

## Supplemental Data

**Title:** Population-based stroke surveillance using Big Data: epidemiological trends in admissions and mortality in Australia.

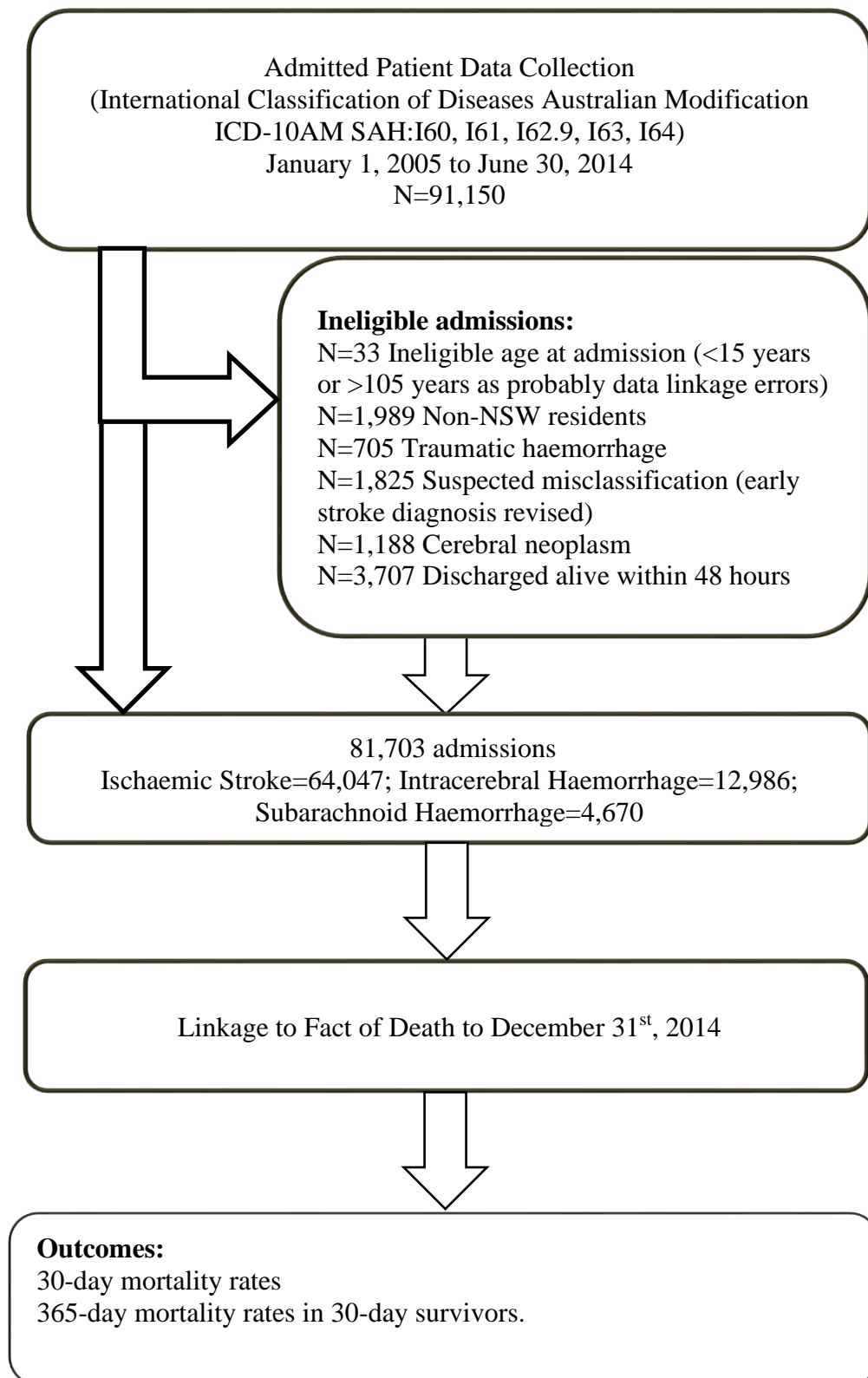
**Authors:** Melina Gattellari\*,<sup>1,2</sup> Chris Goumas,<sup>1,3</sup> Bin Jalaludin,<sup>1,4,5</sup> John M. Worthington<sup>1,2,6</sup>

