

1 **Title:** Measuring stroke outcomes using linked administrative data: Population-based
2 estimates and validation of “home-time” as a surrogate measure of functional status.

3
4 **Authors:**

5 Name: Melina Gattellari*^{a,b}
6 Email: melina.gattellari@health.nsw.gov.au

7
8 Name: Chris Goumas^{a,c}
9 Email: chris.goumas@sydney.edu.au

10
11 Name: Bin Jalaludin^{a,d,e}
12 Email: bin.jalaludin@health.nsw.gov.au

13
14 Name: John Worthington^{a,b,f}
15 Email: John.Worthington@health.nsw.gov.au

16
17 Affiliations: ^aIngham Institute for Applied Medical Research,
18 1 Campbell Street, Liverpool, NSW, 2170, Australia

19
20 ^bDepartment of Neurology, Royal Prince Alfred Hospital,
21 Missenden Road, Camperdown, NSW, 2050, Australia.

22
23 ^cSchool of Public Health, University of Sydney

24
25 ^dPopulation Health Intelligence, Healthy People and Places Unit; South
26 Western Sydney Local Health District, Liverpool, Sydney, New South
27 Wales, Australia

28
29 ^eSchool of Public Health, The University of New South Wales,
30 Kensington, Sydney, New South Wales, Australia

31
32 ^fSouth Western Sydney Clinical School, The University of New South
33 Wales, Liverpool, Sydney, New South Wales, Australia

34
35 *Corresponding author.

36 **Abstract**

37 **Background:** Administrative data offer cost-effective, whole-of-population stroke
38 surveillance yet the lack of validated outcomes is a short-coming. The number of days spent
39 living at home after stroke (“home-time”) is a patient-centred outcome that can be objectively
40 ascertained from administrative data. Population-based validation against both severity and
41 outcome measures and for all subtypes is lacking.

42 **Methods:** Stroke hospitalisations from a state-wide census in New South Wales, Australia,
43 from July 1, 2005 to March 31, 2014 were linked to pre-hospital data, post-stroke admissions
44 and deaths. We calculated correlations between 90-day home-time and Glasgow Coma Scale
45 (GCS) scores, measured upon a patient’s initial contact with paramedics, and Functional
46 Independence Measure (FIM) scores, measured upon entry to rehabilitation after the acute
47 hospital stroke admission. Negative binomial regression models were used to identify
48 predictors of home-time.

49 **Results:** Patients with stroke (N=74,501) spent a median of 53 days living at home after the
50 event. Median home-time was 60 days after ischaemic stroke, 49 days after subarachnoid
51 haemorrhage and 0 days after intracerebral haemorrhage. GCS and FIM scores significantly
52 correlated with home-time (p-values<0.001). Female sex predicted less home-time in
53 ischaemic stroke, while being married increased home time after ischaemic stroke and
54 subarachnoid haemorrhage.

55 **Conclusions:** Home-time measured using administrative data is a robust, replicable and valid
56 patient-centred outcome enabling inexpensive population-based surveillance.

57 **Keywords:** Stroke, outcomes, administrative data, validation, disability, epidemiology

58 **List of Abbreviations:** IS:Ischaemic stroke; ICH: Intracerebral haemorrhage; SAH:
59 Subarachnoid haemorrhage; NIHSS: National Institutes of Health Stroke Scale; FIM:

- 60 Functional Independence Measure; GCS: Glasgow Coma Scale; NSW: New South Wales;
- 61 eMR: electronic Medical Records (Ambulance NSW); PHCR: Patient Health Care Record
- 62 (Ambulance NSW).

63 **1. Introduction**

64 Stroke is a leading cause of death and disability with significant global disparities in
65 incidence, mortality and morbidity [1]. Over the past two decades there have been
66 considerable advances in primary and secondary stroke prevention, particularly in high-
67 income nations [2-5] and, more recently, health services have been re-designed in several
68 countries to facilitate access to multidisciplinary stroke unit care, thrombolysis and
69 mechanical thrombectomy for ischaemic stroke and endovascular coiling for subarachnoid
70 haemorrhage [6-9]. Monitoring stroke outcomes can assist in identifying areas of need,
71 informing implementation strategies to bridge evidence-practice gaps. Administrative data
72 offer a cost-effective alternative to labour-intensive gold-standard epidemiological studies,
73 providing whole-of-population stroke surveillance [10]. However, the absence of validated
74 outcomes after stroke is a significant shortcoming.

75 Home-time measures the number of days spent at home after stroke onset, taking into account
76 patient survival and days spent in hospital, reflecting a patient's capacity to return to their
77 private residence after hospitalisation [11-15]. Home-time therefore can be unobtrusively and
78 objectively measured using routinely collected data that are linked to health records following
79 stroke and is proposed as a patient-centred surrogate measure of functional status. Unlike
80 traditional stroke outcome measures, such as the modified Rankin Scale, the validity of
81 home-time is not threatened by inter-rater variability, attrition of research participants or
82 respondent characteristics (proxy versus patient) [16,17]. Home-time derived from
83 administrative data may therefore be a cost-effective adjunct to traditional patient reported
84 outcomes to minimise selection bias, improve the reliability of outcome assessment and allow
85 continuous large-scale monitoring that otherwise would be impracticable. Deriving home-
86 time from administrative data may also broaden patient representation in epidemiological and
87 health services research. For example, patients whose preferred language is not the dominant

88 community language or those with cognitive impairment may encounter barriers to effective
89 engagement in research and quality assurance activities. Characterised in this way, home-
90 time is inclusive, ascertained for all patients irrespective of their socio-demographic and
91 health characteristics.

92 Few studies validating home-time are population-based or have validated home-time against
93 both severity and outcome measures and for all stroke sub-types [11,14]. Further, there is
94 limited information about prognostic indicators of home-time and validation has yet to be
95 conducted where severity, home-time and functional status are measured wholly using
96 routinely collected data. We analysed a jurisdiction-wide analysis reporting population-based
97 estimates of home-time and validated home-time as an outcome measure for routine stroke
98 surveillance.

99 **2. Methods**

100 Reporting according to the “**RE**porting of studies **C**onducted using **O**bservational **R**outinely-
101 collected health **D**ata (RECORD) Statement” is reported in Supplementary Table 1.

102 **2.1 Patient Selection:** The study was part of the Home to Outcomes (H₂O) project which
103 aimed to report stroke epidemiology and outcomes in New South Wales (NSW), Australia’s
104 most populous state (population ~7.99 million in 2019). The project utilised linked routinely
105 collected pre-hospital, hospital, rehabilitation and mortality data. As previously described
106 [18,19] deterministic and probabilistic linkages were performed by the Centre for Health
107 Record Linkage (CHeReL) using gold-standard privacy preserving protocols and ensuring
108 that the linkage error rate was no greater than 5 per 1,000. Patient identifiers, such as names,
109 addresses, dates of birth and dates of health service delivery (for example, dates of
110 admission) facilitated linkage. Quality assurance of the linkage process undertaken by the
111 CHeReL has been described (<http://www.cherel.org.au/quality-assurance>).

112 Patients over the age of 15 years were identified from the Admitted Patient Data Collection, a
113 census of all hospital separations in NSW recording the principal diagnosis and up to 49
114 secondary diagnoses using the International Classification of Diseases, version 10, Australian
115 Modification (ICD-10AM) [20]. Applying ICD-10AM codes recommended to assign stroke
116 sub-types [21,22], we selected admissions with a principal diagnosis of subarachnoid
117 haemorrhage (ICD-10AM I60), intracerebral haemorrhage (ICD-10AM I61, I62.9) and
118 ischaemic stroke (ICS-10AM I63, I64) from January 1, 2005 to March 31st, 2014 providing a
119 minimum follow-up of 90-days to June 30, 2014, the most recently available data at the time
120 linkage was performed. Strokes recorded in secondary diagnostic positions were included if
121 the primary diagnosis indicated an acute stroke event, where iatrogenic causes were identified
122 or where the stroke was flagged as occurring in hospital. Where primary diagnoses indicated
123 I64 or I62.9 codes (unspecified stroke and unspecified haemorrhagic stroke, respectively), we
124 interrogated records for more specific stroke codes to assign stroke-subtype. Sub-type was
125 defined in this manner for 2,417 or 2.5% of cases; otherwise, we adhered to
126 recommendations to code these cases as ischaemic stroke (that is, I64) [21] or intracerebral
127 haemorrhage (that is, I62.9) [22].

128 As previously reported [18,19], patients were ineligible if resident outside NSW or if cases
129 had concomitantly recorded major head trauma (ICD-10AM S04, S06, S07, S08, S17, S18,
130 S01.7, S02.0, S02.1, S02.6, S02.7, S03.0, S09.0, S09.2, S09, S05.2, S05.7) or primary or
131 metastatic cerebral neoplasms (C71,C70.0,C70.9, C79.3,D33.0, D33.1, D33.2, D33.3,
132 D33.9). We excluded cases whose initial stroke diagnosis was revised upon transfer to
133 another hospital by the end of the next day after presentation as presumed misclassified cases.
134 Early discharges home, defined as completion of care allowing discharge within the first 48
135 hours, were excluded to increase the specificity of case ascertainment consistent with
136 previous approaches [23-25] and remove patients with early symptom resolution suggesting

137 transient ischaemic attack or stroke mimics and cases with unrevised provisional stroke
138 diagnoses. Patients transferred for care outside the state and other transfers in the absence of
139 contiguous records were also excluded as missing hospitalisation data precluded accurate
140 calculation of home-time. The first eligible stroke admission (“index admission”) for each
141 patient during the study period was selected for analysis.

