

## High quality linked data for stroke obtained using non-government clinical registry and routinely collected hospital and death data

Kilkenny, M<sup>1</sup>, Kim, J<sup>1</sup>, Andrew, N<sup>2</sup>, Sundararajan, V<sup>3</sup>, Thrift, A<sup>1</sup>, Katzenellenbogen, J<sup>4</sup>, Flack, F<sup>5</sup>, Gattellari, M<sup>6</sup>, Boyd, J<sup>7</sup>, Anderson, P<sup>8</sup>, Lannin, N<sup>9</sup>, Siphthorp, M<sup>10</sup>, Chen, Y<sup>11</sup>, and Johnston, T<sup>12</sup>

<sup>1</sup>Stroke and Ageing Research, Department of Medicine, School of Clinical Sciences at Monash Health, Monash University

<sup>2</sup>Peninsula Clinical School, Central Clinical School, Monash University

<sup>3</sup>University of Melbourne

<sup>4</sup>Cardiovascular Disease Research Group, School of Population and Global Health, The University of Western Australia

<sup>5</sup>Population Health Research Network

<sup>6</sup>Ingham Institute for Applied Medical Research

<sup>7</sup>Curtin University

<sup>8</sup>Data Linkage Unit, Australian Institute of Health and Welfare

<sup>9</sup>Faculty of Health Sciences, La Trobe University

<sup>10</sup>Centre for Victorian Data Linkage, Department of Health and Human Services

<sup>11</sup>Victorian Agency for Health Information

<sup>12</sup>Statistical Services Branch, Queensland Department of Health

## Introduction

Recent advances in data linkage infrastructure in Australia mean that data can be linked based on various identifiers across datasets. In a first for Australia, we tested the feasibility of linking data between a clinical quality disease registry with Australian and state government health data across multiple jurisdictions.

## Objectives and Approach

To determine whether high quality linked data for stroke can be obtained using a non-government managed registry (Australian Stroke Clinical Registry, AuSCR), national death registry data (Australian government), and hospital admission and emergency presentation data (state governments) to assess the accuracy of consistent variables across the different datasets. We used a cohort design with probabilistic data linkage to merge patient-level records. Descriptive statistics presented for matching concordance and Cohen's kappa for concordance across demographic variables. The sensitivity and specificity of in-hospital deaths collected in the AuSCR was assessed against national death registrations.

## Results

There were 16,214 registrants in the study cohort. Their identifiers in the AuSCR from 2009–2013 were linked with death, emergency department and hospital discharge data from April 2004 to December 2016. In total, 99% of the AuSCR regis-

trants were linked to one or more datasets; 98% were linked with emergency presentation (80%) and/or admission (95%) data. Linkage to national death registrations identified 4,183 death; 1440 of these were identified as in-hospital deaths in both data sets demonstrating that in-hospital death classification in AuSCR had a 98.7% sensitivity and 99.6% specificity. Concordance between common demographic variables was excellent (kappa 0.84 for aboriginal status and kappa 0.99 for sex).

## Conclusion/Implications

The majority of AuSCR registrants were accurately linked to the Australian and state government datasets. Linkage quality was excellent and there was high concordance between common variables. The ability to reliably merge the datasets assures future comprehensive analyses of stroke care, ongoing health care resource utilisation and patient outcomes.

