

International Journal of Population Data Science

Journal Website: www.ijpds.org



Linking population-level education and hospital data in England: who are we missing in ECHILD?

Vincent Nguyen¹, Farzan Ramzan¹, Milagros Ruiz Nishiki¹, Ruth Blackburn¹, Ruth Gilbert¹, and Katie Harron¹

¹UCL Great Ormond Street Institute of Child Health, London, United Kingdom

Introduction

Linkage of National Pupil Database (NPD) to national hospital data has the potential to provide almost universal population coverage of children and young people in England. However, quantifying who is missed is a key requirement for producing unbiased population estimates.

Methods

Deterministic linkage was conducted between longitudinal data from the NPD and secondary healthcare services via the NHS Personal Demographic Service. National Data OptOuts were applied to the linked data. We assessed the proportion of individuals who linked and used modified Poisson regression with robust standard errors to evaluate the risk of non-linkage when accounting for a variety of social determinants.

Results

Of 25,286,870 individuals born between 1984–2020 included in NPD, 84% (21,287,176/25,286,870) linked to a secondary care dataset. Linkage rates improved from 66% (507,820/765,642) for those born in 1985/86 to 98% (306,165/311,121) in 2019/20 and ranged from 89% (908,998/1,026,575) for those with Black ethnicity to 95% (13,644,182/14,401,835) in White groups. Females (IRR: 1.08, 95%CI: 1.08–1.09) and those living in more deprived areas or in London were more likely to be unlinked. Individuals with recorded Special Educational Needs provision (IRR: 0.81, 95%CI: 0.81–0.82) or Free School Meals (IRR: 0.67, 95%CI: 0.67–0.68) were less likely to be unlinked.

Conclusions

Whilst linkage rates have improved over time, there remain important differences in sex, ethnicity, deprivation and geography, between those who are and are not included in the linked dataset. Given high overall linkage rates reported by data providers, these differences are primarily driven by the application of Opt-Outs.

