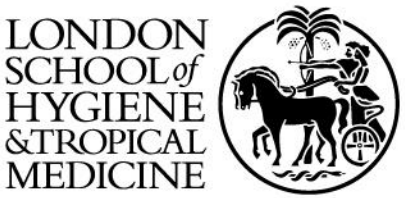
High-dimensional propensity scores in UK Electronic Health Records

John Tazare



Acknowledgements

- **Supervisory Team:** Elizabeth Williamson, Ian Douglas, Stephen Evans & Liam Smeeth

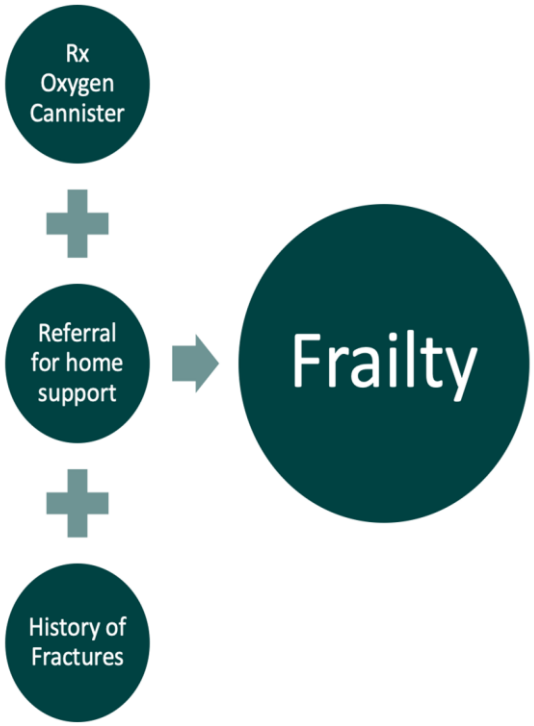
- **Funding:** Medical Research Council Studentship

High-dimensional propensity scores

High-dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data

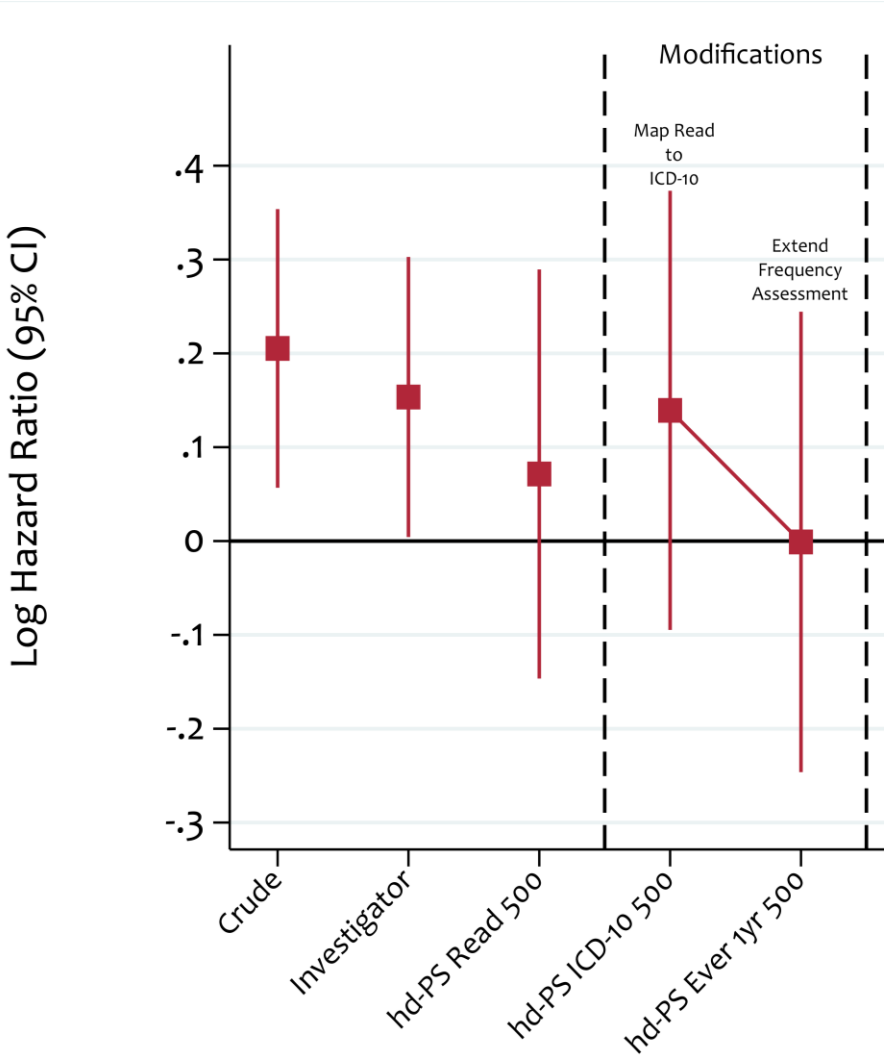
Sebastian Schneeweiss, Jeremy A. Rassen, Robert J. Glynn, Jerry Avorn, Helen Mogun, and M. Alan Brookhart

