

Fairness Is a Safety Outcome: Cross-System Real-World Evaluation of Pharmacoepidemiology, Medication Adherence, Outcomes Research and Population Health

Md. Humayan Kabir

Sep 21, 2025

Abstract

This article treats fairness in pharmacoepidemiology and medication adherence as a question of measurement validity and population health. It explains that dispensing records are only proxies for medication use and that nonadherence may reflect cost, pharmacy access, treatment burden, side effects, health literacy, or beliefs rather than a single patient behavior. These pathways are unevenly distributed across populations, so average adherence estimates can conceal structural differences in treatment opportunity and follow-up. The article recommends separating initiation, implementation, persistence, discontinuation, and reinitiation while using multi-database replication, adherence trajectory modeling, inverse probability weighting, subgroup calibration, missingness analysis, and outcome validation. Income, geography, insurance, multimorbidity, pharmacy access, and continuity of care are highlighted as context variables. Risks include measurement error, unmeasured social determinants, differential loss to follow-up, and algorithmic underperformance in smaller groups. Standardized definitions, reproducible code, traceability, and prospective model validation are emphasized. Overall, fairness is presented as a validity condition for real-world adherence and outcomes research.



The idea of "fairness" in pharmacoeconomics and adherence to medicines is becoming more complex than just the question of "power," "delivery," or "association," as the development of modern medicines increasingly relies on understanding the reasons the result presents and when it is likely to be valid. A mechanistic network that centers on the opportunity for prescription, cost, refill adherence, health literacy, access to pharmacy, comorbidities and outcome measurements is at the core of this change. That network is important because the averages of population might confuse that the access to various services is not the same or that follow-up is incomplete with the real differences in the effectiveness of various drugs. The design variables of mechanism, context and uncertainty are therefore not treated as three separate design steps that are undertaken at three different stages of the development process, but as interconnected design variables.

Although many studies of adherence to medication start with dispensing records, they are merely proxies for actual adherence and an unfilled prescription may indicate cost, access, treatment burden, or beliefs regarding side effects of the medication. Those pathways are not evenly distributed in population, thus fairness issues when considering nonadherence as a patient behaviour. Distinguishing the process of initiation, implementation, persistence, discontinuation, and reinitiation of cross-system analysis, examining the structural conditions associated with each of these stages gives advantages to the analysis. Fairness is thus not just one of the many tables at the end of an outcomes study, but part of the validity of the measure.

The implications for practice are a new discovery process, which focuses on multi-

database replication, adherence trajectory modeling, inverse probability weighting, subgroup calibration, missingness analysis and outcome validation. None of these assays is game changers on its own because each assay only measures a subset of the larger biological/clinical process. When independent measurements are in agreement in terms of the causal narrative across dose, time, model system, and relevant biological context, confidence increases. While this is the case, discordant results are still useful, as inconsistencies provide clues to other processes, assay artifacts, population differences, or exposure limitations that should be explored before more time and money are invested in development.

The key questions for decision making are whether treatment effects and estimates of adherence are reliable across populations and care systems. A definitive answer to this question requires explicit and clear comparison to a "best estimate" of possible alternatives such as off target effects, parallel pathways, differential exposure, or measurement bias. Perturbation experiments are particularly informative as they allow the removal, blocking or restoration of a proposed component to determine if the observed effect is dependent upon this component. A mechanistic confidence is gained when challenge tests are successful, not by repeating the same good assay over and over.

Almost all results in terms of fairness in pharmacoepidemiology and medication adherence are dependent on the context. Variation which is relevant is income, geography, insurance status, multimorbidity, treatment burden, accessibility of pharmacies or continuity of treatment. A good intervention for one set of these conditions may be less effective, more toxic or result in a different biological response, for another set. Stratification is more useful when based on variables related to mechanism or exposure, rather than just descriptions categories, which provide little insight into the reasons for changes in treatment performance.

Risk assessment is also of equal importance due to measurement error in the dispensing data, the lack of measurement of social determinants, differential loss of follow-up, and algorithm underperformance in smaller groups. Early safety work has been more successful when associated with the same mechanistic paradigm as efficacy and unwanted effects can be linked to target biology, carrier behavior, exposure, metabolites or analytical bias. This way, the risk of safety being a late checklist that is not linked to discovery logic is minimised. It also provides more transparent trade-offs on how to achieve greater effectiveness from increased exposure, a wider modulation of pathways, or the more extended persistence in a sensitive tissue.

