

BMJ Open

Variations in the recording of common health conditions in Australian hospitals

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2014-005768
Article Type:	Research
Date Submitted by the Author:	26-May-2014
Complete List of Authors:	Lujic, Sanja; University of Western Sydney, Centre for Health Research Watson, Diane; National Health Performance Authority, Randall, Deborah; University of Western Sydney, Centre for Health Research Simpson, Judy; University of Sydney, School of Public Health Jorm, Louisa; University of Western Sydney, Centre for Health Research; The Sax Institute,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Epidemiology
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

Variations in the recording of common health conditions in Australian hospitals

Sanja Lujic¹, Diane E Watson², Deborah A Randall¹, Judy M Simpson³, Louisa R Jorm^{1,4}

1. Centre for Health Research, University of Western Sydney, Sydney, Australia.

2. National Health Performance Authority, Sydney, Australia.

3. School of Public Health, University of Sydney, Sydney, Australia.

4. The Sax Institute, Sydney Australia.

Correspondence:

Ms Sanja Lujic

Centre for Health Research, University of Western Sydney, Locked Bag 1797, Penrith NSW 2751, Australia

E: s.lujic@uws.edu.au

T: +61 4620 3684

F: +61 4620 3299

Word count: 4,302 words

Abstract

Objectives: To investigate the agreement between self-report and recording of six morbidities in administrative hospital data, quantify the between-hospital variation and identify predictors of positive agreement between the two data sources.

Setting and participants: Retrospective analysis of linked self-report and administrative hospital data for 32,832 participants in the large-scale cohort study (45 and Up Study), who joined the study from 2006-2009 and who were admitted to 313 hospitals in New South Wales, Australia, for an overnight stay, up to a year prior to study entry.

Outcome measures: Agreement between self-report and administrative hospital data and positive agreement among participants who self-reported any of the six morbidities.

Results: Agreement between data sources was good for diabetes ($\kappa=0.79$), moderate for smoking ($\kappa=0.59$), fair for heart disease, stroke and hypertension ($\kappa=0.40$, $\kappa=0.30$, $\kappa=0.24$, respectively) and poor for obesity ($\kappa=0.09$), indicating that a large number of individuals with self-reported morbidities did not have a corresponding diagnosis coded in their hospital records. Significant between-hospital variation was found (ranging from 8% of unexplained variation for diabetes to 22% for heart disease), with higher agreement in public and large hospitals, and hospitals with greater depth of coding.

Conclusions: The recording of six common health conditions in administrative hospital data is highly variable, and for some conditions, very poor. To support more valid performance comparisons, it is important to stratify or control for factors that predict the completeness of recording, including hospital depth of coding and hospital type (public/private), and to increase efforts to standardize recording across hospitals. Studies using these conditions for risk adjustment should also be cautious of their use in smaller hospitals.

Keywords: hospital data, morbidity, multilevel modeling, agreement

Article summary

Article focus

- Explore the agreement between administrative hospital data and self-report information for four clinical conditions, as well as smoking and obesity.
- Quantify the between-hospital variation in the level of recording of these conditions.

Key messages

- Good levels of agreement found only for diabetes, with other conditions exhibiting moderate to poor agreement.
- Better recording was found in public and large hospitals, and hospitals with greater depth of coding.
- Significant between-hospital variation in the levels of agreement, amplified when smaller and private hospitals are included.

Strengths and limitations

- Use of linked data from a large-scale cohort study and advanced multilevel modeling methods to comprehensively evaluate the recording of common health conditions in hospital data, and explore between-hospital variation.
- Limitations include the absence of 'gold standard' such as medical records.

Introduction

Most nations with advanced economies publicly report on the comparative performance of hospitals with a view to accelerating and informing efforts to improve quality and allowing patients to make informed choices. Diagnoses recorded in administrative hospital data are commonly used in the construction and case-mix adjustment of hospital performance metrics, as well as for risk adjustment in epidemiological studies.

The construction of reliable health metrics relies on statistical methods that take into account the degree to which patients treated in different facilities have different morbidity and risk profiles that predispose them to requiring different interventions or to achieving different outcomes. These statistical methods, known as case-mix or risk adjustment, account for patient-related factors that are above and beyond the immediate control of health care professionals.

Thus, properly constructed performance metrics fairly reflect differences in health care experiences, patient outcomes and risks of adverse events. There has been some criticism of case-mix adjustments because they are subject to measurement error, [1] but case-mix adjustment is still considered to be less biased than unadjusted comparisons. [2]

Most methods of case-mix adjustment rely principally on demographic and diagnostic information that is captured in administrative hospital data. [3] This approach may be sub-optimal [4 5] because evidence from many countries suggests that administrative hospital data underreport the morbidity information needed to fully account for differences between hospitals in patient-related factors that predispose them to differences in measured outcomes. [6-13] However, the impact of this underreporting on comparative measures of hospital performance depends on whether it varies systematically among hospitals, because of

1
2
3 differences in factors such as training or practice among coding staff, the comprehensiveness
4
5 of clinicians' notes, or "upcoding" relating to funding models or incentives. [14]
6
7

8 This issue is relatively unexplored, aside from the work by Mohammed et al. [2] which
9
10 reported a non-constant relationship between case-mix variables and mortality among
11
12 hospitals in the UK, explained by differences in clinical coding and admission practices
13
14 across hospitals. These variations in coding accuracy were shown to be related to geographic
15
16 location and bed size, with small rural facilities performing better than large urban hospitals.
17
18 [15 16] In Australia, variations in the reporting and coding of secondary diagnoses have been
19
20 shown to exist in public hospitals among Australian states, [17] and also among hospitals
21
22 within the state of New South Wales (NSW), with worse underreporting in private and rural
23
24 hospitals. [3] However, the relative contributions of patient and hospital factors to these
25
26 variations have not been identified, nor has this variation been formally quantified.
27
28

29
30 This study aimed to further investigate the nature and potential implications of underreporting
31
32 of morbidity information in administrative hospital data, by: 1) measuring the agreement
33
34 between self-reported morbidity information and coded diagnoses; 2) quantifying the amount
35
36 of between-hospital variation in this agreement; and 3) identifying patient and hospital
37
38 characteristics that predict higher or lower levels of agreement. We focused on clinical
39
40 conditions common to case-mix and risk-adjustment models – diabetes, heart disease,
41
42 hypertension and stroke. We also focus on smoking and obesity, due to their impact on health
43
44 trajectories, rapid shifts in prevalence, substantial geographic variation in rates [18] and
45
46 paucity of international evidence on completeness of coding.
47
48
49
50

51 52 53 **Methods**

54 55 **Data sources**

56
57 *The 45 and Up Study*
58
59
60

1
2
3 The 45 and Up Study is a large-scale cohort study involving 267,153 men and women aged
4
5 45 years and over from the general population of NSW, Australia. The study is described in
6
7 detail elsewhere. [19] Briefly, participants in the 45 and Up Study were randomly sampled
8
9 from the database of Australia's universal health insurance provider, Medicare Australia,
10
11 which provides near complete coverage of the population. People 80+ years of age and
12
13 residents of rural and remote areas were oversampled. Participants joined the Study by
14
15 completing a baseline questionnaire (available at [https://www.saxinstitute.org.au/our-](https://www.saxinstitute.org.au/our-work/45-up-study/questionnaires/)
16
17 [work/45-up-study/questionnaires/](https://www.saxinstitute.org.au/our-work/45-up-study/questionnaires/)) between January 2006 and December 2009 and giving
18
19 signed consent for follow-up and linkage of their information to routine health databases.
20
21 About 18% of those invited participated and participants included about 11% of the NSW
22
23 population aged 45 years and over. [19] Exposure-outcome relationships estimated from the
24
25 45 and Up Study data have been shown to be consistent with a large random survey of the
26
27 same population. [20]
28
29
30
31
32
33
34

35 *The NSW Admitted Patient Data Collection (APDC)*

36
37 The APDC includes records of all public and private hospital admissions ending in a
38
39 separation, i.e. discharge, transfer, type-change or death. Diagnoses are coded according to
40
41 the Australian modification of the International Statistical Classification of Diseases and
42
43 Related Problems 10th Revision, ICD-10-AM. [21] Up to 55 diagnoses codes are recorded on
44
45 the APDC, including the principal diagnosis and up to 54 additional diagnoses. Additional
46
47 diagnoses are defined as 'a condition or complaint either coexisting with the principal
48
49 diagnosis or arising during the episode of care' and should be interpreted as conditions that
50
51 affect patient management. [22]
52
53
54
55
56
57
58
59
60

1
2
3 The APDC from 1 July 2000 to 31 December 2010 was linked probabilistically to survey
4 information from the 45 and Up Study by the NSW Centre for Health Record Linkage
5
6
7 (www.cherel.org.au) using the ‘best practice’ protocol for preserving privacy. [23]
8
9

10 11 12 **Study population**

13
14 The study population comprised patients aged 45 years and above who participated in the 45
15 and Up Study and who had an overnight hospitalisation up to 365 days prior to filling out the
16 baseline 45 and Up Study survey. NSW is home to 7.4 million people or one-third of the
17 population of Australia.
18
19
20
21
22
23

24 25 26 27 **Measuring morbidity**

28
29 We examined four health conditions (diabetes, heart disease, hypertension and stroke) and
30 two health risk factors (obesity and smoking), referred to hereafter collectively as
31 “morbidity”. For each participant, these health conditions were measured using self-report
32 and administrative hospital data.
33
34
35
36
37
38

39
40 Self-reported morbidities were ascertained on the basis of responses to questions in the
41 baseline 45 and Up Study survey. Diabetes, hypertension, stroke and heart disease were
42 identified using the question “Has a doctor ever told you that you have [name of condition]?”.
43
44
45 Participants who did not answer the question were excluded from analyses (n=1,242).
46
47
48

49
50 Smoking was classified on the basis of answering “yes” to both of the questions “Have you
51 ever been a regular smoker?” and “Are you a regular smoker now?”. Participants’ responses
52 to the questions “How tall are you without shoes?” and “About how much do you weigh?”
53
54
55 were used to derive body mass index (BMI), defined as body weight divided by height
56
57
58
59
60

1
2
3 squared (kg/m^2). The World Health Organization's [24] classification system was used to
4
5 categorize individuals as obese ($\text{BMI} \geq 30\text{kg}/\text{m}^2$).
6
7

8 Morbidity information in administrative hospital data was ascertained using all 55 diagnosis
9
10 codes in the APDC records (ICD-10-AM: E10-E16 for diabetes, I20-I52 for heart disease,
11
12 I60-I69, G45, G46 for stroke, I10-I15 and R03.0 for hypertension, F17.2 or Z72.0 for
13
14 smoking and E66 for obesity).
15
16

17 18 19 20 **Predictors of agreement**

21
22 We explored both patient- and hospital-level factors as predictors of agreement between the
23
24 two data sources.
25
26

27 Patient-level factors were self-reported in the 45 and Up Study baseline survey and included
28
29 age, sex, education, country of birth, income and functional limitation. Functional limitation
30
31 was measured using the Medical Outcomes Study – Physical Functioning scale, [25] and
32
33 classified into 5 groups: no limitation (score of 100), minor limitation (score 95-99), mild
34
35 limitation (score 85-94), moderate limitation (60-84) and severe limitation (score 0-59).
36
37

38 Facility-level factors were type of hospital (public/private), hospital peer group (akin to
39
40 hospital size defined by number of case-mix weighted separations, [26] which includes
41
42 hospital remoteness in the classification), remoteness of hospital and depth of coding.
43
44

45 Remoteness of the Statistical Local Area in which the hospital was located was classified
46
47 according to the Accessibility/Remoteness Index of Australia (ARIA+), grouped into four
48
49 categories (major city, inner regional, outer regional, remote/very remote). [27] Depth of
50
51 hospital coding was the mean number of additional diagnoses coded for each hospital,
52
53 calculated using all overnight hospitalizations for the full 45 and Up Study cohort from 2000
54
55 to 2010, and divided into four groups at the 25th, 50th and 75th percentile. Hospital peer
56
57
58
59
60

1
2
3 groups were divided into 5 categories: principal referral ($\geq 25,000$ separations per year), major
4 (10,000-24,999 separations per year), district (2,000-9,999 separations per year), community
5 (up to 2,000 separations per year) and other (non-acute, un-peered hospitals). Missing
6 information was treated as a separate category for any variables with missing data.
7
8
9
10

11 12 13 14 15 **Statistical methods**

16
17 We examined patient-level agreement between data sources for each of the six morbidities
18 individually, as well as for their 15 two-way combinations. We compared the self-reported
19 responses (yes/no) with all the diagnoses provided in the hospital records both for 'index'
20 admissions and for the 'lookback' period admissions. [28] The 'index' admission was the
21 overnight hospital stay with admission date closest to the survey completion date and no
22 longer than a year prior. Morbidity was coded as 'yes' if any of the diagnoses during that stay
23 contained a mention of that morbidity. The 'lookback' admissions included all overnight
24 stays in the 365-day period that preceded and included the 'index' admission. Morbidity was
25 coded as 'yes' if any of the diagnoses from any lookback admissions contained a mention of
26 that morbidity.
27
28
29
30
31
32
33
34
35
36
37
38

39 Agreement between the two data sources (yes/no) was measured using Cohen's kappa
40 statistic (κ). Kappa values above 0.75 denote excellent agreement, 0.40 to 0.75 fair to good
41 agreement and below 0.45 poor agreement. [29] Agreement was computed for all 313
42 hospitals in the state, regardless of size, as well as for the 82 largest public hospitals, for
43 which performance metrics are publicly reported.
44
45
46
47
48
49
50

51 Multilevel logistic regression was used to estimate odds ratios (OR) with 95% confidence
52 intervals (CI) for patient- and hospital-level factors that predicted positive agreement between
53 the two data sources. Multilevel models were chosen because of the clustering of patients
54
55
56
57
58
59
60

1
2
3 within hospitals. Models were run for each of the six morbidities separately. These analyses
4
5 were constrained to only those participants who self-reported the morbidity of interest, and
6
7 the outcome was whether the index hospital record contained a mention of the morbidity or
8
9 not. Addition of the hospital-level characteristics was done one at a time, due to the
10
11 collinearity between variables. All ORs presented are adjusted for all other demographic
12
13 variables in the model.
14

15
16 Variation at the hospital level was expressed as a median odds ratio (MOR), which is the
17
18 median of the odds ratios of pair-wise comparisons of patients taken from randomly chosen
19
20 hospitals, calculated as $\exp^{0.954 \times \sqrt{\text{variance}}}$; [30] and the intraclass correlation coefficient (ICC),
21
22 which is the percentage of the total variance attributable to the hospital level. [31] Large ICCs
23
24 indicate that differences among hospitals account for a considerable part of the variation in
25
26 the outcome, whereas a small ICC means that the hospital effect on the overall variation is
27
28 minimal. The relative influence of the hospital on reporting of morbidity was calculated using
29
30 a variance partitioning coefficient expressed as a percentage of the total variance using the
31
32 Snijders and Bosker latent variable approach. [31]
33
34
35

36
37 All data management was done using SAS 9.2 [32] and multilevel modeling using MLwiN
38
39 2.24. [33]
40

41
42 The conduct of the 45 and Up Study was approved by the University of New South Wales
43
44 Human Research Ethics Committee (HREC), while ethical approval for this particular study
45
46 was provided by the NSW Population and Health Services Research Ethics Committee and
47
48 the University of Western Sydney HREC.
49

50 51 52 53 54 **Results**

55 56 57 *Descriptive characteristics*

1
2
3 A total of 32,832 study participants were admitted to 313 hospitals up to a year prior to
4 completing the 45 and Up Study baseline survey. Just over half of the index admissions
5 (53%) were planned stays in hospital, and 57% were to a public hospital. Around one-third of
6 the index admissions occurred within the three months before study entry, and the mean
7 length of stay was 4.8 days (median = 3 days). Just under half of the sample (47%) reported
8 having hypertension, with heart disease or obesity reported by 25%, and current smoking by
9 6.1% of the sample. One-third (34%) of participants had two or more morbidities (data not
10 shown). Other characteristics of the sample at their index admission are shown in Table 1.
11
12
13
14
15
16
17
18
19
20
21
22
23

24 Table 1 about here

25 26 27 28 29 *Concordance between self-report and hospital records*

30
31
32 Overall, reporting of morbidity differed between the two data sources with 23,257 (71%)
33 participants having at least one of the six self-reported morbidities, and 11,977 (36.5%) and
34 14,335 (43.7%) of the sample having at least one morbidity recorded on their index or
35 lookback hospital admissions, respectively.
36
37
38
39

40
41 Table 2 gives the summary concordance measures for each morbidity and two-way morbidity
42 combination. For the index admission, good agreement was found for diabetes ($\kappa=0.79$),
43 moderate agreement for smoking ($\kappa=0.59$), fair agreement for heart disease ($\kappa=0.40$), and
44 poor agreement for stroke ($\kappa=0.30$), hypertension ($\kappa=0.24$) and obesity ($\kappa=0.09$). In two-way
45 combinations, moderate levels of agreement were found only for diabetes combinations (with
46 smoking, hypertension and heart disease).
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 2 about here

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Incorporating a one-year lookback period increased the numbers of participants with a morbidity recorded in a hospital record, with average relative increases in the kappa values of 20% (ranging from 2% increase for smoking, to 41% increase for obesity). Good to excellent level of agreements were still found only for diabetes ($\kappa = 0.83$) and smoking ($\kappa = 0.60$).