142 **2.2 Calculation of 90-day home-time:** In keeping with other research [14], home-time was
143 defined as the sum of full days not spent in hospital in the 90 days after the stroke admission.
144 For patients dying or newly discharged to a nursing home within 90-days, zero days were
145 counted from the date of death or discharge, respectively. Ninety-day “home-time” was
146 calculated by linking stroke cases to subsequent hospital records and the state death register
147 to June 30, 2014. Hospital records incorporated the acute stroke admission including inter-
148 hospital transfers, any contiguously recorded in-patient rehabilitation following the acute
149 period of stay and all subsequent acute and sub-acute readmissions for any reason, including
150 rehabilitation. Hospital encounters also included same-day admissions, which included out-
151 patient rehabilitation.

152 Prior to applying eligibility and exclusion criteria, we undertook data cleaning for cases with
153 illogical death record data to facilitate valid calculation of 90-day home-time. Specifically,
154 there were 70 death records with duplicate records indicating more than one probable linkage
155 to mortality data. We selected the record with information consistent with data recorded in
156 the other available linked data-sets (for example, consistent age and sex between death and
157 admissions records). There were a further 38 records with dates of death preceding hospital
158 discharge by more than one day. Of these, eleven were deemed clerical errors and accepted as
159 valid, while the remaining were considered erroneously linked death records.

160 **2.3 Measures validating home-time**

161 **2.31 Proxy stroke severity measurement**

162 As described previously [19], we developed a proxy measure of stroke severity incorporating
163 the Glasgow Coma Scale (GCS) and mode of arrival (ambulance or private transport) by
164 linking admissions to pre-hospital data obtained from Ambulance NSW, the sole provider of
165 emergency services for the state. Total GCS scores were obtained from the initial pre-hospital
166 vital signs survey recorded at first contact with paramedics. Pre-determined cut-offs
167 categorised patients as fully alert with a maximum score of 15 or with mild, moderate or
168 severe levels of lowered consciousness [26,27]. Self-presenting patients arriving to hospital
169 by private transport were considered “ambulant”. Both arrival mode and GCS were near
170 universally available, measured prospectively at first contact with health services and
171 ascertained blinded to patient outcomes. The GCS and mode of arrival strongly predict stroke
172 mortality [26-29], providing de-facto validity for the composite measure. Within the H₂O
173 study cohort, we have published a validation of the proxy severity measure in ischaemic
174 stroke (N>17,000 patients) demonstrating a strong association with 30-day mortality with
175 cross-validation of findings (N>18,000 patients) [19].

176 **2.32 Functional Status:** We obtained the 18-item Functional Independence Measure (FIM)
177 [30] from the NSW Subacute and Non-Acute Data Collection, recorded upon patient’s first
178 admission to inpatient rehabilitation within 90-days of the stroke admission. FIM scores were
179 available from July 1, 2009 for the sub-set of publicly funded hospitals, which provided
180 ~70% of inpatient stroke rehabilitation in Australia during the study period [31]. The FIM
181 comprises 19 items, with 13 measuring motor function and five measuring cognitive function.
182 Each item was measured using a seven-point ordinal scale with higher scores indicating
183 greater independence with everyday tasks, such as mobility, eating, bathing, memory and
184 problem solving. Total scores ranged from 18 to 126, while motor and cognitive subscale
185 scores ranged from 13 to 91 and 5 to 35, respectively.

186 **2.4 Statistical analysis:** As home-time had a skewed distribution, we report median values
187 and the interquartile range (IQR) for all strokes and by stroke sub-type. Mean values (\pm
188 Standard Deviation) were calculated to enable comparisons with other studies. We calculated
189 Spearman correlations between home-time and GCS and FIM scores. These descriptive
190 analyses were repeated excluding patients dying in hospital who, by definition, spent zero
191 days at home. We also report mean and median values for “home-time” for each level of the
192 proxy stroke severity measure incorporating the GCS categories and mode of arrival.

193 To identify predictors of home-time, multivariate negative binomial regression was carried
194 out to take into account the skewed distribution of home-time and its measurement in counts
195 of whole days. We used generalised estimation equations to account for inter-dependencies of
196 data clustered on hospital of presentation. Modelling was carried out using PROC GENMOD
197 (SAS 9.4, SAS institute Inc, Cary, USA) and an exchangeable working correlation. Adjusted
198 means were derived from the model output. Potential predictors included age, sex, year of
199 admission, marital status and our proxy stroke severity measure. In addition, we ascertained
200 stroke history and Charlson comorbidities via linkage to prior hospitalisations up to July 1,
201 2001, providing 4.5 years of historical admissions data for all cases. To allow for departure
202 from linearity, we included both linear and quadratic terms for age and modelled each
203 Charlson comorbidity characterised by applying ICD-10 codes as reported by Quan,
204 Sundararjan, Halfon et al [32]. Hemiplegia and cerebrovascular disease were removed as
205 redundant [33] while the effect of HIV/AIDs was not assessed given its very low prevalence.

206 We modelled a measure of socio-economic status, the Index of Relative Socio-economic
207 Disadvantage (IRSD) [34] reflecting the degree of disadvantage within the geographical area
208 of patient residence based on area-aggregated indicators such as income, educational status
209 and occupation collected from a mandatory, national census carried out once every five years
210 (>95% participation rate of eligible households) [35]. IRSD scores are categorised into

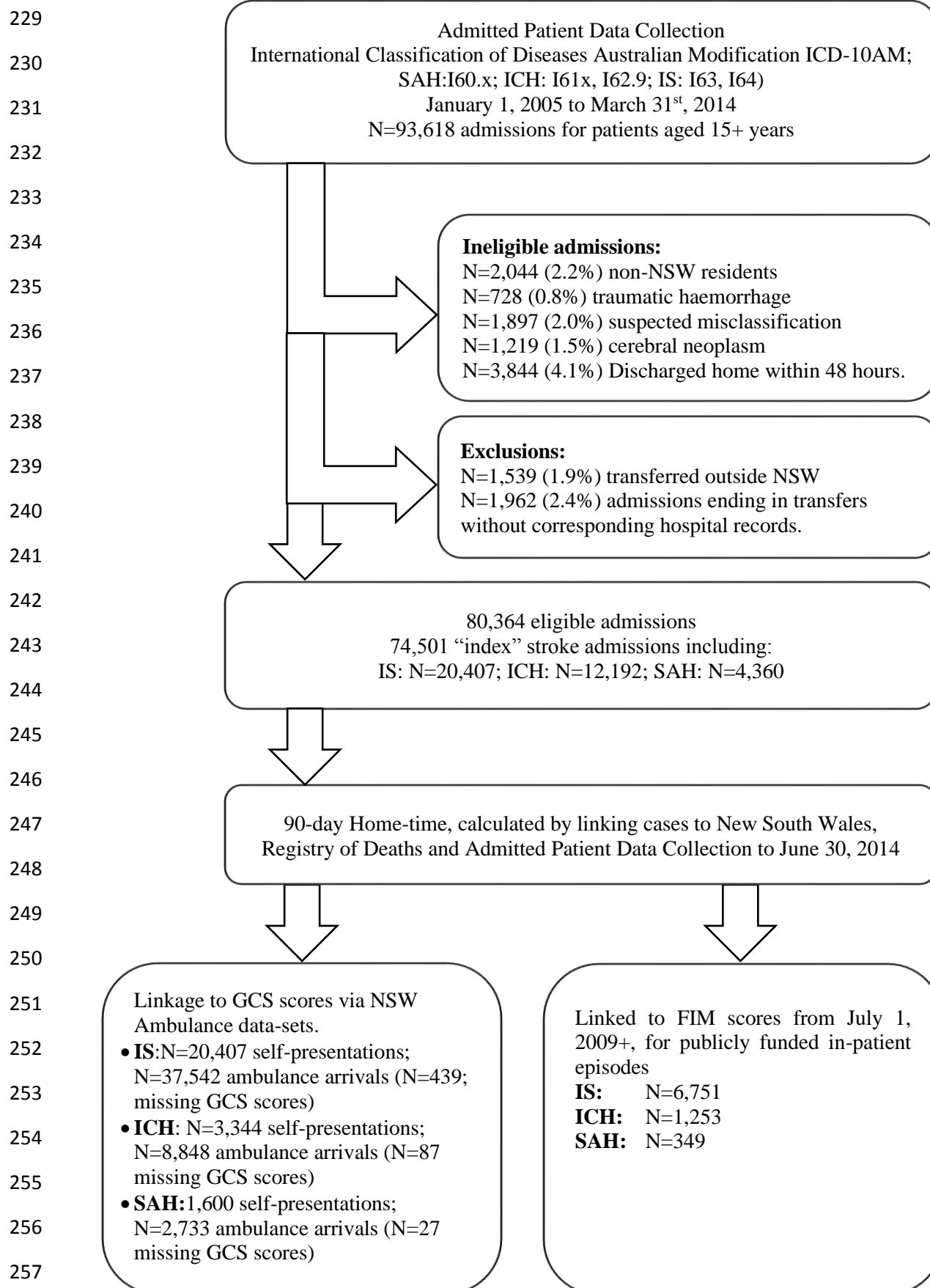
211 quintile cut-offs with lower scores indicating greater relative disadvantage. Multivariate
212 analyses also adjusted for the length of time patients spent in hospital in the twelve-months
213 prior to the stroke admission to account for pre-existing comorbidity burden and health
214 service utilisation proximal to the stroke event. Analyses were undertaken separately for each
215 of the individual stroke sub-types given the well-established differences in patient
216 characteristics, risk factors and prognosis. All analyses were carried out using SAS software
217 (SAS 9.4, SAS institute Inc, Cary, USA).

