Kolloquium „Statistische Methoden in der empirischen Forschung“

Wann: 23. Januar 2024, 17:00 – 18:30 Uhr

Wo: Campus Charité Mitte | Paul-Ehrlich-Hörsaal | Virchowweg 4, 10117 Berlin

Achtung: abweichender Hörsaal

Online-Übertragung: der Link wird auf der Website zur Verfügung gestellt

Heiner Stuke (Robert Koch-Institut)

Interpretable Machine Learning to Simulate Public Health Interventions for Mental Health in Adolescents

Adolescence is a critical time period for the development of mental health, holding a lasting influence into adulthood. Converging evidence indicates an adolescent mental health crisis in Western societies that has developed and exacerbated over the past decade. This cross-national phenomenon requires urgent and targeted public health interventions. Yet, the causal interplay of personal, socioeconomic, and environmental drivers for mental health is still unclear.

To support this debate, we implemented different machine learning (ML) pipelines to examine risk and resilience factors in early adolescence (age 11 - 14) that predict the future development of mental health across adolescence (until age 17). The ML pipelines were applied to predict the development of mental health of adolescents in two comprehensive population-based cohort studies from the United Kingdom and Germany and evaluated with repeated 10-fold cross validation. Based on these predictive models, we performed permutation-based feature importance analyses. Moreover, we carried out a series of counterfactual simulations to test potential public health interventions by changing the values of modifiable risk factors and estimating the hypothetical effect on prospective mental health assessments.

Gradient boosting regression models with nested hyperparameter optimization achieved the best predictive performance (mean squared error and Pearson correlation) and were hence used for further model interpretation. Feature importance analyses indicated a strong impact of pre-existing mental health and weaker impacts of sex, socioeconomic and family factors, physical health, and lifestyle. Counterfactual intervention simulations of increased physical exercise and improving income for disadvantaged families showed a positive impact on mental health development during adolescence. Simulated reductions in screen time only led to an improved development of mental health, when additional related changes such as improved peer relationships were assumed.

The validity of the simulated intervention effects depends on an approximately correct specification of the causal model, including the modeling of potential confounding variables and mediators, which will be explored in more detail during the presentation. We argue that robust (i.e., consistent over different cohorts, and modeling assumptions) ML-based intervention simulations can be used to assess public health interventions that cannot be evaluated with randomized trials.