Target trial emulation of physical activity and cardiovascular disease risk: What is impact of the exposure assessment method?

Matthew Ahmadi, Borja del Pozo Cruz, Raaj Biswas, Armando Teixeira-Pinto, Dot Dumuid, Marla Beauchamp, Joanne McVeigh, Mark Hamer, Emmanuel Stamatakis

Background: Target trial emulation (TTE) strengthens causal inference in observational research, but it is unclear how vulnerable TTE designs are to error when the emulated intervention is measured imprecisely. In physical activity epidemiology, correcting for confounding via TTE without addressing substantial self-report measurement error (e.g., questionnaires capturing only partial behaviour with low precision) raises concerns about dominant Type II error. High-resolution wearable devices capture most movement with far less measurement error, yet no study has examined how measurement method influences causal inference in this field.

Objectives: To compare estimated effects of an emulated physical activity intervention on incident CVD using self-report versus wearable-based exposure assessment.

Methods: We emulated a randomised controlled trial within UK Biobank by identifying physically inactive adults (<150 moderate-to-vigorous physical activity [MVPA] min/week) with repeat self-report and wearable assessments. At follow-up, participants were categorised as intervention (adopted ≥150 MVPA min/week) or control (remained inactive). Propensity score matching balanced demographic, lifestyle, and health factors. Cumulative CVD risk was evaluated using survival curves, hazard ratios, risk ratios, Fine-Gray subdistribution models, and Poisson regression.

Results: The wearable sample included 490 participants (245/arm; mean follow-up 4.4 years) and the self-report sample 11,302 participants (5,651/arm; mean 6.3 years). With wearables, guideline-adherent participants showed substantially lower cumulative CVD risk (8.0% vs 17.0%; hazard ratio 0.59 [0.36–0.98]; relative risk 0.45 [0.28–0.72]). In contrast, self-report assessments showed near-identical risks (21.6% vs 21.2%; hazard ratio 0.98 [0.89–1.08]; relative risk 0.92 [0.84–1.00]). Sensitivity analyses supported these findings.

Conclusion: Self-reported activity in TTE studies may obscure true intervention effects through non-differential misclassification, substantially increasing Type II error. Wearable-based exposure assessment appears essential for valid causal inference in TTE evaluations of physical activity and CVD risk. Future TTE studies should prioritise objective measures to better inform public health guidelines.