218 **2.5 Ethics**

219 The NSW Population and Health Services Ethics Committee approved the study
220 (HREC/14/CIPHS/17).

221 **3. Results**

222 **3.1 Patient selection and characteristics:** Of the 83,866 potentially eligible stroke
223 admissions amongst NSW residents, there were 80,365 (95.8%) cases without known missing
224 home-time data. Of these, 74,501 were index admissions including 57,949 patients with
225 ischaemic stroke (77.8%), 12,192 with intracerebral haemorrhage (16.4%) and 4,360 with
226 subarachnoid haemorrhage (5.9%) for which home-time was calculated (Figure 1). Almost
227 half were female (N=37,097; 49.8%). One in four patients were older than 85 years of age
228 (24.6%), while 7.0% were younger than 50 years (Supplementary Table 1).



258 **Figure 1: Linkage and patient selection flow chart.**

259 **3.2 Descriptive analysis of home-time estimates:** Median home-time in the 90-days after
260 stroke was 53 days (IQR=0 to 80 days). Median home-time was longer after an ischaemic
261 stroke (Median=60-days; IQR=0 to 81) and shortest after an intracerebral haemorrhage
262 (Median=0 days; IQR=0 to 63). Median home-time after a subarachnoid haemorrhage was
263 49 days (IQR=0 to 80 days). Almost 20% of patients (N=25,987) spent zero days at home in
264 the 90-days after stroke. Just under one-third of ischaemic stroke patients (N=17,575; 30.3%)
265 did not spend any full days at home compared with 55.2% (N=6,731) with intracerebral
266 haemorrhage and 38.6% (N=1,681) with subarachnoid haemorrhage. When the study
267 population was restricted to the sub-set of patients who survived their stroke admission, the
268 estimated median home-time was similar for patients with ischaemic stroke and subarachnoid
269 haemorrhage (Median=70 days and 69 days, respectively) and lower for those with an
270 intracerebral haemorrhage (Median=55 days (IQR=5 to 78) (Table 1).

271 **3.2.1 Descriptive statistics for home-time according to age and sex** Median home-time was
272 15 days longer for males compared with females after an ischaemic stroke and nine days
273 longer for men after a subarachnoid haemorrhage (Table 1). While the median home-time
274 value for intracerebral haemorrhage was zero days for both sexes, the 75th percentile value for
275 home-time was 67 days for men and 59 days for women. When restricted to patients who
276 survived their stroke admission at discharge, median values for males remained higher for all
277 subtypes although differences were attenuated.

278 Increasing age was associated with a reduction in home-time for all sub-types. The decrease
279 in median time across age was relatively gradual after ischaemic stroke ranging from 80 days
280 in patients younger than 50 years of age to 56 days in those aged 75 to 84 years. Home-time
281 reduced to zero days in patients older than 85 years. After intracerebral haemorrhage, median
282 home-time decreased from 54 days in patients younger than 50 to 40 days in those aged 50-
283 64 and then eight days in patients aged 65 to 74 years. After subarachnoid haemorrhage,

284 median home-time decreased from 69 days in patients younger than 50 to 25 days in those
285 aged 65-74 years. In both haemorrhagic stroke subtypes, median home-time was zero for
286 patients older than 75 years.

287 **Table 1. Descriptive statistics for days at home “home-time” within 90 days of diagnosis**
 288 **by sex and age**

Sub-type	All cases			Excluding in-hospital deaths		
	N	Mean (±SD)	Median (IQR)	N	Mean (±SD)	Median (IQR)
Sub-type						
IS	57,949	47±36	60 (0-81)	48,153	56±32	70 (33-83)
ICH	12,192	27±34	0 (0-63)	7,147	46±33	55 (5-78)
SAH	4,360	40±35	49 (0-74)	3,082	56±28	69 (41-77)
<i>Total</i>	<i>74,501</i>	<i>43±36</i>	<i>53 (0-80)</i>	<i>58,382</i>	<i>55±32</i>	<i>69 (30-82)</i>
Subtype by sex						
IS						
Male	29,653	51±35	66 (0-82)	25,431	59±31	74 (41-83)
Female	28,296	42±36	51 (0-80)	22,722	53±33	66 (21-82)
ICH						
Male	6,059	29±35	0 (0-67)	3,670	48±33	57 (12-80)
Female	6,133	25±33	0 (0-59)	3,477	43±33	52 (0-76)
SAH						
Male	1,692	42±35	55 (0-76)	1,237	58±28	70 (43-78)
Female	2,668	38±35	46 (0-73)	1,845	55±29	68 (38-76)
Subtype by age						
IS						
<50 years	2,925	66±28	80 (60-85)	2,777	70±24	80 (67-85)
50-64 years	8,348	62±31	78 (46-84)	7,775	66±27	79 (57-85)
65-74 years	11,668	55±33	72 (26-83)	10,505	62±29	76 (47-84)
75-84 years	19,870	45±35	56 (0-80)	16,471	54±32	67 (28-82)
85+ years	15,138	30±35	0 (0-67)	10,625	43±34	52 (0-76)
ICH						
<50 years	967	44±36	54 (0-79)	765	56±31	69 (35-82)
50-64 years	1,796	39±36	40 (0-78)	1,282	54±31	66 (28-82)
65-74 years	2,350	32±35	8 (0-69)	1,499	49±32	58 (17-81)
75-84 years	4,205	23±32	0 (0-53)	2,277	43±33	49 (0-74)
85+ years	2,874	17±29	0 (0-28)	1,324	36±33	38 (0-68)
SAH						
<50 years	1,313	53±32	69 (25-77)	1,122	62±25	72 (56-78)
50-64 years	1,376	46±34	63 (0-75)	1,096	58±27	69 (48-76)
65-74 years	674	34±34	25 (0-69)	448	50±30	61 (25-76)
75-84 years	656	21±31	0 (0-48)	311	44±31	51 (6-74)
85+ years	341	11±26	0 (0-0)	105	37±35	39 (0-73)

290 IS=Ischaemic Stroke

291 ICH=Intracerebral Haemorrhage

292 SAH=Subarachnoid Haemorrhage

293 **3.3 Validity of home-time**

294 ***3.31 Association between home-time and proxy severity measure***

295 Amongst self-presenting “ambulant” patients who did not require paramedic transport, the
296 median home-time was 76 days (IQR=39-84) after ischaemic stroke, 53 days (IQR=0 to 81
297 days) after intracerebral haemorrhage and 70 days (IQR=28 to 78) after subarachnoid
298 haemorrhage. GCS scores were available for 48,597 out of 49,150 (98.9%) patients
299 transported via ambulance. After either an intracerebral or subarachnoid haemorrhage,
300 median home-time was zero for patients with any decreased level of consciousness (GCS
301 scores less than 15) (Table 1). Median home-time was zero after ischaemic stroke in patients
302 with GCS scores less than 13. Amongst patients surviving their stroke admission, home-time
303 decreased with decreasing levels of consciousness. However, in patients with haemorrhagic
304 stroke, median home-time was higher for patients whose GCS scores indicated “very severe”
305 brain injury (that is, unconscious patients), than for those with “severe” brain injury.

306 **3.32 Home-time and GCS and FIM scores**

307 Higher GCS scores indicating higher levels of consciousness were moderately and
308 significantly correlated with more home-time in ischaemic stroke and intracerebral
309 haemorrhage ($\rho=0.35$ and $\rho=0.39$) ($p<0.001$) (Table 3). GCS scores more strongly
310 correlated with home-time in subarachnoid haemorrhage ($\rho=0.51$) ($p<0.001$). The
311 correlation between GCS scores and home-time remained significant in the sub-group of
312 patients who survived their stroke admission; however, the association was reduced (Table
313 3).

314 **Table 2: Descriptive statistics for days at home “home-time” within 90 days of diagnosis**
 315 **by sex and age**

	Overall			Excluding in-hospital deaths		
	N	Mean±SD	Median (IQR)	N	Mean±SD	Median (IQR)
IS						
Ambulant	20,407	59±32	76 (39-84)	18,443	65±27	78 (56-84)
GCS=15	21,735	50±34	63 (1-81)	19,190	56±31	69 (36-82)
GCS=13-14	7,209	35±35	28 (0-73)	5,812	44±34	52 (0-77)
GCS=9-12	5,652	22±32	0 (0-51)	3,532	36±34	32 (0-70)
GCS=3-8	2,507	11±25	0 (0-0)	866	33±33	23 (0-67)
ICH						
Ambulant	3,344	43±37	53 (0-81)	2,489	58±31	72 (39-83)
GCS=15	3,694	33±34	23 (0-68)	2,638	46±32	53 (9-76)
GCS=13-14	1,792	22±31	0 (0-50)	1,101	37±32	40 (0-66)
GCS=9-12	1,567	11±24	0 (0-0)	640	27±31	12 (0-55)
GCS=3-8	1,708	4±16	0 (0-0)	235	30±33	15 (0-62)
SAH						
Ambulant	1,600	54±32	70 (28-78)	1,351	64±24	73 (61-80)
GCS=15	1,297	48±32	64 (4-75)	1,087	57±27	68 (45-76)
GCS=13-14	392	29±33	0 (0-65)	260	43±31	52 (0-73)
GCS=7-12	459	16±26	0 (0-28)	226	33±30	28 (0-62)
GCS=3-6	585	8±21	0 (0-0)	145	33±30	34 (0-63)