Agreement was only slightly higher among the 82 large public hospitals (see Supplementary Table 1) with relative kappa values higher by 4%, on average.

Patient- and hospital-level predictors of positive agreement

The patient factors which predicted positive agreement between the two data sources differed between morbidities (Table 3). Male sex was associated with better agreement for diabetes (OR=1.37, 95% CI 1.19 – 1.58), heart disease (OR=1.30, 95% CI 1.17 – 1.44) and hypertension (OR=1.28, 95% CI 1.18 – 1.38) (Supplementary Table 2).

Older patients were significantly less likely to have smoking (80+ years OR=0.48, 95% CI 0.31 – 0.74) and obesity (OR=0.14, 95% CI 0.08 – 0.26) recorded in their hospital records, and significantly more likely to have hypertension recorded (OR=1.32, 95% CI 1.16 – 1.49), compared to younger patients (45 – 59 years). People with higher levels of functional limitation were significantly more likely to have hypertension, diabetes and obesity recorded on their most recent hospital stay. Planned admissions to hospital had lower odds of having any of the six conditions recorded, as did medical admissions (for diabetes, smoking and obesity only). Agreement did not vary significantly for any other patient factors.

Table 3 about here

The four hospital-level covariates (hospital type, hospital peer group, hospital remoteness and depth of coding) were added to multilevel models (including a random intercept for hospital) one at a time, separately. Positive agreement between self-report and hospital records was significantly lower for hospitals with lower depth of coding across all morbidities. The odds of recording were also lower among private hospitals for all six morbidities, with this difference being statistically significant for hypertension, heart disease and stroke only. Records from smaller hospitals (district and community peer groups) were significantly less likely to agree with self-reported data on hypertension, diabetes and heart disease. Positive agreement did not vary significantly with remoteness of hospital, with the exceptions of diabetes (lower agreement for outer regional, remote and very remote hospitals) and smoking (lower agreement for remote and very remote hospitals) (Supplementary Table 3).

Quantifying variation between hospitals

Before any hospital-level variables were added into the multilevel model, the intraclass correlation coefficient indicated that between 8% (diabetes) and 22% (heart disease) of the residual (unexplained) variation in agreement was attributable to the hospital, after adjustment for the patient-level factors (Table 4). This equated to median odds ratios (MORs) of 1.64 and 2.48, respectively, indicating that a patient in one hospital had an average of between 64% and 148% higher odds of having a particular morbidity recorded than a patient in a hospital with lower levels of recording. Less variation at the hospital level was found for the recording of diabetes, smoking and stroke, while more variation at the hospital level was found for the recording of hypertension, heart disease and obesity. When the analyses were

1
2
3 restricted to 82 large public hospitals only, the between-hospital variation decreased to
4
5 between 2% (stroke) and 13% (hypertension), or MOR of 1.24 and 1.94 (Figure 1). This
6
7 between-hospital variation was still significant for all morbidities except for stroke. Between-
8
9 hospital variation was further reduced once lookback admissions were used to identify
10
11 morbidities.
12
13

14
15
16
17 Figure 1 about here
18
19

20
21
22 The addition of hospital-level variables to multilevel models, one at the time, separately,
23
24 helped ascertain which factors explained the variation between hospitals (Table 4). The
25
26 addition of at least one of the four hospital-level factors contributed to explaining the residual
27
28 variation between hospitals for all conditions, except obesity. For the other morbidities,
29
30 differences in the depth of coding explained from 16% (smoking) to 42% (hypertension) of
31
32 residual variation between hospitals, while hospital type (public/private) explained from 0%
33
34 (smoking) to 59% (stroke), and hospital peer group explained from 10% (hypertension) to
35
36 27% (diabetes) residual variation between hospitals.
37
38
39
40
41
42
43

44 Table 4 about here
45
46
47
48

49 Discussion

50
51
52 Our study found that the concordance of administrative hospital and self-reported data varied
53
54 between the six morbidities examined, with agreement ranging from good for diabetes,
55
56 moderate for smoking, through to fair for heart disease, and poor for hypertension, stroke and
57
58
59
60

1
2
3 obesity. We demonstrated considerable between-hospital variation in the recording of these
4
5 common health conditions. Smaller, but still significant, between-hospital variation was
6
7 found when restricting the analyses to the largest public hospitals in the state.
8
9

10 Previous studies have validated information recorded in NSW administrative hospital data for
11
12 demographic factors, [34 35] and recording of perinatal conditions, [36-39] but there have
13
14 been limited studies of the accuracy of the recording of health conditions commonly used for
15
16 case-mix or risk-adjustment. Our findings regarding agreement for the recording of diabetes
17
18 ($\kappa=0.83$) were similar to previous Australian studies [3 10], while agreement for hypertension
19
20 ($\kappa=0.30$) and heart disease ($\kappa=0.47$) was considerably lower in our study. These differences
21
22 may be due to the fact that both previous studies used medical records as a 'gold standard',
23
24 while we used self-report. Higher sensitivities reported in a study from the state of Victoria
25
26 [10] could also be attributable to the differences in public hospital funding models between
27
28 the two states. Specifically, Victoria has used activity- based funding since 1993, while this
29
30 method of funding was introduced in NSW and other Australian states only subsequent to our
31
32 study period. [40] Introduction of activity-based funding has been shown to increase
33
34 recording of additional diagnoses and procedures in Europe. [41]
35
36
37

38
39 Some of the apparent discrepancies in the levels of coding between conditions can be
40
41 attributed to the coding rules that govern whether or not a diagnosis is recorded in
42
43 administrative hospital data. Additional diagnoses are coded only if they affect the patient's
44
45 treatments received, investigations required and/or resources used during the hospital stay.
46
47 Thus, diagnoses that relate to an earlier episode, and which have no bearing on the current
48
49 hospital stay, are excluded. Therefore, it is not surprising that (managed) hypertension, in
50
51 particular, might not be recorded in hospital data relating to, for example, elective surgery.
52
53 On the other hand, we found that diabetes is well recorded, suggesting that it is considered to
54
55 affect patient management in most hospital stays.
56
57
58
59
60

1
2
3 As well as looking at single morbidities, ours is the first study to our knowledge to explore
4 the variations of recording of multiple conditions in hospital data. Concordance of two-way
5 condition combinations was very low, with best results found for combinations of diseases
6 involving diabetes, which had the highest single-condition level of agreement with self-
7 reported data ($\kappa = 0.83$). Agreement measures for two-way combinations were found to be fair
8 to good at best, with agreement on three-way condition combinations (not investigated here)
9 expected to be even lower. These findings have implications for research into multimorbidity
10 (the co-occurrence of multiple chronic or acute diseases and medical conditions within one
11 person [42]). We suggest that researchers who use administrative data for research into
12 multimorbidity should use linked data to increase ascertainment, and, if possible, supplement
13 this information from other data sources, such as physician claims data or self-reported data.
14
15

16
17 We identified considerable between-hospital variability in the levels of recording of common
18 health conditions, with between 8% and 22% of the variation attributable to hospital-level
19 factors, after adjustment for patient factors. This was similar in magnitude to the variability
20 previously reported for performance measures (varying from patient satisfaction, mortality,
21 length of stay to quality of care) clustered at the facility level (0-51%) [43] and hospital-level
22 variations in the use of services. [44-46] Significant between-hospital variation was still
23 present after constraining the analyses to the 82 largest public hospitals in the state.
24
25

26
27 The recording of hypertension and heart disease was particularly variable between hospitals,
28 those with better reporting having on average 2.3 and 2.5 times, respectively, the odds of
29 recording these conditions than those with lower levels of reporting. The corresponding
30 figures were 1.9 and 1.6 times for the 82 largest hospitals in the state. These findings indicate
31 the potential for reporting bias to influence comparisons of health performance indicators
32 between hospitals, especially for indicators that use conditions such as heart disease or
33 hypertension for case-mix adjustment. To our knowledge, no previous studies have provided
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 detailed information about how the validity of morbidity reporting varies among hospitals
4
5 after accounting for patient factors.
6

7
8 Further, we have shown that variations in the accuracy of morbidity reporting between
9
10 hospitals are predominantly driven by the hospital's depth of coding – concordance between
11
12 self-reported and hospital data is lower in hospitals with a lower average number of
13
14 additional diagnoses recorded. Up to 42% of the variation in recording at the hospital level
15
16 could be attributed to differences in hospital depth of coding. Even though the measure of
17
18 depth of coding we used was crude, and related to hospital size, it still helps in highlighting
19
20 the impact of coding practices on variations among hospitals. Other research using the same
21
22 depth of coding measure has shown that the lower depth of coding can disproportionately
23
24 disadvantage hospitals' standardised mortality ratios, one of the commonly reported measures
25
26 of hospital performance. [2] It will be important to track changes in the levels of the depth of
27
28 coding across Australian states, and to consider the implications of these for state-based
29
30 performance comparisons, following the national rollout of activity-based funding and
31
32 comparative performance reporting.
33
34
35

36
37 Several factors might explain variation in depth of coding between hospitals. Clinical coders
38
39 can code only information that has been recorded in the patient's medical record, so varying
40
41 level of details recorded by clinicians will influence what gets coded. The training and
42
43 professional development opportunities for coding staff might also influence the depth of
44
45 coding. Also, casemix funding systems, such as the Diagnosis Related Group (DRG)
46
47 classification, are prone to 'upcoding' in order for services to receive higher reimbursement
48
49 costs. [14]
50
51

52
53 We found that the reporting of conditions varied with hospital size, larger metropolitan
54
55 hospitals having higher concordance, with kappa values higher by 7% on average when
56
57 comparing large tertiary with smaller urban hospitals. This finding echoes those of Powell et
58
59
60

1
2
3 al. [3] in NSW, Australia during 1996 – 1998 and Rangachari et al. [16] in the US, during
4
5 2000 – 2004. Our study showed that large tertiary hospitals had better concordance for the
6
7 recording of hypertension and heart disease than smaller urban hospitals, but the reverse was
8
9 true for stroke and smoking. Our finding that between-hospital variation in the recording of
10
11 morbidities was up to two times higher when all hospitals, rather than just the largest ones,
12
13 were included has implications for further research using data from smaller hospitals. This
14
15 high variability in concordance among smaller hospitals may mean that morbidity-adjusted
16
17 comparisons are not as valid as for larger hospitals. Researchers using information from these
18
19 hospitals are encouraged to supplement their data with either self-report information and/or
20
21 data linkage. The value-add of incorporating previous hospitalizations was also highlighted in
22
23 our results for stroke and obesity, with 43% – 47% more patients identified using lookback
24
25 admissions than from a single admission only.
26
27
28

29
30 A particular strength of our study lies in the use of linked data from a large-scale cohort study
31
32 to comprehensively evaluate the recording of common conditions in hospital data, and
33
34 explore the variation in recording among hospitals. The 45 and Up Study contains records for
35
36 one in every 10 persons aged 45 and over in NSW, so it provides a rich resource to answer
37
38 research questions. Additionally, we used advanced multilevel modeling methods to quantify
39
40 the amount of between-hospital variation in the level of recording of common health
41
42 conditions, a finding which is of importance for both research and policy paradigms due to its
43
44 impact on adjusted comparisons among hospitals and the highlighted need to improve
45
46 consistency of recording in hospitals across the State. To date, hospital-level variation has
47
48 only been explored with a set outcome (e.g. mortality, readmission) in mind.
49
50

51
52 A potential limitation of our study was its use of self-reported information to explore
53
54 concordance, in the absence of another ‘gold standard’, such as medical records. Access to
55
56 medical records was not possible given the de-identified nature of our data, and the large
57
58
59
60

1
2
3 number of records in the dataset. Moreover, studies that have examined accuracy of self-
4
5 reported conditions against medical records have found high levels of agreement, ranging
6
7 from 81% [47] to 87% [48] for hypertension, 66% [49] to 96% [47 48] for diabetes and 60%
8
9 [47] to 98% [50] for acute myocardial infarction. Validation studies in the 45 and Up Study
10
11 cohort have reported strong correlations and excellent levels of agreement between self-
12
13 reported and measured height and weight, and derived BMI [51] as well as self-reported
14
15 diabetes. [52]
16
17

18 19 **Conclusion**

20
21 The recording of common comorbid conditions in routine hospital data is highly variable,
22
23 and, for some conditions, very poor. Recording varies considerably among hospitals,
24
25 presenting the potential to introduce bias into risk-adjusted comparisons of hospital
26
27 performance, especially for indicators that use heart disease or hypertension for risk
28
29 adjustment. Furthermore, between-hospital variation is amplified when smaller and private
30
31 hospitals are included in the analyses. Stratification of analyses according to factors that
32
33 predict the completeness of recording, including hospital depth of coding and hospital type
34
35 and size, supplementing morbidity information with linked data from previous
36
37 hospitalizations and increases in efforts to standardize recording across hospitals, all offer
38
39 potential for increasing the validity of risk-adjusted comparisons.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgement

This work was supported by the National Health and Medical Research Council [grant number 1036858].

This research was completed using data collected through the 45 and Up Study (www.saxinstitute.org.au). The 45 and Up Study is managed by the Sax Institute in collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; *beyondblue*; Ageing, Disability and Home Care, Department of Family and Community Services; the Australian Red Cross Blood Service; and UnitingCare Ageing. We thank the many thousands of people participating in the 45 and Up Study.

We would like to acknowledge the NSW Ministry of Health for allowing access to the data, and the Centre for Health Record Linkage for conducting the probabilistic linkage of records.

We are grateful to Dr Fiona Blyth and Dr Kris Rogers on their advice at the early stages of the project.

Funding statement

The Assessing Preventable Hospitalisation Indicators (APHID) study is funded by a National Health and Medical Research Council Partnership Project Grant (#1036858) and by partner agencies the Australian Commission on Safety and Quality in Health Care, the Agency for Clinical Innovation and the NSW Bureau of Health Information.