- Successful mitigation of confounding bias can rely on capturing hard to measure concepts, e.g. surrounding frailty and disease severity
- Method for generating and empirically ranking proxy covariates for confounder adjustment
- Aims to minimise residual confounding by adjusting for a large number of these HDPS covariate/proxies (e.g. 200 – 500)



Application of HDPS to UK EHRs

- Proposed modifications to capture disparities between claims data and UK EHRs
- Applied modified-HDPS to CPRD study of clopidogrel users investigating a possible drug interaction with proton pump inhibitors (PPI).
- Analyses estimated the hazard ratio for myocardial infarction between PPI users and non-users using a propensity score weighted cox model.
- Our modifications highlight the potential for HDPS to improve confounder adjustment in EHRs



Diagnostics

- Transparency of reporting and thorough use of diagnostics and sensitivity analyses is a key issue

Diagnostic Tools

- The HDPS is often described as a 'black-box' approach

- Diagnostic tools can give important insights and enhance knowledge of the factors driving treatment decisions https://github.com/johntaz/HDPS-Diagnostics

Sensitivity Analyses

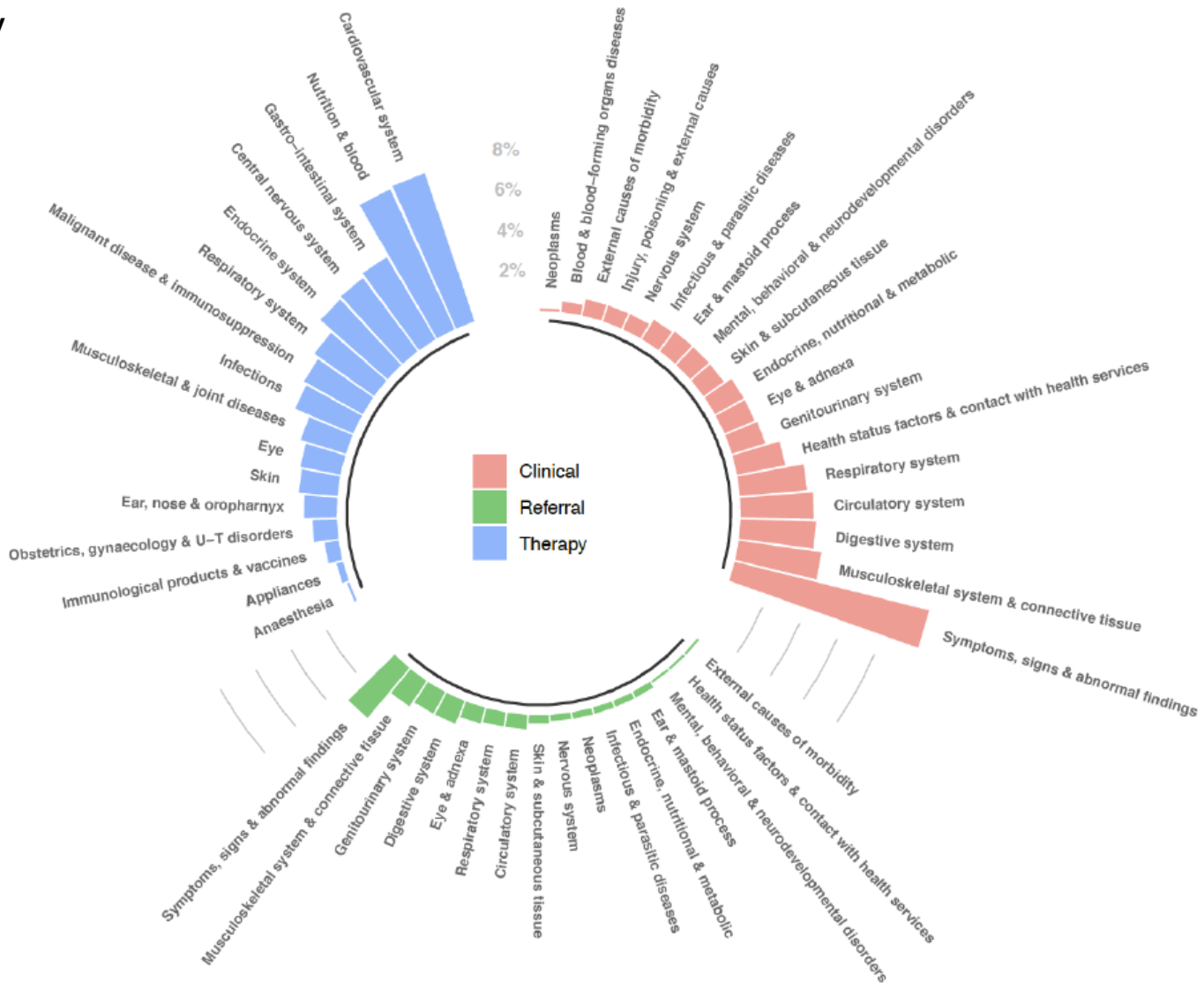
- Existing studies highlight the lack of robustness to investigator decisions e.g. a key decision surrounds the no. of covariates selected