317 IS=Ischaemic Stroke

318 ICH=Intracerebral Haemorrhage

319 SAH=Subarachnoid Haemorrhage

320 **Table 3: Spearman correlation (rho) between GCS and FIM scores and**
 321 **“Home-Time”**

		Overall	Excluding in-hospital deaths	
Sub-type	N	Rho	N	Rho
GCS scores				
IS	37,103	0.35	29,400	0.24
ICH	8,761	0.39	4,614	0.21
SAH	2,733	0.51	1,718	0.34
FIM scores				
Motor subscale				
IS	6,751	0.64	6,608	0.64
ICH	1,253	0.62	1,229	0.63
SAH	349	0.68	346	0.68
Cognitive subscale				
IS	6,751	0.43	6,608	0.42
ICH	1,253	0.36	1,229	0.36
SAH	349	0.47	346	0.47
Total FIM score				
IS	6,751	0.66	6,608	0.65
ICH	1,253	0.63	1,229	0.63
SAH	349	0.69	346	0.69

324 IS=Ischaemic Stroke

325 ICH=Intracerebral Haemorrhage

326 SAH=Subarachnoid Haemorrhage

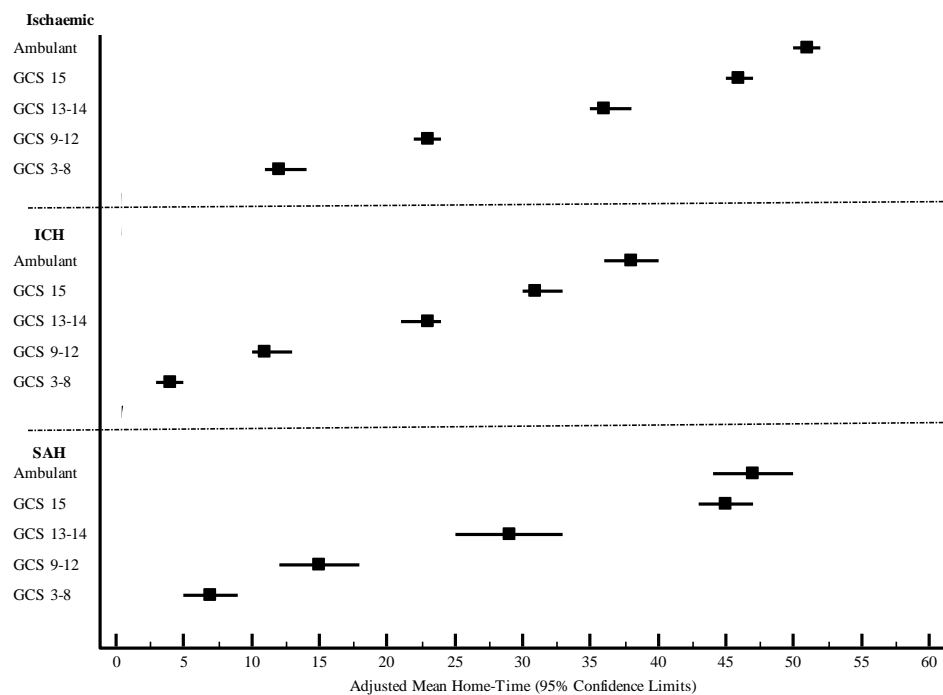
327 All Spearman rho-values were statistically significant (p-values<0.001).

328 FIM scores were available from 2009 onwards for 8,353 patients undergoing publicly funded
329 in-patient rehabilitation. FIM scores correlated with home-time for all sub-types ($r_{\text{values}}=0.63$
330 to 0.69) ($p\text{-values}<0.001$). Motor scores were more strongly correlated with home-time than
331 cognitive scores ($r_{\text{values}}=0.62$ to 0.68 versus $r_{\text{values}}=0.36\text{-}0.47$) ($p\text{-values}<0.001$) (Table 3) and
332 were largely unchanged when restricted to patients surviving their stroke admission.

333 **3.4 Predictors of home-time**

334 Increasing stroke severity assessed using our proxy composite measure independently
335 predicted fewer days spent at home after stroke for all subtypes in multivariate analyses (p -
336 values <0.001) (Figure 2; Supplementary Tables 2-5).

337 Amongst ischaemic stroke patients, married patients spent significantly more days at home
338 than single patients (Adjusted mean values=43 versus 39 days) ($p<0.001$) (Supplementary
339 Table 3). The effect of socio-economic status was near significant ($p=0.05$). Patients
340 residing in areas of least socio-economic disadvantage spent the greatest number of days at
341 home (42 days) although differences in adjusted means between categories were small (one
342 to two days). Home-time was reduced in the presence of all Charlson comorbidities tested,
343 except connective tissue disease. Dementia (31 versus 41 days), metastatic malignancy (27
344 versus 41 days) and moderate or severe liver disease (30 versus 41 days) had the greatest
345 impact on home time after ischaemic stroke. The adjusted effect of sex was significant
346 ($p=0.02$). While adjusted mean home-time values for males and females were equivalent (41
347 days) the confidence intervals suggesting a higher range of estimated values for males (95%
348 CI=41-42 versus 95% CI=40 to 41).



349

350 IS=Ischaemic stroke

351 ICH=Intracerebral haemorrhage

352 SAH=Subarachnoid haemorrhage

353

354 **Figure 2: Adjusted Mean Home-time by stroke severity category**

355 After intracerebral haemorrhage, patients spent significantly less time at home if they had a
356 documented diagnosis of myocardial infarction (18 versus 21 days), congestive heart failure
357 (17 versus 21 days), peripheral vascular disease (18 versus 21 days), renal disease (18 versus
358 21) and non-metastatic malignancy (17 versus 21 days). Patients with uncomplicated diabetes
359 spent more time at home (23 versus 20 days) (Supplementary Table 4). Patients with
360 subarachnoid haemorrhage spent significantly more time at home if they were married (32
361 versus 28 days) or if they had a previous stroke (40 versus 30 days) ($p=0.02$). The history of
362 renal disease and moderate or severe liver disease were associated with less home-time after
363 stroke (23 versus 31 days, $p=0.03$; 16 versus 31 days, $p=0.03$, respectively) (Supplementary
364 Table 5). There was significant variation in home-time estimates according to year of
365 admission for all sub-types, however, no clear linear association was evident.

366 **4. Discussion**

367 **4.1 Main findings:**

368 Home-time was reduced with increasing age, more severe disease and after intracerebral and
369 subarachnoid haemorrhage, all of which are known to adversely affect stroke outcome. Less
370 functional independence and greater comorbidity burden also reduced home-time. These
371 findings underscore the immediate adverse impact of stroke on people's ability to live at
372 home after the event and provide further evidence that home-time serves as a useful proxy for
373 disability.

374 The large cohort afforded reliable insights into the impact of stroke in relatively young
375 patients of which there were 5,205 cases aged younger than 50 years. Younger patients with
376 intracerebral and subarachnoid haemorrhage spent less time at home compared with patients
377 older than 75 after an ischaemic stroke. More than half of patients over the age 75 years after
378 haemorrhagic stroke and half of those older than 85 years after an ischaemic stroke did not

379 spend any time at home highlighting the considerable burden of stroke affecting elderly
380 members of the community.

381 Females with ischaemic stroke spent fewer days at home as reported elsewhere [11,36].

382 While sex differences either narrowed or did not persist in multivariate analyses accounting
383 for age, these findings demonstrate women's vulnerability to poor outcomes due to their
384 relatively advanced age, consistent with results from gold-standard incident stroke studies
385 [37]. Patients residing in areas of greater socio-economic disadvantage and single patients
386 with ischaemic stroke or subarachnoid haemorrhage had lower home-time estimates. Socio-
387 demographic characteristics have also been associated with traditional stroke outcome
388 measures [38]. These results either implicate patient characteristics in post-stroke morbidity
389 or, alternatively, indicate that home-time measures constructs other than disability. For
390 example, socio-demographic variables, may mediate their accessibility to health services
391 post-stroke.

392 These findings inform the validity of the composite surrogate severity measure incorporating
393 GCS and mode of arrival. The association between GCS and home-time was strongest for
394 subarachnoid haemorrhage corroborating GCS as a valid component of standardised severity
395 measures for this sub-type [27]. The correlation between GCS and home-time attenuated
396 when in-patient deaths were removed from the cohort consistent with patterns of correlations
397 reported when using gold-standard stroke severity assessment [14].

398 Previous results demonstrate that home-time correlates with the modified Rankin Scale or the
399 Barthel Index [11-15,36,39]. Here, we demonstrate a strong correlation with another measure
400 of functional status, the FIM, adding weight to the construct validity of home-time. Total
401 FIM and motor subscale scores were strongly associated with home-time, while the
402 correlation between cognitive subscale scores and days spent at home was more moderate.

403 Home-time may therefore underestimate disability due to cognitive deficits, a shortcoming
404 shared with traditional stroke outcome measures, such as the modified Rankin Scale score
405 [40].