Contributorship

SL had overall responsibility for the design of this study, data management, statistical analysis and drafting this paper. DW and LJ contributed to the conception and design of the study. LJ helped with data acquisition, and provided oversight for all analyses. DR and JS provided oversight and advice for the design and interpretation of the statistical analyses. All

1
2
3 authors contributed to the interpretation of the findings, the writing of the paper and approved
4 the final draft.
5

6 **Data sharing** 7

8
9 This study used the data from the Assessing Preventable Hospitalisation InDicators
10 (APHID) project. The data has been constructed with the permission of each of the
11 custodians of the respective source datasets and with specific ethical approval. These data
12 are available to researchers on request and subject to approval from the relevant data
13 custodians and ethics committees, as outlined on the 45 and Up Study website
14 (<https://www.saxinstitute.org.au/our-work/45-up-study/for-researchers/#application-forms>)
15 as well as the NSW Centre for Health Record Linkage website (www.cherel.org.au). More
16 information about these approvals is available from the authors on request.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

1. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *The Lancet* 2004;**363**(9415):1147-54
2. Mohammed MA, Deeks JJ, Girling A, et al. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ: British Medical Journal* 2009;**338**
3. Powell H, Lim LL, Heller RF. Accuracy of administrative data to assess comorbidity in patients with heart disease: an Australian perspective. *J Clin Epidemiol* 2001;**54**(7):687-93
4. Pine M, Jordan HS, Elixhauser A, et al. Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA* 2007;**297**(1):71-76 doi: 10.1001/jama.297.1.71[published Online First: Epub Date].
5. Pine M, Norusis M, Jones B, et al. Predictions of Hospital Mortality Rates: A Comparison of Data Sources. *Annals of Internal Medicine* 1997;**126**(5):347-54 doi: 10.7326/0003-4819-126-5-199703010-00002[published Online First: Epub Date].
6. Hawker GA, Coyte PC, Wright JG, et al. Accuracy of administrative data for assessing outcomes after knee replacement surgery. *J Clin Epidemiol* 1997;**50**(3):265-73
7. Kieszak SM, Flanders WD, Kosinski AS, et al. A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *J Clin Epidemiol* 1999;**52**(2):137-42
8. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Medical care* 2002;**40**(8):675-85 doi: 10.1097/01.MLR.0000020927.46398.5D[published Online First: Epub Date].
9. Preen DB, Holman CAJ, Lawrence DM, et al. Hospital chart review provided more accurate comorbidity information than data from a general practitioner survey or an administrative database. *J Clin Epidemiol* 2004;**57**(12):1295-304
10. Henderson T, Shephard J, Sundararajan V. Quality of diagnosis and procedure coding in ICD-10 administrative data. *Medical care* 2006;**44**(11):1011-9 doi: 10.1097/01.mlr.0000228018.48783.34[published Online First: Epub Date].
11. Quach S, Blais C, Quan H. Administrative data have high variation in validity for recording heart failure. *Canadian Journal of Cardiology* 2010;**26**(8):e306-e12
12. Leal J, Laupland K. Validity of ascertainment of co-morbid illness using administrative databases: a systematic review. *Clinical Microbiology and Infection* 2010;**16**(6):715-21
13. Chong WF, Ding YY, Heng BH. A comparison of comorbidities obtained from hospital administrative data and medical charts in older patients with pneumonia. *BMC health services research* 2011;**11**(1):105
14. Steinbusch PJ, Oostenbrink JB, Zuurbier JJ, et al. The risk of upcoding in casemix systems: A comparative study. *Health Policy* 2007;**81**(2):289-99
15. Lorence DPHIA. Benchmarking Variation in Coding Accuracy Across the United States. *Journal of Health Care Finance* 2003;**29**(4):29
16. Rangachari P. Coding for quality measurement: the relationship between hospital structural characteristics and coding accuracy from the perspective of quality measurement. *Perspectives in health information management/AHIMA, American Health Information Management Association* 2007;**4**

17. Coory M, Cornes S. Interstate comparisons of public hospital outputs using DRGs: Are they fair? *Australian and New Zealand Journal of Public Health* 2005;**29**(2):143-48 doi: 10.1111/j.1467-842X.2005.tb00064.x[published Online First: Epub Date].
18. National Health Performance Authority. *Healthy Communities: Overweight and obesity rates across Australia, 2011–12 (In Focus)*. 2013.
19. Banks E, Redman S, Jorm L, et al. Cohort profile: the 45 and up study. *International Journal of Epidemiology* 2008;**37**(5):941
20. Mealing NM, Banks E, Jorm LR, et al. Investigation of relative risk estimates from studies of the same population with contrasting response rates and designs. *BMC Med Res Methodol* 2010;**10**:26
21. National Centre for Classification in Health. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification (ICD-10-AM), Australian Classification of Health Interventions (ACHI)*. Sydney: National Centre for Classification in Health, 2006.
22. Australian Institute of Health and Welfare. *National Health Data Dictionary. Version 16*. Canberra: Australian Institute of Health and Welfare, 2012.
23. Kelman CW, Bass AJ, Holman CDJ. Research use of linked health data — a best practice protocol. *Australian and New Zealand Journal of Public Health* 2002;**26**(3):251-55 doi: 10.1111/j.1467-842X.2002.tb00682.x[published Online First: Epub Date].
24. WHO. *Obesity: preventing and managing the global epidemic. World Health Organization technical report series*, 2000.
25. Stewart AL, Ware JE. *Measuring functioning and well-being: The medical outcomes study approach*: Duke University Press, 1992.
26. NSW Health. *NSW Health Services Comparison Data Book 2008/09. Secondary NSW Health Services Comparison Data Book 2008/09* 2010.
http://www0.health.nsw.gov.au/pubs/2010/pdf/yellowbook_09.pdf.
27. Australian Institute of Health and Welfare. *Regional and remote health: A guide to remoteness classifications* Canberra
Australian Institute of Health and Welfare, 2004.
28. Zhang JX, Iwashyna TJ, Christakis NA. The performance of different lookback periods and sources of information for Charlson comorbidity adjustment in Medicare claims. *Medical care* 1999;**37**(11):1128-39
29. Fleiss JL, Levin B, Paik MC. The measurement of interrater agreement. *Statistical methods for rates and proportions* 1981;**2**:212-36
30. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health* 2006;**60**(4):290-97 doi: 10.1136/jech.2004.029454[published Online First: Epub Date].
31. Snijders T, Bosker R. *Multilevel analysis*. 2nd ed. London: Sage, 2012.
32. SAS Version 9.3 [software] [program]. Cary, North Carolina, 2010.
33. MLwiN Version 2.25 [software] [program]. Centre for Multilevel Modelling, University of Bristol, Bristol., 2012.
34. Bentley JP, Taylor LK, Brandt PG. Reporting of Aboriginal and Torres Strait Islander peoples on the NSW Admitted Patient Data Collection: the 2010 data quality survey. *New South Wales public health bulletin* 2012;**23**(1-2):17-20 doi: 10.1071/NB11034[published Online First: Epub Date].

- 1
2
3 35. Tran DT, Jorm L, Lujic S, et al. Country of birth recording in Australian hospital morbidity data:
4 accuracy and predictors. *Australian and New Zealand Journal of Public Health*
5 2012;**36**(4):310-16
- 6
7 36. Taylor LK, Travis S, Pym M, et al. How useful are hospital morbidity data for monitoring
8 conditions occurring in the perinatal period? *Aust N Z J Obstet Gynaecol* 2005;**45**(1):36-41
- 9
10 37. Roberts CL, Bell JC, Ford JB, et al. The accuracy of reporting of the hypertensive disorders of
11 pregnancy in population health data. *HYPERTENS PREGNANCY* 2008;**27**(3):285-97 doi:
12 10.1080/10641950701826695[published Online First: Epub Date].
- 13
14 38. Hadfield RM, Lain SJ, Cameron CA, et al. The prevalence of maternal medical conditions during
15 pregnancy and a validation of their reporting in hospital discharge data. *Aust N Z J Obstet*
16 *Gynaecol* 2008;**48**(1):78-82
- 17
18 39. Lain SJ, Hadfield RM, Raynes-Greenow CH, et al. Quality of data in perinatal population health
19 databases: a systematic review. *Medical care* 2012;**50**(4):e7-20
- 20
21 40. Eagar KM. What is activity-based funding? ABF Information Series No. 1. Secondary What is
22 activity-based funding? ABF Information Series No. 1 2011.
23 <http://ro.uow.edu.au/gsbpapers/47/>.
- 24
25 41. O'Reilly J, Busse R, Häkkinen U, et al. Paying for hospital care: the experience with
26 implementing activity-based funding in five European countries. *Health Economics Policy*
27 *and Law* 2012;**7**(1):73
- 28
29 42. van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity: what's in a name? A
30 review of literature. *European Journal of General Practice* 1996;**2**(2):65-70
- 31
32 43. Fung V, Schmittziel JA, Fireman B, et al. Meaningful variation in performance: a systematic
33 literature review. *Medical care* 2010;**48**(2):140-48
- 34
35 44. Bennett-Guerrero E, Zhao Y, O'Brien SM, et al. Variation in use of blood transfusion in coronary
36 artery bypass graft surgery. *JAMA: the journal of the American Medical Association*
37 2010;**304**(14):1568-75
- 38
39 45. Seymour CW, Iwashyna TJ, Ehlenbach WJ, et al. Hospital-Level Variation in the Use of Intensive
40 Care. *Health Services Research* 2012;**47**(5):2060-80
- 41
42 46. Randall DA, Jorm LR, Lujic S, et al. Disparities in Revascularization Rates After Acute
43 Myocardial Infarction Between Aboriginal and Non-Aboriginal People in Australia
44 Clinical
45 Perspective. *Circulation* 2013;**127**(7):811-19
- 46
47 47. Merkin S, Cavanaugh K, Longenecker JC, et al. Agreement of self-reported comorbid conditions
48 with medical and physician reports varied by disease among end-stage renal disease patients.
49 *J Clin Epidemiol* 2007;**60**(6):634
- 50
51 48. Barber J, Muller S, Whitehurst T, et al. Measuring morbidity: self-report or health care records?
52 *Family practice* 2010;**27**(1):25-30
- 53
54 49. Okura Y, Urban LH, Mahoney DW, et al. Agreement between self-report questionnaires and
55 medical record data was substantial for diabetes, hypertension, myocardial infarction and
56 stroke but not for heart failure. *J Clin Epidemiol* 2004;**57**(10):1096
- 57
58 50. Machón M, Arriola L, Larrañaga N, et al. Validity of self-reported prevalent cases of stroke and
59 acute myocardial infarction in the Spanish cohort of the EPIC study. *Journal of epidemiology*
60 *and community health* 2013;**67**(1):71-75
- 51
52 51. Ng SP, Korda R, Clements M, et al. Validity of self-reported height and weight and derived body
53 mass index in middle-aged and elderly individuals in Australia. *Australian and New Zealand*
54 *Journal of Public Health* 2011;**35**(6):557-63
- 55
56 52. Comino EJ, Tran DT, Haas M, et al. Validating self-report of diabetes use by participants in the 45
57 and up study: a record linkage study. *BMC health services research* 2013;**13**(1):481

1
2
3
4 **List of tables and figures**
5

6 **Table 1.** Characteristics of the study sample at their index admission
7

8 **Table 2.** Agreement measures between self-report and hospital data, index and lookback
9 admissions, all public and private hospitals in New South Wales, Australia (n=313)
10

11 **Table 3.** Factors that predict positive agreement between self-report and hospital data, using
12 multilevel modeling, all public and private hospitals in New South Wales, Australia (n=313)
13

14 **Table 4.** Variance and intraclass correlation coefficient for hospital-level random effects
15 from multilevel logistic regression, all public and private hospitals in New South Wales,
16 Australia (n=313)
17

18 **Figure 1.** Variance for hospital-level random effects from multilevel logistic regression, for
19 index and lookback admissions, by hospital size
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Characteristics of the study sample at their index admission

Characteristics	All participants		All hospitals ^a	
	(N = 32,832)		(N = 313)	
	N	%	N	%
<i>Demographic characteristics</i>				
Sex				
Male	16,812	51.2		
Female	16,020	48.8		
Age				
45-59	9,666	29.4		
60-79	16,624	50.6		
80+	6,540	19.9		
Country of birth				
Australia	25,001	76.2		
Other	7,448	22.7		
Unknown	383	1.2		
Highest education level				
No school	5,196	15.8		
Year 10 or equivalent	7,894	24.0		
Year 12 or equivalent	2,975	9.1		
Trade	4,270	13.0		
Certificate	6,109	18.6		
University degree	5,662	17.3		
Unknown	726	2.2		
Household income (\$, per annum)				
<20,000	9,077	27.7		
20,000 - <50,000	8,223	25.1		
50,000 - <70,000	2,560	7.8		
70,000+	5,042	15.4		
Not disclosed	6,003	18.3		
Missing	1,927	5.9		
Functional status				
No limitation	4,915	15.0		
Mild limitation	6,011	18.3		
Moderate limitation	8,701	26.5		
Severe limitation	10,121	30.8		
Missing	3,084	9.4		
<i>Admission characteristics</i>				
Admission type				
Surgical	15,464	47.1		
Other	1,439	4.4		
Medical	15,929	48.5		
Emergency status				
Emergency	13,484	41.1		

1					
2					
3	Planned	17,544	53.4		
4	Other	1,803	5.5		
5	Hospital characteristics				
6	Hospital type				
7					
8	Public	18,734	57.1	224	71.6
9	Private	14,096	42.9	88	28.1
10	Hospital remoteness				
11					
12	Major city	19,754	60.2	124	39.6
13	Inner regional	8,424	25.7	72	23.0
14	Outer regional	4,137	12.6	94	30.0
15	Remote/very remote	363	1.1	20	6.4
16	Hospital depth of coding				
17					
18	1 - least comprehensive	1,629	5.0	48	15.3
19	2	8,803	26.8	91	29.1
20	3	11,543	35.2	89	28.4
21	4 - most comprehensive	10,857	33.1	85	27.2
22	Hospital peer group				
23					
24	Principal referral	6,329	19.3	14	4.5
25	Major	11,052	33.7	33	10.5
26	District	6,862	20.8	51	16.3
27	Community	7,018	21.4	121	38.7
28	Other	1,571	4.8	94	30.0
29	<hr/>				
30	^a for comparisons of hospital characteristics				

Table 2. Agreement measures between self-report and hospital data, index and lookback admissions, all public and private hospitals in New South Wales, Australia (n=313)

Morbidities ^a	<u>Index admission</u>					<u>Lookback admissions</u>						
	<u>45 and Up Yes:</u>		<u>45 and Up No:</u>		Kappa %	<u>45 and Up Yes:</u>		<u>45 and Up No:</u>		Kappa %		
	APDC yes	APDC no	APDC yes	APDC no		95% CI	APDC yes	APDC no	APDC yes		APDC no	95% CI
Hypertension	4,767	10,512	1,434	16,119	24.0	(22.9-25.0)	6,260	9,019	2,051	15,502	30.2	(29.1-31.2)
Heart disease	3,639	4,668	1,942	22,583	40.3	(39.0-41.5)	4,673	3,634	2,697	21,828	47.0	(45.8-48.2)
Diabetes	3,560	1,234	347	27,691	79.1	(78.1-80.1)	3,928	866	479	27,559	83.0	(82.1-83.9)
Stroke	541	1,939	306	30,046	29.8	(27.0-32.6)	776	1,704	488	29,864	38.3	(35.8-40.8)
Smoking	1,205	804	727	30,096	58.7	(56.7-60.7)	1,411	598	1,076	29,747	60.1	(58.2-61.9)
Obesity	551	7,611	114	24,556	9.1	(7.3-10.9)	810	7,352	209	24,461	12.8	(11.1-14.6)
Hypertension + heart disease	1,172	3,481	1,270	26,909	25.8	(23.8-27.7)	1,807	2,846	2,008	26,171	34.3	(32.6-36.0)
Hypertension + diabetes	1,819	1,238	759	29,016	61.3	(59.6-62.9)	2,186	871	1,021	28,754	66.6	(65.2-68.1)
Hypertension + stroke	203	1,317	189	31,123	19.7	(15.7-23.7)	329	1,191	340	30,972	28.0	(24.5-31.5)
Hypertension + smoking	133	598	180	31,921	24.5	(19.2-29.7)	199	532	319	31,782	30.6	(26.0-35.2)
Hypertension + obesity	234	4,574	93	27,931	7.4	(4.9-9.8)	383	4,425	183	27,841	11.5	(9.2-13.9)
Heart disease + diabetes	646	1,154	404	30,628	43.0	(40.3-45.8)	904	896	661	30,371	51.2	(48.9-53.6)
Heart disease + stroke	76	973	126	31,657	11.2	(6.1-16.4)	149	900	261	31,522	19.0	(14.4-23.5)
Heart disease + smoking	76	294	222	32,240	22.0	(15.3-28.6)	118	252	373	32,089	26.5	(20.8-32.2)
Heart disease + obesity	79	1,938	79	30,736	6.4	(2.5-10.4)	151	1,866	169	30,646	11.4	(7.7-15.2)
Diabetes + stroke	85	555	58	32,134	21.1	(15.0-27.3)	140	500	119	32,073	30.4	(24.9-35.8)
Diabetes + smoking	143	161	108	32,420	51.1	(45.3-56.9)	171	133	176	32,352	52.1	(46.7-57.4)
Diabetes + obesity	232	1,701	65	30,834	19.5	(15.9-23.2)	351	1,582	120	30,779	27.5	(24.2-30.9)
Stroke + smoking	13	142	28	32,649	13.1	(0.1-26.1)	23	132	57	32,620	19.3	(7.8-30.8)
Stroke + obesity	6	558	9	32,259	2.0	(0.0-10.0)	13	551	21	32,247	4.2	(0.0-11.9)
Smoking + obesity	27	447	29	32,329	9.9	(1.9-17.9)	38	436	47	32,311	13.2	(5.5-20.9)

^a ICD-10-AM codes: hypertension (I10-I15, R03.0), heart disease (I20-I25, I26-I28, I30-I52), diabetes (E10-E14), stroke (I60-I69, G45, G46), smoking (F17.2, Z72.0), obesity (E66)

Table 3. Factors that predict positive agreement between self-report and hospital data, using multilevel modelling, all public and private hospitals in New South Wales, Australia (n=313)

