LINKAGE: Factors in selecting a data linkage approach

Jones, K[1][*] and Ford, D[1]

[1]Swansea University

Background

The great benefits of linking health datasets for research in the public interest have long been demonstrated. More recently, we are seeing an increase in the availability of wider administrative data, such as employment, education and housing, to add new opportunities for population data science. However, there are challenges to be overcome in selecting a data linkage approach.

Main aim

We set out to examine various data linkage approaches, and to formulate some high level questions to inform decision-making.

Methods

We used published literature to review various data linkage methods in theory and in practical settings. The study was commissioned by the UK Government Statistical Service and a key focus was privacy and confidentiality in data linkage.

Results

The questions we formulated are based on: Legislative position; Information systems; Nature of datasets; Knowledge-base; Aims and purposes; Ground truth; and Environment.

Conclusion

There are many factors influencing the selection of a data linkage approach. While not exhaustive, our set of questions covers some of the major ones. The findings of the study are being taken forward by UK Government Statistical Service and government departments to inform decision-making on options for data linkage research and the greater availability of their datasets.

*Corresponding Author:
Email Address: k.h.jones@swansea.ac.uk (K Jones)