology and Pharmacovigilance for Pharmacy Students. Porto: CIPF; 2025.

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

Lesson 1: Foundations of Pharmacoepidemiology

Contents

1.1 Introduction to Pharmacoepidemiology

Pharmacoepidemiology represents the confluence of pharmacotherapy and epidemiology, applying epidemiological principles to study the utilization, effectiveness, and safety of drugs in large populations. Unlike clinical pharmacology, which examines drug mechanisms in controlled settings, or clinical pharmacy, which optimizes pharmacotherapy to a given patient's characteristics, pharmacoepidemiology addresses the real-world complexities of drug use, including long-term outcomes, polypharmacy, and heterogeneous patient populations. This discipline emerged from the necessity to bridge the gap between pre-marketing clinical trials - which prioritize internal validity under idealized conditions - and the **external validity** required to inform clinical practice, regulatory decisions, and public health policy.

The term "pharmacoepidemiology" was first coined in the 1970s, reflecting the growing recognition that post-marketing surveillance and observational research were critical to understanding drug effects beyond the confines of randomized controlled trials (RCTs). By leveraging large-scale datasets and robust epidemiological methods, pharmacoepidemiology provides insights into rare adverse drug reactions (ADRs), drug utilization patterns, and the comparative effectiveness of therapies in diverse populations. Its applications span regulatory science, healthcare policy, and clinical practice, making it indispensable in the modern pharmaceutical landscape.

1.2 Historical Evolution of Pharmacoepidemiology

1.2.1 Pre-Modern Era: Early Observations and Disasters

The origins of pharmacoepidemiology could be traced to ancient civilizations, where herbal remedies were documented for efficacy and toxicity. However, systematic drug safety evaluation began in the 19th century alongside industrialization and the rise of synthetic chemistry. Early pharmacovigilance relied on anecdotal reports, such as the 1877 documentation of chloral hydrate addiction. The absence of regulatory frameworks led to catastrophic public health events, most notably the **sulfanilamide disaster of 1937**, resulting in 107 deaths in the U.S. due to renal failure. This tragedy exposed the lack of pre-market safety regulations and prompted the 1938 Federal Food, Drug, and Cosmetic Act, **mandating proof of safety before drug approval**.

Box 1.1: The Sulfanilamide Disaster of 1937

The 1937 Sulfanilamide Disaster stands as one of the most consequential pharmaceutical tragedies in U.S. history, catalyzing fundamental reforms in drug regulation. The crisis began when the S.E. Massengill Company, a Tennessee-based pharmaceutical firm, sought to market a liquid formulation of sulfanilamide, a widely used antibacterial drug. Sulfanilamide had proven effective in treating streptococcal infections but was available only in tablet or powder form, which some patients found difficult to swallow. To meet demand for a palatable alternative, the company's chief chemist, Harold Cole Watkins, dissolved sulfanilamide in **diethylene glycol (DEG)** - a sweet-tasting industrial solvent commonly used in antifreeze - and added raspberry flavoring, branding the product "Elixir Sulfanilamide"

Unbeknownst to Watkins, DEG was highly toxic to humans, causing severe kidney failure, neurological damage, and death. At the time, premarketing safety testing was not legally required, and Massengill distributed the elixir nationwide without evaluating its toxicity in animals or humans

By October 1937, physicians began reporting deaths linked to the drug, primarily among children treated for sore throats. Symptoms included agonizing abdominal pain, vomiting, convulsions, and renal shutdown, with victims often succumbing within days. The American Medical Association (AMA) and the U.S. Food and Drug Administration (FDA) launched urgent investigations, identifying DEG as the lethal agent

The FDA, led by Commissioner Walter Campbell, mobilized a nationwide recall, but its authority was limited to prosecuting the company for misbranding (the elixir contained no alcohol), violating the technical definition of an "elixir"

Over 240 gallons of the toxic preparation were recovered, but at least **107 deaths** (including 34 children) were documented across 15 states. Public outrage intensified as grieving families wrote letters to President Franklin D. Roosevelt, detailing their loved ones' suffering

Box 1.1: The Sulfanilamide Disaster of 1937 (continues)

Legacy and Regulatory Impact

The disaster exposed critical gaps in the 1906 Pure Food and Drugs Act, which did not require safety testing or mandate disclosure of drug ingredients. Public pressure and advocacy by FDA officials spurred Congress to pass the **1938 Federal Food, Drug, and Cosmetic Act**, signed into law by President Roosevelt. This landmark legislation:

- Required premarket safety testing for new drugs.
- Empowered the FDA to ban dangerous products and enforce labeling standards.
- Established penalties for noncompliance, laying the foundation for modern drug regulation

The tragedy also underscored the importance of pharmacovigilance and ethical responsibility in drug development. Frances Oldham Kelsey, later famed for blocking thalidomide's U.S. approval, contributed to early research confirming DEG's toxicity

The Sulfanilamide Disaster remains a pivotal case study in pharmaceutical ethics, emphasizing the lifesaving role of rigorous safety protocols and regulatory oversight.