	Hypertension (N = 15,279)	Diabetes (N = 4,794)	Heart disease (N = 8,307)	Stroke (N = 2,480)	Smoking (N = 2,099)	Obesity (N = 8,162)
Person-level variables						
Sex ¹	**	**	**			
Age ¹	**				**	**
Education ¹			*	**	**	
Country of birth ¹						
Functional limitation ¹	**	**				**
Income ¹						
Admission type ²	**	**	**	**	**	**
Emergency status ²	**	**	**	**		**
Hospital-level variables						
Hospital type (public/private) ³	**		**	**		
Hospital remoteness ³					*	
Hospital depth of coding ³	**	**	**	**	**	**
Hospital peer group ³	**	**	**		**	

* Significant at 5% level

** Significant at 1% level

1 – Model 0: adjusted for demographic factors + random intercept for hospital

2 – Model 0 + admission type + emergency status

3 – Model 0 + hospital-level variables (entered one at a time)

Table 4. Variance and intraclass correlation coefficient for hospital-level random effects from multilevel logistic regression, all public and private hospitals in New South Wales, Australia (n=313)

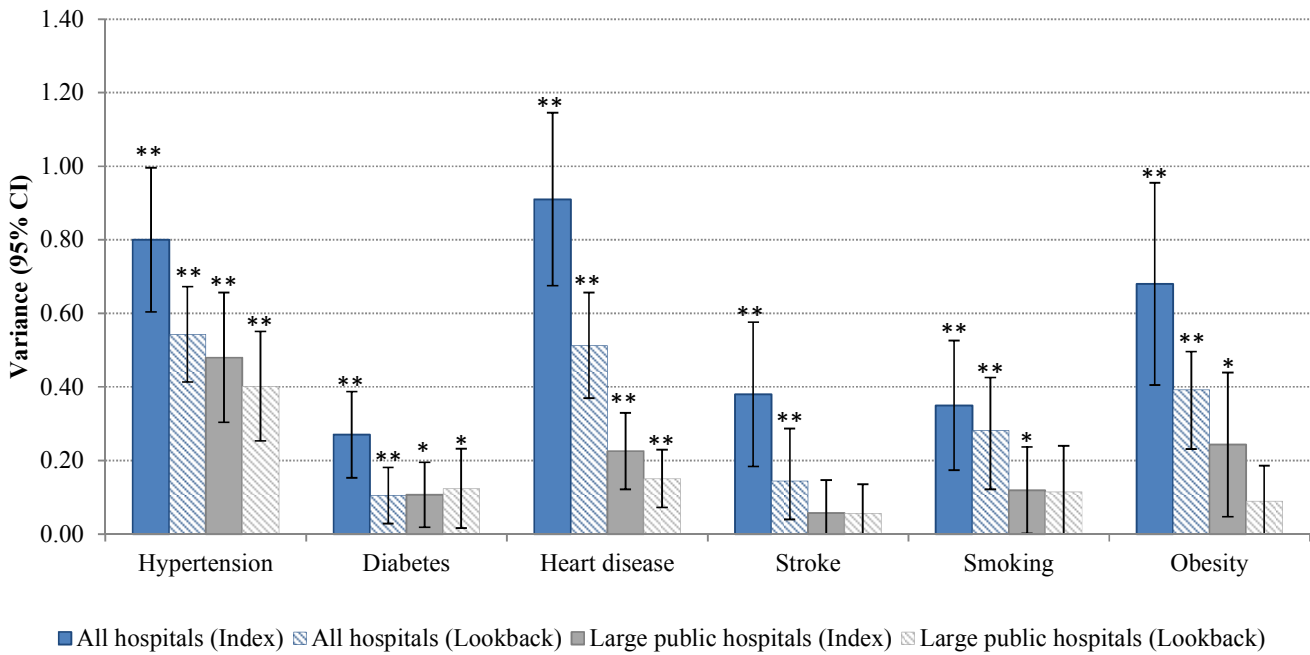
	Hypertension (N [†] = 15,279)	Diabetes (N [†] = 4,794)	Heart disease (N [†] = 8,307)	Stroke (N [†] = 2,480)	Smoking (N [†] = 2,099)	Obesity (N [†] = 8,162)
Hospital-level variance (SE)*						
Model 0. Patient factors	0.80 (0.10)	0.27 (0.06)	0.91 (0.12)	0.38 (0.10)	0.35 (0.09)	0.68 (0.14)
Model 1. Model 0 + hospital type (public/private)	0.65 (0.08)	0.27 (0.06)	0.71 (0.10)	0.16 (0.06)	0.35 (0.09)	0.69 (0.14)
Model 2. Model 0 + hospital remoteness	0.77 (0.09)	0.25 (0.05)	0.92 (0.12)	0.37 (0.10)	0.33 (0.08)	0.68 (0.14)
Model 3. Model 0 + hospital depth of coding	0.46 (0.06)	0.20 (0.05)	0.56 (0.08)	0.26 (0.08)	0.29 (0.08)	0.68 (0.14)
Model 4. Model 0 + hospital peer group	0.72 (0.09)	0.21 (0.05)	0.75 (0.10)	0.34 (0.09)	0.31 (0.08)	0.67 (0.14)
Intraclass correlation coefficient (ICC)**	19.5%	7.6%	21.6%	10.4%	9.6%	17.1%
Median odds ratio (MOR)**	2.34	1.64	2.48	1.80	1.76	2.19

[†] N = number of patients who self-reported condition

* Patient-level variance in a logistic regression is set at $\pi^2/3=3.29$ [31]

** ICC and MOR calculated from Model 0 [ICC = hospital-level variance divided by total variance (hospital-level + patient-level); MOR is calculated as $\exp^{0.954 \times \sqrt{\text{variance}}}$] [30]

Figure 1. Variance for hospital-level random effects from multilevel logistic regression, for index and lookback admissions, by hospital size



* Significantly different from 0 at 5% level
 ** Significantly different from 0 at 1% level

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

For peer review only

1
2
3 **Supplementary Table 1.** Agreement measures between self-report and hospital data, index and lookback admissions, large public
4 hospitals in New South Wales, Australia (n=82)
5
6

7 **Supplementary Table 2.** Adjusted ORs for patient-level variables from the multilevel logistic regression with random intercept for
8 hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)
9

10 **Supplementary Table 3.** Adjusted ORs for hospital-level variables from the multilevel logistic regression with random intercept for
11 hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Supplementary Table 1. Agreement measures between self-report and hospital data, index and lookback admissions, large public hospitals in New South Wales, Australia (n=82)

Morbidities ^a	<u>Index admission</u>						<u>Lookback admissions</u>					
	45 and Up Yes:		45 and Up No:		Kappa %	95% CI	45 and Up Yes:		45 and Up No:		Kappa %	95% CI
	APDC yes	APDC no	APDC yes	APDC no			APDC yes	APDC no	APDC yes	APDC no		
Hypertension	3,061	4,803	983	7,634	28.1	(26.6-29.6)	3,829	4,035	1,339	7,278	33.7	(32.2-35.1)
Heart disease	2,306	2,455	1,309	10,411	40.1	(38.5-41.8)	2,910	1,851	1,710	10,010	46.9	(45.4-48.5)
Diabetes	2,168	661	214	13,438	80.1	(78.8-81.4)	2,355	474	289	13,363	83.3	(82.1-84.4)
Stroke	414	1,210	213	14,644	33.1	(29.8-36.4)	563	1,061	311	14,546	41	(38.0-44.0)
Smoking	820	507	468	14,686	59.5	(57.0-62.0)	948	379	692	14,462	60.4	(58.1-62.7)
Obesity	265	3,857	61	12,298	8.6	(6.1-11.1)	414	3,708	114	12,245	12.9	(10.4-15.3)
Hypertension + heart disease	799	1,878	893	12,911	27.4	(25.0-29.9)	1,159	1,518	1,327	12,477	34.7	(32.5-36.9)
Hypertension + diabetes	1,129	670	518	14,164	61.5	(59.4-63.6)	1,317	482	662	14,020	65.8	(63.9-67.7)
Hypertension + stroke	160	825	145	15,351	22.6	(17.9-27.3)	238	747	237	15,259	29.9	(25.6-34.1)
Hypertension + smoking	106	399	135	15,841	27	(20.9-33.1)	154	351	237	15,739	32.6	(27.2-37.9)
Hypertension + obesity	157	2,291	62	13,971	9.6	(6.2-13.0)	251	2,197	117	13,916	14.5	(11.3-17.7)
Heart disease + diabetes	452	686	293	15,050	45	(41.7-48.3)	620	518	442	14,901	53.2	(50.4-56.1)
Heart disease + stroke	61	641	93	15,686	12.9	(6.8-19.1)	107	595	171	15,608	19.9	(14.4-25.4)
Heart disease + smoking	65	209	163	16,044	24.8	(17.2-32.3)	98	176	280	15,927	28.7	(22.2-35.1)
Heart disease + obesity	63	1,145	62	15,211	8.2	(3.2-13.2)	117	1,091	119	15,154	14.1	(9.5-18.8)
Diabetes + stroke	66	374	38	16,003	23.5	(16.2-30.8)	110	330	69	15,972	34.5	(28.2-40.9)
Diabetes + smoking	107	112	83	16,179	51.7	(45.0-58.5)	130	89	132	16,130	53.4	(47.3-59.5)
Diabetes + obesity	150	994	39	15,298	20.9	(16.3-25.6)	229	915	76	15,261	29.5	(25.3-33.8)
Stroke + smoking	13	106	23	16,339	16.5	(2.1-30.8)	19	100	41	16,321	20.8	(7.8-33.9)
Stroke + obesity	4	368	7	16,102	2	(0.0-11.8)	9	363	13	16,096	4.3	(0.0-13.9)
Smoking + obesity	17	292	18	16,154	9.5	(0.0-19.5)	25	284	29	16,143	13.3	(3.8-22.8)

^a ICD-10-AM codes: hypertension (I10-I15, R03.0), heart disease (I20-I25, I26-I28, I30-I52), diabetes (E10-E14), smoking (F17.2, Z72.0), obesity (E66)

Supplementary Table 2. Adjusted ORs^a for patient-level variables from the multilevel logistic regression with random intercept for hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)

Patient characteristics	Hypertension (N = 15,279)		Diabetes (N = 4,794)		Heart disease (N = 8,307)		Stroke (N = 2,480)		Smoking (N = 2,099)		Obesity (N = 8,162)	
	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b
Sex												
Female	1		1		1		1		1		1	
Male	1.28	(1.18,1.38)	1.37	(1.19,1.58)	1.30	(1.17,1.44)	1.13	(0.91,1.40)	1.14	(0.94,1.40)	0.85	(0.70,1.04)
Age												
45-59	1		1		1		1		1		1	
60-79	1.27	(1.15,1.41)	0.97	(0.81,1.16)	0.94	(0.82,1.09)	1.08	(0.78,1.52)	0.80	(0.65,0.99)	0.57	(0.47,0.70)
80+	1.32	(1.16,1.49)	1.00	(0.80,1.25)	1.01	(0.86,1.19)	1.02	(0.72,1.46)	0.48	(0.31,0.74)	0.14	(0.08,0.26)
Education												
None	1		1		1		1		1		1	
Trade	0.90	(0.79,1.03)	0.93	(0.73,1.18)	0.80	(0.68,0.94)	1.08	(0.76,1.54)	0.67	(0.48,0.94)	1.42	(1.01,2.02)
School certificate	0.96	(0.86,1.07)	1.00	(0.82,1.23)	0.90	(0.78,1.05)	1.22	(0.90,1.66)	0.87	(0.66,1.16)	1.04	(0.77,1.41)
HSC	0.99	(0.85,1.15)	0.89	(0.68,1.17)	0.91	(0.75,1.11)	2.23	(1.51,3.30)	0.53	(0.37,0.76)	1.24	(0.84,1.83)
Diploma	0.96	(0.84,1.09)	0.90	(0.72,1.14)	0.87	(0.74,1.03)	1.08	(0.75,1.56)	1.00	(0.73,1.37)	1.15	(0.83,1.59)
University	0.85	(0.74,0.98)	0.90	(0.70,1.16)	0.72	(0.60,0.86)	1.23	(0.83,1.81)	0.54	(0.37,0.80)	1.25	(0.88,1.79)
County of birth												
Australia	1		1		1		1		1		1	
Overseas	1.00	(0.91,1.09)	0.95	(0.81,1.11)	1.10	(0.98,1.23)	1.19	(0.94,1.51)	1.17	(0.92,1.48)	0.89	(0.69,1.14)
Functional limitation												
No limitation	1		1		1		1		1		1	
Mild	1.07	(0.91,1.25)	0.91	(0.68,1.23)	1.02	(0.82,1.28)	0.82	(0.48,1.42)	0.92	(0.65,1.30)	1.07	(0.72,1.60)
Moderate	1.23	(1.07,1.42)	1.14	(0.87,1.50)	0.92	(0.75,1.13)	0.68	(0.42,1.11)	0.79	(0.57,1.09)	1.06	(0.73,1.53)
Severe	1.53	(1.33,1.76)	1.54	(1.18,2.01)	0.97	(0.79,1.19)	0.84	(0.53,1.33)	0.82	(0.60,1.12)	2.27	(1.59,3.24)
Income												
<20,000	1		1		1		1		1		1	
20-50,000	0.89	(0.81,0.99)	0.95	(0.79,1.14)	1.03	(0.91,1.17)	1.14	(0.87,1.49)	1.17	(0.90,1.53)	0.76	(0.57,1.00)

1												
2												
3												
4	50-70,000	0.89 (0.75,1.05)	0.87 (0.64,1.19)	1.11 (0.89,1.38)	1.16 (0.68,1.99)	1.37 (0.93,2.02)	0.89 (0.60,1.30)					
5	>70,000	0.86 (0.74,1.00)	1.03 (0.77,1.38)	1.24 (1.02,1.50)	1.07 (0.63,1.82)	0.95 (0.66,1.36)	1.15 (0.83,1.59)					
6	Not disclosed	1.00 (0.90,1.12)	1.04 (0.86,1.27)	1.14 (0.99,1.31)	1.18 (0.89,1.56)	1.36 (1.02,1.80)	1.07 (0.81,1.41)					
7												
8	Admission type^c											
9	Surgical	1	1	1	1	1	1					
10	Other	1.45 (1.23,1.72)	1.01 (0.72,1.42)	2.34 (1.91,2.87)	0.47 (0.16,1.37)	0.69 (0.41,1.14)	0.62 (0.36,1.09)					
11	Medical	1.14 (1.03,1.27)	0.66 (0.55,0.80)	0.97 (0.84,1.11)	4.36 (3.02,6.29)	0.50 (0.38,0.65)	0.64 (0.50,0.84)					
12												
13	Emergency status^c											
14	Emergency	1	1	1	1	1	1					
15	Planned	0.63 (0.56,0.71)	0.64 (0.52,0.77)	0.42 (0.36,0.49)	0.65 (0.48,0.88)	0.86 (0.65,1.13)	0.58 (0.44,0.78)					
16	Other	0.96 (0.80,1.15)	0.98 (0.72,1.33)	1.02 (0.81,1.28)	1.19 (0.80,1.76)	1.03 (0.65,1.62)	0.80 (0.50,1.28)					
17												

^a Odds ratio of a hospital record of a condition, among those that self-reported having a condition. Adjusted for age, sex, income, education, country of birth and functional limitation

^b Confidence interval

^c Model included both admission type and emergency status together with other listed patient characteristics

Supplementary Table 3. Adjusted ORs^a for hospital-level variables from the multilevel logistic regression with random intercept for hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)

