

International Journal of Population Data Science

Journal Website: www.ijpds.org



Explaining Data Linkage Quality to the Public: Workshops on Bias, Exclusion and Trust

Georgina Ferguson-Glover¹, Giulia Mantovani², Leah Maizey³, Katie Harron⁴, and Joseph Lam⁵

¹Expert Public Contributor, NA, United Kingdom

²NHS England, London, United Kingdom

³Office for National Statistics, Newport, United Kingdom

⁴UCL Great Ormond Street Institute of Child Health, London, United Kingdom

⁵UCL Great Ormond Street Institute of Child Health, London, United Kingdom; NHS England, London, United Kingdom

Linked administrative and health datasets hold potential to improve health service planning and delivery. However, strategies to communicate to the public about linkage quality, uncertainty, and bias remains under-developed. Poor linkage quality impacts research and operational use differently. We developed and piloted workshops to communicate what is record linkage, linkage quality and bias to the public, in context of using linked data for health service delivery. We ran three online workshops with UK-wide coverage. Sessions introduced linkage concepts, developed a glossary, and discussed the principles for linking and using personal data. We shared how linkage quality and bias are assessed and assured, including consequences when linkage goes wrong. Materials foregrounded plain language and visual explanations. Questionnaire and qualitative feedback were summarised to evaluate

changes in participants' knowledge and attitudes about trust and use of linked data for service delivery. Participants began with uneven knowledge of linkage, misunderstandings about population coverage, and uncertainty about disclosure risks. Plain-language explanations increased confidence and willingness to consent to using health data. Acceptability hinged on quality assurance, independent oversight, and transparency; persistent concerns focused on ongoing linkage bias, exclusion, and potential commercial use. Communication about linkage quality must be a core responsibility, not an afterthought: processes should embed transparency, reporting, and avenues for change. Increased transparency about methods, uncertainties, and safeguards offers a concrete way for the public to see how their data are handled in practice, and is central to building and sustaining trust in the use of linked data.