1.2.2 The Thalidomide Crisis (1950s–1960s)

Thalidomide, marketed in the late 1950s as a sedative and antiemetic for pregnancy-related nausea, became a divide moment in drug safety. By 1961, over 10,000 infants worldwide were born with severe congenital malformations, including phocomelia (limb shortening), linked to thalidomide exposure in utero. The crisis revealed critical flaws in drug approval processes and the inadequacy of pre-marketing safety data. Regulatory reforms followed, including the 1962 Kefauver-Harris Amendment in the U.S., which required proof of efficacy, adverse event reporting, and informed consent for clinical trials. Globally, the tragedy spurred the establishment of pharmacovigilance systems, such as the WHO Program for International Drug Monitoring (1968).

Box 1.2: The Thalidomide Crisis (1950s–1960s)

The Thalidomide Crisis remains one of the most catastrophic failures in pharmaceutical history, exposing critical gaps in drug regulation and testing. Developed in the 1950s by the German company Chemie Grünenthal, thalidomide was marketed as a "wonder drug" under trade names like Contergan (Germany) and Distaval (UK). Touted as a safe sedative, anti-anxiety medication, and remedy for morning sickness, it gained rapid approval across 46 countries between 1957 and 1961, bypassing rigorous teratogenicity testing due to lax regulatory standards of the era.

Rapid Approval and Widespread Use

Thalidomide's path to market was alarmingly swift. Preclinical studies in **rodents** suggested low acute toxicity, and no tests were conducted on pregnant animals, as **teratogenicity screening was not yet mandated**. In Germany, the drug was approved in 1957 as an **over-the-counter** medication, becoming a bestseller due to its perceived safety and effectiveness for pregnant women. By 1960, an estimated 14.6 tons of thalidomide were sold in Germany alone, with global distribution facilitated by licensing agreements.

The Emergence of a Crisis

By 1961, pediatricians Wilhelm Lenz (Germany) and William McBride (Australia) independently linked thalidomide to severe congenital malformations, including phocomelia (limb shortening), internal organ defects, and neurological damage. The drug's teratogenic effects depended on the timing of exposure: a single dose during early pregnancy (20–42 days gestation) could disrupt fetal development, with limb deformities occurring as late as day 28.

An estimated 10,000–20,000 infants were affected globally, with approximately 40% dying shortly after birth. Survivors faced lifelong disabilities, including blindness, deafness, and cardiac anomalies. In the UK, 466 survivors were documented by 2010, while Germany reported over 5,000 cases.

Regulatory Failures and Legacy

Thalidomide's withdrawal in 1961–1962 came too late for thousands of families. The crisis revealed systemic flaws:

- **Lack of teratogenicity testing**: No guidelines required fetal safety studies, and animal models (rodents) failed to predict human harm.
- **Inadequate post-marketing surveillance**: Adverse events were dismissed until independent clinicians sounded alarms.

The U.S. narrowly avoided disaster due to pharmacologist Frances Oldham Kelsey at the FDA, who blocked approval despite pressure from the manufacturer, Richardson-Merrell. Her insistence on robust data spared the U.S. widespread harm, though 17 cases arose from clinical trials.

Global Regulatory Reforms

The tragedy catalyzed sweeping changes:

- **1962 Kefauver-Harris Amendment (U.S.)**: Mandated premarket proof of efficacy and safety, including fetal risk assessments.
- **1968 Medicines Act (UK)**: Classified drugs into prescription-only, pharmacy-sold, and general sales categories.
- **International harmonization**: The WHO established stricter pharmacovigilance protocols, requiring teratogenicity testing and pregnancy-specific trials.

Thalidomide's legacy endures as a stark reminder of the ethical imperative for rigorous drug safety evaluation. Its story underpins modern regulatory frameworks, ensuring that "never again" remains a guiding principle in pharmaceutical development.

References:

- Science Museum. Thalidomide [Internet]. 2023 [cited 2024 May 14]. Available from: https://www.sciencemuseum.org.uk/objects-and-stories/medicine/thalidomide
- Vargesson N. Thalidomide-induced teratogenesis: History and mechanisms. Birth Defects Res. 2015;105(2):140–56.
- Kim JH, Scialli AR. Thalidomide: The tragedy of birth defects and the effective treatment of disease. Toxicol Sci. 2011;122(1):1–6.