Hospital characteristics	Hypertension (N = 15,279)		Diabetes (N = 4,794)		Heart disease (N = 8,307)		Stroke (N = 2,480)		Smoking (N = 2,099)		Obesity (N = 8,162)	
	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b
Hospital type^c												
Public	1		1		1		1		1		1	
Private	0.49	(0.38,0.63)	0.98	(0.78,1.23)	0.35	(0.26,0.47)	0.31	(0.22,0.43)	0.99	(0.72,1.35)	0.91	(0.64,1.31)
Hospital remoteness^c												
Major city	1		1		1		1		1		1	
Inner regional	0.89	(0.64,1.23)	0.86	(0.67,1.11)	1.01	(0.70,1.47)	1.29	(0.91,1.83)	1.04	(0.75,1.45)	0.91	(0.60,1.38)
Outer regional	0.75	(0.55,1.02)	0.69	(0.52,0.91)	0.97	(0.67,1.41)	1.09	(0.72,1.67)	0.82	(0.57,1.18)	0.91	(0.58,1.44)
Remote/very remote	1.05	(0.57,1.94)	0.53	(0.28,1.00)	1.70	(0.81,3.59)	0.66	(0.22,1.98)	0.33	(0.15,0.71)	0.52	(0.16,1.68)
Hospital size^c												
Principal referral	1		1		1		1		1		1	
Major	0.59	(0.34,1.01)	0.89	(0.62,1.27)	0.76	(0.44,1.34)	0.93	(0.58,1.47)	1.03	(0.66,1.61)	1.10	(0.59,2.05)
District	0.45	(0.27,0.76)	0.83	(0.58,1.19)	0.45	(0.26,0.78)	0.97	(0.60,1.55)	0.73	(0.47,1.15)	1.02	(0.55,1.91)
Community	0.41	(0.25,0.68)	0.61	(0.43,0.87)	0.35	(0.20,0.59)	0.57	(0.35,0.94)	0.89	(0.56,1.39)	0.88	(0.47,1.62)
Other	0.52	(0.30,0.89)	0.44	(0.29,0.68)	0.35	(0.19,0.65)	1.19	(0.66,2.14)	0.39	(0.22,0.68)	1.22	(0.59,2.53)
Depth of coding^c												
1 - least comprehensive	0.17	(0.11,0.27)	0.26	(0.17,0.40)	0.09	(0.04,0.17)	0.38	(0.17,0.82)	0.22	(0.12,0.42)	0.28	(0.12,0.65)
2	0.29	(0.22,0.38)	0.66	(0.52,0.85)	0.41	(0.29,0.56)	0.31	(0.21,0.48)	0.74	(0.52,1.06)	0.59	(0.38,0.92)
3	0.58	(0.45,0.76)	0.85	(0.66,1.08)	0.75	(0.55,1.02)	0.66	(0.48,0.91)	0.89	(0.65,1.24)	0.65	(0.43,0.99)
4 - most comprehensive	1		1		1		1		1		1	

^a Odds ratio of a hospital record of a condition, among those that self-reported having a condition. Adjusted for age, sex, income, education, country of birth and functional limitation

^b Confidence interval

^c Hospital-level covariates added one at a time, separately

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8
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	6-8
Bias	9	Describe any efforts to address potential sources of bias	9
Study size	10	Explain how the study size was arrived at	7, 10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	9-10
		(b) Describe any methods used to examine subgroups and interactions	9
		(c) Explain how missing data were addressed	8
		(d) If applicable, describe analytical methods taking account of sampling strategy	n/a
		(e) Describe any sensitivity analyses	9
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10-11, 25-26
		(b) Indicate number of participants with missing data for each variable of interest	25-26
Outcome data	15*	Report numbers of outcome events or summary measures	11, 27,28,29,31-34
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10, 12-13, 32-34
		(b) Report category boundaries when continuous variables were categorized	30,32-34
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-13,30
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
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	18-19
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-19
Generalisability	21	Discuss the generalisability (external validity) of the study results	18-19
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	20

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Variation in the recording of common health conditions in routine hospital data: study using linked survey and administrative data in New South Wales, Australia

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2014-005768.R1
Article Type:	Research
Date Submitted by the Author:	14-Jul-2014
Complete List of Authors:	Lujic, Sanja; University of Western Sydney, Centre for Health Research Watson, Diane; National Health Performance Authority, Randall, Deborah; University of Western Sydney, Centre for Health Research Simpson, Judy; University of Sydney, School of Public Health Jorm, Louisa; University of Western Sydney, Centre for Health Research; The Sax Institute,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Epidemiology
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

1
2
3 **Variation in the recording of common health conditions in routine hospital data: study**
4 **using linked survey and administrative data in New South Wales, Australia**
5
6
7
8
9

10
11 Sanja Lujic¹, Diane E Watson², Deborah A Randall¹, Judy M Simpson³, Louisa R Jorm^{1,4}
12
13

- 14
15
16
17 1. Centre for Health Research, University of Western Sydney, Sydney, Australia.
18 2. National Health Performance Authority, Sydney, Australia.
19 3. School of Public Health, University of Sydney, Sydney, Australia.
20 4. The Sax Institute, Sydney Australia.
21
22
23
24
25

26 Correspondence:

27 Ms Sanja Lujic
28 Centre for Health Research, University of Western Sydney, Locked Bag 1797, Penrith NSW
29 2751, Australia
30 E: s.lujic@uws.edu.au
31 T: +61 4620 3684
32 F: +61 4620 3299
33
34
35
36

37 Word count: 4,625 words
38
39

40 **Keywords:** hospital data, morbidity, multilevel modeling, agreement
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objectives: To investigate the nature and potential implications of underreporting of morbidity information in administrative hospital data.

Setting and participants: Retrospective analysis of linked self-report and administrative hospital data for 32,832 participants in the large-scale cohort study (45 and Up Study), who joined the study from 2006-2009 and who were admitted to 313 hospitals in New South Wales, Australia, for at least an overnight stay, up to a year prior to study entry.

Outcome measures: Agreement between self-report and recording of six morbidities in administrative hospital data, and between-hospital variation and predictors of positive agreement between the two data sources.

Results: Agreement between data sources was good for diabetes ($\kappa=0.79$), moderate for smoking ($\kappa=0.59$), fair for heart disease, stroke and hypertension ($\kappa=0.40$, $\kappa=0.30$, $\kappa=0.24$, respectively) and poor for obesity ($\kappa=0.09$), indicating that a large number of individuals with self-reported morbidities did not have a corresponding diagnosis coded in their hospital records. Significant between-hospital variation was found (ranging from 8% of unexplained variation for diabetes to 22% for heart disease), with higher agreement in public and large hospitals, and hospitals with greater depth of coding.

Conclusions: The recording of six common health conditions in administrative hospital data is highly variable, and for some conditions, very poor. To support more valid performance comparisons, it is important to stratify or control for factors that predict the completeness of recording, including hospital depth of coding and hospital type (public/private), and to increase efforts to standardize recording across hospitals. Studies using these conditions for risk adjustment should also be cautious of their use in smaller hospitals.

Keywords: hospital data, morbidity, multilevel modeling, agreement

Article summary

Article focus

- Explore the agreement between administrative hospital data and self-report information for four clinical conditions, as well as smoking and obesity.
- Quantify the between-hospital variation in the level of recording of these conditions.

Key messages

- Good levels of agreement found only for diabetes, with other conditions exhibiting moderate to poor agreement.
- Better recording was found in public and large hospitals, and hospitals with greater depth of coding.
- Significant between-hospital variation in the levels of agreement, amplified when smaller and private hospitals are included.

Strengths and limitations

- Use of linked data from a large-scale cohort study and advanced multilevel modeling methods to comprehensively evaluate the recording of common health conditions in hospital data, and explore between-hospital variation.
- Limitations include the absence of 'gold standard' such as medical records.

Introduction

Most nations with advanced economies publicly report on the comparative performance of hospitals with a view to accelerating and informing efforts to improve quality and allowing patients to make informed choices. Diagnoses recorded in administrative hospital data are commonly used in the construction and case-mix adjustment of hospital performance metrics, as well as for risk adjustment in epidemiological studies.

The construction of reliable health metrics relies on statistical methods that take into account the degree to which patients treated in different facilities have different morbidity and risk profiles that predispose them to requiring different interventions or to achieving different outcomes. These statistical methods, known as case-mix or risk adjustment, account for patient-related factors that are above and beyond the immediate control of health care professionals.

Thus, properly constructed performance metrics fairly reflect differences in health care experiences, patient outcomes and risks of adverse events. There has been some criticism of case-mix adjustments because they are subject to measurement error, [1] but case-mix adjustment is still considered to be less biased than unadjusted comparisons. [2]

Most methods of case-mix adjustment rely principally on demographic and diagnostic information that is captured in administrative hospital data collections. The hospital data is collected and recorded in a database for administrative purposes, with clinical coders coding diagnostic information based on the patient's medical records. [3] This approach may be sub-optimal [4 5] because evidence from many countries suggests that administrative hospital data underreport the morbidity information needed to fully account for differences between hospitals in patient-related factors that predispose them to differences in measured outcomes. [6-13] However, the impact of this underreporting on comparative measures of hospital

1
2
3 performance depends on whether it varies systematically among hospitals, because of
4 differences in factors such as training or practice among coding staff, the comprehensiveness
5 of clinicians' notes, or "upcoding" relating to funding models or incentives. [14]
6
7

8
9
10 This issue is relatively unexplored, aside from the work by Mohammed et al. [2] which
11 reported a non-constant relationship between case-mix variables and mortality among
12 hospitals in the UK, explained by differences in clinical coding and admission practices
13 across hospitals. These variations in coding accuracy were shown to be related to geographic
14 location and bed size, with small rural facilities performing better than large urban hospitals.
15 [15 16] In Australia, variations in the reporting and coding of secondary diagnoses in
16 administrative hospital data have been shown to exist in public hospitals among Australian
17 states, [17] and also among hospitals within the state of New South Wales (NSW), with
18 greater underreporting in private and rural hospitals. [3] However, the relative contributions
19 of patient and hospital factors to these variations have not been identified, nor has this
20 variation been formally quantified.
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35 This study, using data-linkage of survey and administrative data, aimed to further investigate
36 the nature and potential implications of underreporting of morbidity information in
37 administrative hospital data by: 1) measuring the agreement between self-reported morbidity
38 information and coded diagnoses; 2) quantifying the amount of between-hospital variation in
39 this agreement; and 3) identifying patient and hospital characteristics that predict higher or
40 lower levels of agreement. We focused on clinical conditions common to case-mix and risk-
41 adjustment models – diabetes, heart disease, hypertension and stroke. We also focus on
42 smoking and obesity, due to their impact on health trajectories, rapid shifts in prevalence,
43 substantial geographic variation in rates [18] and paucity of international evidence on
44 completeness of coding.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Methods

Data sources

The 45 and Up Study

The 45 and Up Study is a large-scale cohort study involving 267,153 men and women aged 45 years and over from the general population of NSW, Australia. The study is described in detail elsewhere. [19] Briefly, participants in the 45 and Up Study were randomly sampled from the database of Australia's universal health insurance provider, Medicare Australia, which provides near complete coverage of the population. People 80+ years of age and residents of rural and remote areas were oversampled. Participants joined the Study by completing a baseline questionnaire (available at <https://www.saxinstitute.org.au/our-work/45-up-study/questionnaires/>) between January 2006 and December 2009 and giving signed consent for follow-up and linkage of their information to routine health databases. About 18% of those invited participated and participants included about 11% of the NSW population aged 45 years and over. [19]

The NSW Admitted Patient Data Collection (APDC)

The APDC includes records of all public and private hospital admissions ending in a separation, i.e. discharge, transfer, type-change or death. Each separation is referred to as an episode of care. Diagnoses are coded according to the Australian modification of the International Statistical Classification of Diseases and Health Related Problems 10th Revision, ICD-10-AM. [20] Up to 55 diagnoses codes are recorded on the APDC, including the principal diagnosis and up to 54 additional diagnoses. Additional diagnoses are defined as 'a condition or complaint either coexisting with the principal diagnosis or arising during the episode of care' in the Australian Coding Standards and should be interpreted as

1
2
3 conditions that affect patient management. [21] Assignment of diagnosis codes is done by
4
5 trained clinical coders, using information from the patient's medical records.
6

7
8 The APDC from 1 July 2000 to 31 December 2010 was linked probabilistically to survey
9
10 information from the 45 and Up Study by the NSW Centre for Health Record Linkage
11
12 (www.cherel.org.au) using the 'best practice' protocol for preserving privacy. [22]
13
14

15 16 17 18 **Study population**

19
20 The study population comprised patients aged 45 years and above who participated in the 45
21
22 and Up Study and who had a hospitalization lasting at least one night in the period up to 365
23
24 days prior to filling out the baseline 45 and Up Study survey. Day stay patients were
25
26 excluded from the analysis to make the study more robust and generalizable beyond NSW
27
28 and Australia, as there are differences in admission practices for same day patients between
29
30 Australia and most other comparable countries.[23] NSW is home to 7.4 million people or
31
32 one-third of the population of Australia.
33
34
35
36
37
38

39 **Measuring morbidity**

40
41 We examined four health conditions (diabetes, heart disease, hypertension and stroke) and
42
43 two health risk factors (obesity and smoking), referred to hereafter collectively as
44
45 "morbidity". For each participant, these health conditions were measured using self-report
46
47 and administrative hospital data.
48
49

50
51 Self-reported morbidities were ascertained on the basis of responses to questions in the
52
53 baseline 45 and Up Study survey. Diabetes, hypertension, stroke and heart disease were
54
55 identified using the question "Has a doctor ever told you that you have [name of condition]?"
56
57
58 Participants who did not answer the question were excluded from analyses (n=1,242).
59
60

1
2
3 Smoking was classified on the basis of answering “yes” to both of the questions “Have you
4 ever been a regular smoker?” and “Are you a regular smoker now?”. Participants’ responses
5 to the questions “How tall are you without shoes?” and “About how much do you weigh?”
6 were used to derive body mass index (BMI), defined as body weight divided by height
7 squared (kg/m^2). The World Health Organization’s [24] classification system was used to
8 categorize individuals as obese ($\text{BMI} \geq 30\text{kg}/\text{m}^2$).
9

10
11
12 Morbidity information in administrative hospital data was ascertained using all 55 diagnosis
13 codes in the APDC records (ICD-10-AM: E10-E14 for diabetes, I20-I52 for heart disease,
14 I60-I69, G45, G46 for stroke, I10-I15 and R03.0 for hypertension, F17.2 or Z72.0 for
15 smoking and E66 for obesity). The inclusion of broader ICD-10-AM codes for heart disease
16 and stroke was chosen because of the broad definition of disease type in the self-reported
17 data. Thus, heart disease codes were inclusive of coronary heart disease, pulmonary heart
18 disease, and other forms of heart diseases including heart failure and arrhythmias. Stroke
19 codes included cerebrovascular diseases without infarction among others.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 **Predictors of agreement**

39
40 We explored both patient- and hospital-level factors as predictors of agreement between the
41 two data sources.
42
43

44
45 Patient-level factors were self-reported in the 45 and Up Study baseline survey and included
46 age, sex, education, country of birth, income and functional limitation. Functional limitation
47 was measured using the Medical Outcomes Study – Physical Functioning scale, [25] and
48 classified into 5 groups: no limitation (score of 100), minor limitation (score 95-99), mild
49 limitation (score 85-94), moderate limitation (60-84) and severe limitation (score 0-59).
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Facility-level factors were type of hospital (public/private), hospital peer group (akin to
4 hospital size defined by number of case-mix weighted separations, [26] which includes
5 hospital remoteness in the classification), remoteness of hospital and depth of coding.
6
7 Remoteness of the Statistical Local Area in which the hospital was located was classified
8 according to the Accessibility/Remoteness Index of Australia (ARIA+), grouped into four
9 categories (major city, inner regional, outer regional, remote/very remote). [27] Depth of
10 hospital coding was the mean number of additional diagnoses coded per episode of care for
11 each hospital, calculated using all overnight hospitalizations for the full 45 and Up Study
12 cohort from 2000 to 2010, and divided into four groups at the 25th, 50th and 75th percentile.
13
14 Hospital peer groups were divided into 5 categories: principal referral ($\geq 25,000$ separations
15 per year), major (10,000-24,999 separations per year), district (2,000-9,999 separations per
16 year), community (up to 2,000 separations per year) and other (non-acute, un-peered
17 hospitals). Missing information was treated as a separate category for any variables with
18 missing data.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 **Statistical methods**

38
39
40 We examined patient-level agreement between data sources for each of the six morbidities
41 individually, as well as for their 15 two-way combinations. We compared the self-reported
42 responses (yes/no) with all the diagnoses provided in the hospital records both for 'index'
43 admissions and for the 'lookback' period admissions. [28] The 'index' admission was the
44 overnight hospital stay with admission date closest to the survey completion date and no
45 longer than a year prior. Morbidity was coded as 'yes' if any of the diagnoses during that stay
46 contained a mention of that morbidity. The 'lookback' admissions included all overnight
47 stays in the 365-day period that preceded and included the 'index' admission. Morbidity was
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 coded as 'yes' if any of the diagnoses from any lookback admissions contained a mention of
4
5 that morbidity.
6

7
8 Agreement between the two data sources (yes/no) was measured using Cohen's kappa
9
10 statistic (κ). Kappa values above 0.75 denote excellent agreement, 0.40 to 0.75 fair to good
11
12 agreement and below 0.45 poor agreement. [29] Agreement was computed for all 313
13
14 hospitals in the state, regardless of size, as well as for the 82 largest public hospitals, for
15
16 which performance metrics are publicly reported.
17