406 **4.2 Comparison with previous work:** Two other population-based studies have published
407 home-time estimates [11,14]. A nation-wide Scottish study of over 101,000 patients defined
408 home-time as we have done and reported estimates comparable to those observed here [14].
409 Home-time was not reported separately for haemorrhagic and ischaemic stroke and the gold-
410 standard NIHSS severity measure was ascertained in around 5% of cases. Standardised
411 outcome measures were unavailable. Another study identified ~15,000 stroke cases
412 representing all subtypes as here and counted overnight hospital stays only [11]. Validation of
413 home-time was carried out for 552 patients with ischaemic stroke for whom the modified
414 Rankin Scale scores were available.

415 In a multi-centre study of 1,866 subarachnoid haemorrhage patients with gold-standard case
416 ascertainment, readmission data were not available [12]. Home-time strongly correlated with
417 the modified Rankin Scale. Our results strengthen evidence for the validity of home-time in
418 subarachnoid haemorrhage with more robust home-time assessment.

419 **4.3 Strengths and limitations:** Strengths include cost-effective, near universal home-time
420 and proxy severity ascertainment, and population-based validation of home-time against
421 baseline and outcome measures for all stroke sub-types with granular reporting of home-time
422 by patient characteristics. Another strength was the capacity to capture same-day admissions
423 including out-patient rehabilitation offered by hospitals and short-stay admissions as ignoring
424 these admissions will artificially inflate home-time for patients with residual, ongoing
425 disability.

426 Coding inaccuracy is a potential limitation although accuracy is often high [21]. As
427 previously discussed [18], stroke coding in Australia has been validated in several studies
428 showing high levels of accuracy in terms of positive predictive value, inter-rater coder
429 reliability and sensitivity. Comorbidities are only coded if these require treatment in hospital
430 and therefore it is likely comorbidities were under-numerated, although applying a look-back
431 period reduces under-numeration [41]. The apparent protective effects of stroke history in
432 subarachnoid haemorrhage and uncomplicated diabetes in intracerebral haemorrhage may
433 reflect less severe aetiology (for example, non-aneurysmal aetiology for subarachnoid
434 haemorrhage) or differential ascertainment of comorbidities in patients with dissimilar
435 prognoses. Specifically, under-ascertainment may be more likely in patients who die within
436 hours of admission or who are acutely unwell as the management of chronic comorbidities is
437 not central to the care of critically ill patients. In contrast, comorbidities are relevant to
438 relatively well patients and may be preferentially recorded in surviving patients.

439 FIM scores were obtained from in-patient rehabilitation records for the sub-group of patients
440 receiving publicly funded rehabilitation who, by definition, had disability. These results may
441 not generalise to patients undergoing rehabilitation as an outpatient or at private facilities.

442 Home-time was not calculated for the 4.2% of patients either residing or transferred interstate
443 or for those with missing hospitalisation data, and the impact of this source of missing data
444 could not be assessed here. Anticipated national data linkage may eliminate shortcomings
445 associated with cross-jurisdictional patient flows [42] while improved data quality will
446 mitigate against data loss. We could not identify cases entering residential care who later
447 returned to a private residence or patients admitted to supported care facilities after discharge
448 to home, although the short 90-day follow-up period may have minimised such
449 misclassification. We did not assess stroke severity using gold-standard measures as these are
450 not available using administrative data, although the GCS is integral to measuring severity in

451 subarachnoid haemorrhage [27] and is strongly predictive of mortality for ischaemic stroke
452 [19, 26] and intracerebral haemorrhage [26].

453 Loss of consciousness may be an inexact proxy for severity. Very low GCS at first point of
454 contact may indicate temporary loss of consciousness due to syncope or seizure, rather than
455 stroke severity. This may account for the higher home-time estimates observed in
456 haemorrhagic stroke patients – only in those surviving their admission – whose initial GCS
457 scores fell within the most severe range compared with those whose scores were within the
458 penultimate severity category. While GCS and mode of arrival predict mortality after stroke,
459 the GCS is a weaker correlate of home-time than the National Institutes of Health Stroke
460 Scale (NIHSS) [14], likely reflecting its less sensitive assessment of neurological deficits.

461 **4.4 Implications for future research:** Definitions of home-time have been adapted to suit
462 the scope of administrative data sources at hand. For example, here and previously [14],
463 home-time was defined as the number of full days not spent in hospital while others have
464 calculated home-time as the number of nights spent at home [11]. Where available,
465 residential care data provide reliable information about supportive care arrangements [11]
466 whereas we and others [14] relied on hospital disposition coding to identify residential care
467 admissions. Efforts are needed to harmonise the definition of home-time to aid multicentre
468 and multinational studies and comparisons.

469 As noted by others [15], home-time may not accurately reflect disability for patients who
470 receive assistance in their private residence nor does it differentiate between patients
471 according to the level of support required in residential care. Patients discharged home who
472 then die during the follow-up period may have equivalent home-time values as those who
473 survive to 90-days with lengthy in-hospital stays. Further, discharge to nursing home care,
474 death during the stroke admission and hospitalisation for the entire follow-up period yield

475 equivalent home-time estimates of zero days. Whether the validity of home-time can be
476 improved by adding qualitative information to achieve a more nuanced disability measure
477 merits future research. Large differences in median values between the main cohort and the
478 sub-set of patients surviving their stroke admission highlights the sizeable influence deaths
479 have on home-time estimates. For this reason, we and others report descriptive analyses
480 excluding in-patient deaths.

481 **4.5 Conclusion:** In conclusion, these findings support home-time ascertained using
482 administrative data as a robust, inexpensive, replicable and valid patient-centred outcome that
483 facilitates population-based stroke surveillance.

484 **Acknowledgements:** We thank the Centre for Health Record Linkage, for data linkage
485 services and the NSW Ministry of Health and Ambulance NSW for supplying data for these
486 analyses.

487 **Funding:** This work was supported by the New South Wales Ministry of Health, Office for
488 Health and Medical Research, New South Wales Neurological Conditions Translational
489 Research Grants Program NSW Neurological Conditions Translational Research Grants
490 Program. The study sponsor had no role in the study design, in the collection, analysis and
491 interpretation of data, in the writing of the report and in the decision to submit the article for
492 publication.

493 **Declaration of Interests:** JW serves as a Board Member, NSW Ministry of Health, Bureau
494 of Health Information. The Bureau of Health Information utilises linked and unlinked
495 routinely collected health data to evaluate patient outcomes in NSW hospitals. JW was a paid
496 consultant for the Agency for Clinical Innovation, NSW Ministry of Health (2014-2016) to
497 lead a project to reduce unwarranted clinical variation in stroke across the NSW Health
498 Service.

499 **Data statement:** Data analysed as part of this study are third-party owned. Signed
500 agreements between the researchers and data custodians preclude sharing of individual level
501 data in adherence to ethical and legal restrictions that apply to the use of these data-sets.
502 Requests to access these third-party owned data-sets by qualified researchers can be directed
503 to the Centre for Health Record Linkage (www.cherel.org.au). Requests for the programming
504 code can be directed to the corresponding author.

505 **References:**

- 506 1. Global Burden of Disease Collaborators. Global, regional, and national burden of
507 stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study
508 2016. *Lancet Neurology*, 2019;18:439-58. [http://dx.doi.org/10.1016/S1474-](http://dx.doi.org/10.1016/S1474-4422(19)30034-1)
509 [4422\(19\)30034-1](http://dx.doi.org/10.1016/S1474-4422(19)30034-1).
- 510 2. NCD Risk Factor Collaboration (NCD-RisC). Long-term and recent trends in
511 hypertension awareness, treatment, and control in 12 high-income countries: an
512 analysis of 123 nationally representative survey. *Lancet*, 2019;
513 [http://dx.doi.org/10.1016/S0140-6736\(19\)31145-6](http://dx.doi.org/10.1016/S0140-6736(19)31145-6)
- 514 3. Sluggett JK, Caughey GE, Ward MB, Gilbert AL. Use of secondary stroke prevention
515 medicines in Australia: national trends, 2003-2009. *Medical Journal of Australia*
516 2014; 201:54-57. doi:10.5694/mja13.00186
- 517 4. Marshall IJ, Wang Y, McKevitt C, Rudd AG, Wolfe CDA. Trends in risk factor
518 prevalence and management before first stroke. Data from the South London Stroke
519 Register 1995-2011. *Stroke* 2013;44:1809-1816.
- 520 5. Bergström L, Irewall AL, Söderström L, Ögren J, Laurell K, Moee T. One-year
521 incidence, time trends and predictors of recurrent ischaemic stroke in Sweden from
522 1998 to 2010. An observational study. *Stroke*, 2017; 48:2046-2051.
523 <https://doi.org/10.1161/STROKEAHA.117.016815>