1.2.3 Institutionalization and Methodological Advances (1970s–1990s)

The 1970s marked the formalization of pharmacoepidemiology as a scientific discipline. Key developments included:

- Automated Healthcare Databases: The U.K. General Practice Research Database (GPRD, 1987) and U.S. Medicaid data enabled large-scale observational studies.
- Landmark Studies: The 1989 case-control study linking aspirin to Reye's syndrome demonstrated pharmacoepidemiology's capacity to identify rare ADRs.
- Regulatory Integration: The U.S. FDA began requiring post-marketing surveillance (Phase IV studies) for high-risk drugs.

1.2.4 Modern Era (2000s–Present)

The **2004 withdrawal of rofecoxib**, a COX-2 inhibitor, due to cardiovascular risks identified through pharmacoepidemiologic studies, emphasized the field's centrality to drug safety. Advances in data analytics, **electronic health records** (EHRs), and **real-world evidence** (RWE)[1] have further expanded its scope. Today, pharmacoepidemiology informs **risk evaluation and mitigation strategies** (REMS), vaccine safety monitoring, and personalized medicine.

[1] clinical evidence derived from the analysis of real-world data (RWD), which refers to data collected from routine clinical practice or other non-research settings. JAMA. 2019;321(14):1359-1360. doi: 10.1001/jama.2019.4064

Box 1.3: The Withdrawal of Rofecoxib in 2004

The withdrawal of rofecoxib (Vioxx), a widely prescribed COX-2 inhibitor, marked a pivotal moment in pharmacovigilance and exposed systemic failures in post-marketing drug safety monitoring. Marketed by Merck & Co. in 1999, Vioxx was initially hailed for its efficacy in treating osteoarthritis and acute pain with reduced gastrointestinal risks compared to traditional NSAIDs. By 2003, annual sales exceeded $2.5 billion, with over 84 million prescriptions worldwide.

Approval and Emerging Safety Concerns

Rofecoxib gained FDA approval based on the VIGOR trial (2000), which demonstrated superior gastrointestinal safety over naproxen. However, the same trial revealed a fivefold increase in myocardial infarction risk among Vioxx users (0.4% vs. 0.1% for naproxen). Merck attributed this disparity to naproxen's purported cardioprotective effects, a claim later challenged by independent analyses.

The APPROVe Trial and Market Withdrawal

In 2001, Merck initiated the APPROVe trial to assess Vioxx's efficacy in preventing colorectal polyps. By 2004, interim data showed a 2.24-fold increased risk of cardiovascular events (e.g., heart attack, stroke) in patients taking Vioxx for ≥18 months compared to placebo. On September 30, 2004, Merck voluntarily withdrew Vioxx globally, acknowledging the irreversible damage to its safety profile. An estimated 2 million patients were using the drug at the time of withdrawal.

Regulatory and Scientific Fallout

- **FDA Criticisms**: Epidemiologist David Graham estimated that Vioxx caused 88,000–139,000 excess heart attacks in the U.S. alone, 30–40% fatal. The FDA's delayed response highlighted weaknesses in post-marketing surveillance systems.
- **Litigation**: Merck faced over 27,000 lawsuits, settling claims for $4.85 billion by 2007 without admitting liability.
- **Meta-Analysis Revelations**: A 2004 *Lancet* review of 18 trials confirmed elevated cardiovascular risks across all Vioxx doses, disputing claims of safety within 18 months.

Legacy and Regulatory Reforms

The Vioxx scandal catalyzed critical changes:

- **FDA Amendments Act (2007)**: Strengthened post-marketing surveillance, mandating risk evaluation and mitigation strategies (REMS) for high-risk drugs.
- **COX-2 Inhibitor Reassessment**: Celecoxib (Celebrex) remained on the market with black-box warnings, while valdecoxib (Bextra) was withdrawn in 2005.
- **Transparency Demands**: Increased scrutiny of industry-sponsored trials and emphasis on independent replication of safety data.

Vioxx remains a cautionary tale of the ethical and scientific imperative to prioritize patient safety over commercial interests, reshaping global drug regulatory frameworks.

1.2.5 Ineffectiveness as a Problem

Pharmacoepidemiology is often associated with the detection of drug safety issues, but ineffectiveness, when drugs fail to achieve their intended benefits in real-world practice, is an equally significant concern. Drug ineffectiveness can arise from multiple points in the "therapeutic chain," including inappropriate drug selection, suboptimal dosing, poor adherence, medication errors, or intrinsic lack of efficacy. Social and economic factors, such as aggressive marketing or regulatory shortcomings, may also contribute to widespread use of less effective or even ineffective drugs.

