

Big Data and Health: An Australian Perspective Which Highlights the Importance of Data Linkage to Support Health Research at a National Level

Authors : James Semmens, James Boyd, Anna Ferrante, Katrina Spillsbury, Sean Randall, Adrian Brown

Abstract : 'Big data' is a relatively new concept that describes data so large and complex that it exceeds the storage or computing capacity of most systems to perform timely and accurate analyses. Health services generate large amounts of data from a wide variety of sources such as administrative records, electronic health records, health insurance claims, and even smart phone health applications. Health data is viewed in Australia and internationally as highly sensitive. Strict ethical requirements must be met for the use of health data to support health research. These requirements differ markedly from those imposed on data use from industry or other government sectors and may have the impact of reducing the capacity of health data to be incorporated into the real time demands of the Big Data environment. This 'big data revolution' is increasingly supported by national governments, who have invested significant funds into initiatives designed to develop and capitalize on big data and methods for data integration using record linkage. The benefits to health following research using linked administrative data are recognised internationally and by the Australian Government through the National Collaborative Research Infrastructure Strategy Roadmap, which outlined a multi-million dollar investment strategy to develop national record linkage capabilities. This led to the establishment of the Population Health Research Network (PHRN) to coordinate and champion this initiative. The purpose of the PHRN was to establish record linkage units in all Australian states, to support the implementation of secure data delivery and remote access laboratories for researchers, and to develop the Centre for Data Linkage for the linkage of national and cross-jurisdictional data. The Centre for Data Linkage has been established within Curtin University in Western Australia; it provides essential record linkage infrastructure necessary for large-scale, cross-jurisdictional linkage of health related data in Australia and uses a best practice 'separation principle' to support data privacy and security. Privacy preserving record linkage technology is also being developed to link records without the use of names to overcome important legal and privacy constraint. This paper will present the findings of the first 'Proof of Concept' project selected to demonstrate the effectiveness of increased record linkage capacity in supporting nationally significant health research. This project explored how cross-jurisdictional linkage can inform the nature and extent of cross-border hospital use and hospital-related deaths. The technical challenges associated with national record linkage, and the extent of cross-border population movements, were explored as part of this pioneering research project. Access to person-level data linked across jurisdictions identified geographical hot spots of cross border hospital use and hospital-related deaths in Australia. This has implications for planning of health service delivery and for longitudinal follow-up studies, particularly those involving mobile populations.

Keywords : data integration, data linkage, health planning, health services research

Conference Title : ICBDM 2017 : International Conference on Big Data Management

Conference Location : Rome, Italy

Conference Dates : May 04-05, 2017