Blaming patients for nonadherence is an additional hurdle, and another layer is added to the challenge of translational development when it comes to population-health interventions that target barriers. Preclinical success is more informative if experiments are conducted under clinically relevant conditions with relevant concentration ranges, biological barriers, treatment history and sources of biological heterogeneities. It is not so helpful to have a simple progression from a convenient model to a complex model, but rather a series of models designed to answer particular uncertainty questions. Therefore for each step, there must be a clear decision that will be made: to continue, to redesign, to narrow the target population, to change dosing, or to gather another sort of mechanistic information.

Manufacturing and analytical control are both critical as this influences the reliability of the findings across systems, with the infrastructure of the analytics, standardized definitions of adherence, and reproducible code. If the material being tested changes from batch to batch or the inputs into the assay are poorly defined, a biological explanation is not useful. Therefore critical attributes must have specific ranges, be sensitive to measure and link process conditions to biological performance. This discipline is particularly relevant for complex molecules, nanomaterials, biomaterials, and studies involving large amounts of data, which may result in very small changes to exposure, signaling, and/or response to treatment.

In the field of data analysis in the context of fairness in pharmacoepidemiology and medication adherence, a separation between prediction and confirmation is beneficial. While ranking hypotheses, incorporating high dimensions of measurement and revealing uncertainty, computational models, machine learning, causal estimators, and mechanistic simulations are not the same as biological truth. Useful models set up assumptions, measure calibration, mark down over-framed areas and suggest experiments that will provide information to alter a decision. Prospective validation is still required, as there is a tendency for patterns that are successful in the past to "pass the test" in retrospect, but fail the test in the future under new populations, new batches or new biological conditions.

Another key change relates to time as many processes critical to fairness in pharmacoepidemiology and adherence to medication are dynamic after exposure. Even if there is a single endpoint that is favorable, the early response, adaptation, clearance, immune programming, tissue remodeling, adherence, and/or resistance may take different trajectories. Longitudinal measurement determines if the intervention has an effect that is stable, upheaval, or liability. Another advantage of time-resolved data is that it can enhance the causal interpretation of the proposed mechanism, since it will appear

prior to the downstream phenotype, not afterward.

The more efficient the experimental design, the more efficient the study will be when associated with a concrete uncertainty budget. The questions may be ranked by their ability to affect the validity of the estimate of treatment effects and adherence, not just their novelty. Replication then is performed across complementary models to see if the conclusion is dependent on one assay, one laboratory condition, or one population. Such discipline would preclude wasted optimization and informative failures as each failure reduces the design space of what may be happening and the next design iteration would be more successful.

There is also a need for traceability between claims and measurements for regulatory and clinical credibility. In order to be fair in pharmacoepidemiology and medication adherence, a development package gains strength when the raw observations, processing decisions, model assumptions, deviations from the protocol, and attributes of the materials are linked from early experiments to the studies for the translation to practice. Such traceability aids in the detection of true biological variation as opposed to noise, and facilitates a further evaluation of safety claims and/or effectiveness claims at a later stage. It also establishes a clearer foundation to modify conclusions should new data lead to a change in an initial assumption.

Precisely, communication in this field is also in need of refinement. The following terms: targeted, selective, durable, safe, predictive, or personalized are of little value unless accompanied by a measurement, a comparator, a population, and a range of uncertainty. Having negative reports is more useful and avoids having false-positive findings reported as mechanisms. Such precision is a common language to determine if a claim is based on direct measurement or calibrated inference or on an untested assumption for researchers, developers, clinicians and regulators.

Such tight coupling of chemistry/formulation with mechanistic measurement, quantitative modelling, safety, manufacturing control and clinically meaningful context is the strongest direction of fairness in pharmacoepidemiology and medication adherence, rather than any one new technology. The rate of progress increases if failed assumptions are uncovered early, and successful mechanisms are the subject of conscious testing with difficulty. The logic of fairness is used as a validity criterion for real world outcomes research. The most important innovations, as a whole field, are systems that learn from their experiments, but also have causation, reproducibility and applicability in the real world as core parts of their decisions.