

International Journal of Population Data Science

Journal Website: www.ijpds.org



Feasibility research into using multiple administrative data sources and predictive modelling to produce disability status estimates for England

Matthew Minifie¹, Charlotte Standeven¹, Jesse Ransley¹, and Sarah Wood¹

¹National Statistician's Office, Newport, United Kingdom

Official estimates of population by disability status for England and Wales are produced as part of the census once every ten years, but there is a pressing user need for more frequent, robust disability statistics. This presentation outlines ongoing novel research that responds to that challenge by examining the feasibility of producing population-level estimates for disability status in England using a predictive model applied to a suite of linked administrative data. We have linked nine administrative data sources from different government departments, including benefits, education, employment and health datasets. The de-identified data are linked via a population spine (the England-based residents on the Office for National Statistics' 2021 Statistical Population Dataset with a valid disability status from Census 2021), with a sample size of

48.9 million. Machine learning algorithms were then used to train a model on a subsample of data. The binary outcome variable is Census 2021 disability status (non-disabled or disabled). Predictor variables include age, sex, geography and various benefit, education, employment and health variables from the administrative data. The performance of the model was tested and evaluated on a different subsample of the data to that on which the model was trained. Various evaluation metrics were employed to assess the performance of the model, along with predicted and observed disability prevalence rates by socio-demographic factors to further assess the coherence between the predictive model and Census 2021. Additionally, the predictive model was applied to earlier years for limited analysis of a short time series.