A key distinction in pharmacoepidemiology is between efficacy (the benefit demonstrated in randomized controlled trials, RCTs) and effectiveness (the benefit observed in routine clinical practice). While RCTs are considered the gold standard for establishing efficacy due to their randomization and control of confounders, their highly selected populations and controlled environments may not reflect real-world settings. Observational post-marketing studies, which assess effectiveness, can reveal discrepancies—sometimes showing that a drug is less effective or even ineffective in broader, more diverse populations.

While randomized controlled trials (RCTs) establish a drug's efficacy, they often overestimate the benefits seen in real-world practice due to their controlled environments, selective patient populations, and shorter follow-up periods. Observational post-marketing studies, which include broader and more diverse patient groups, frequently reveal that effectiveness in routine care is lower than initially suggested by RCTs[2].

A notable example is the ALLHAT trial, which challenged the prevailing belief, based on earlier RCTs and clinical wisdom, that thiazide diuretics were inferior to first-line antihypertensive agents[3]. ALLHAT, a large post-marketing effectiveness study, demonstrated that thiazides were at least as effective as newer agents like ACE inhibitors and calcium channel blockers, and even slightly better at preventing cardiovascular events in real-world populations. Another example is the CATIE study in elderly patients with dementia, where RCTs had suggested some advantages among atypical antipsychotics, but post-marketing research found these drugs to be equally ineffective and poorly tolerated in this population[4].

These cases illustrate that RCTs can be too optimistic, and post-marketing pharmacoepidemiological studies are crucial for identifying when effectiveness in routine care falls short of expectations, guiding more appropriate and evidence-based medication use.

---

[2] Schneeweiss S. Developments in post-marketing comparative effectiveness research. Clin Pharmacol Ther. 2007;82(2):143-156. doi:10.1038/sj.clpt.6100249

[3] ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). JAMA. 2002;288(23):2981-2997. doi:10.1001/jama.288.23.2981

[4] Schneider LS, Tariot PN, Dagerman KS, et al. Effectiveness of atypical antipsychotic drugs in patients with Alzheimer's disease. N Engl J Med. 2006;355(15):1525-1538. doi:10.1056/NEJMoa061240

Box 1.4: Treatments for COVID-19?

A 2025 systematic review and network meta-analysis (NMA) of 187 randomized trials (166,230 patients) evaluated 40 drug treatments for mild to moderate COVID-19, prioritizing patient-important outcomes and assessing evidence certainty using GRADE. Nirmatrelvir-ritonavir and remdesivir demonstrated **moderate certainty** evidence for reducing hospital admissions, with absolute risk differences of 25 and 21 fewer admissions per 1,000 patients, respectively. Molnupiravir and systemic corticosteroids showed potential benefits (**low certainty**), though their effects were less precise.

For symptom resolution, azithromycin (**moderate certainty**) shortened duration by a mean of 4 days, while systemic corticosteroids, favipiravir, molnupiravir, and umifenovir also reduced symptom duration with **moderate to high certainty**. Notably, nirmatrelvir-ritonavir and remdesivir did not significantly improve symptom resolution. Safety profiles varied: lopinavir-ritonavir increased adverse events leading to discontinuation (41 more per 1,000; **high certainty**), while most other interventions, including hydroxychloroquine and ivermectin, showed no meaningful benefits.

No drug convincingly reduced mortality, mechanical ventilation, or thromboembolic events compared to standard care (**high/moderate certainty**). Prophylactic-dose anticoagulation had no net benefit for venous thromboembolism or bleeding risk (**moderate certainty**). Doxycycline probably reduced hospital stay duration by 1.3 days (**moderate certainty**), whereas lopinavir-ritonavir prolonged stays (**high certainty**).

So, nirmatrelvir-ritonavir and remdesivir are supported by moderate-certainty evidence for preventing hospitalization, while azithromycin and several antivirals improve symptom duration. Most other therapies, including widely debated agents like hydroxychloroquine and ivermectin, lacked efficacy or had unfavorable risk-benefit profiles. **These findings underscore the importance of evidence-based prioritization in COVID-19 management, guided by rigorous NMA and GRADE assessments**.