- Reporting of investigator decisions and sensitivity analyses is inconsistent -> undermines transparency

- Reporting guidance developed

Diagnostic tools

- High-level summary of covariates included:
  - BNF Chapters
  - ICD-10 Chapters

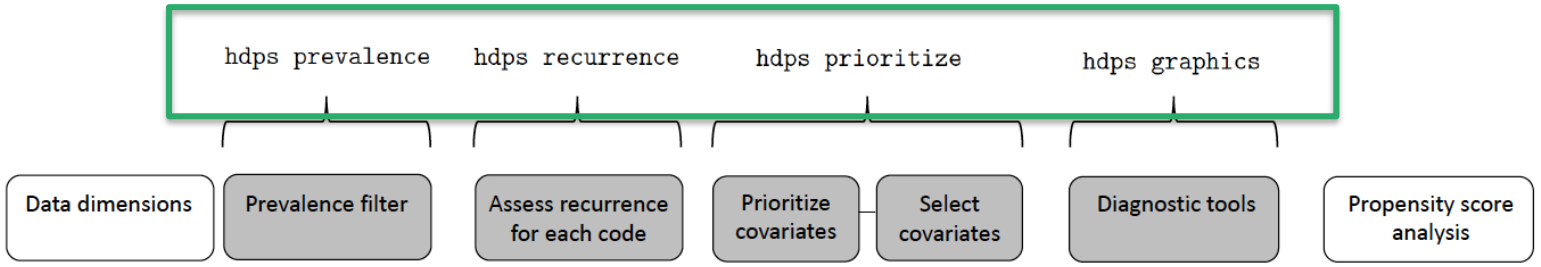


Sensitivity analyses – varying no. of covariates

- Impact of incrementally adjusting for top 750 covariates
- Stabilization of treatment effect vs reporting a range of estimates

Implementing HDPS in Stata

- Developed 5 commands and help documentation for applying the standard and modified-HDPS in Stata
  - All key tuning parameters in HDPS procedure can be varied, e.g. method of code prioritization and the number of covariates chosen
- Investigate properties of selected covariates using graphical diagnostic tools
- Command is hosted on GitHub (github.com/johntaz/hdps)
- Simulated data and example analysis scripts are available on GitHub (github.com/johntaz/HDPS-Stata-Demo)



Thank you

John Tazare
john.tazare1@lshtm.ac.uk
@JohnTStats

References

- J. Tazare, L. Smeeth, S.J.W. Evans, E.J. Williamson, and I.J. Douglas, “Implementing high-dimensional propensity score principles to improve confounder adjustment in UK electronic health records”, 2020.
  Published in Pharmacoepidemiology & Drug Safety.