Why real-world evidence studies often overstate drug benefits

Written by The Life Science Feed Editorial Team · Edited by Mara Voss

When a new drug hits the market, the initial excitement often stems from meticulously controlled randomised clinical trials. These trials, while robust, represent a highly selected patient population. The real world, however, is far messier, leading many to seek insights from real-world evidence (RWE) studies.
But the transition from controlled trial to real-world data is fraught with peril. One of the most insidious pitfalls in RWE is confounding by indication, a bias that can make a drug appear more or less effective than it truly is, simply because of who receives it.

KEY TAKEAWAYS

- The Pivot: Real-world evidence, while valuable for generalisability, introduces significant biases not present in randomised controlled trials.
- The Data: Confounding by indication can inflate or deflate a drug's apparent efficacy by associating treatment choice with baseline patient characteristics and prognosis.
- The Action: Clinicians must critically evaluate RWE studies for robust methods to address confounding, particularly when treatment decisions are driven by disease severity or comorbidities.

Randomised controlled trials (RCTs) are the gold standard for evaluating drug efficacy and safety. They achieve this by randomly assigning patients to treatment or placebo, ensuring that, on average, all known and unknown prognostic factors are equally distributed between groups. This randomisation eliminates confounding, allowing for a direct comparison of outcomes attributable solely to the intervention. But RCTs are expensive, time-consuming, and often exclude patients with complex comorbidities or those on multiple medications, limiting their generalisability to everyday clinical practice.
Real-world evidence studies, conversely, leverage data from electronic health records, insurance claims, registries, and other sources reflecting routine clinical care. These studies offer insights into how drugs perform in diverse, unselected patient populations. They can illuminate long-term outcomes, identify rare adverse events, and assess effectiveness in subgroups often excluded from trials. But the very strength of RWE, its reflection of real-world practice, is also its greatest weakness when it comes to causal inference.

The insidious nature of confounding by indication

Confounding by indication occurs when the reason a patient receives a particular treatment is itself a prognostic factor for the outcome being measured. This is not merely a correlation; it is a causal pathway where the indication for treatment directly influences the outcome, independent of the treatment's effect. Consider a scenario where a newer, more expensive

drug is reserved for patients who have failed multiple prior therapies or who present with more severe disease. If this drug then appears to perform worse than an older, cheaper alternative, is it because the drug is truly inferior, or because it was given to sicker patients with a poorer prognosis from the outset?

Conversely, a drug might appear superior if it is preferentially prescribed to healthier patients, perhaps those with fewer comorbidities or an earlier stage of disease, simply because they are deemed more likely to tolerate it or respond well. The observed benefit would then be an artefact of patient selection, not the drug's inherent efficacy. This bias is particularly prevalent in observational studies where treatment assignment is not random but rather a clinical decision based on patient characteristics, disease severity, and physician judgment. For example, in the context of managing chronic conditions, a clinician might prescribe a more aggressive or novel therapy to a patient whose disease is progressing rapidly, while a patient with stable disease continues on standard care. If the rapidly progressing patient then experiences a worse outcome, it would be incorrect to attribute this solely to the new therapy without accounting for their baseline severity.

Spotting the bias in real-world data

Identifying confounding by indication requires a keen eye for baseline differences between treatment groups. The first step is to examine the baseline characteristics of patients receiving different treatments. Are the groups balanced in terms of age, sex, comorbidities, disease severity, prior treatments, and other factors known to influence the outcome? If there are significant imbalances, confounding is likely at play. A well-designed RWE study will present detailed tables of baseline characteristics for each treatment arm, allowing for this critical comparison.

But simply noting imbalances is not enough. Researchers must employ statistical methods to adjust for these differences. Common techniques include propensity score matching or weighting, instrumental variables, and multivariable regression. Propensity scores, for instance, estimate the probability of a patient receiving a particular treatment based on their baseline characteristics. Patients with similar propensity scores but different treatments can then be compared, mimicking randomisation. But these methods can only adjust for measured confounders. Unmeasured confounders, such as a physician's subjective assessment of frailty or a patient's adherence to lifestyle modifications, remain a persistent challenge, as discussed in our coverage of the challenges of real-world data in pharmacovigilance.

Another red flag is when the observed effect size in an RWE study dramatically deviates from what was seen in well-conducted RCTs for the same intervention. While some variation is expected due to differences in patient populations, a stark discrepancy should prompt suspicion of unaddressed confounding. Clinicians should also scrutinise the rationale for treatment assignment. If the study population includes patients for whom the treatment is typically reserved as a last resort, or conversely, those who are unusually healthy, the results should be interpreted with caution. The Oxford Handbook of General Practice offers practical guidance on evaluating evidence in a primary care setting, emphasising the importance of understanding study design.

Mitigation strategies and the limits of observation

Researchers employ several strategies to minimise confounding by indication. Active comparator designs, where a new drug is compared against another active treatment rather than placebo, can help. This is particularly useful when a placebo arm would be unethical or impractical. But even with active comparators, the choice of which active drug to use can still be influenced by patient characteristics. For example, a clinician might choose a newer, more potent anti-hypertensive for a patient with higher baseline blood pressure, making it appear more effective than an older drug given to patients

with milder hypertension.