Reference:

Ibrahim S, Siemieniuk RAC, Oliveros MJ, et al. Drug treatments for mild or moderate covid-19: systematic review and network meta-analysis. *BMJ*. 2025;389:e081165. doi:10.1136/bmj-2024-081165

1.3 Pharmacoepidemiology: Definition, Scope, and Theoretical Foundations

Pharmacoepidemiology is best understood as the **scientific field that applies advanced epidemiological methods to the quantitative and qualitative study of drug use, effectiveness, and safety in large, heterogeneous populations**. Its emergence as a distinct discipline was not simply a response to the limitations of clinical pharmacology and randomized controlled trials, but also a reflection of the growing complexity of modern therapeutics, the diversity of patient populations, and the pressing need for post-marketing evidence that could inform both clinical and regulatory decision-making. The field’s epistemological roots lie in the recognition that while clinical pharmacology elucidates mechanisms of drug action at the molecular and individual level, classical epidemiology investigates the distribution and determinants of diseases, neither alone suffices to capture **the real-world impact of pharmaceuticals on public health**.

The object of pharmacoepidemiology is not merely the drug itself, but the intricate web of **interactions between the medicinal product, the patient, the prescriber, and the health system, all situated within a broader socio-economic context**. This complexity is evident in the discipline's dual orientation: on one hand, it seeks to describe and quantify the patterns of drug utilization, adherence, and persistence across populations and healthcare settings; on the other, it rigorously investigates the causal relationships between drug exposures and a wide array of clinical outcomes, ranging from therapeutic benefits to rare and delayed adverse effects. The discipline is thus inherently transdisciplinary, drawing not only from epidemiology and pharmacology, but also from biostatistics, informatics, and health economics.

A central theoretical challenge in pharmacoepidemiology is the **distinction between efficacy and effectiveness**. **Efficacy**, as established in the controlled environment of randomized clinical trials, refers to the **performance of a drug under idealized conditions-typically in narrowly defined patient populations**, with strict inclusion and exclusion criteria, and close monitoring of adherence and concomitant therapies. **Effectiveness**, by contrast, **is the measure of a drug's impact in routine clinical practice**, where patients are older, sicker, more diverse, and subject to a myriad of real-world influences such as polypharmacy, comorbidities, and variable adherence. **Pharmacoepidemiology is uniquely positioned to evaluate effectiveness**, as it leverages observational data from electronic health records, insurance claims, and disease registries to capture the true heterogeneity of clinical practice.

The field's methodological sophistication is exemplified by its approach to **confounding**, **bias**, and **causal inference**. Unlike clinical trials, where randomization serves as a powerful tool to balance both known and unknown confounders, pharmacoepidemiologic studies must contend with the inherent limitations of non-randomized, observational data. **Confounding by indication**, for example, arises when the underlying reason for prescribing a drug is itself associated with the outcome of interest - a challenge that has historically led to erroneous conclusions, such as the initial belief that beta-blockers increased mortality in heart failure, when in fact they were preferentially prescribed to the sickest patients. To address such complexities, pharmacoepidemiologists have developed and refined advanced analytical techniques, including propensity score matching, instrumental variable analysis, and self-controlled case series designs, each with its own assumptions, strengths, and limitations.

The scope of pharmacoepidemiology extends far beyond the mere detection of adverse drug reactions. While **pharmacovigilance - the science of detecting, assessing, and preventing adverse effects** - remains a core component, the broader field encompasses comparative effectiveness research, drug utilization studies, **post-authorization safety studies** (PASS), and the evaluation of **risk minimization measures** and regulatory interventions. For instance, the withdrawal of rofecoxib from the market in 2004, following pharmacoepidemiologic evidence of increased cardiovascular risk, illustrates the discipline's capacity to influence regulatory policy and safeguard public health. Similarly, pharmacoepidemiologic studies have been helpful in elucidating the real-world **risk-benefit profiles** of new oral anticoagulants compared to warfarin, informing clinical guidelines and prescribing practices.

A further dimension of pharmacoepidemiology's theoretical foundation is its engagement with the concept of **external validity**. The discipline is fundamentally concerned with the **generalizability of findings**: can the results of a clinical trial, conducted in a highly selected population, be extrapolated to the broader, more complex populations encountered in everyday practice? Pharmacoepidemiology addresses this question by generating evidence from large, representative **cohorts**, often encompassing millions of patients, and by employing sophisticated statistical methods to account for confounding, effect modification, and selection bias. This commitment to external validity is particularly crucial in the context of rare diseases, pediatric and geriatric populations, and in the evaluation of long-term safety and effectiveness, where clinical trials are often underpowered or infeasible.

Moreover, pharmacoepidemiology is characterized by its responsiveness to emerging challenges in therapeutics and healthcare delivery. The increasing availability of real-world data, advances in data science and machine learning, and the globalization of pharmaceutical markets have all expanded the horizons of pharmacoepidemiology. Contemporary research now routinely integrates data from multiple sources (e.g., electronic health records, claims databases, disease registries, biobanks, and even patient-reported outcomes) to construct rich, multidimensional profiles of drug exposure and clinical outcomes. This integrative approach enables the study of complex phenomena such as drug-drug interactions, off-label use, and the impact of health policy interventions on prescribing behavior and patient outcomes.

