# Death Predictions from Social Determinants: Holistically and precisely with Explainable AI

Jiani Yan[1,2,3]

[1]Leverhulme Centre for Demographic Science, University of Oxford
[2]Wolfson College, University of Oxford
[3]Max Planck Institute for Demographic Research

October 27, 2023

***Keywords***: *Social Determinants of Health, Machine Learning, Explainable AI, Predictive Precision*

## Extended Abstract

As Link and Phelan (1995) argue in the Fundamental Causes of Diseases theory, health outcomes are never insulated from social determinants. Evidence has been accruing through the past years. For example, both stages of the Whitehall study identified a positive association between health and socioeconomic conditions (Marmot et al., 1991). Similarly, the Commission on Social Determinants of Health (2008) highlighted that health status consistently correlates with socioeconomic circumstances throughout one's life. In England and Wales, professional workers exhibit a risk of serious illness and premature death at least 50% lower compared to that of unskilled manual workers. Death per se as the termination of life and the inexorable demise of health, also intertwines with diverse social factors. However, in both sociological and biological research on health, the majority of the work which attempts to understand death at the individual level concentrates on a singular perspective. It is surprisingly understudied in a holistic approach that integrates information from different disciplines.

Such scant evidence also leaves the question of how predictable death is unanswered, as accurate predictions rely on the quality of the information which trains the algorithm, not to mention explorations from solely the sociological perspective. Therefore, underpinned by theories such as intersectionality (Crenshaw, 1989), we embrace a more comprehensive perspective

**For Correspondence:** Jiani Yan, Email: jiani.yan@sociology.ox.ac.uk. **Code Availability**: Codes can be found at GitHub. Please see the readme.md file within that repository for a data availability statement. **Acknowledgements**: Funding is gratefully acknowledged from the ESRC Grand Union DTP.

in our investigation of the social determinants of health. One of the seminal attempts is from Puterman et al. (2020), where they have integrated information from 57 variables to access the hazard levels of those factors. The study itself does not adequately address model performance and inferences are built upon precarious models. Instead of traditional survival analyses or regressions, we implement a predictive framework to accurately investigate the existing frontier of the predictive capability of various social determinants (holistically) in estimating death predictivity (predictive precisely), with advanced machine learning algorithms and Explainable AI methods to surface variable importance (explainability). We also contribute to several methodological discussions in quantifying seed variability and exploring predictive asymptotics.

Specifically, there are three main contributions from this work, including predicting, holistic comparisons and quantifying precision. Specifically, using advanced data linkage in the field of ageing including the Health and Retirement Study in the US (HRS), the Survey of Health, Ageing and Retirement in Europe (SHARE) and the English Longitudinal Study of Ageing in the UK (ELSA), we estimate death predictability in a holistic manner across those three different regions. We extract information on seven health-related domains including Adulthood Psychological Diathesis, Adulthood Socioeconomic, Childhood Adversity, Adulthood Adverse Experience, Health Behaviours, Social Connections and Demography. Risk factors are selected on the grounds of sociological theories such as Fetal Conditions (Barker, 1995), Fundamental Causes (Link & Phelan, 1995), Life Course Theory (Ben-Shlomo & Kuh, 2002) and Biopsychosocial model (Engel, 1977), contending the strength of social determinants of health. Prediction accuracy is powered and guaranteed by the Super Learner algorithm, a state-of-the-art ensemble learning algorithm to calculate the predictability of death. We then 'unravel the black box' of prediction via the use of Shapley values, allowing us to better 'surface' variables at the single risk factor level. We also develop a new method which allows us to further understand risk factors' relative importance at the domain level. Both of those calculations are performed at the single dataset level to facilitate cross-sectional comparisons and at the combined dataset level to further explore the predictive ceiling of death.

Lastly, we also investigate predictive precision: more methodological and technical aspects of prediction to quantify precision. This is achieved in two stages. Firstly, we develop the foundations of tools which allow us to consider the 'asymptotics of predictive power', estimating the impact of varying predictor numbers and training dataset size on the predictive power. Outcomes will be derived from the mean of five random subsetting procedures on the predictors and training set. We then introduce a new heuristic – the median of a large number of trials of seeds – which allows us to account for the randomness inherent in manifold works. The workflow is summarised in Figure 1. Our conclusion is that when approached as a predictive endeavour, the accuracy of death prediction is notably high. With the augmentation of datasets, there is a plausible potential to eradicate reducible error, thereby approaching the 'predictive ceiling'.