**Affiliations:** <sup>1</sup>Ingham Institute for Applied Medical Research, Liverpool (Sydney), New South Wales, Australia; <sup>2</sup>Department of Neurology, Royal Prince Alfred Hospital, Camperdown (Sydney), New South Wales, Australia; <sup>3</sup>School of Public Health, The University of Sydney; <sup>4</sup>Population Health Intelligence, Healthy People and Places Unit, South Western Sydney Local Health District, Liverpool, Sydney, New South Wales, Australia; <sup>5</sup>School of Public Health and Community Medicine, The University of New South Wales, Kensington, Sydney, New South Wales, Australia; <sup>6</sup>South Western Sydney Clinical School, The University of New South Wales, Liverpool, Sydney, New South Wales, Australia

Words: 2,984 words (Limit 3000)

\*Corresponding author: Department of Neurology, Missenden Road, Camperdown, NSW 2052; Ph: +61 2 9515 3251; Email:

[Melina.Gattellari@health.nsw.gov.au](mailto:Melina.Gattellari@health.nsw.gov.au)



**Figure S1: Linkage and patient selection flow chart.**

**Table S1: Annual percentage change (APC) in admission rates by sex and age**

Subtype	APC (95% CI)	p	p-interaction
<i>Ischaemic</i>			
Males	-1.3 (-2.9 to 0.2)	0.09	
Females	-1.2 (-2.8 to 0.4)	0.15	0.93
<i>Intracerebral Haemorrhage</i>			
Males	-2.5 (-4.0 to -0.9)	0.003	
Females	-1.9 (-3.6 to -0.1)	0.04	0.64
<i>Subarachnoid Haemorrhage</i>			
Males	-0.5 (-2.2 to 1.3)	0.6	
Females	-0.1 (-1.7 to 1.5)	0.91	0.77
<i>Ischaemic &amp; Intracerebral Haemorrhage</i>			
Males	-1.6 (-3.0 to -0.1)	0.03	
Females	-1.4 (-2.9 to 0.2)	0.08	0.89
<i>Ischaemic</i>			
15-44 years	1.5 (-4.4 to 7.7)	0.63	
45-64 years	-0.5 (-5.2 to 4.4)	0.84	
65-74 years	-3.1 (-6.3 to 0.2)	0.06	
75-84 years	-3.9 (-6.7 to -1.0)	0.009	
85+ years	-2.0 (-3.4 to -0.5)	0.007	0.60
<i>Intracerebral haemorrhage</i>			
15-44 years	-3.2 (-7.7 to 1.4)	0.17	
45-64 years	-1.7 (-5.6 to 2.3)	0.39	
65-74 years	-1.2 (-4.7 to 2.4)	0.50	
75-84 years	-2.9 (-5.6 to -0.0)	0.05	
85+ years	-2.0 (-3.3 to -0.6)	0.006	0.95
<i>Subarachnoid Haemorrhage</i>			
15-44 years	-0.1 (-6.2 to 6.3)	0.97	
45-64 years	0.4 (-1.4 to 2.2)	0.70	
65-74 years	-1.7 (-4.4 to 1.1)	0.23	
75-84 years	2.0 (-1.5 to 5.6)	0.27	
85+ years	-3.4 (-7.8 to 1.1)	0.13	0.73
<i>Ischaemic &amp; Intracerebral haemorrhage</i>			
15-44 years	0.2 (-5.0 to 5.7)	0.94	
45-64 years	-0.7 (-5.2 to 3.9)	0.76	
65-74 years	-2.8 (-6.0 to 0.5)	0.10	
75-84 years	-3.7 (-6.5 to -0.9)	0.01	
85+ years	-2.1 (-3.4 to -0.8)	0.002	0.80

