

International Journal of Population Data Science

Journal Website: www.ijpds.org



Investigating the quality of NHS number linkage of COVID 19 vaccines data to Census 2021.

Hannah O'Dair¹, Sarah Cummins¹, and Rachel Huck¹

¹Office for National Statistics

Objectives

Due to the availability of data and the speed of delivery required, NHS number (Health ID) was used during the COVID-19 pandemic to join different health datasets to census 2021. This research sought to assess the quality of linkage done by NHS number, between COVID-19 vaccinations data and census 2021, by comparing to the quality that could be achieved with deterministic linkage using Personal Identifiable Information (PII). Specifically, we wanted to assure against concern that unlinked records may have been misclassified as unvaccinated due to inaccurate recording of NHS number at time of vaccination.

datasets can increase the link rate and plays a vital role in estimating linkage error.

Approach

We used deterministic (rules-based) linkage to identify extra links that were not made on NHS number. We also investigated how many links were incongruent between NHS number linkage and PII deterministic linkage.

Results

We were able to link 98.5% of census records to vaccines data, with a precision of 98.94% and a recall of 99.71%. This identified 1,833,626 additional links made using PII linkage. Furthermore, analysis of the linked data was re-run with the additional links.

Conclusion

We were able to assure that the NHS number linkage used to produce COVID-19 vaccinations statistics, was valid. Whilst 1,833,626 more links were made, the findings resulting from the linked data were unchanged.

Implications

Whilst this research demonstrates the utility of NHS number-based linkage, acquisition and use of PII on health

