

International Journal of Population Data Science

Journal Website: www.ijpds.org



What is needed to make administrative data ‘research ready’ for causal inference? A case study of the Education and Child Health Insights from Linked Data (ECHILD) database of 15 million children in England

Ruth Gilbert¹, Bianca De Stavola¹, Vincent Nguyen¹, and Katie Harron¹

¹University College London, London, United Kingdom

Introduction

Many jurisdictions have linked education and health data throughout childhood to address policy questions relating to equity and effectiveness of services. We use the example of special educational needs and disability (SEND) policy to show that causal evidence of effectiveness, based on administrative data, can be limited and potentially misleading. We discuss what needs to change.

and generalisable SEND practices. These lessons are likely relevant to other practitioner-based interventions.

Methods

We derived health phenotypes to group children likely to benefit from SEND provision and used the target trial emulation framework to define causal contrasts, reduce design biases and guide analytic methods. We estimated the effectiveness of SEND provision at age 5/6 on outcomes recorded in health and education data for two phenotypes (cerebral palsy and cleft lip/palate).

Results

SEND provision did not improve rates of unplanned hospital admissions or educational attainment but did reduce rates of unauthorised school absences for both phenotypes. These results are likely biased due to unmeasured confounding and imprecise measures of SEND provision and relevant outcomes. We failed to find instrumental variables to account for unmeasured confounding. Despite the detailed information on health and school characteristics in ECHILD, there is insufficient information on which children are selected for which SEND interventions or comparators.

Conclusions

Causal evaluations of SEND provision using observational analyses of ECHILD need linkage to additional measures of need, interventions and outcomes. Adding similar measures to linked health-education data across jurisdictions, and wider use of experimental designs, would help to quantify effective