Pharmacoepidemiology's theoretical rigor is further demonstrated by its critical engagement with issues of **measurement validity**, **data quality**, and **reproducibility**. Pharmacoepidemiologists must fight with the limitations of administrative data, such as **misclassification of exposures and outcomes**, incomplete capture of over-the-counter medications, and the challenges of linking disparate data sources. The discipline has responded with the development of standardized definitions, validation studies, and the adoption of transparent reporting standards such as the **RECORD**[5] and **STROBE**[6] **guidelines** for observational research.

Finally, pharmacoepidemiology is distinguished by its translational orientation. The ultimate goal is not merely the accumulation of knowledge, but the translation of evidence into actionable insights that **improve patient care**, **inform regulatory decisions**, and **optimize the allocation of healthcare resources**. This translational imperative is evident in the pharmacoepidemiology engagement with health technology assessment, pharmacoeconomics, and the development of clinical guidelines. For example, pharmacoepidemiologic evidence has shaped national vaccination policies, informed the design of risk management programs for high-risk medications, and guided the prioritization of drugs for inclusion in essential medicines lists.

1.4 Methodological Foundations

The methodological rigor of pharmacoepidemiology is grounded in its adaptation and innovation of **epidemiological study designs**, analytical strategies, and **data sources** to address the

[5] REporting of studies Conducted using Observational Routinely-collected Data (RECORD). https://www.record-statement.org/

[6] Strengthening the Reporting of Observational Studies in Epidemiology (STROBE). https://www.equator-network.org/reporting-guidelines/strobe/

unique complexities of drug research in real-world populations. Unlike randomized controlled trials, which benefit from random allocation and strict protocol adherence, pharmacoepidemiologic investigations must contend with **non-randomized exposures**, **confounding variables**, and the inherent **heterogeneity of clinical practice**. The validity and interpretability of pharmacoepidemiologic findings depend on the careful selection of study design, meticulous control of bias and confounding, and the critical appraisal of data quality.

1.4.1 Observational Study Designs

Pharmacoepidemiology primarily relies on **observational designs**, as randomized assignment to potentially harmful exposures is neither ethical nor feasible post-marketing. The three principal observational designs are **cohort studies, case-control studies, and self-controlled designs**, each with distinct strengths, limitations, and optimal contexts of application (See Lesson 2).

Cohort Studies

Cohort studies follow groups of individuals defined by their exposure status (e.g., users versus non-users of a specific drug) over time to compare the incidence of outcomes. These studies can be prospective, with data collected forward in time, or historical, utilizing existing records. Cohort designs are particularly powerful for estimating **absolute and relative risks**, and for studying multiple **outcomes** of a single **exposure**. For example, the association between hormone replacement therapy and breast cancer risk was elucidated through large-scale cohort studies such as the Nurses' Health Study. However, cohort studies are resource-intensive, may require prolonged follow-up, and are susceptible to **loss to follow-up** and **confounding by indication**.

Case-Control Studies

Case-control studies are especially efficient for investigating rare outcomes or those with long latency periods. In this design, **cases** (individuals with the outcome of interest) are compared to **controls** (without the outcome), and prior exposure histories are ascertained. The classic example is the identification of the association between aspirin use and Reye's syndrome in children. Case-control studies are less costly and faster than cohort studies, but they are vulnerable to **recall bias**, **selection bias**, and difficulties in establishing temporal relationships between exposure and outcome.

Self-Controlled Designs

Self-controlled designs, such as the **self-controlled case series** (SCCS) and **case-crossover** studies, use each individual as their own control, comparing periods of drug exposure and non-exposure within the same person. This approach inherently controls for time-invariant confounders, such as genetics and chronic comorbidities. SCCS designs have been instrumental in vaccine safety studies, while case-crossover designs are suited for exposures with transient effects on acute outcomes, such as the risk of myocardial infarction following non-steroidal anti-inflammatory drug (NSAID) use. Nevertheless, these designs require careful consideration of time-varying confounders and are not suitable for exposures with cumulative or permanent effects.

1.4.2 Data Sources

The expansion of pharmacoepidemiology has been propelled by the increasing availability and sophistication of large-scale health data. The choice of data source profoundly influences the scope, validity, and generalizability of findings.

Electronic Health Records (EHRs)

EHRs provide comprehensive, longitudinal data on diagnoses, prescriptions, laboratory results, and clinical outcomes. They enable the construction of real-world cohorts and facilitate the study of rare events across diverse populations. For instance, the FDA Sentinel Initiative leverages EHRs from multiple health systems to monitor drug safety on a national scale. However, EHRs are subject to data entry errors, missing information, and variability in coding practices.

