

Quantifying representativeness in randomized clinical trials using machine learning fairness metrics

Miao Qi¹, Owen Cahan², Morgan A. Foreman³, Daniel M. Gruen², Amar K. Das³, and Kristin P. Bennett²

¹Rensselaer Polytechnic Institute, Department of Computer Science, Troy, New York, USA, ²Rensselaer Polytechnic Institute, Department of Mathematical Sciences, Troy, New York, USA and ³IBM Research, Cambridge, Massachusetts, USA

Corresponding Author: Kristin P. Bennett, Rensselaer Polytechnic Institute, 110 8th Street, CII 3129, Troy, NY 12180, USA (bennek@rpi.edu, (518) 276-4400)

Keywords: population representativeness, machine learning, randomized clinical trials, patient subgroup, health equity

Abbreviations

Randomized Clinical Trial (RCT), Machine Learning (ML), National Health and Nutrition Examination Survey (NHANES), Action to Control Cardiovascular Risk in Diabetes (ACCORD), Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), Systolic Blood Pressure Intervention Trial (SPRINT)