

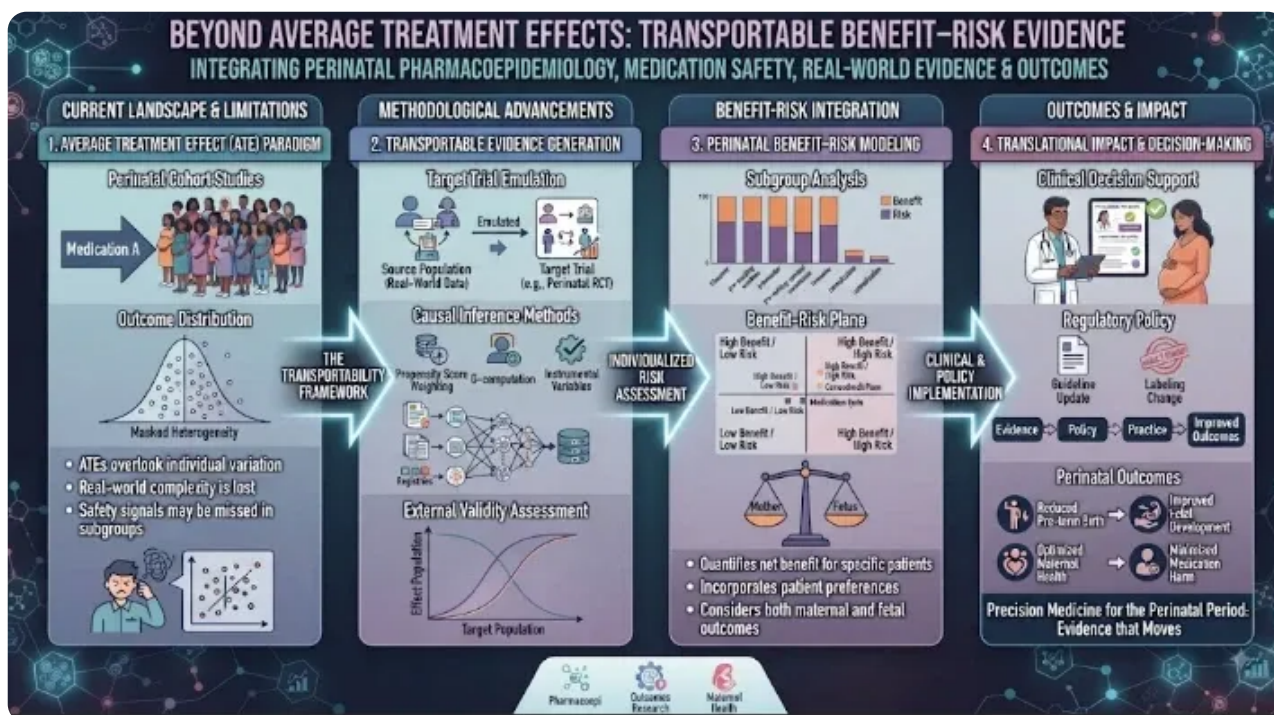
# Beyond Average Treatment Effects: Transportable Benefit–Risk Evidence for Perinatal Pharmacoepidemiology, Medication Safety, Real-World Evidence and Outcomes

Md. Humayan Kabir

Oct 19, 2025

## Abstract

*This article examines how benefit–risk evidence in perinatal pharmacoepidemiology can be transported beyond the original database or healthcare system. It emphasizes gestational timing, treatment indication, maternal disease severity, fetal development, treatment continuation, and competing outcomes as interconnected variables that can be hidden by average estimates. Time-related bias, exposure misclassification, live-birth selection, and confounding by indication are identified as major threats to valid pregnancy research. Target-trial emulation is presented as an organizing framework for defining eligibility, treatment strategies, time zero, follow-up, outcomes, and analysis before estimation begins. Active-comparator designs, time-varying exposure models, negative controls, quantitative bias analysis, and external validation strengthen transportability. The article also stresses context differences in gestational age, comorbidity, dose, concomitant therapy, and access to care. Data linkage quality, coding consistency, calibration, longitudinal measurement, and transparent uncertainty are treated as essential. Overall, the article argues that perinatal benefit–risk evidence should move from a single population average toward explicitly tested, transportable causal estimates.*



