International Journal of Population Data Science

Journal Website: www.ijpds.org

Linking the National Down Syndrome Cytogenetic Register to Hospital Episode Statistics: Methods and challenges

Doidge, J^{1*}, Morris, J², and Gilbert, R¹

¹ADRC-E / UCL

²Queen Mary University of London

Background

Linking national disease and congenital anomaly registers to other routinely collected data could provide multiple benefits. It could allow researchers to examine long-term outcomes for people with registered conditions and to explore synergies between services that are provided in different sectors, such as health and education. Linkage could also allow for the quality of diagnostic codes in electronic health records to be assessed using the national registers coverage and, conversely, for coverage of the national registers to be assessed using diagnostic codes.

Objectives

To demonstrate the feasibility and value of creating a linked data resource to support research on health, social and educational services and outcomes for children with Down's Syndrome.

Methods

We are currently linking the National Down's Syndrome Cytogenetic Register (NDSCR), which includes all births with Down's syndrome from 1989 to 2014 in England and Wales, to Hospital Episode Statistics (HES) for England and ONS Mortality data, and to the National Pupil Database (NPD). This requires a complex sequence of new linkages: (1) linkage of mothers to babies in HES, (2) linkage of mother-baby pairs to the NDSCR to identify a population cohort, (3) linkage of the cohort to the Personal Demographic Service to update address histories, and (4) linkage of updated cohort to NPD.

Findings

Establishment of this linked data resource has met with significant challenges. The main administrative challenge has been incompatible accreditation of information governance across government departments. Technical challenges have included overcoming issues with the internal linkage of contributing datasets, requiring additional linkage steps to be added, and the propagation of error and uncertainty across multiple linkages, which involves complex data structures that require novel methods for analysis.

Conclusions

We hope this project will demonstrate that while the challenges to intersectoral data linkage are great, they are surmountable and the benefits outweigh the costs.

*Corresponding Author:

Email Address: j.doidge@ucl.ac.uk (J Doidge)