- 524 6. Opeolu A, Nyström KV, Yavagal DR, Luciano J, Nogueira RG, Zorowitz RD,
525 KHalessi AA, Bushness C, Barsan WG, Panagos P, Alberts MJ, Tiner C, Schwamm
526 LH, Jauch EC. Recommendations for the establishment of stroke systems of care: A
527 2019 Update. *Stroke*, 2019;50:e187-e210.
528 <https://doi.org/10.1161/STR.0000000000000173>
- 529 7. Campbell BCV. *Advances in stroke medicine*. 2019;210:doi:10.5694/mja2.50137
- 530 8. Norrving B, Barrick J, Davalos A, Dichgan M, Cordonnier C, Guekht A, Kutluk K,
531 Mikulik R, Wardlaw J, Richard E, Nabav D, Molina C, Bath PM, Sunnerhagen KS,
532 Rudd A, Drummond A, PPlanas A, Caso V. Action plan for stroke in Europe 2018-
533 2030. *European Stroke Journal*,2018;3:309-336.
- 534 9. Connolly ES, Rabinstein AA, Carhuapoma R, Derdeyn CP, Dion J, Higashida RT,
535 Hoh BL, Kirkness CJ, Naidech AM, Ogilby CS, Patel AB, Thompson G, Vespa P.
536 Guidelines for the management of aneurysmal subarachnoid hemorrhage. *Stroke*.
537 2012;43:1711-1737.
- 538 10. Kleindorfer, D. The bad news. Stroke incidence is stable. *Lancet Neurology*, 2007; 6:
539 470-471. DOI:10.1016/S1474-4422(07)70111-4
- 540 11. Yu AYX, Rogers E, Wang M, Sajobi TT, Coutts SB, et al. Population-based study of
541 home-time by stroke type and correlation with modified Rankin Score. *Neurology*,
542 2017;89:1970-1976. DOI:<https://doi.org/10.1212/WNL.00000000000004631>.
- 543 12. Stienen MN, Smoll NR, Fung C, Goldberg J, Bervini D, et al. Home-time as a surrogate
544 marker for functional outcome after aneurysmal subarachnoid haemorrhage. *Stroke*
545 2018;49:3081-3084. DOI:10.1161/STROKEAHA.118.022808.
- 546 13. Quinn TJ, Dawson J, Lees JS, Chang TP, Walters MR et al. Time spent at home
547 poststroke: “Home-Time” a meaningful and robust outcome measure for stroke trials.
548 *Stroke*, 2008;39:231-233. DOI:10.1161/STROKEAHA.107.493320.

- 549 14. McDermid I, Barber M, Dennis M, Langhorne P, Macleod MJ, et al. Home-time is a
550 feasible and valid stroke outcome measure in national datasets. *Stroke* 2019;50:00-00.
551 Doi:10.1161/StrokeAHA.118.023916.
- 552 15. Fonarow GC, Liang L, Thomas L, Xian Y, Saver JL et al. Assessment of home-time
553 after acute ischaemic stroke in Medicare beneficiaries. *Stroke*, 2016;47:836-842.
554 DOI:10:1161/STROKEAHA.115.011599
- 555 16. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the Modified Rankin
556 Scale: a systematic review. *Stroke*, 2009;40:3393-3395.
557 <https://doi.org/10.1161/STROKEAHA.109.557256>
- 558 17. McArthur K, Beagan MLC, Degnan A, Howarth RC, Mitchell KA, McQuaige FB,
559 Shannon MAC, Stott DJ, Quinn TJ. Properties of proxy-derived modified Rankin
560 Scale assessments. *International Journal of Stroke*, 2012; 8;403-407.
561 <https://doi.org/10.1111/j.1747-4949.2011.00759.x>
- 562 18. Worthington JM, Gattellari M, Goumas C, Jalaludin B. Differentiating incident from
563 recurrent stroke using administrative data: the impact of varying lengths of look-back
564 periods on the risk of misclassification. *Neuroepidemiology* 2017;48:111-118 doi:
565 10.1159/000478016
- 566 19. Gattellari M, Goumas C, Jalaludin B, Worthington J. The impact of disease severity
567 adjustment on hospital standardised mortality ratios: Results from a service-wide
568 analysis of ischaemic stroke admissions using linked pre-hospital, admissions and
569 mortality data. *PLoS ONE* 2019; 14:e0216325.
570 <https://doi.org/10.1371/journal.pone.0216325>
- 571 20. Australian Consortium for Classification Development. International Statistical
572 Classification of Diseases and Health Related Problems and Australian Coding
573 Standards for International Statistical Classification of Diseases and Related Health

- 574 Problems. Tenth Revision, Australian Modification, 2015. Independent Pricing
575 Authority (Ninth Edition).
- 576 21. McCormick N, Bhole V, Lacaille D, Avina-Zubieta JA. Validity of diagnostic codes
577 for acute stroke in administrative databases: a systematic review. PLoS ONE 2015; 10:
578 e0135834. doi:10.1371/journal.pone.0135834
- 579 22. Hemphill JC, Adeoye OM, Alexander DN, Amin-Hanjani S, Cushman M, George
580 MG, LeRoux PD, Mayer SA, Qureshi AI, Saver JL, Schwamm LH, Sheth KN,
581 Tirschwell D. Clinical performance measures for adults hospitalized with
582 intracerebral haemorrhage: performance measures for healthcare professionals from
583 the American Heart Association/American Stroke Association. Stroke; 49:e243-e261.
584 <https://doi.org/10.1161/STR.0000000000000171>
- 585 23. Lichtman JH, Jones SB, Leifheit-Limson EC, Wang Y, Goldstein LB. 30-day
586 mortality and readmission after hemorrhagic stroke among Medicare beneficiaries in
587 Joint Commission primary stroke center-certified and noncertified hospitals. Stroke
588 2011;42:3387-3391 doi: 10.1161/STROKEAHA.111.622613
- 589 24. Bardach NS, Zhao S, Gress DR, Lawton MT, Johnston C. Association between
590 subarachnoid hemorrhage outcomes and number of cases treated at California
591 Hospitals. Stroke, 2002; 33:1851-1856.
592 <https://doi.org/10.1161/01.STR.0000019126.43079.7B>
- 593 25. Krumholz HM, Brindis RG, Brush JE, Cohen DJ, Epstein AJ, Furie K, et al.
594 Standards for statistical models used for public reporting of health outcomes. An
595 American Heart Association Scientific Statement from the Quality of Care and
596 Outcomes Research Interdisciplinary Writing Group. Circulation, 2006;113:456-62
597 doi: 10.1161/CIRCULATIONAHA.105.170769

- 598 26. Weir CJ, Bradford AP, Lees KR. The prognostic value of the components of the
599 Glasgow Coma Scale following acute stroke. *QJM*, 2003;96:67-74.
600 <https://doi.org/10.1093/qjmed/hcg008>
- 601 27. Report of World Federation of Neurological Surgeons Committee on aneurysmal
602 Subarachnoid Hemorrhage Grading Scale. *J Neurosurg*. 1988;68:985–986
- 603 28. Fonarow GC, Pan W, Saver JL, Smith EE, Broderick JP et al. Relationship of
604 National Institutes of Health Stroke Scale to 30-day mortality in Medicare
605 Beneficiaries with Acute Ischemic Stroke. *JAHA*;2012;00:1-9. doi:
606 10.1161/JAHA.111.000034
- 607 29. Simpson AN, Wilmskoetter J, Hong I, Li, C-Y, Juach EC, Bonilha HS, Anderson K,
608 Harvey J, Simpson KN. Stroke administrative severity index: using administrative
609 data for 30-day post-stroke outcomes prediction. *Journal of Comparative*
610 *Effectiveness Research*, 2018;7:2930304.
- 611 30. Kidd D, Stewart G, Baldry J, Johnson J, Rossiter D et al. The functional independence
612 measure: a comparative validity and reliability study. *Disability Rehabilitation*.
613 1995;17:10–14
- 614 31. Australasian Rehabilitation Outcomes Centre. The Annual AROC Report: the state of
615 Rehabilitation. 2009, 2010, 2011, 2012, 2013.
616 <https://ahsri.uow.edu.au/aroc/archivedreports/index.html> last accessed July 28, 2019.
- 617 32. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding
618 algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data.
619 *Medical Care* 2005;43:1130-9 doi: 10.1097/01.mlr.0000182534.19832.83
- 620 33. Goldstein LB, Samsa GP, Matchar DB, Horner RD. Charlson Index Comorbidity
621 Adjustment for Ischemic Stroke Outcome Studies. *Stroke* 2004;35:1941-1945. doi:
622 10.1161/01.STR.0000135225.80898.1c

- 623 34. Australian Bureau of Statistics. Socioeconomic Indexes for Areas (SEIFA). Technical
624 Paper 2006. Canberra, Commonwealth of Australia, 2008.:
625 <http://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0.55.001/>.
- 626 35. Australian Bureau of Statistics.. Available from:
627 <http://www.abs.gov.au/websitedbs/censushome.nsf/home/nonresponseratesLast>
628 accessed 28th July 2019.
- 629 36. O'Brien EC, Xian Y, Xu H, Wu J, Saver JL, Smith EE, Schwamm LH, Peterson ED,
630 Reeves MJ, Bhatt DL, Maisch L, Hannah D, Lindholm B, Olson D, Bettger JP,
631 Pencina M, Hernandez AF, Fonarow GC. Hospital variation in home-time after acute
632 ischemic stroke: insights from the PROPSEER study (Patient-Centred Researcher Into
633 Outcomes Stroke Patients Prefer and Effectiveness Research). *Stroke* 2016;47:2627-
634 2633. DOI:10.1161/STROKEAHA.116.013563.
- 635 37. Gall SL, DOnnan G, Dewey HM, Macdonnell R, Sturm J, Gilligan A, Srikanth V,
636 Thrift AG. Sex differences in presentation, severity, and management of stroke in a
637 population-based study. *Neurology* 2010, 74:975-81.
638 doi:10.1212/WNL.0b013e3181d5a48f.
- 639 38. Banks JL, Marotta CA. Outcomes validity and reliability of the Modified Rankin
640 scale: Implications for stroke clinical trials. *Stroke*, 2007;38:1091-1096.
641 <https://doi.org/10.1161/01.STR.0000258355.23810.c6>Stroke. 2007;38:1091–1096
- 642 39. Mishra NK, Shuaib A, Lyden P, Diener HC, Grotta J, Davis S, Davalos A, Ashwood
643 T, Wasiewski W, Lees KR. Home time is extended in patients with ischemic stroke
644 who receive thrombolytic therapy: a validation study of home time as an outcome
645 measure. *Stroke* 2011;42:1046-1050. doi: 10.1161/STROKEAHA.110.601302
- 646 40. Kapoor A, Lanctôt KL, Bayley M, Kiss A, Herrmann N. “Good outcome” isn’t good
647 enough. Cognitive impairment, depressive symptoms, and social restrictions in

648 physically recovered stroke patients. *Stroke*, 2017;48:1688-1690.