18
19 Multilevel logistic regression was used to estimate odds ratios (OR) with 95% confidence
20
21 intervals (CI) for patient- and hospital-level factors that predicted positive agreement between
22
23 the two data sources. Multilevel models were chosen because of the clustering of patients
24
25 within hospitals. Models were run for each of the six morbidities separately. These analyses
26
27 were constrained to only those participants who self-reported the morbidity of interest, and
28
29 the outcome was whether the index hospital record contained a mention of the morbidity or
30
31 not. Addition of the hospital-level characteristics was done one at a time, due to the
32
33 collinearity between variables. All ORs presented are adjusted for all other demographic
34
35 variables in the model.
36
37

38
39 Variation at the hospital level was expressed as a median odds ratio (MOR), which is the
40
41 median of the odds ratios of pair-wise comparisons of patients taken from randomly chosen
42
43 hospitals, calculated as $\exp^{0.954 \times \sqrt{\text{variance}}}$; [30] and the intraclass correlation coefficient (ICC),
44
45 which is the percentage of the total variance attributable to the hospital level. [31] Large ICCs
46
47 indicate that differences among hospitals account for a considerable part of the variation in
48
49 the outcome, whereas a small ICC means that the hospital effect on the overall variation is
50
51 minimal. The relative influence of the hospital on reporting of morbidity was calculated using
52
53 a variance partitioning coefficient expressed as a percentage of the total variance using the
54
55 Snijders and Bosker latent variable approach. [31]
56
57
58
59
60

1
2
3 All data management was done using SAS 9.2 [32] and multilevel modeling using MLwiN
4
5 2.24. [33]
6

7
8 The conduct of the 45 and Up Study was approved by the University of New South Wales
9
10 Human Research Ethics Committee (HREC), while ethical approval for this particular study
11
12 was provided by the NSW Population and Health Services Research Ethics Committee and
13
14 the University of Western Sydney HREC.
15

16 17 18 19 20 **Results**

21 22 *Descriptive characteristics*

23
24
25 A total of 32,832 study participants were admitted to 313 hospitals up to a year prior to
26
27 completing the 45 and Up Study baseline survey. Just over half of the index admissions
28
29 (53%) were planned stays in hospital, and 57% were to a public hospital. Around one-third of
30
31 the index admissions occurred within the three months before study entry, and the mean
32
33 length of stay was 4.8 days (median = 3 days). Just under half of the sample (47%) reported
34
35 having hypertension, with heart disease or obesity reported by 25%, and current smoking by
36
37 6.1% of the sample. One-third (34%) of participants had two or more morbidities (data not
38
39 shown). Other characteristics of the sample at their index admission are shown in Table 1.
40
41
42
43 Characteristics of hospitals are summarized in Table 2.
44
45
46
47

48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 about here

Table 2 about here

Concordance between self-report and hospital records

Overall, reporting of morbidity differed between the two data sources with 23,257 (71%) participants having at least one of the six self-reported morbidities, and 11,977 (36.5%) and 14,335 (43.7%) of the sample having at least one morbidity recorded on their index or lookback hospital admissions, respectively.

Table 3 gives the summary concordance measures for each morbidity and two-way morbidity combination. For the index admission, good agreement was found for diabetes ($\kappa=0.79$), moderate agreement for smoking ($\kappa=0.59$), fair agreement for heart disease ($\kappa=0.40$), and poor agreement for stroke ($\kappa=0.30$), hypertension ($\kappa=0.24$) and obesity ($\kappa=0.09$). In two-way combinations, moderate levels of agreement were found only for diabetes combinations (with smoking, hypertension and heart disease).

Table 3 about here

Incorporating a one-year lookback period increased the numbers of participants with a morbidity recorded in a hospital record, with average relative increases in the kappa values of 20% (ranging from 2% increase for smoking, to 41% increase for obesity). Good to excellent level of agreements were still found only for diabetes ($\kappa=0.83$) and smoking ($\kappa=0.60$).

Agreement was only slightly higher among the 82 large public hospitals (see Supplementary Table 1) with relative kappa values higher by 4%, on average.

Patient- and hospital-level predictors of positive agreement

1
2
3 The patient factors which predicted positive agreement between the two data sources differed
4
5 between morbidities (Table 4). Male sex was associated with better agreement for diabetes
6
7 (OR=1.37, 95% CI 1.19 – 1.58), heart disease (OR=1.30, 95% CI 1.17 – 1.44) and
8
9 hypertension (OR=1.28, 95% CI 1.18 – 1.38) (Supplementary Table 2).

10
11
12 Older patients were significantly less likely to have smoking (80+ years OR=0.48, 95% CI
13
14 0.31 – 0.74) and obesity (OR=0.14, 95% CI 0.08 – 0.26) recorded in their hospital records,
15
16 and significantly more likely to have hypertension recorded (OR=1.32, 95% CI 1.16 – 1.49),
17
18 compared to younger patients (45 – 59 years). People with higher levels of functional
19
20 limitation were significantly more likely to have hypertension, diabetes and obesity recorded
21
22 on their most recent hospital stay. Planned admissions to hospital had lower odds of having
23
24 any of the six conditions recorded, as did medical admissions (for diabetes, smoking and
25
26 obesity only). Agreement did not vary significantly for any other patient factors.
27
28
29
30
31
32

33 Table 4 about here

34
35
36
37
38
39 The four hospital-level covariates (hospital type, hospital peer group, hospital remoteness and
40
41 depth of coding) were added to multilevel models (including a random intercept for hospital)
42
43 one at a time, separately. Positive agreement between self-report and hospital records was
44
45 significantly lower for hospitals with lower depth of coding across all morbidities. The odds
46
47 of recording were also lower among private hospitals for all six morbidities, with this
48
49 difference being statistically significant for hypertension, heart disease and stroke only.
50
51
52 Records from smaller hospitals (district and community peer groups) were significantly less
53
54 likely to agree with self-reported data on hypertension, diabetes and heart disease. Positive
55
56 agreement did not vary significantly with remoteness of hospital, with the exceptions of
57
58
59
60

1
2
3 diabetes (lower agreement for outer regional, remote and very remote hospitals) and smoking
4
5 (lower agreement for remote and very remote hospitals) (Supplementary Table 3).
6
7
8
9

10 *Quantifying variation between hospitals*

11
12 Before any hospital-level variables were added into the multilevel model, the intraclass
13 correlation coefficient indicated that between 8% (diabetes) and 22% (heart disease) of the
14 residual (unexplained) variation in agreement was attributable to the hospital, after
15 adjustment for the patient-level factors (Table 5). This equated to median odds ratios (MORs)
16 of 1.64 and 2.48, respectively, indicating that a patient in one hospital had an average of
17 between 64% and 148% higher odds of having a particular morbidity recorded than a patient
18 in a hospital with lower levels of recording. Less variation at the hospital level was found for
19 the recording of diabetes, smoking and stroke, while more variation at the hospital level was
20 found for the recording of hypertension, heart disease and obesity. When the analyses were
21 restricted to 82 large public hospitals only, the between-hospital variation decreased to
22 between 2% (stroke) and 13% (hypertension), or MOR of 1.24 and 1.94 (Figure 1). This
23 between-hospital variation was still significant for all morbidities except for stroke. Between-
24 hospital variation was further reduced once lookback admissions were used to identify
25 morbidities.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 Figure 1 about here

50
51
52
53 The addition of hospital-level variables to multilevel models, one at the time, separately,
54 helped ascertain which factors explained the variation between hospitals (Table 5). The
55 addition of hospital-level factors contributed to explaining (i.e. decreasing) the residual
56
57
58
59
60

1
2
3 variation for all conditions, except obesity. For the other morbidities, differences in the depth
4 of coding explained from 16% (smoking) to 42% (hypertension) of residual variation
5 between hospitals, while hospital type (public/private) explained from 0% (smoking) to 59%
6 (stroke), and hospital peer group explained from 10% (hypertension) to 27% (diabetes)
7 residual variation between hospitals.
8
9
10
11
12
13

14
15
16
17 Table 5 about here
18
19

20 21 22 23 **Discussion**

24
25 Our study found that the concordance of administrative hospital and self-reported data varied
26 between the six morbidities examined, with agreement ranging from good for diabetes,
27 moderate for smoking, through to fair for heart disease, and poor for hypertension, stroke and
28 obesity. We demonstrated considerable between-hospital variation in the recording of these
29 common health conditions. Smaller, but still significant, between-hospital variation was
30 found when restricting the analyses to the largest public hospitals in the state.
31
32
33
34
35
36
37
38

39 Previous studies have validated information recorded in NSW administrative hospital data for
40 demographic factors, [34 35] and recording of perinatal conditions, [36-39] but there have
41 been limited studies of the accuracy of the recording of health conditions commonly used for
42 case-mix or risk-adjustment. Our findings regarding agreement for the recording of diabetes
43 ($\kappa=0.83$) were similar to previous Australian studies [3 10], while agreement for hypertension
44 ($\kappa=0.30$) and heart disease ($\kappa=0.47$) was considerably lower in our study. These differences
45 may be due to the fact that both previous studies used medical records as a 'gold standard',
46 while we used self-report. Lower agreement rates for heart disease could be due to the
47 broader range of heart disease types included in our study, with known lower levels of
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 agreement for heart failure compared to myocardial infarction.[9],[40] Higher sensitivities
4
5 reported in a study from the state of Victoria [10] could also be attributable to the differences
6
7 in public hospital funding models between the two states. Specifically, Victoria has used
8
9 activity- based funding since 1993, while this method of funding was introduced in NSW and
10
11 other Australian states only subsequent to our study period. [41] Introduction of activity-
12
13 based funding has been shown to increase recording of additional diagnoses and procedures
14
15 in Europe. [42]

16
17
18 Some of the apparent discrepancies in the levels of coding between conditions can be
19
20 attributed to the coding rules that govern whether or not a diagnosis is recorded in
21
22 administrative hospital data. Additional diagnoses, recorded on administrative hospital data,
23
24 are coded only if they affect the patient's treatments received, investigations required and/or
25
26 resources used during the hospital stay. Thus, diagnoses that relate to an earlier episode, and
27
28 which have no bearing on the current hospital stay, are not coded for that particular stay.
29
30 Therefore, it is not surprising that (managed) hypertension, in particular, might not be
31
32 recorded in hospital data relating to, for example, elective surgery. On the other hand, we
33
34 found that diabetes is well recorded, suggesting that it is considered to affect patient
35
36 management in most hospital stays, and possibly reflecting the impact of changes to the
37
38 Australian Coding Standards for diabetes such that between 2008 and 2010 diabetes with
39
40 complications could be coded even where there was no established cause and effect
41
42 relationship between diabetes and the complication. [43] It is for these reasons that
43
44 researchers using administrative datasets are encouraged to incorporate information from
45
46 previous hospitalizations, to increase the likelihood of capturing morbidity, as demonstrated
47
48 in this as well as other Australian studies.[44]As well as looking at single morbidities, ours is
49
50 the first study to our knowledge to explore the variations of recording of multiple conditions
51
52 in hospital data. Concordance of two-way condition combinations was very low, with best
53
54
55
56
57
58
59
60

1
2
3 results found for combinations of diseases involving diabetes, which had the highest single-
4
5 condition level of agreement with self-reported data ($\kappa=0.83$). Agreement measures for two-
6
7 way combinations were found to be fair to good at best, with agreement on three-way
8
9 condition combinations (not investigated here) expected to be even lower. These findings
10
11 have implications for research into multimorbidity (the co-occurrence of multiple chronic or
12
13 acute diseases and medical conditions within one person [45]). We suggest that researchers
14
15 who use administrative data for research into multimorbidity should use linked data to
16
17 increase ascertainment, and, if possible, supplement this information from other data sources,
18
19 such as physician claims data or self-reported data.
20
21

22
23 We identified considerable between-hospital variability in the levels of recording of common
24
25 health conditions, with between 8% and 22% of the variation attributable to hospital-level
26
27 factors, after adjustment for patient factors. This was similar in magnitude to the variability
28
29 previously reported for performance measures (varying from patient satisfaction, mortality,
30
31 length of stay to quality of care) clustered at the facility level (0-51%) [46] and hospital-level
32
33 variations in the use of services. [47-49] Significant between-hospital variation was still
34
35 present after constraining the analyses to the 82 largest public hospitals in the state.
36
37

38
39 The recording of hypertension and heart disease was particularly variable between hospitals,
40
41 those with better reporting having on average 2.3 and 2.5 times, respectively, the odds of
42
43 recording these conditions than those with lower levels of reporting. The corresponding
44
45 figures were 1.9 and 1.6 times for the 82 largest hospitals in the state. These findings indicate
46
47 the potential for reporting bias to influence comparisons of health performance indicators
48
49 between hospitals, especially for indicators that use conditions such as heart disease or
50
51 hypertension for case-mix adjustment. To our knowledge, no previous studies have provided
52
53 detailed information about how the validity of morbidity reporting varies among hospitals
54
55 after accounting for patient factors.
56
57
58
59
60

1
2
3 Further, we have shown that variations in the accuracy of morbidity reporting between
4 hospitals are predominantly driven by the hospital's depth of coding – concordance between
5 self-reported and hospital data is lower in hospitals with a lower average number of
6 additional diagnoses recorded. Up to 42% of the variation in recording at the hospital level
7 could be attributed to differences in hospital depth of coding. Even though the measure of
8 depth of coding we used was crude, and related to hospital size, it still helps in highlighting
9 the impact of coding practices on variations among hospitals. Other research using the same
10 depth of coding measure has shown that the lower depth of coding can disproportionately
11 disadvantage hospitals' standardised mortality ratios, one of the commonly reported measures
12 of hospital performance. [2] It will be important to track changes in the levels of the depth of
13 coding across Australian states, and to consider the implications of these for state-based
14 performance comparisons, following the national rollout of activity-based funding and
15 comparative performance reporting.

16
17
18 Several factors might explain variation in depth of coding between hospitals. Clinical coders
19 can code only information that has been recorded in the patient's medical record, so varying
20 level of details recorded by clinicians will influence what gets coded. The training and
21 professional development opportunities for coding staff might also influence the depth of
22 coding. Also, casemix funding systems, such as the Diagnosis Related Group (DRG)
23 classification, are prone to 'upcoding' in order for services to receive higher reimbursement
24 costs. [14]

25
26
27 We found that the reporting of conditions varied with hospital size, larger metropolitan
28 hospitals having higher concordance, with kappa values higher by 7% on average when
29 comparing large tertiary with smaller urban hospitals. This finding echoes those of Powell et
30 al. [3] in NSW, Australia during 1996 – 1998 and Rangachari et al. [16] in the US, during
31 2000 – 2004. Our study showed that large tertiary hospitals had better concordance for the

1
2
3 recording of hypertension and heart disease than smaller urban hospitals, but the reverse was
4
5 true for stroke and smoking. Our finding that between-hospital variation in the recording of
6
7 morbidities was up to two times higher when all hospitals, rather than just the largest ones,
8
9 were included has implications for further research using data from smaller hospitals. This
10
11 high variability in concordance among smaller hospitals may mean that morbidity-adjusted
12
13 comparisons are not as valid as for larger hospitals. Researchers using information from these
14
15 hospitals are encouraged to supplement their data with either self-report information and/or
16
17 data linkage. The value-add of incorporating previous hospitalizations was also highlighted in
18
19 our results for stroke and obesity, with 43% – 47% more patients identified using lookback
20
21 admissions than from a single admission only.
22
23
24

25
26 A particular strength of our study lies in the use of linked data from a large-scale cohort study
27
28 to comprehensively evaluate the recording of common conditions in hospital data, and
29
30 explore the variation in recording among hospitals. The 45 and Up Study contains records for
31
32 one in every 10 persons aged 45 and over in NSW, so it provides a rich resource to answer
33
34 research questions. Additionally, we used advanced multilevel modeling methods to quantify
35
36 the amount of between-hospital variation in the level of recording of common health
37
38 conditions, a finding which is of importance for both research and policy paradigms due to its
39
40 impact on adjusted comparisons among hospitals and the highlighted need to improve
41
42 consistency of recording in hospitals across the State. To date, hospital-level variation has
43
44 only been explored with a set outcome (e.g. mortality, readmission) in mind.
45
46
47

48
49 A potential limitation of our study was its use of self-reported information to explore
50
51 concordance, in the absence of another ‘gold standard’, such as medical records. Access to
52
53 medical records was not possible given the de-identified nature of our data, and the large
54
55 number of records in the dataset. Moreover, studies that have examined accuracy of self-
56
57 reported conditions against medical records have found high levels of agreement, ranging
58
59
60

1
2
3 from 81% [50] to 87% [51] for hypertension, 66% [40] to 96% [50 51] for diabetes and 60%
4
5 [50] to 98% [52] for acute myocardial infarction. Validation studies in the 45 and Up Study
6
7 cohort have reported strong correlations and excellent levels of agreement between self-
8
9 reported and measured height and weight, and derived BMI [53] as well as self-reported
10
11 diabetes. [54] Although the 45 and Up Study had a response rate of 18%, the study sample is
12
13 very large and has excellent heterogeneity. Furthermore, exposure-outcome relationships
14
15 estimated from the 45 and Up Study data have been shown to be consistent with a large
16
17 ‘representative’ population survey of the same population.[55]
18
19

20 21 **Conclusion**

22
23 The recording of common comorbid conditions in routine hospital data is highly variable,
24
25 and, for some conditions, very poor. Recording varies considerably among hospitals,
26
27 presenting the potential to introduce bias into risk-adjusted comparisons of hospital
28
29 performance, especially for indicators that use heart disease or hypertension for risk
30
31 adjustment. Furthermore, between-hospital variation is amplified when smaller and private
32
33 hospitals are included in the analyses. Stratification of analyses according to factors that
34
35 predict the completeness of recording, including hospital depth of coding and hospital type
36
37 and size, supplementing morbidity information with linked data from previous
38
39 hospitalizations and increases in efforts to standardize recording across hospitals, all offer
40
41 potential for increasing the validity of risk-adjusted comparisons.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

We thank all of the men and women who participated in the 45 and Up Study. The 45 and Up Study is managed by the Sax Institute (www.saxinstitute.org.au) in collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; *beyondblue*; Ageing, Disability and Home Care, Department of Family and Community Services; the Australian Red Cross Blood Service; and UnitingCare Ageing.

