**Household effects in Big Data research: privacy-preserving methods for grouping individuals into household groups, and its use in evaluating a Welsh Government fuel poverty scheme**

Karen Tingay, Administrative Data Research Centre – Wales, Swansea University

Sian Morrison-Rees, Administrative Data Research Centre – Wales, Swansea University

Sarah Lowe, Knowledge and Analytical Services, Welsh Government

Charles Musselwhite, Centre for Innovative Ageing, Swansea University

The effect of the social environment on physical health and wellbeing has long been an area of study. Extrapolating this would potentially reduce confounding effects in health-related research. While household-level data can be obtained through surveys, doing so using routine administrative linked datasets, such as those held in the Secure Anonymised Information Linkage (SAIL) databank, held at Swansea University, allows for a broader population to be measured using a wider variety of metrics. However, routine datasets often have issues with data quality. Additionally, grouping individuals into households, especially when linking to other data sources, increases the risk of re-identification. A method for securely and accurately modelling household composition on the population as a whole, using routinely collected administrative data, are of benefit to public health research.

The presentation will describe the methods used in creating the households, issues with using routine administrative data to link anonymised individuals in this way, and how the methods were used the evaluate a Welsh Government intervention to improve the health of those living in fuel poverty: the Warm Homes Nest Scheme. Research findings demonstrated a positive effect on respiratory health, and suggested an improvement on cardiovascular conditions, although the latter figures were too small to be significant and will be revisited at a later stage. These findings were used to inform the Warm Home Nest successor scheme. More broadly, these methods have the potential to inform more effectively focused government schemes across a wide spectrum of health and social areas.

Key Words: Linked population data; anonymisation; evaluation; households; fuel poverty

Contact Details for Karen Tingay:

Telephone: 01792 606435

Email: k.s.tingay@swansea.ac.uk

Contact Details for Author 2:

Telephone: 01792 606847

Email: S.Morrison-Rees@Swansea.ac.uk

Contact Details for Sarah Lowe:

Telephone: 03000 625 229

Email: [Sarah.Lowe@wales.gsi.gov.uk](mailto:Sarah.Lowe@wales.gsi.gov.uk)

Contact Details for Charles Musselwhite:

Telephone: 01792 518696

Email: c.b.a.musselwhite@swansea.ac.uk