**Result 1:** With a comprehensive predictive framework, the models demonstrate notable performance across most datasets, where the ROC-AUC scores in the HRS, SHARE, ELSA and Combined datasets are 0.822, 0.796, 0.908 and 0.947 respectively. Considering that all datasets are unbalanced, we also evaluate the model performance with the PR-AUC score, where the results are 0.689 (0.293),0.519 (0.202),0.246 (0.057) and 0.883 (0.225) for those datasets. Numbers in the parenthesis represent the in-sample prevalence of death for each dataset. They serve as the inference base (random guess) for the PR-AUC score.

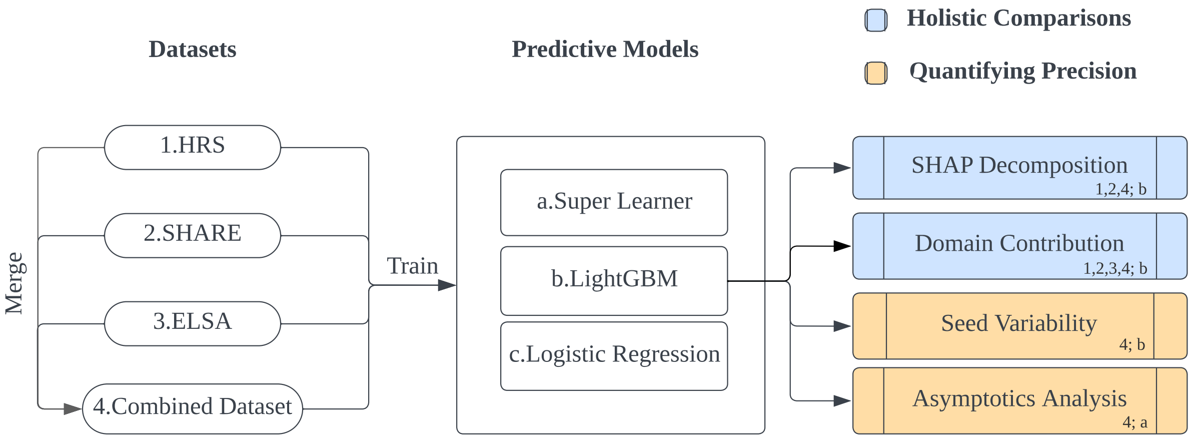

Figure 1: Research Design

**Result 2:** Competing with all risk factors, the top important risk factors from SHAP cover most of the domains, with differences across different datasets. In common, Age and Gender are the most important predictors in all datasets. For HRS, outstanding factors concentrate on socioeconomic and health behaviour domains, including History of Smoking, Low/no Vigorous Activity, Income, Trait Anxiety, Wealth, and Lower Occupational Status (importance rank from high to low). For SHARE, only two other top factors include Wealth, Current Smoker stand out. For the combined data, Low/No Vigorous Activity, Current Smoker and Low/No Vigorous Activity are the other top. Results are illustrated in Figure 2. Giving the sparse nature of the dependent variable in the ELSA dataset (5% prevalence), we refrain from interpreting the SHAP values within the UK context, as inconsistent model performance may yield deceptive insights.

**Result 3:** Similar domain rank patterns emerge across different datasets, as shown in Figure 3. Predominantly, the demography (including Age and Gender) and socioeconomic domain play pivotal roles across all datasets. Intriguingly, in the US context, the psychological domain supersedes others in the rest domains, whereas in the SHARE dataset, behaviours assumes greater significance.

**Result 4:** Persistent and consistent evidence of the effect of dataset size on model performance is observed across all evaluation metrics in Figure 4. Models trained with larger training datasets outperform other models with smaller datasets when predicting the same out-of-sample set, facilitating the understanding of reducible errors in machine learning predictions.

**Result 5:** As shown in Figure 5, prediction accuracy varies greatly with different choices of seed in the train-test set split. Using LightGBM and the combined dataset, the predictive accuracy for each evaluating metric spreads over about three standard deviations away from the mean value to each side (i.e. metric span is circa six to seven times s.d). For example, the mean, max, min and std of PR-AUC score across 10,000 different seeds in train-test splitting is 0.625, 0.597, 0.651 and 0.007 respectively. Albeit normal-like distributions, we call for attention to the range as the choice of seed is completely subjectively random and the resulting single outcome will be an arbitrary point within the span and thus affecting the prediction precision substantially, rather than the accuracy.