We would like to acknowledge the Sax Institute and the NSW Ministry of Health for allowing access to the data, and the Centre for Health Record Linkage for conducting the probabilistic linkage of records.

We are grateful to Dr Fiona Blyth and Dr Kris Rogers on their advice at the early stages of the project.

Funding statement

The Assessing Preventable Hospitalisation Indicators (APHID) study is funded by a National Health and Medical Research Council Partnership Project Grant (#1036858) and by partner agencies the Australian Commission on Safety and Quality in Health Care, the Agency for Clinical Innovation and the NSW Bureau of Health Information.

Contributorship

SL had overall responsibility for the design of this study, data management, statistical analysis and drafting this paper. DW and LJ contributed to the conception and design of the study. LJ helped with data acquisition, and provided oversight for all analyses. DR and JS

1
2
3 provided oversight and advice for the design and interpretation of the statistical analyses. All
4
5 authors contributed to the interpretation of the findings, the writing of the paper and approved
6
7 the final draft.
8
9

10
11
12
13 **Competing interests:** None
14

15
16
17
18 **Data sharing statement**
19

20
21 No additional data are available.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

1. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004;**363**(9415):1147-54
2. Mohammed MA, Deeks JJ, Girling A, et al. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* 2009;**338**
3. Powell H, Lim LL, Heller RF. Accuracy of administrative data to assess comorbidity in patients with heart disease: an Australian perspective. *J Clin Epidemiol* 2001;**54**(7):687-93
4. Pine M, Jordan HS, Elixhauser A, et al. Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA* 2007;**297**(1):71-76
5. Pine M, Norusis M, Jones B, et al. Predictions of Hospital Mortality Rates: A Comparison of Data Sources. *Ann Intern Med* 1997;**126**(5):347-54
6. Hawker GA, Coyte PC, Wright JG, et al. Accuracy of administrative data for assessing outcomes after knee replacement surgery. *J Clin Epidemiol* 1997;**50**(3):265-73
7. Kieszak SM, Flanders WD, Kosinski AS, et al. A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *J Clin Epidemiol* 1999;**52**(2):137-42
8. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Med Care* 2002;**40**(8):675-85
9. Preen DB, Holman CAJ, Lawrence DM, et al. Hospital chart review provided more accurate comorbidity information than data from a general practitioner survey or an administrative database. *J Clin Epidemiol* 2004;**57**(12):1295-304
10. Henderson T, Shephard J, Sundararajan V. Quality of diagnosis and procedure coding in ICD-10 administrative data. *Med Care* 2006;**44**(11):1011-9
11. Quach S, Blais C, Quan H. Administrative data have high variation in validity for recording heart failure. *Can J Cardiol* 2010;**26**(8):e306-e12
12. Leal J, Laupland K. Validity of ascertainment of co-morbid illness using administrative databases: a systematic review. *Clin Microbiol Infect* 2010;**16**(6):715-21
13. Chong WF, Ding YY, Heng BH. A comparison of comorbidities obtained from hospital administrative data and medical charts in older patients with pneumonia. *BMC Health Serv Res* 2011;**11**(1):105
14. Steinbusch PJ, Oostenbrink JB, Zuurbier JJ, et al. The risk of upcoding in casemix systems: A comparative study. *Health Policy* 2007;**81**(2):289-99
15. Lorence DP, Ibrahim IA. Benchmarking Variation in Coding Accuracy Across the United States. *J Health Care Finance* 2003;**29**(4):29
16. Rangachari P. Coding for quality measurement: the relationship between hospital structural characteristics and coding accuracy from the perspective of quality measurement. *Perspectives in health information management/AHIMA, American Health Information Management Association* 2007;**4**
17. Coory M, Cornes S. Interstate comparisons of public hospital outputs using DRGs: Are they fair? *Aust N Z J Public Health* 2005;**29**(2):143-48
18. National Health Performance Authority. Healthy Communities: Overweight and obesity rates across Australia, 2011–12 (In Focus). 2013.

19. Banks E, Redman S, Jorm L, et al. Cohort profile: the 45 and up study. *Int J Epidemiol* 2008;**37**(5):941
20. National Centre for Classification in Health. International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification (ICD-10-AM), Australian Classification of Health Interventions (ACHI). Sydney: National Centre for Classification in Health, 2006.
21. Australian Institute of Health and Welfare. National Health Data Dictionary. Version 16. Canberra: AIHW, 2012.
22. Kelman CW, Bass AJ, Holman CDJ. Research use of linked health data — a best practice protocol. *Aust N Z J Public Health* 2002;**26**(3):251-55
23. Eagar K. Counting acute inpatient care. ABF Information Series No 5 2010. <https://ahsri.uow.edu.au/content/groups/public/@web/@chsd/documents/doc/uow082637.pdf> (accessed July 2014).
24. World Health Organization. *Obesity: preventing and managing the global epidemic*: World Health Organization, 2000.
25. Stewart AL, Ware JE. *Measuring functioning and well-being: The medical outcomes study approach*: Duke University Press, 1992.
26. NSW Health. NSW Health Services Comparison Data Book 2008/09. 2010. http://www0.health.nsw.gov.au/pubs/2010/pdf/yellowbook_09.pdf (accessed April 2014).
27. Australian Institute of Health and Welfare. Regional and remote health: A guide to remoteness classifications Canberra: AIHW, 2004.
28. Zhang JX, Iwashyna TJ, Christakis NA. The performance of different lookback periods and sources of information for Charlson comorbidity adjustment in Medicare claims. *Med Care* 1999;**37**(11):1128-39
29. Fleiss JL, Levin B, Paik MC. The measurement of interrater agreement. *Statistical methods for rates and proportions* 1981;**2**:212-36
30. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health* 2006;**60**(4):290-97
31. Snijders T, Bosker R. *Multilevel analysis*. 2nd ed. London: Sage, 2012.
32. SAS Version 9.3 [software] [program]. Cary, North Carolina, 2010.
33. MLwiN Version 2.25 [software] [program]. Centre for Multilevel Modelling, University of Bristol, Bristol., 2012.
34. Bentley JP, Taylor LK, Brandt PG. Reporting of Aboriginal and Torres Strait Islander peoples on the NSW Admitted Patient Data Collection: the 2010 data quality survey. *N S W Public Health Bull* 2012;**23**(1-2):17-20
35. Tran DT, Jorm L, Lujic S, et al. Country of birth recording in Australian hospital morbidity data: accuracy and predictors. *Aust N Z J Public Health* 2012;**36**(4):310-16
36. Taylor LK, Travis S, Pym M, et al. How useful are hospital morbidity data for monitoring conditions occurring in the perinatal period? *Aust N Z J Obstet Gynaecol* 2005;**45**(1):36-41
37. Roberts CL, Bell JC, Ford JB, et al. The accuracy of reporting of the hypertensive disorders of pregnancy in population health data. *Hypertens Pregnancy* 2008;**27**(3):285-97
38. Hadfield RM, Lain SJ, Cameron CA, et al. The prevalence of maternal medical conditions during pregnancy and a validation of their reporting in hospital discharge data. *Aust N Z J Obstet Gynaecol* 2008;**48**(1):78-82

- 1
2
3 39. Lain SJ, Hadfield RM, Raynes-Greenow CH, et al. Quality of data in perinatal population health
4 databases: a systematic review. *Med Care* 2012;**50**(4):e7-20
- 5
6 40. Okura Y, Urban LH, Mahoney DW, et al. Agreement between self-report questionnaires and
7 medical record data was substantial for diabetes, hypertension, myocardial infarction and
8 stroke but not for heart failure. *J Clin Epidemiol* 2004;**57**(10):1096
- 9
10 41. Eagar KM. What is activity-based funding? ABF Information Series No. 1. 2011.
11 <http://ro.uow.edu.au/gsbpapers/47/> (accessed April 2014).
- 12
13 42. O'Reilly J, Busse R, Häkkinen U, et al. Paying for hospital care: the experience with
14 implementing activity-based funding in five European countries. *Health Economics Policy
15 and Law* 2012;**7**(1):73
- 16
17 43. Knight L, Halech R, Martin C, et al. Impact of changes in diabetes coding on Queensland hospital
18 principal diagnosis morbidity data. 2011.
19 http://www.health.qld.gov.au/hsu/tech_report/techreport_9.pdf (accessed July 2014).
- 20
21 44. Preen DB, Holman CDAJ, Spilsbury K, et al. Length of comorbidity lookback period affected
22 regression model performance of administrative health data. *J Clin Epidemiol*
23 2006;**59**(9):940-46
- 24
25 45. van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity: what's in a name? A
26 review of literature. *Eur J Gen Pract* 1996;**2**(2):65-70
- 27
28 46. Fung V, Schmittiel JA, Fireman B, et al. Meaningful variation in performance: a systematic
29 literature review. *Med Care* 2010;**48**(2):140-48
- 30
31 47. Bennett-Guerrero E, Zhao Y, O'Brien SM, et al. Variation in use of blood transfusion in coronary
32 artery bypass graft surgery. *JAMA* 2010;**304**(14):1568-75
- 33
34 48. Seymour CW, Iwashyna TJ, Ehlenbach WJ, et al. Hospital-Level Variation in the Use of Intensive
35 Care. *Health Serv Res* 2012;**47**(5):2060-80
- 36
37 49. Randall DA, Jorm LR, Lujic S, et al. Disparities in Revascularization Rates After Acute
38 Myocardial Infarction Between Aboriginal and Non-Aboriginal People in Australia: Clinical
39 Perspective. *Circulation* 2013;**127**(7):811-19
- 40
41 50. Merkin S, Cavanaugh K, Longenecker JC, et al. Agreement of self-reported comorbid conditions
42 with medical and physician reports varied by disease among end-stage renal disease patients.
43 *J Clin Epidemiol* 2007;**60**(6):634
- 44
45 51. Barber J, Muller S, Whitehurst T, et al. Measuring morbidity: self-report or health care records?
46 *Fam Pract* 2010;**27**(1):25-30
- 47
48 52. Machón M, Arriola L, Larrañaga N, et al. Validity of self-reported prevalent cases of stroke and
49 acute myocardial infarction in the Spanish cohort of the EPIC study. *J Epidemiol Community
50 Health* 2013;**67**(1):71-75
- 51
52 53. Ng SP, Korda R, Clements M, et al. Validity of self-reported height and weight and derived body
53 mass index in middle-aged and elderly individuals in Australia. *Aust N Z J Public Health*
54 2011;**35**(6):557-63
- 55
56 54. Comino EJ, Tran DT, Haas M, et al. Validating self-report of diabetes use by participants in the 45
57 and up study: a record linkage study. *BMC Health Serv Res* 2013;**13**(1):481
- 58
59 55. Mealing NM, Banks E, Jorm LR, et al. Investigation of relative risk estimates from studies of the
60 same population with contrasting response rates and designs. *BMC Med Res Methodol*
2010;**10**:26

1
2
3 **List of tables and figures**
4

5 **Table 1.** Characteristics of the study sample at their index admission
6

7 **Table 2.** Characteristics of the hospital of admission
8

9 **Table 3.** Agreement measures between self-report and hospital data, index and lookback
10 admissions, all public and private hospitals in New South Wales, Australia (n=313)
11

12 **Table 4.** Factors that predict positive agreement between self-report and hospital data, using
13 multilevel modeling, all public and private hospitals in New South Wales, Australia (n=313)
14

15 **Table 5.** Variance and intraclass correlation coefficient for hospital-level random effects
16 from multilevel logistic regression, all public and private hospitals in New South Wales,
17 Australia (n=313)
18
19

20
21
22
23 **Figure 1 Title.** Variance for hospital-level random effects from multilevel logistic regression,
24 for index and lookback admissions, by hospital size
25

26 **Figure 1 Legend:**

- 27 * Significantly different from 0 at 5% level
28 ** Significantly different from 0 at 1% level
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Characteristics of the study sample at their index admission

	All participants	
	(N = 32,832)	
	N	%
<i>Demographic characteristics</i>		
Sex		
Male	16,812	51.2
Female	16,020	48.8
Age		
45-59	9,666	29.4
60-79	16,624	50.6
80+	6,540	19.9
Country of birth		
Australia	25,001	76.2
Other	7,448	22.7
Unknown	383	1.2
Highest education level		
No school	5,196	15.8
Year 10 or equivalent	7,894	24.0
Year 12 or equivalent	2,975	9.1
Trade	4,270	13.0
Certificate	6,109	18.6
University degree	5,662	17.3
Unknown	726	2.2
Household income (\$, per annum)		
<20,000	9,077	27.7
20,000 - <50,000	8,223	25.1
50,000 - <70,000	2,560	7.8
70,000+	5,042	15.4
Not disclosed	6,003	18.3
Missing	1,927	5.9
Functional status		
No limitation	4,915	15.0
Mild limitation	6,011	18.3
Moderate limitation	8,701	26.5
Severe limitation	10,121	30.8
Missing	3,084	9.4
<i>Admission characteristics</i>		
Admission type		
Surgical	15,464	47.1
Other	1,439	4.4
Medical	15,929	48.5
Emergency status		
Emergency	13,484	41.1

1			
2			
3	Planned	17,544	53.4
4	Other	1,803	5.5
5	<i>Hospital of admission</i>		
6			
7	Hospital type		
8	Public	18,734	57.1
9	Private	14,096	42.9
10			
11	Hospital remoteness		
12	Major city	19,754	60.2
13	Inner regional	8,424	25.7
14	Outer regional	4,137	12.6
15	Remote/very remote	363	1.1
16			
17	Hospital depth of coding		
18	1 - least comprehensive	1,629	5.0
19	2	8,803	26.8
20	3	11,543	35.2
21	4 - most comprehensive	10,857	33.1
22			
23	Hospital peer group		
24	Principal referral	6,329	19.3
25	Major	11,052	33.7
26	District	6,862	20.8
27	Community	7,018	21.4
28	Other	1,571	4.8
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			

Table 2. Characteristics of the hospital of admission

	All hospitals^a	
	(N = 313)	
	N	%
Hospital type		
Public	224	71.6
Private	88	28.1
Hospital remoteness		
Major city	124	39.6
Inner regional	72	23.0
Outer regional	94	30.0
Remote/very remote	20	6.4
Hospital depth of coding		
1 - least comprehensive	48	15.3
2	91	29.1
3	89	28.4
4 - most comprehensive	85	27.2
Hospital peer group		
Principal referral	14	4.5
Major	33	10.5
District	51	16.3
Community	121	38.7
Other	94	30.0

Table 3. Agreement measures between self-report and hospital data, index and lookback admissions, all public and private hospitals in New South Wales, Australia (n=313)