Administrative Claims Databases

Claims data, such as those from Medicare & Medicaid (US), Infarmed/SPMS (Portugal), or private organizations (HMR/ANF)[7], offer large sample sizes and detailed information on healthcare utilization, including prescriptions, procedures, and hospitalizations. These databases are invaluable for pharmacoepidemiologic research on drug utilization patterns, adherence, and health outcomes. Their limitations include lack of clinical detail (e.g., laboratory values, lifestyle factors), potential misclassification of exposures and outcomes, and absence of over-the-counter drug use.

Disease Registries and Specialized Databases

Disease registries collect granular data on specific conditions (e.g., cancer, congenital anomalies) and are often linked to pharmacy and hospitalization records. The EUROCAT registry[8], for example, has been pivotal in identifying teratogenic risks associated with antiepileptic drugs. Or the Portuguese **Registro Oncologico Nacional (RON)**[9] is useful to follow cancer incidence and treatment outcomes. While registries offer high-quality, disease-specific data, their generalizability may be limited, and linkage to prescription data is not always available.

1.4.3 Analytical Challenges

The interpretation of pharmacoepidemiologic studies axes on the rigorous control of **confounding, bias, and measurement error**.

Confounding by Indication and Channeling Bias

Confounding by indication arises when the underlying reason for prescribing a drug is itself a risk factor for the outcome. For example, patients prescribed anticoagulants may have a higher baseline risk of bleeding due to their underlying conditions. **Channeling bias** occurs when new drugs are preferentially prescribed to patients with specific characteristics, distorting comparisons with established therapies. Advanced statistical methods, such as **propensity**

[7] https://www.hmr.co.com/about-hmr/
[8] European network of population-based registries for the epidemiological surveillance of congenital anomalies. https://eu-rd-platform.jrc.ec.europa.eu/eurocat_en
[9] https://ron.min-saude.pt/pt/biblioteca/publicacoes-ron/

score matching, **stratification**, and **instrumental variable analysis**, are employed to mitigate these biases. Propensity scores estimate the probability of treatment assignment based on observed covariates, allowing for balanced comparison groups.

Information and Misclassification Bias

Accurate measurement of exposures, outcomes, and covariates is critical. Misclassification can be non-differential (biasing results toward the null) or differential (potentially inflating or deflating associations). **Validation studies**, **sensitivity analyses**, and the use of standardized definitions (e.g., validated ICD codes for outcomes) are essential to ensure data quality.

Time-Related Biases

Immortal time bias, **time-window bias**, and **latency bias** can arise if the timing of exposure and outcome assessment is not appropriately aligned. For instance, immortal time bias occurs when periods during which an outcome cannot occur (e.g., before drug initiation) are misclassified as exposed, artificially lowering risk estimates. Proper time alignment and advanced modeling techniques, such as time-dependent Cox regression, are necessary to avoid these pitfalls.

1.4.4 Strengths and Limitations of Observational Pharmacoepidemiology

While observational studies offer unparalleled opportunities to study drug effects in real-world settings, they are inherently limited by the potential for residual confounding and bias. Nevertheless, when rigorously designed and analyzed, they provide critical evidence that complements randomized trials, particularly for rare events, long-term outcomes, and understudied populations. The integration of multiple data sources, triangulation of findings, and transparency in reporting further enhance the reliability and impact of pharmacoepidemiologic research.

1.5 Regulatory and Ethical Context

The practice of pharmacoepidemiology is inseparable from the complex regulatory and ethical frameworks that govern the development, authorization, and post-marketing surveillance of medicinal products. These frameworks are designed to ensure that drugs are not only effective but also safe for use in the general population, and that the rights, dignity, and privacy of individuals whose data are used in pharmacoepidemiologic studies are rigorously protected. Understanding the interplay between regulatory requirements and ethical obligations is essential for any advanced practitioner in this field.

1.5.1 Drug Approval and Post-Marketing Surveillance

The regulatory lifecycle of a medicinal product is characterized by a series of rigorous checkpoints, beginning with preclinical studies and randomized controlled trials (RCTs) and extending into the post-marketing period, where pharmacoepidemiology plays a central role. Regulatory agencies such as the U.S. Food and Drug Administration (FDA), the European Medicines Agency (EMA), and Japan's Pharmaceuticals and Medical Devices Agency (PMDA) have established **harmonized guidelines** (e.g., ICH E2E, GVP modules) that require continuous assessment of a drug's benefit-risk profile throughout its market life.