**Table S2: Age standardised 30-day fatality rates (ASR per 100,000) by year**

Year	Population	Cases	Crude rate	WHO ASR	
<b>Ischaemic Stroke</b>					
2005	5,373,880	1,153	21.5	11.6	(10.9-12.3)
2006	5,423,700	1,242	22.9	12.1	(11.4-12.8)
2007	5,507,371	1,174	21.3	11.1	(10.5-11.8)
2008	5,606,333	1,212	21.6	10.9	(10.3-11.6)
2009	5,704,008	1,165	20.4	10.2	(9.6-10.8)
2010	5,783,499	1,028	17.8	8.9	(8.3-9.5)
2011	5,851,202	1,205	20.6	9.8	(9.3-10.4)
2012	5,926,116	1,165	19.7	9.3	(8.8-9.9)
2013	6,009,142	1,071	17.8	8.4	(7.9-9.0)
<b>Intracerebral Haemorrhage</b>					
2005	5,373,880	520	9.7	5.8	(5.3-6.3)
2006	5,423,700	596	11.0	6.5	(5.9-7.0)
2007	5,507,371	533	9.7	5.6	(5.1-6.1)
2008	5,606,333	536	9.6	5.5	(5.0-6.0)
2009	5,704,008	540	9.5	5.4	(4.9-5.8)
2010	5,783,499	545	9.4	5.3	(4.8-5.7)
2011	5,851,202	569	9.7	5.3	(4.8-5.8)
2012	5,926,116	602	10.2	5.6	(5.2-6.1)
2013	6,009,142	570	9.5	5.0	(4.6-5.5)
<b>Subarachnoid Haemorrhage</b>					
2005	5,373,880	123	2.3	1.6	(1.3-1.9)
2006	5,423,700	150	2.8	1.9	(1.6-2.3)
2007	5,507,371	131	2.4	1.6	(1.4-1.9)
2008	5,606,333	134	2.4	1.6	(1.4-2.0)
2009	5,704,008	142	2.5	1.6	(1.4-1.9)
2010	5,783,499	127	2.2	1.6	(1.3-1.8)
2011	5,851,202	134	2.3	1.5	(1.2-1.8)
2012	5,926,116	120	2.0	1.4	(1.1-1.6)
2013	6,009,142	137	2.3	1.5	(1.2-1.7)
<b>Ischaemic Stroke and Intracerebral Haemorrhage</b>					
2005	5,373,880	1,673	31.1	17.5	(16.6-18.3)
2006	5,423,700	1,838	33.9	18.6	(17.7-19.5)
2007	5,507,371	1,707	31.0	16.7	(15.9-17.5)
2008	5,606,333	1,748	31.2	16.4	(15.6-17.2)
2009	5,704,008	1,705	29.9	15.6	(14.8-16.3)
2010	5,783,499	1,573	27.2	14.2	(13.4-14.9)
2011	5,851,202	1,774	30.3	15.1	(14.4-15.9)
2012	5,926,116	1,767	29.8	15.0	(14.3-15.7)
2013	6,009,142	1,641	27.3	13.5	(12.8-14.2)

**Table S3: Annual percentage change (APC) in fatality rates by sex and age**

Subtype	APC (95% CI)	p	p- interaction
Ischaemic			
Males	-3.1 (-5.8 to -0.4)	0.02	
Females	-3.4 (-6.3 to -0.4)	0.03	0.84
ICH			
Males	-1.8 (-4.5 to 0.9)	0.19	
Females	-2.1 (-4.9 to 0.8)	0.15	0.93
SAH			
Males	-1.0 (-4.8 to 3.0)	0.62	
Females	-3.6 (-6.5 to -0.5)	0.02	0.31
Ischaemic & ICH			
Males	-2.5 (-4.9 to 0.0)	0.05	
Females	-3.0 (-5.5 to -0.3)	0.03	0.75
Ischaemic			
15-44	5.0 (-6.1 to 17.5)	0.39	
45-64	-2.9 (-9.1 to 3.8)	0.39	
65-74	-6.0 (-10.9 to -0.8)	0.02	
75-84	-4.9 (-9.5 to 0.0)	0.05	
85+	-2.2 (-5.8 to 1.4)	0.23	0.38
ICH			
15-44	-6.8 (-15.2 to 2.4)	0.14	
45-64	-0.4 (-5.7 to 5.2)	0.88	
65-74	-2.0 (-7.5 to 3.9)	0.50	
75-84	-2.7 (-6.8 to 1.6)	0.22	
85+	-0.9 (-3.0 to 1.2)	0.39	0.59
SAH			
15-44	-1.0 (-10.2 to 9.2)	0.85	
45-64	-0.6 (-4.9 to 3.9)	0.80	
65-74	-8.2 (-13.8 to -2.2)	0.008	
75-84	0.5 (-4.7 to 6.0)	0.85	
85+	-4.6 (-9.7 to 0.7)	0.09	0.20
Ischaemic & ICH			
15-44	-2.3 (-10.0 to 6.1)	0.58	
45-64	-1.9 (-7.2 to 3.7)	0.50	
65-74	-4.3 (-9.3 to 0.9)	0.10	
75-84	-4.1 (-8.5 to 0.5)	0.08	
85+	-2.3 (-5.2 to 0.7)	0.14	0.94

