COMMENT ON CUROVIC ET AL.

# Optimization of Albuminuria-Lowering Treatment in Diabetes by Crossover Rotation to Four Different Drug Classes: A Randomized Crossover Trial. Diabetes Care 2023;46:593–601

*Leila R. Zelnick*

The recent work by Curovic et al. (1) highlights the need to individualize novel therapies in heterogeneous conditions such as diabetic kidney disease. However, the trial design used by the authors has serious methodological flaws that threaten the validity of the comparisons made between alternative treatments. In their article, the authors examine the albuminuria-lowering performance of four drug classes in the first four periods of a randomized crossover trial. Each individual is then re-exposed to the drug class with the greatest response (the "winner") in a confirmatory period; the trial's primary outcome is the difference between the response in this confirmatory period and the mean response of the three "losers" from the first four periods of the trial.

The authors are to be commended for implementing a confirmatory fifth period to obtain an unbiased estimate of response to the winner drug class and for showing that the initially identified winner response remains reproducible, if attenuated, in the confirmatory period. However, comparing this response against the mean response of the losers, obtained from the very periods that defined them as losers, gives a falsely low estimate of their performance ("random low bias") and thus will tend to overestimate the difference in response between the winner and the remaining drug classes.

This design illustrates the perils of selective inference, wherein selection (which therapies lose for an individual) and estimation (the difference in responses to winner and loser drug classes for an individual) cannot be performed using the same data without adverse statistical consequences. A simple simulation of the trial shows the pitfalls of this approach (https://github.com/leilazelnick/Diabetes_Care_letter/). Simulating five normally distributed responses centered at zero under the null hypothesis of no response to any drug, and following the selection and estimation approach used by Curovic et al. (1), rejection of the (true) null hypothesis of no difference between therapeutic response to the winner and the remaining therapies occurs an eye-popping 67% of the time instead of the putative type I error rate of 5%. Said another way, a researcher using this study design would conclude there was a significantly better drug class for patients 67% of the time when there was no true difference among classes, exposing patients unnecessarily to the risks of drugs that do not help the patient better than any other.

This subtle but fundamental analytical issue should be handled differently in similar future trials. Assuming the focus of the trial is finding an individual's best therapy (versus identifying population-level differences in response), one simple albeit expensive solution is to decouple the selection and estimation steps by adding additional trial periods in which the performance of the losers could similarly be confirmed (i.e., crossover periods 6–8). A comparison of the response to the winner class from period 5 and to the loser classes from periods 6–8 would give an unbiased estimate of their difference.

As always in trial design, the need to identify efficacious treatments must be balanced against the risks of exposing patients to therapies that do not work.

## Reference

1. Curovic VR, Jongs N, Kroonen MYAM, et al. Optimization of albuminuria-lowering treatment in diabetes by crossover rotation to four different drug classes: a randomized crossover trial. Diabetes Care 2023;46:593–601

*Kidney Research Institute and Division of Nephrology, University of Washington, Seattle, WA*

*Corresponding author: Leila R. Zelnick, lzelnick@uw.edu*