The task of transporting benefit–risk evidence in perinatal pharmacoepidemiology is not just one of potency, delivery or association, but a more complex one, in which it is crucial to grasp the rationale behind the result and its reliability at a given time. This is around a mechanistic network of gestational timing, indication, severity of maternal disease, fetal development, treatment continuation and competing clinical outcomes. This network is important because of the clinically relevant variations that exist within each stage of pregnancy and patient population that can be hidden by an average estimate from one database or health care system. A more powerful development strategy then regards mechanism, context and uncertainty as interrelated design variables instead of three design steps that happen in different phases.

Pharmacoepidemiology of the perinatal period is associated with unique issues of timing, as gestational age determines exposure opportunity, fetal development, therapy decisions and ascertainment of the outcome of the therapy at the same time. Prior to 2025, methodological research identified time-related bias, exposure misclassification, live-birth selection and confounding by indication as common threats in pregnancy research. Target-trial emulation is a convenient organizing principle that establishes the eligibility criteria, therapeutic interventions, time at which treatment is expected to begin, when to follow up, what outcomes to measure, and the analysis to be conducted before estimation begins. Transportability introduces a second question: Will the benefit–risk ratio obtained be credible for populations with different prescribing patterns, baseline risks, and healthcare systems?

The result is a new approach to the discovery workflow that includes target-trial emulation, active-comparator designs, time-varying exposure models, negative controls, quantitative bias analysis and external validation. None of these tools can be used for a specific decision in isolation – each only measures a portion of a larger biological or clinical process. Confidence increases with consistent conclusions drawn from independent measurements across the dose, time, model system and relevant biological context. It is important to note that discordant results do not necessarily mean that they are of no value, as the lack of consistency can sometimes be due to hidden mechanisms, assay artefacts, population differences or exposure constraints, and should be explored before further development of the assay becomes too costly.

When examining decisions, it will be important to determine if the association observed is strong enough and can be generalized to other patients than those in the original study. In order to answer this question, you have to make a clear comparison between the intended mechanism and possible alternative mechanisms such as off-target activity, alternative pathways, differential exposure, or measurement bias. Perturbation experiments are particularly useful since the removal, blocking or restoration of a proposed component can determine whether the observed effect is dependent on that component. The mechanistic confidence is gained when a challenge test can be successfully performed, and not when the test is repeated in one set of favorable assay conditions.

In the field of transportable benefit–risk evidence in perinatal pharmacoepidemiology, the meaning of nearly all results depends on the context. Variation that is relevant to the study includes access to care, dose, disease severity, concomitant therapy, and gestational age and comorbidity. The efficacy of an intervention can be as low or more toxic under some set of these conditions and lead to a different biological response under another set. When the stratification is based on variables associated with mechanism or exposure, then it is more useful because descriptive categories alone don't account for differences in treatment performance.

Risk assessment must also be considered as a potential source of confounding, of exposure misclassification, of live-birth bias, of time-related bias and of incomplete linkage of maternal and infant records. A correlation with the same mechanistic basis as efficacy studies, so unwanted effects can be linked to target biology, carrier behaviour, exposure, metabolites or analytical bias, is best. This way, there is a lower risk that safety will be a late checklist that's unrelated to discovery logic. It also allows for more transparent trade-offs for more effective efficacy when either more exposure is needed, more pathway modulation is needed, or a longer persistence in sensitive tissue is required.

There is an additional layer of translational development as there is greater counseling clarity regarding maternal benefit, fetal risk, untreated-disease risk, and treatment alternatives. Preclinical success is more informative when experiments are performed with clinically relevant concentrations, time course, barriers, history of treatment and heterogeneity sources. It is not especially helpful to obtain a "complex" model from a "simple" model, but it is helpful to obtain a series of models that will provide answers to specific uncertainty questions. A clear decision to either continue, redesign, restrict the target population, change the dose, or gather a new kind of mechanistic data is required for each step.

