NOTEWORTHY

THERE IS NO AVERAGE PATIENT

6.05.2026

I am collecting material for an article questioning EBM (evidence based medicine) while coming across an interesting preprint by Zach Shahn “Trust me, I’m a doctor”.

> Suppose that outcomes under usual care, e.g., collected from hospital health records, outperform the outcomes in both arms of a randomized experiment conducted in the same population. A textbook example concerning lung cancer patients comes from Hernan and Robins [2024], see also Sarvet and Stensrud [2025]. Then, Deaton and Cartwright’s argument that one should trust their physician over a trial is validated. In this case, a next step is to find the criteria that physicians are using to make personalized decisions.

He continues to examine study settings in which a randomized trial is nested within an observational cohort, so that outcomes are observed under treatment, control, and usual care while I am following up here his reference to Sarveed & Stensrud Unfortunately the abstract of this paper is poor – it should have explained the two definitions of “harm” in personalized medicine. So I try it on my own.

Counterfactual harm – a patient is harmed if they received a treatment whose outcome is worse than what *would have happened* under the alternative – requires knowing unobservable potential outcomes / principal strata. The interventionist harm (the authors’ preference): a patient is harmed if their *expected* outcome under the assigned treatment is worse than under the alternative, conditional on their measured features which requires only experimentally identifiable quantities. The counterfactual approach is practically problematic because principal strata are metaphysical objects that can never be verified, require non-experimental data and partial identification. The interventionist approach is transparent, observable, and doesn’t coerce commitment to unverifiable metaphysics.

The paper does not explain how to do this in practice – I think this could be just a well-designed RCT with pre-specified effect modification analysis. The workflow would be; pre-register → stratified randomization → interaction-term analysis or causal forest → decision rule

by argmax of expected outcome. You never need to ask “what would have happened to this patient under the other treatment” — you only ask “what does the evidence say about patients like this one.”