ICH=Intracerebral Haemorrhage; SAH=Subarachnoid Haemorrhage

**The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.**

	<b>Item No.</b>	<b>STROBE items</b>	<b>Location in manuscript where items are reported</b>	<b>RECORD items</b>	<b>Location in manuscript where items are reported</b>
<b>Title and abstract</b>					
	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Title, Abstract	<p>RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included.</p> <p>RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract.</p> <p>RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.</p>	<p>Title, Abstract</p> <p>Abstract</p> <p>Title and Abstract</p>
<b>IntroductionTot;</b>					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Pages 4 and 5		
Objectives	3	State specific objectives, including any prespecified hypotheses	Last paragraph of Introduction		
<b>Methods</b>					
Study Design	4	Present key elements of study design early in the paper	Pages 5 to 7, Supplemental Figure 1.		
Setting	5	Describe the setting, locations, and relevant dates, including	Pages 5 to 7		

		periods of recruitment, exposure, follow-up, and data collection			
Participants	6	<p>(a) <i>Cohort study</i> - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up</p> <p><i>Case-control study</i> - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls</p> <p><i>Cross-sectional study</i> - Give the eligibility criteria, and the sources and methods of selection of participants</p> <p>(b) <i>Cohort study</i> - For matched studies, give matching criteria and number of exposed and unexposed</p> <p><i>Case-control study</i> - For matched studies, give matching criteria and the number of controls per case</p>		<p>RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.</p> <p>RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.</p> <p>RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.</p>	<p>Pages 5 and 6, Supplemental Figure 1</p> <p>Citation to reference 9</p> <p>Supplemental Figure 1.</p>
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Pages 5 to 7	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Page 5 and 6
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement).	Page 7		

		Describe comparability of assessment methods if there is more than one group			
Bias	9	Describe any efforts to address potential sources of bias	Page 6, Case Selection sub-heading		
Study size	10	Explain how the study size was arrived at	Page 5		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Page 7, Statistical Analysis sub-heading		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> - If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> - If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Page 7, Statistical Analysis sub-heading.		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the	Page 5 and 6



				<p>database population used to create the study population.</p> <p>RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.</p>	Page 6, Case Selection subheading.
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Page 6, second paragraph
<b>Results</b>					
Participants	13	<p>(a) Report the numbers of individuals at each stage of the study (<i>e.g.</i>, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed)</p> <p>(b) Give reasons for non-participation at each stage.</p> <p>(c) Consider use of a flow diagram</p>	Supplemental Figure 1, Pages 6, Page 7, Cohort selection and characteristics sub-heading.	RECORD 13.1: Describe in detail the selection of the persons included in the study ( <i>i.e.</i> , study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	
Descriptive data	14	<p>(a) Give characteristics of study participants (<i>e.g.</i>, demographic, clinical, social) and information on exposures and potential confounders</p> <p>(b) Indicate the number of participants with missing data for each variable of interest</p> <p>(c) <i>Cohort study</i> - summarise follow-up time (<i>e.g.</i>, average and total amount)</p>	Table 1, Pages 7-10.		

Outcome data	15	<p><i>Cohort study</i> - Report numbers of outcome events or summary measures over time</p> <p><i>Case-control study</i> - Report numbers in each exposure category, or summary measures of exposure</p> <p><i>Cross-sectional study</i> - Report numbers of outcome events or summary measures</p>	Tables 1 and 2		
Main results	16	<p>(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included</p> <p>(b) Report category boundaries when continuous variables were categorized</p> <p>(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period</p>	Tables 2 and 3; Supplemental Tables 1,2 and 3.		
Other analyses	17	Report other analyses done— e.g., analyses of subgroups and interactions, and sensitivity analyses	Table 2, Results.		
<b>Discussion</b>					
Key results	18	Summarise key results with reference to study objectives	Page 11 and Page 12.		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Pages 14 and 15	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing	

				data, and changing eligibility over time, as they pertain to the study being reported.	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Pages 11 to 16		
Generalisability	21	Discuss the generalisability (external validity) of the study results	Pages 14 to 15		
<b>Other Information</b>					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Page 16 under funding sub-heading		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Acknowledgements, Page 16.

\*Reference: Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

\*Checklist is protected under Creative Commons Attribution