# References

Barker, D. J. (1995). Fetal origins of coronary heart disease. *Bmj*, *311*(6998), 171–174.

Ben-Shlomo, Y., & Kuh, D. (2002). A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology*, *31*(2), 285–293. https://doi.org/10.1093/ije/31.2.285

Commission on Social Determinants of Health. (2008). Closing the gap in a generation : Health equity through action on the social determinants of health : Final report of the commission on social determinants of health, 247.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *U. Chi. Legal F.*, *1989*, 139.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science (New York, N.Y.)*, *196*(4286), 129–136. https://doi.org/10.1126/science.847460

Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of health and social behavior*, 80–94.

Marmot, M. G., Smith, G. D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., & Feeney, A. (1991). Health inequalities among british civil servants: The whitehall II study. *Lancet (London, England)*, *337*(8754), 1387–1393. https://doi.org/10.1016/0140-6736(91)93068-k

Puterman, E., Weiss, J., Hives, B. A., Gemmill, A., Karasek, D., Mendes, W. B., & Rehkopf, D. H. (2020). Predicting mortality from 57 economic, behavioral, social, and psychological factors [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *117*(28), 16273–16282. https://doi.org/10.1073/pnas.1918455117

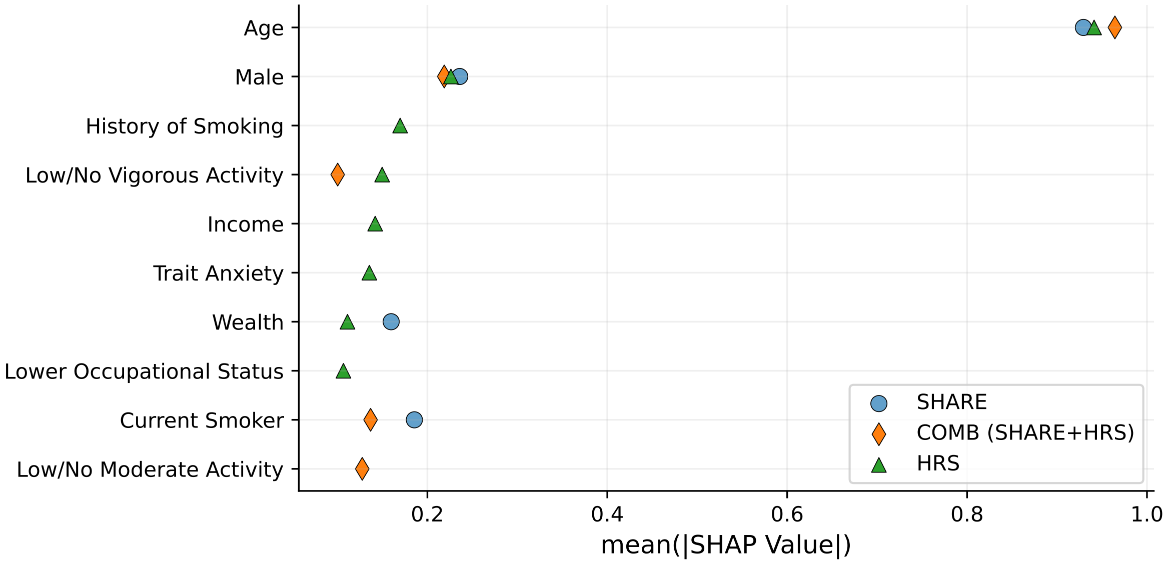

Figure 2: Top Important Risk Factors of Three Datasets from SHAP

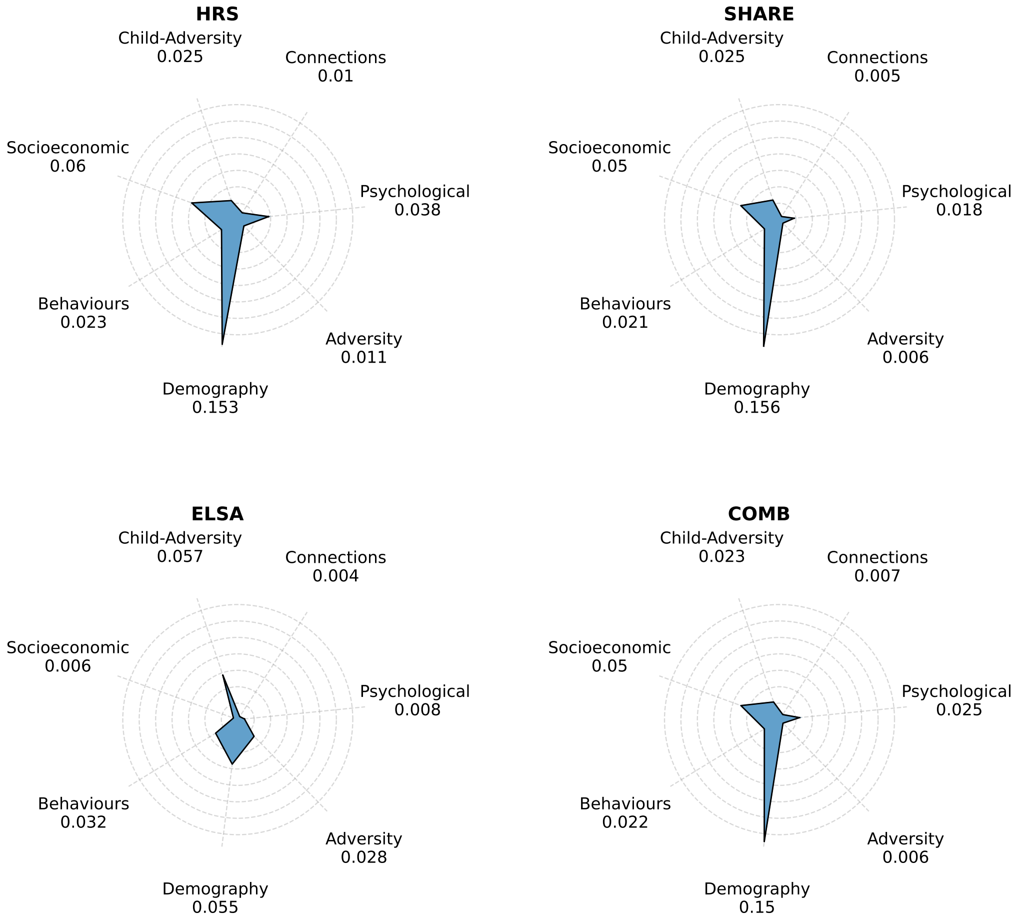

Figure 3: Domain-level Prediction Contribution of Three Datasets

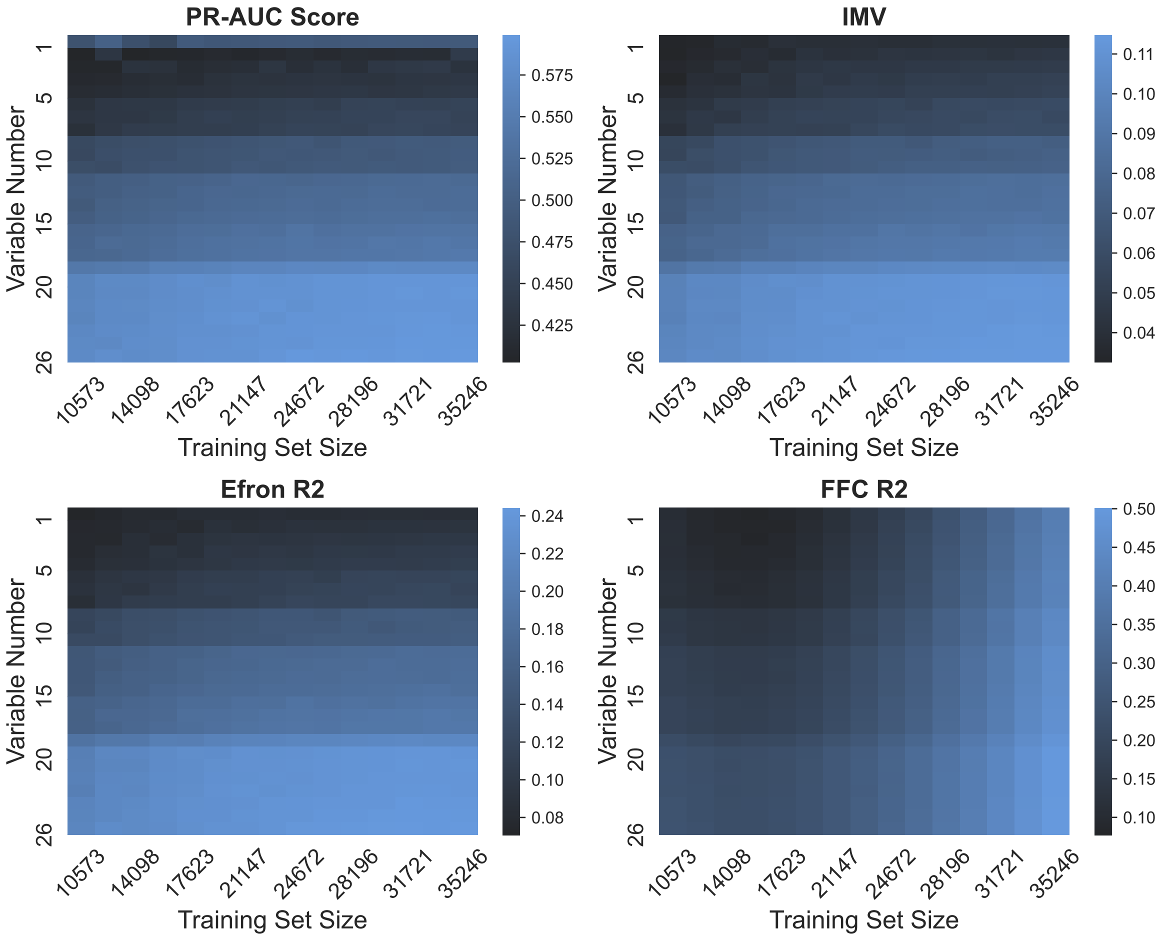

Figure 4: Predictive Asymptotics with Super Learner on varying predictor and training set size

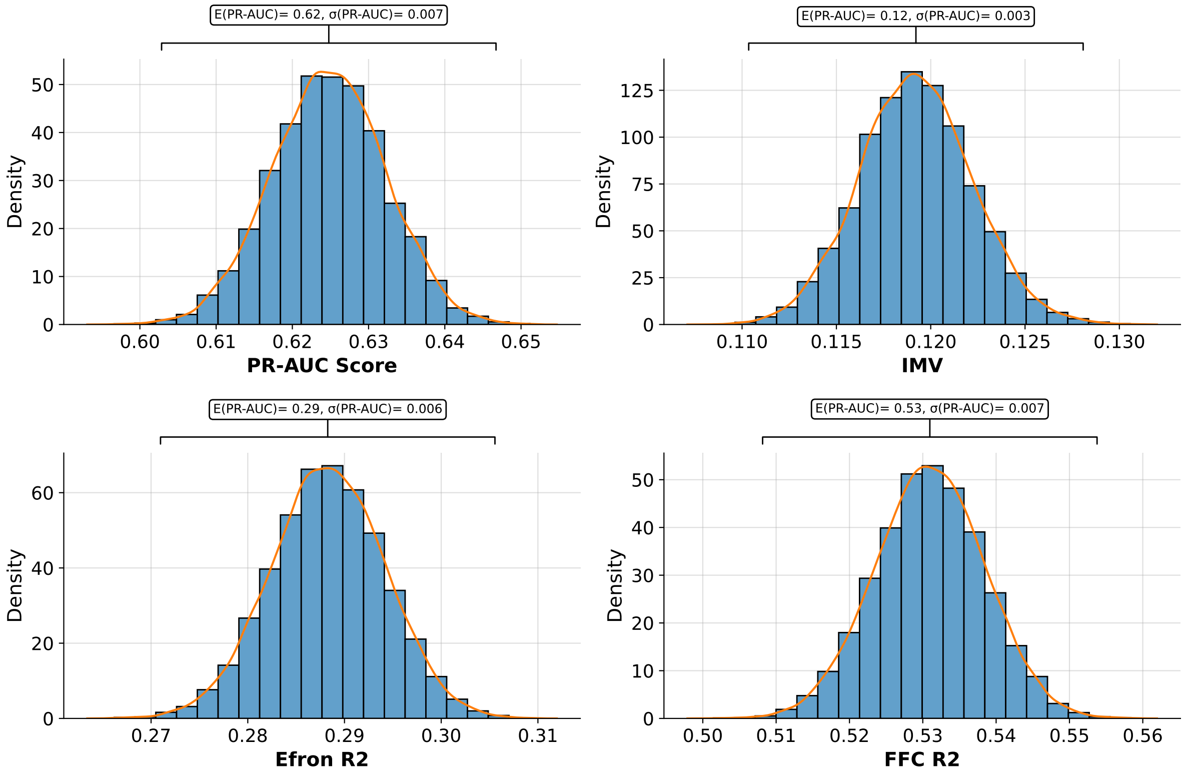

Figure 5: Model performance distribution over 10000 seeds in train-test split