649 <https://doi.org/10.1161/STROKEAHA.117.016728>

650 41. Preen DB, Holman CD, SPilsbury K, Semmens JB, Brameld KJ. Length of

651 comorbidity lookback period affected regression model performance of administrative

652 health data. *Journal of Clinical Epidemiology* 2006; 59:940-6.

653 DOI:10.1016/j.jclinepi.2005.12.013

654 42. Australian Institute of Health and Welfare 2018. Australian Institute of Health and

655 Welfare Corporate Plan 2018–19 to 2021–22. Cat. no. AUS 224. Canberra: AIHW.

656 **Supplementary Material**

657 **Title:** Measuring stroke outcomes using linked administrative data: Population-based
658 estimates and validation of “home-time” as a surrogate measure of functional status.

659

660

661 **Authors:**

662 Name: Melina Gattellari*^{a,b}

663 Email: melina.gattellari@health.nsw.gov.au

664

665 Name: Chris Goumas^{a,c}

666 Email: chris.goumas@sydney.edu.au

667

668 Name: Bin Jalaludin^{a,d,e}

669 Email: bin.jalaludin@health.nsw.gov.au

670

671 Name: John Worthington^{a,b,f}

672 Email: John.Worthington@health.nsw.gov.au

673

674 Affiliations: ^aIngham Institute for Applied Medical Research,
675 1 Campbell Street, Liverpool, NSW, 2170, Australia

676

677 ^bDepartment of Neurology, Royal Prince Alfred Hospital,
678 Missenden Road, Camperdown, NSW, 2050, Australia.

679

680 ^cUniversity of Sydney, School of Public Health, University of Sydney

681

682 ^dPopulation Health Intelligence, Healthy People and Places Unit; South
683 Western Sydney Local Health District, Liverpool, Sydney, New South
684 Wales, Australia

685

686 ^eSchool of Public Health, The University of New South Wales,
687 Kensington, Sydney, New South Wales, Australia

688

689 ^fSouth Western Sydney Clinical School, The University of New South
690 Wales, Liverpool, Sydney, New South Wales, Australia

691

692 *Corresponding author.

693 **Supplementary Table 1: The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported**
 694 **in observational studies using routinely collected health data.**

695

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	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 (administrative data is specified).</p> <p>Abstract, Line 43.</p> <p>Title; Abstract, lines 43-44.</p>
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction, Lines 64-98.		Introduction, Lines 64-98.
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction, Lines 92-98.		

Methods					
Study Design	4	Present key elements of study design early in the paper	Methods, Lines 102-111		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods, Lines 112-120.		
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>Not applicable: this is neither a cohort nor a case-control study</p> <p>Not applicable: this is neither a cohort nor a case-control study.</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>Methods, lines 112-141.</p> <p>Methods, Lines 115-127; citation to relevant references 18,19,21,22</p> <p>Figure 1.</p>

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Methods, Lines 142-151 (outcome described); Lines 162-175, incorporating reference 19 describing previous validation; Lines 176 to 185; Lines 203 to 207.	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.	Methods, Lines 142-151 (outcome described); Lines 162-175, incorporating reference 19 describing previous validation; Lines 176 to 185; Lines 203 to 207.
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods, Lines 142-151 (outcome described); Lines 162-175, incorporating reference 19 describing previous validation; Lines 176 to 185; Lines 203 to 207.		
Bias	9	Describe any efforts to address potential sources of bias	Lines 132 to 141.		
Study size	10	Explain how the study size was arrived at	Methods, line 112 to 114 (all available cases)		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Methods, Statistical analysis. Lines 186 to 192; Lines 210 to 216.		

Statistical methods	12	<p>(a) Describe all statistical methods, including those used to control for confounding</p> <p>(b) Describe any methods used to examine subgroups and interactions</p> <p>(c) Explain how missing data were addressed</p> <p>(d) <i>Cohort study</i> - If applicable, explain how loss to follow-up was addressed</p> <p><i>Case-control study</i> - If applicable, explain how matching of cases and controls was addressed</p> <p><i>Cross-sectional study</i> - If applicable, describe analytical methods taking account of sampling strategy</p> <p>(e) Describe any sensitivity analyses</p>	<p>Methods, Statistical Analysis, Lines 186 to 217.</p> <p>Figure 1.</p> <p>Not applicable.</p> <p>Not applicable</p> <p>Not applicable</p> <p>e) Lines 189-192.</p>		.
Data access and cleaning methods		..		<p>RECORD 12.1: Authors should describe the extent to which the investigators had access to the 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>	<p>Methods, Line 104 to 111.</p> <p>Data statement, Lines 499 to 503.</p> <p>Methods, Line</p>
Linkage		..		RECORD 12.3: State whether the study included person-level,	Figure 1, Lines 105 to 111.

				institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	
Results					
Participants	13	(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) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	a) Figure 1. b) Lines 222 to 228. c) Figure 1.	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.	Figure 1, Lines 112 to 136.
Descriptive data	14	(a) Give characteristics of study participants (<i>e.g.</i> , demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) <i>Cohort study</i> - summarise follow-up time (<i>e.g.</i> , average and total amount)	a) Supplementary Tables 2-5. b) Figure 1, Supplementary Tables 2-5. c) Not applicable		
Outcome data	15	<i>Cohort study</i> - Report numbers of outcome events or summary measures over time	Tables 1 and 2.		

		<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>			
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 1 to 3. Supplementary Tables 2 to 5.		
Other analyses	17	Report other analyses done— e.g., analyses of subgroups and interactions, and sensitivity analyses	Tables 1 to 3 (removing patients who died in hospital).		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Lines 361 to 405.		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision.	Lines 419 to 460	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include	Lines 426 to 460. Lines 461 to 480.

		Discuss both direction and magnitude of any potential bias		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	Discussion.		
Generalisability	21	Discuss the generalisability (external validity) of the study results	Lines 406 to 418; 461 to 480.		
Other Information					
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	Lines 487 to 492.		
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.	Lines 499 to 504.

696

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

700

701 *Checklist is protected under Creative Commons Attribution ([CC BY](#)) license.

702 **Supplementary Table 2: Patient characteristics (N=74,501)**

Characteristics	N	%
Sub-type		
Ischaemic Stroke	57,949	77.8
Intracerebral Haemorrhage	12,192	16.4
Subarachnoid Haemorrhage	4,360	5.9
Sex		
Male	37,404	50.2
Female	37,097	49.8
Age at diagnosis (years)		
<50	5,205	7.0
50-64	11,520	15.5
65-74	14,692	19.7
75-84	24,731	33.2
85+	18,353	24.6
Stroke Severity (IS and ICH); N=70,141**		
Ambulant	23,751	33.9
GCS score=15	25,429	36.3
GCS score=13-14	9,001	12.8
GCS score=9-12	7,219	10.3
GCS score=3-8	4,215	6.0
Missing*	526	0.7
Stroke Severity (SAH); N=4,360**		
Ambulant	1,600	36.7
GCS score=15	1,297	29.7
GCS score=13-14	392	9.0
GCS score=7-12	459	10.5
GCS score=3-6	585	13.4
Missing*	27	0.6

703 *Patients arriving by ambulance with missing GCS (Glasgow Coma Scale) scores

704 **Denominator used for calculation of percentages

705 **Supplementary Table 3: Adjusted* mean days spent at home: Ischaemic**
 706 **Stroke (n=56,445)***

Variable	N	Mean (95% CI)	P-value
Age (years; quadratic effect)	56,445	-	<0.002
Sex			
Males	28,952	41 (41-42)	
Females	27,493	41 (40-41)	0.02
Year			
2005	6,563	38 (37-39)	
2006	6,266	41 (39-42)	
2007	6,094	41 (40-43)	
2008	6,119	41 (40-42)	
2009	5,992	41 (40-43)	
2010	5,787	42 (41-44)	
2011	6,041	41 (39-42)	
2012	6,157	41 (40-42)	
2013	6,053	43 (41-44)	
2014	1,373	42 (40-44)	<0.001
SES (IRSD)*			
Q1 (lowest)	13,420	41 (39-42)	
Q2	11,429	40 (39-41)	
Q3	11,302	41 (39-42)	
Q4	9,760	41 (40-43)	
Q5 (highest)	10,534	42 (41-43)	0.05
Marital status			
Married/de facto	28,526	43 (42-43)	
Other	27,919	39 (38-40)	<0.001
Previous stroke			
No	53,007	41 (40-42)	
Yes	3,438	43 (41-44)	0.07
Stroke severity			
Ambulant	20,036	51 (50-52)	
GCS score=15	21,416	46 (45-47)	
GCS score=13-14	7,056	36 (35-38)	
GCS score=9-12	5,498	23 (22-24)	
GCS score=3-8	2,439	12 (11-14)	<0.001