Newer methods, such as target trial emulation, attempt to design observational studies that closely mimic the structure of an RCT, specifying eligibility criteria, treatment assignment, and follow-up periods as if a trial were being conducted. This structured approach helps to clarify the causal question and identify potential sources of bias. But even with sophisticated statistical adjustments and rigorous design, RWE studies can never fully eliminate confounding by indication, especially from unmeasured factors. This is a fundamental limitation of observational research. For instance, a patient's socioeconomic status, which can influence both treatment access and health outcomes, might not be adequately captured in administrative datasets.

The open-label design inherent in most RWE is an obvious caveat. Patients and clinicians know which treatment is being administered, which can introduce performance bias and detection bias. Patients might report symptoms differently, and clinicians might assess outcomes with a different lens, depending on their expectations of the treatment. This is a stark contrast to the double-blind nature of many RCTs, where neither patient nor investigator knows the treatment assignment. When evaluating RWE, it is essential to consider whether the outcomes measured are objective (e.g., mortality, hospitalisation) or subjective (e.g., patient-reported pain scores, quality of life), as subjective outcomes are more susceptible to these biases. Our previous article on the role of inflammation in CAD highlighted how complex biological pathways can be difficult to disentangle from treatment effects in observational data.

The challenge of confounding by indication underscores the need for a balanced perspective on RWE. While it provides invaluable insights into real-world effectiveness and safety, it should always be interpreted in conjunction with evidence from RCTs. RWE can generate hypotheses, identify patient subgroups that might benefit or be harmed, and inform clinical guidelines, but it rarely provides definitive proof of causality on its own. Clinicians should approach RWE with a healthy dose of skepticism, always asking: could the observed effect be explained by differences in the patients receiving the treatment, rather than the treatment itself? This critical appraisal is vital for making informed decisions in patient care, particularly when considering new therapies or off-label uses. The Harrison's Principles of Internal Medicine provides a comprehensive overview of evidence-based medicine, detailing the hierarchy of evidence and the strengths and weaknesses of various study designs.

The next generation of RWE studies will need to integrate more granular patient-level data, including genetic information and social determinants of health, to better account for unmeasured confounders. Without such advancements, the shadow of confounding by indication will continue to obscure the true picture of drug effectiveness in the real world.

> **CLINICAL IMPLICATIONS**
>
> Clinicians must maintain a critical stance when presented with real-world evidence, especially when it appears to contradict or significantly amplify findings from randomised trials. The allure of generalisability is strong, but the methodological compromises are often substantial. A drug that looks remarkably effective in an observational study might simply be prescribed to patients who were already destined for a better outcome. Conversely, a therapy that appears to underperform could be a victim of being reserved for the sickest patients, those for whom other options have failed. Understanding the baseline characteristics of the treated population is paramount. If the study does not adequately describe and adjust for these differences, its conclusions are, at best, speculative.

> For pharmaceutical companies, the temptation to highlight favourable RWE is understandable, but responsible dissemination requires transparency about potential biases. Regulators, too, are increasingly considering RWE, but they must demand the same rigour in addressing confounding as they do for other aspects of study design. The ultimate goal is to ensure that treatment decisions are based on genuine efficacy, not statistical artefacts.

EDITORIAL & AI STANDARDS

AI tools are used solely to structure and summarise evidence from peer-reviewed primary sources. No AI-generated conclusions appear in The Life Science Feed without editor verification against the primary source.

REFERENCES

1. 1. Euser AM, Zoccali C, Jager KJ, Dekker FW. Cohort studies: prospective versus retrospective. Nephron Clin Pract. 2009;113(3):c214-7. doi:10.1159/000235241
2. 2. Lee J, Chang SM. Confounding by Indication in Studies of Selective Serotonin Reuptake Inhibitors. Psychiatry Investig. 2022;19(11):873-883. doi:10.30773/pi.2022.0209
3. 3. Patel T. Cholangiocarcinoma--controversies and challenges. Nat Rev Gastroenterol Hepatol. 2011;8(4):189-200. doi:10.1038/nrgastro.2011.20
4. 4. Bower KL, Lollar DI, Williams SL, Adkins FC, Luyimbazi DT, Bower CE. Small Bowel Obstruction. Surg Clin North Am. 2018;98(5):945-971. doi:10.1016/j.suc.2018.05.007
5. 5. Sendor R, Stürmer T. Core concepts in pharmacoepidemiology: Confounding by indication and the role of active comparators. Pharmacoepidemiol Drug Saf. 2022;31(3):261-269. doi:10.1002/pds.5407
6. 6. Vaezi MF, Yang YX, Howden CW. Complications of Proton Pump Inhibitor Therapy. Gastroenterology. 2017;153(1):35-48. doi:10.1053/j.gastro.2017.04.047

LICENCE & RIGHTS

MEDICAL DISCLAIMER

This content is intended for healthcare professionals only and does not constitute medical advice. Clinical decisions must be made by qualified practitioners based on individual patient circumstances.

FOLLOW THE LIFE SCIENCE FEED

Instagram @lifesciencefeed **LinkedIn** linkedin.com/company/thelifesciencefeed

thelifesciencefeed.com