As well as manufacturing, analytical control is also important as the quality of the analytics is determined by the engineering of the data and not the manufacturing of the product, particularly in terms of linkage quality, consistency of coding, and windows of exposure. If the material tested between batches is different, or there is poor definition of assay inputs, then a biological explanation is of less value. Therefore, critical attributes must be set to a range, measured with a high level of sensitivity and documented to link process conditions to biological activity. It is particularly relevant to complex molecules, nanoparticles, biomaterials and data intensive studies where variation in exposure or signaling, or apparent treatment response, can be important.

As with other areas of evidence in perinatal pharmacoepidemiology, the separation of prediction from confirmation is a useful feature of data analysis in transportable benefit-risk evidence. While computational models, machine learning, causal estimators and mechanistic simulations can rank hypotheses, combine high-dimensional measurements, and quantify uncertainties, the output of models is not a form of truth. Useful models state assumptions, quantify calibration, flag regions with weak support, and guide the next experiment to information that will cause a change in a decision. Prospective validation is still crucial as retrospective fit can be beneficial for a pattern that doesn't take hold in new populations, new batches, or new biotic conditions.

Another change is that of time, as many processes in important to transportable benefit-risk evidence in perinatal pharmacoepidemiology are time evolving after initial exposure. In the case of a single endpoint that appears desirable, the response, adaptation, clearance, immune program, remodeling, adherence or resistance can take various paths early on. Longitudinal measurement determines if an event creates a state or a perturbation or a liability in the future. Time-resolved data also offer great advantages for causal interpretation, because it becomes clear whether the proposed mechanism has occurred before the downstream phenotype or only after the downstream phenotype is set.

When all the studies are coupled with a specific uncertainty budget, the experimental design becomes more efficient. The relevance of questions to the observed association, not just their novelty, can be gauged by their ability to influence whether that association is strong enough and applicable to other cohorts to guide decisions about medication treatment. Replication across complementary models is then used to determine the robustness of the conclusion, in the sense that it does not rely on one assay, one laboratory condition or one population. This discipline avoids over-optimization and enables a negative result to be informative, as each unsuccessful prediction reduces the candidates in the “mechanism space” and thus makes the next design cycle more effective.

Traceability between claims and measurements is also crucial for regulatory and clinical credibility. In perinatal pharmacoepidemiology, the transportable benefit-risk evidence strength is enhanced by linking the raw observations, processing decisions, model assumptions, protocol deviations and material attributes from initial experiments to translational studies. The traceability helps to differentiate between true biological variation and technical noise and makes it easier to review safety or effectiveness claims later. It also provides a more transparent framework for modification of the conclusions should there be a change in an assumption in the light of new data.

There's also a need for more precision in communicating around this field. Unless accompanied by a measure, comparator, population, and uncertainty range, terms like targeted, selective, durable, safe, predictive, or personalized are meaningless. Getting results back is better because it means those that turn out negatively are more useful and exploratory results aren't presented as mechanisms. This precision provides a common ground for making judgments among researchers, developers, clinicians, and regulators whether a claim is based on actual measurement, calibrated inference, or an assumption that has not been tested.

A tighter loop of transportable benefit-risk evidence in perinatal pharmacoepidemiology, therefore, is not to be found in a single new technology, but in the sequential wrapping of chemistry or formulation, mechanistic measurement, quantitative modeling, safety, manufacturing control, and finally, clinically meaningful context. It's easier to make progress if assumptions that were wrong are identified early and mechanisms that were successful are deliberately tested. This logic has led to a paradigm change in benefit-risk evaluation, from a population average to structured transportability. The most useful innovation in the field is a development system that is able to learn from each experiment, while maintaining a focus on causal interpretation, reproducibility, and relevance to the real world throughout the development process.