

International Journal of Population Data Science

Journal Website: www.ijpds.org



Strategies for ascertaining timing of address changes in administrative data

Joseph Lam¹, Mario Cortina-Borja¹, Peter Christen^{2,3}, Richard Thomas⁴, Robert Aldridge⁵, Ruth Blackburn¹, Ruth Gilbert¹, Andy Boyd⁴, and Katie Harron¹

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

²University of Edinburgh, Edinburgh, United Kingdom

³Australian National University, Canberra, Australia

⁴University of Bristol, Bristol, United Kingdom

⁵Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA

There is growing interest in incorporating the timing of place-based exposures into administrative health data to examine the impacts of the home environment on population health. We compared the accuracy of three methods for estimating the timing of changes in Lower Super Output Areas (LSOA; geographical output areas with around 400 and 1,200 households) in deidentified hospital records against self-reported address in cohort studies.

In hospital records, addresses are updated after patients move to a new address when they next used health service. We compared three approaches to ascertain periods of time where individuals were staying at a particular address: 1) inferring the end date of the current address as the start date of the next reported address minus 1 day (N-1 method); 2) using median date between current and next start date as the end date for current address, and update the start date for the next address (Median method); 3) generating address end dates as a function of beta distribution between current and next start dates (random method), assuming that most people update their addresses not too long after they moved.

We compared LSOA derived from addresses recorded in the UK Longitudinal Linkage Collaboration cohorts (Cohort, $n = 40,963$) with linked hospital records (NHSD, $n = 40,102$) from Jan 1989 to Apr 2023. Cross-sectionally, 39,216 (95.7%) Cohort members had at least 1 matching LSOA reported in both Cohort and NHSD data. Of the matching LSOA, 47% of the NHSD recorded the same LSOAs dated two years before or after Cohort recorded dates. All three methods correctly represented around 78% of each individual's LSOA across the period, with negligible differences across methods.

A number of approaches are available for ascertaining timing of address changes in administrative data, with similar levels of accuracy. Researchers should consider assumptions and implications of each method, and if possible, formally test and justify their approach for processing dates recorded in administrative records.