Variable	N	Mean (95% CI)	P-value
Charlson Comorbidities†			
Myocardial infarction			
No	49,429	42 (41-42)	
Yes	7,016	37 (36-38)	<0.001
Congestive heart failure			
No	47,579	42 (42-43)	
Yes	8,866	34 (33-35)	<0.001
Peripheral vascular disease			
No	51,315	41 (40-42)	
Yes	5,130	39 (37-40)	<0.001
Dementia			
No	50,060	43 (42-43)	
Yes	6,385	31 (29-32)	<0.001
Chronic pulmonary disease			
No	50,590	41 (40-42)	
Yes	5,855	39 (38-41)	0.002
Connective tissue disease			
No	55,382	41 (40-42)	
Yes	1,063	39 (37-42)	0.12
Peptic ulcer			
No	55,040	41 (40-42)	
Yes	1,405	36 (34-38)	<0.001
Mild liver disease			
No	55,589	41 (40-42)	
Yes	856	38 (36-40)	0.01
Diabetes without complications			
No	47,436	41 (40-42)	
Yes	9,009	40 (38-41)	0.01
Diabetes with complications			
No	45,663	41 (40-42)	
Yes	10,782	40 (39-41)	0.001
Renal disease			
No	49,977	42 (41-42)	
Yes	6,468	36 (35-38)	<0.001

Variable	N	Mean (95% CI)	P-value
Non-metastatic malignancy			
No	51,775	41 (40-42)	
Yes	4,670	38 (37-40)	<0.001
Moderate/severe liver disease			
No	56,179	41 (40-42)	
Yes	266	30 (26-36)	<0.001
Metastatic malignancy			
No	54,925	41 (41-42)	
Yes	1,520	27 (25-30)	<0.001

707 *Adjusted for all other variables in table and days spent in hospital in the 12 months before the
708 stroke admission. Analyses adjusted for interdependencies of data due to clustering of cases within
709 hospitals. Complete case analysis undertaken; excludes 1,504 (2.6%) cases with missing covariate
710 data.

711 **SES=socioeconomic status measured using IRSD=Index of Relative Socio-economic
712 Disadvantage³⁴

713 †Charlson comorbidities excluded cerebrovascular diseases and hemiplegia as redundant following
714 example of Goldstein et al³³; HIV/AIDs was also given low prevalence.

715 **Supplementary Table 4: Adjusted* mean days spent at home: ICH**
 716 **(n=11,793)**

Variable	N	Mean (95% CI)	P-value
Age (quadratic)			<0.001
Sex			
Males	5,888	21 (20-22)	
Females	5,905	21 (20-22)	0.89
Year			
2005	1,284	17 (15-19)	
2006	1,355	21 (19-24)	
2007	1,188	20 (18-23)	
2008	1,285	23 (21-25)	
2009	1,265	22 (20-25)	
2010	1,267	23 (20-25)	
2011	1,294	21 (19-24)	
2012	1,295	20 (18-22)	
2013	1,261	20 (19-23)	
2014	299	18 (15-22)	<0.001
SES (IRSD)**			
Q1 (lowest)	2,846	21 (19-23)	
Q2	2,217	21 (19-23)	
Q3	2,259	21 (20-23)	
Q4	2,062	20 (18-22)	
Q5 (highest)	2,409	21 (19-23)	0.85
Marital status			
Married/de facto	6,360	21 (20-22)	
Other	5,433	21 (19-22)	0.53
Previous stroke			
No	10,590	21 (20-22)	
Yes	1,203	22 (20-25)	0.28
Stroke severity			
Ambulant	3,265	38 (36-40)	
GCS score=15	3,615	31 (30-33)	
GCS score=13-14	1,744	23 (21-24)	
GCS score=9-12	1,524	11 (10-13)	
GCS score=3-8	1,645	4 (3-5)	<0.001

Variable	N	Mean (95% CI)	P-value
Charlson Comorbidities†			
Myocardial infarction			
No	10,693	21 (20-22)	
Yes	1,100	18 (15-20)	0.006
Congestive heart failure			
No	10,530	21 (20-22)	
Yes	1,263	17 (14-20)	0.007
Peripheral vascular disease			
No	10,926	21 (20-22)	
Yes	867	18 (16-20)	0.01
Dementia			
No	10,361	21 (20-22)	
Yes	1,432	19 (16-23)	0.42
Chronic pulmonary disease			
No	10,741	21 (20-22)	
Yes	1,052	20 (17-23)	0.50
Connective tissue disease			
No	11,596	21 (20-22)	
Yes	197	20 (14-28)	0.81
Peptic ulcer			
No	11,510	21 (20-22)	
Yes	283	22 (17-30)	0.60
Mild liver disease			
No	11,418	21 (20-22)	
Yes	375	20 (14-27)	0.70
Diabetes without complications			
No	10,251	20 (19-21)	
Yes	1,542	23 (21-26)	0.03
Diabetes with complications			
No	9,986	21 (20-22)	
Yes	1,807	20 (18-23)	0.74
Renal disease			
No	10,621	21 (20-22)	
Yes	1,172	18 (15-21)	0.05

Variable	N	Mean (95% CI)	P-value
Non-metastatic malignancy			
No	10,750	21 (20-22)	
Yes	1,043	17 (15-19)	<0.001
Moderate/severe liver disease			
No	11,636	21 (20-22)	
Yes	157	17 (11-27)	0.35
Metastatic malignancy			
No	11,531	21 (20-22)	
Yes	262	27 (19-38)	0.16

717 *Adjusted for all other variables in table and days spent in hospital in the 12 months before the
718 stroke admission. Analyses adjusted for interdependencies of data due to clustering of cases within
719 hospitals. Complete case analysis undertaken; excludes 399 (3.3%) cases with missing covariate
720 data.
721 **SES=socioeconomic status measured using IRSD=Index of Relative Socio-economic
722 Disadvantage³⁴
723 †Charlson comorbidities excluded cerebrovascular diseases and hemiplegia as redundant following
724 example of Goldstein et al;³³ HIV/AIDs was also excluded given low prevalence.

725 **Supplementary Table 5: Adjusted* mean days spent at home: SAH (n=4,222)**

Variable	N	Mean (95% CI)	P-value
Age (quadratic)			<0.001
Sex			
Males	1,644	31 (29-33)	
Females	2,578	30 (29-32)	0.38
Year			
2005	401	25 (22-29)	
2006	438	31 (26-36)	
2007	446	31 (27-35)	
2008	476	33 (30-37)	
2009	464	30 (26-34)	
2010	430	30 (27-34)	
2011	476	29 (26-32)	
2012	479	32 (29-35)	
2013	489	33 (29-38)	
2014	123	30 (23-39)	0.04
SES (IRSD)**			
Q1 (lowest)	929	32 (28-36)	
Q2	709	30 (27-34)	
Q3	842	31 (28-35)	
Q4	839	29 (27-31)	
Q5 (highest)	903	30 (27-33)	0.53
Marital status			
Married/de facto	2,553	32 (30-34)	
Other	1,669	28 (27-30)	0.02
Previous stroke			
No	4,035	30 (29-32)	
Yes	187	40 (32-49)	0.02
Stroke severity			
Ambulant	1,562	47 (44-50)	
GCS score=15	1,277	45 (43-47)	
GCS score=13-14	380	29 (25-33)	
GCS score=9-12	446	15 (12-18)	
GCS score=3-8	557	7 (5-9)	<0.001

Variable	N	Mean (95% CI)	P-value
Charlson Comorbidities†			
Myocardial infarction			
No	4,007	31 (29-32)	0.12
Yes	215	25 (20-32)	
Congestive heart failure			
No	3,958	31 (29-32)	0.48
Yes	264	28 (21-36)	
Peripheral vascular disease			
No	4,010	31 (29-32)	0.88
Yes	212	30 (22-41)	
Dementia			
No	4,080	30 (29-32)	0.89
Yes	142	31 (21-48)	
Chronic pulmonary disease			
No	3,976	31 (30-32)	0.09
Yes	246	25 (19-32)	
Connective tissue disease			
No	4,169	30 (29-32)	0.11
Yes	53	38 (29-50)	
Peptic ulcer			
No	4,165	31 (29-32)	0.21
Yes	57	24 (17-35)	
Mild liver disease			
No	4,142	31 (29-32)	0.93
Yes	80	30 (21-42)	
Diabetes without complications			
No	3,955	30 (29-32)	0.15
Yes	267	36 (28-47)	
Diabetes with complications			
No	3,901	31 (29-32)	0.53
Yes	321	28 (23-35)	
Renal disease			
No	4,054	31 (29-32)	0.03
Yes	168	23 (18-29)	

Variable	N	Mean (95% CI)	P-value
Non-metastatic malignancy			
No	4,041	31 (29-32)	0.46
Yes	181	28 (21-36)	
Moderate/severe liver disease			
No	4,196	31 (29-32)	0.03
Yes	26	16 (8-29)	
Metastatic malignancy			
No	4,165	30 (29-32)	0.69
Yes	57	33 (22-49)	

726 *Adjusted for all other variables in table and days spent in hospital in the 12 months before the
727 stroke admission. Analyses adjusted for interdependencies of data due to clustering of cases within
728 hospitals. Complete case analysis undertaken; excludes 138 (3.2%) cases with missing covariate
729 data.
730 **SES=socioeconomic status measured using IRSD=Index of Relative Socio-economic
731 Disadvantage³⁴
732 †Charlson comorbidities excluded cerebrovascular diseases and hemiplegia as redundant following
733 example of Goldstein et al;³³ HIV/AIDs was also excluded given low prevalence.