While pre-marketing trials are designed to establish efficacy and initial safety in selected populations, they are inherently limited by their sample size, duration, and exclusion criteria. As a result, rare adverse drug reactions (ADRs), long-term risks, and effectiveness in special populations often remain undetected until after widespread use. Pharmacoepidemiologic studies, leveraging real-world data from electronic health records, claims databases, and spontaneous reporting systems, are thus indispensable in the detection and quantification of such risks.

Post-marketing surveillance is formalized through mechanisms such as **Risk Evaluation and Mitigation Strategies (REMS) in the United States, Risk Management Plans (RMPs) in the European Union**, and Periodic Safety Update Reports (PSURs) required globally. These instruments mandate the continuous collection, analysis, and reporting of safety data, and may require the implementation of additional risk minimization measures, such as prescriber education, patient monitoring, or restricted distribution programs. The withdrawal of rofecoxib (Vioxx) and the regulatory response to the thalidomide tragedy are paradigmatic examples of how post-marketing pharmacoepidemiology can drive major policy changes and protect public health.

1.5.2 Pharmacovigilance Systems and International Collaboration

Modern pharmacovigilance is built on the foundation of international cooperation and data sharing. The World Health Organization (WHO) Program for International Drug Monitoring, coordinated by the **Uppsala Monitoring Centre**[10] in Sweden, aggregates adverse drug reaction reports from more than 130 countries into the global **VigiBase database**. This global surveillance infrastructure enables the rapid detection of safety signals that may not be apparent within a single country or region.

National pharmacovigilance centers, such as the **FDA's MedWatch**[11] program and the **EMA's EudraVigilance** system, operate in parallel, collecting spontaneous reports from healthcare professionals, patients, and manufacturers. Increasingly, these systems are being augmented by active surveillance networks that utilize automated health data to identify and investigate potential safety concerns in near real-time. The **Sentinel Initiative** in the United States and the **EU's ACCESS** project exemplify this shift toward proactive, data-driven pharmacovigilance.

1.5.3 Ethical Considerations in Pharmacoepidemiologic Research

The ethical conduct of pharmacoepidemiologic research is governed by the principles **of respect for persons, beneficence, and justice**, as articulated in foundational documents such as the **Declaration of Helsinki** and the **Belmont Report**. These principles are operationalized through requirements for informed consent, privacy protection, data security, and independent ethical review.

One of the central ethical challenges in pharmacoepidemiology arises from the use of large-scale, routinely collected health data, where obtaining individual informed consent is often impractical or impossible. Regulatory frameworks such as the **European Union's General Data Protection Regulation** (GDPR)[12] and the **U.S. Health Insurance Portability and Accountability Act** (HIPAA)[13] provide detailed guidance on the conditions under which personal health data may be used for research without explicit consent. These include requirements for data

[10] https://who-umc.org/
[11] https://www.fda.gov/safety/medwatch-fda-safety-information-and-adverse-event-reporting-program
[12] https://gdpr-info.eu/
[13] https://www.cdc.gov/phlp/php/resources/health-insurance-portability-and-accountability-act-of-1996-hipaa.html

anonymization or pseudonymization, minimization of data collection, and robust safeguards against unauthorized access or re-identification.

Institutional review boards (IRBs) or ethics committees play a critical role in overseeing pharmacoepidemiologic studies, ensuring that the potential benefits outweigh the risks, that data are handled responsibly, and that vulnerable populations are protected. Transparency in study design, analysis, and reporting is also essential to maintain public trust and scientific integrity.

1.5.4 Transparency, Data Sharing, and Public Engagement

Recent years have seen increasing demands for transparency and data sharing in pharmacoepidemiology, driven by high-profile drug safety scandals and the recognition that open science accelerates discovery and improves reproducibility. Regulatory agencies now encourage or require the registration of post-marketing studies, public disclosure of protocols, and sharing of de-identified data for independent analysis. The EMA's policy on clinical data publication and the FDA's open access initiatives are leading examples.

Public engagement is also recognized as a critical component of ethical pharmacoepidemiology. Patients and advocacy groups are increasingly involved in the design, conduct, and dissemination of research, ensuring that studies address real-world concerns and that findings are communicated in accessible, actionable ways.

1.5.5 The Evolving Regulatory Landscape

The regulatory and ethical landscape of pharmacoepidemiology is dynamic, continually adapting to new scientific, technological, and societal challenges. The rise of big data analytics, artificial intelligence, and global data networks presents unprecedented opportunities for drug safety monitoring, but also raises new questions about privacy, consent, and the governance of health information. Advanced practitioners must remain vigilant, not only in applying current standards, but also in anticipating and shaping the ethical and regulatory frameworks of the future.