Morbidities ^a	Index admission						Lookback admissions					
	45 and Up Yes:		45 and Up No:		Kappa %	95% CI	45 and Up Yes:		45 and Up No:		Kappa %	95% CI
	APDC yes	APDC no	APDC yes	APDC no			APDC yes	APDC no	APDC yes	APDC no		
Hypertension	4,767	10,512	1,434	16,119	24.0	(22.9-25.0)	6,260	9,019	2,051	15,502	30.2	(29.1-31.2)
Heart disease	3,639	4,668	1,942	22,583	40.3	(39.0-41.5)	4,673	3,634	2,697	21,828	47.0	(45.8-48.2)
Diabetes	3,560	1,234	347	27,691	79.1	(78.1-80.1)	3,928	866	479	27,559	83.0	(82.1-83.9)
Stroke	541	1939	306	30,046	29.8	(27.0-32.6)	776	1,704	488	29,864	38.3	(35.8-40.8)
Smoking	1,205	804	727	30,096	58.7	(56.7-60.7)	1,411	598	1,076	29,747	60.1	(58.2-61.9)
Obesity	551	7,611	114	24,556	9.1	(7.3-10.9)	810	7,352	209	24,461	12.8	(11.1-14.6)
Hypertension + heart disease	1,172	3,481	1,270	26,909	25.8	(23.8-27.7)	1,807	2,846	2,008	26,171	34.3	(32.6-36.0)
Hypertension + diabetes	1,819	1,238	759	29,016	61.3	(59.6-62.9)	2,186	871	1,021	28,754	66.6	(65.2-68.1)
Hypertension + stroke	203	1,317	189	31,123	19.7	(15.7-23.7)	329	1,191	340	30,972	28.0	(24.5-31.5)
Hypertension + smoking	133	598	180	31,921	24.5	(19.2-29.7)	199	532	319	31,782	30.6	(26.0-35.2)
Hypertension + obesity	234	4,574	93	27,931	7.4	(4.9-9.8)	383	4,425	183	27,841	11.5	(9.2-13.9)
Heart disease + diabetes	646	1,154	404	30,628	43.0	(40.3-45.8)	904	896	661	30,371	51.2	(48.9-53.6)
Heart disease + stroke	76	973	126	31,657	11.2	(6.1-16.4)	149	900	261	31,522	19.0	(14.4-23.5)
Heart disease + smoking	76	294	222	32,240	22.0	(15.3-28.6)	118	252	373	32,089	26.5	(20.8-32.2)
Heart disease + obesity	79	1,938	79	30,736	6.4	(2.5-10.4)	151	1,866	169	30,646	11.4	(7.7-15.2)
Diabetes + stroke	85	555	58	32,134	21.1	(15.0-27.3)	140	500	119	32,073	30.4	(24.9-35.8)
Diabetes + smoking	143	161	108	32,420	51.1	(45.3-56.9)	171	133	176	32,352	52.1	(46.7-57.4)
Diabetes + obesity	232	1,701	65	30,834	19.5	(15.9-23.2)	351	1,582	120	30,779	27.5	(24.2-30.9)
Stroke + smoking	13	142	28	32,649	13.1	(0.1-26.1)	23	132	57	32,620	19.3	(7.8-30.8)
Stroke + obesity	6	558	9	32,259	2.0	(0.0-10.0)	13	551	21	32,247	4.2	(0.0-11.9)
Smoking + obesity	27	447	29	32,329	9.9	(1.9-17.9)	38	436	47	32,311	13.2	(5.5-20.9)

^a ICD-10-AM codes: hypertension (I10-I15, R03.0), heart disease (I20-I52), diabetes (E10-E14), stroke (I60-I69, G45, G46), smoking (F17.2, Z72.0), obesity (E66)

Table 4. Factors that predict positive agreement between self-report and hospital data, using multilevel modelling, all public and private hospitals in New South Wales, Australia (n=313)

	Hypertension (N = 15,279)	Diabetes (N = 4,794)	Heart disease (N = 8,307)	Stroke (N = 2,480)	Smoking (N = 2,099)	Obesity (N = 8,162)
Person-level variables						
Sex ¹	**	**	**			
Age ¹	**				**	**
Education ¹			*	**	**	
Country of birth ¹						
Functional limitation ¹	**	**				**
Income ¹						
Admission type ²	**	**	**	**	**	**
Emergency status ²	**	**	**	**		**
Hospital-level variables						
Hospital type (public/private) ³	**		**	**		
Hospital remoteness ³					*	
Hospital depth of coding ³	**	**	**	**	**	**
Hospital peer group ³	**	**	**		**	

* Significant at 5% level

** Significant at 1% level

1 – Model 0: adjusted for demographic factors + random intercept for hospital

2 – Model 0 + admission type + emergency status

3 – Model 0 + hospital-level variables (entered one at a time)

Table 5. Variance and intraclass correlation coefficient for hospital-level random effects from multilevel logistic regression, all public and private hospitals in New South Wales, Australia (n=313)

	Hypertension (N [†] = 15,279)	Diabetes (N [†] = 4,794)	Heart disease (N [†] = 8,307)	Stroke (N [†] = 2,480)	Smoking (N [†] = 2,099)	Obesity (N [†] = 8,162)
Hospital-level variance (SE)*						
Model 0. Patient factors	0.80 (0.10)	0.27 (0.06)	0.91 (0.12)	0.38 (0.10)	0.35 (0.09)	0.68 (0.14)
Model 1. Model 0 + hospital type (public/private)	0.65 (0.08)	0.27 (0.06)	0.71 (0.10)	0.16 (0.06)	0.35 (0.09)	0.69 (0.14)
Model 2. Model 0 + hospital remoteness	0.77 (0.09)	0.25 (0.05)	0.92 (0.12)	0.37 (0.10)	0.33 (0.08)	0.68 (0.14)
Model 3. Model 0 + hospital depth of coding	0.46 (0.06)	0.20 (0.05)	0.56 (0.08)	0.26 (0.08)	0.29 (0.08)	0.68 (0.14)
Model 4. Model 0 + hospital peer group	0.72 (0.09)	0.21 (0.05)	0.75 (0.10)	0.34 (0.09)	0.31 (0.08)	0.67 (0.14)
Intraclass correlation coefficient (ICC)**	19.5%	7.6%	21.6%	10.4%	9.6%	17.1%
Median odds ratio (MOR)**	2.34	1.64	2.48	1.80	1.76	2.19

[†] N = number of patients who self-reported condition

* Patient-level variance in a logistic regression is set at $\pi^2/3=3.29$ [31]

** ICC and MOR calculated from Model 0 [ICC = hospital-level variance divided by total variance (hospital-level + patient-level); MOR is calculated as $\exp^{0.954 \times \sqrt{\text{variance}}}$] [30]

1
2
3
4
5
6
7 **Variations in the recording of common health conditions in Australian-routine hospital**
8 **datas: study using linked survey and administrative data in New South Wales, Australia**
9
10

11
12
13 Sanja Lujic¹, Diane E Watson², Deborah A Randall¹, Judy M Simpson³, Louisa R Jorm^{1,4}
14
15

- 16
17
18 1. Centre for Health Research, University of Western Sydney, Sydney, Australia.
19 2. National Health Performance Authority, Sydney, Australia.
20 3. School of Public Health, University of Sydney, Sydney, Australia.
21 4. The Sax Institute, Sydney Australia.
22
23
24

25
26
27 Correspondence:

28 Ms Sanja Lujic
29 Centre for Health Research, University of Western Sydney, Locked Bag 1797, Penrith NSW
30 2751, Australia
31 E: s.lujic@uws.edu.au
32 T: +61 4620 3684
33 F: +61 4620 3299
34
35

36
37 Word count: 4,3024,625 words
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objectives: To investigate the nature and potential implications of underreporting of morbidity information in administrative hospital data. ~~To investigate the agreement between self-report and recording of six morbidities in administrative hospital data, quantify the between-hospital variation and identify predictors of positive agreement between the two data sources.~~

Setting and participants: Retrospective analysis of linked self-report and administrative hospital data for 32,832 participants in the large-scale cohort study (45 and Up Study), who joined the study from 2006-2009 and who were admitted to 313 hospitals in New South Wales, Australia, for at least an overnight stay, up to a year prior to study entry.

Outcome measures: Agreement between self-report and recording of six morbidities in administrative hospital data. ~~and between-hospital variation and predictors of~~ ~~and~~ positive agreement ~~among participants who self-reported any of the six morbidities~~ between the two data sources.

Results: Agreement between data sources was good for diabetes ($\kappa=0.79$), moderate for smoking ($\kappa=0.59$), fair for heart disease, stroke and hypertension ($\kappa=0.40$, $\kappa=0.30$, $\kappa=0.24$, respectively) and poor for obesity ($\kappa=0.09$), indicating that a large number of individuals with self-reported morbidities did not have a corresponding diagnosis coded in their hospital records. Significant between-hospital variation was found (ranging from 8% of unexplained variation for diabetes to 22% for heart disease), with higher agreement in public and large hospitals, and hospitals with greater depth of coding.

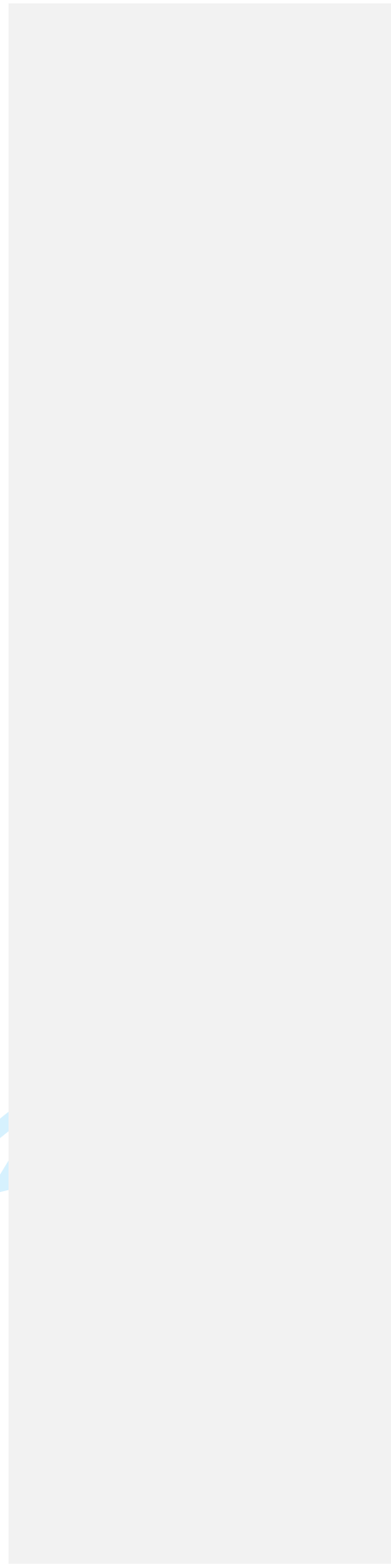
Conclusions: The recording of six common health conditions in administrative hospital data is highly variable, and for some conditions, very poor. To support more valid performance comparisons, it is important to stratify or control for factors that predict the completeness of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

recording, including hospital depth of coding and hospital type (public/private), and to increase efforts to standardize recording across hospitals. Studies using these conditions for risk adjustment should also be cautious of their use in smaller hospitals.

Keywords: hospital data, morbidity, multilevel modeling, agreement

For peer review only



Article summary

Article focus

- Explore the agreement between administrative hospital data and self-report information for four clinical conditions, as well as smoking and obesity.
- Quantify the between-hospital variation in the level of recording of these conditions.

Key messages

- Good levels of agreement found only for diabetes, with other conditions exhibiting moderate to poor agreement.
- Better recording was found in public and large hospitals, and hospitals with greater depth of coding.
- Significant between-hospital variation in the levels of agreement, amplified when smaller and private hospitals are included.

Strengths and limitations

- Use of linked data from a large-scale cohort study and advanced multilevel modeling methods to comprehensively evaluate the recording of common health conditions in hospital data, and explore between-hospital variation.
- Limitations include the absence of 'gold standard' such as medical records.

Introduction

Most nations with advanced economies publicly report on the comparative performance of hospitals with a view to accelerating and informing efforts to improve quality and allowing patients to make informed choices. Diagnoses recorded in administrative hospital data are commonly used in the construction and case-mix adjustment of hospital performance metrics, as well as for risk adjustment in epidemiological studies.

The construction of reliable health metrics relies on statistical methods that take into account the degree to which patients treated in different facilities have different morbidity and risk profiles that predispose them to requiring different interventions or to achieving different outcomes. These statistical methods, known as case-mix or risk adjustment, account for patient-related factors that are above and beyond the immediate control of health care professionals.

Thus, properly constructed performance metrics fairly reflect differences in health care experiences, patient outcomes and risks of adverse events. There has been some criticism of case-mix adjustments because they are subject to measurement error, [1] but case-mix adjustment is still considered to be less biased than unadjusted comparisons. [2]

Most methods of case-mix adjustment rely principally on demographic and diagnostic information that is captured in administrative hospital data collections. The hospital data is collected and recorded in a database for administrative purposes, with clinical coders coding diagnostic information based on the patient's medical records. [3]. This approach may be sub-optimal [4 5] because evidence from many countries suggests that administrative hospital data underreport the morbidity information needed to fully account for differences between hospitals in patient-related factors that predispose them to differences in measured outcomes. [6-13] However, the impact of this underreporting on comparative measures of hospital

1
2
3
4
5
6
7 performance depends on whether it varies systematically among hospitals, because of
8 differences in factors such as training or practice among coding staff, the comprehensiveness
9 of clinicians' notes, or "upcoding" relating to funding models or incentives. [14]
10
11

12 This issue is relatively unexplored, aside from the work by Mohammed et al. [2] which
13 reported a non-constant relationship between case-mix variables and mortality among
14 hospitals in the UK, explained by differences in clinical coding and admission practices
15 across hospitals. These variations in coding accuracy were shown to be related to geographic
16 location and bed size, with small rural facilities performing better than large urban hospitals.
17
18
19
20
21

22 [15 16] In Australia, variations in the reporting and coding of secondary diagnoses in
23 administrative hospital data have been shown to exist in public hospitals among Australian
24 states, [17] and also among hospitals within the state of New South Wales (NSW), with
25 worse-greater underreporting in private and rural hospitals. [3] However, the relative
26 contributions of patient and hospital factors to these variations have not been identified, nor
27 has this variation been formally quantified.
28
29
30
31
32
33

34 This study, using data-linkage of survey and administrative data, aimed to further investigate
35 the nature and potential implications of underreporting of morbidity information in
36 administrative hospital data by: 1) measuring the agreement between self-reported morbidity
37 information and coded diagnoses; 2) quantifying the amount of between-hospital variation in
38 this agreement; and 3) identifying patient and hospital characteristics that predict higher or
39 lower levels of agreement. We focused on clinical conditions common to case-mix and risk-
40 adjustment models – diabetes, heart disease, hypertension and stroke. We also focus on
41 smoking and obesity, due to their impact on health trajectories, rapid shifts in prevalence,
42 substantial geographic variation in rates [18] and paucity of international evidence on
43 completeness of coding.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Methods

Data sources

The 45 and Up Study

The 45 and Up Study is a large-scale cohort study involving 267,153 men and women aged 45 years and over from the general population of NSW, Australia. The study is described in detail elsewhere. [19] Briefly, participants in the 45 and Up Study were randomly sampled from the database of Australia's universal health insurance provider, Medicare Australia, which provides near complete coverage of the population. People 80+ years of age and residents of rural and remote areas were oversampled. Participants joined the Study by completing a baseline questionnaire (available at <https://www.saxinstitute.org.au/our-work/45-up-study/questionnaires/>) between January 2006 and December 2009 and giving signed consent for follow-up and linkage of their information to routine health databases.

About 18% of those invited participated and participants included about 11% of the NSW population aged 45 years and over. [19] ~~Exposure outcome relationships estimated from the 45 and Up Study data have been shown to be consistent with a large random survey of the same population. [20]~~

The NSW Admitted Patient Data Collection (APDC)

The APDC includes records of all public and private hospital admissions ending in a separation, i.e. discharge, transfer, type-change or death. ~~Each separation is referred to as an episode of care.~~ Diagnoses are coded according to the Australian modification of the International Statistical Classification of Diseases and ~~Health~~ Related Problems 10th Revision, ICD-10-AM. [20] Up to 55 diagnoses codes are recorded on the APDC, including the principal diagnosis and up to 54 additional diagnoses. Additional diagnoses are defined

1
2
3
4
5
6
7 as 'a condition or complaint either coexisting with the principal diagnosis or arising during
8 the episode of care' [in the Australian Coding Standards](#) and should be interpreted as
9 conditions that affect patient management. [21] [Assignment of diagnosis codes is done by](#)
10 [trained clinical coders, using information from the patient's medical records.](#)
11
12

13
14 The APDC from 1 July 2000 to 31 December 2010 was linked probabilistically to survey
15 information from the 45 and Up Study by the NSW Centre for Health Record Linkage
16 (www.cherel.org.au) using the 'best practice' protocol for preserving privacy. [22]
17
18
19

20 21 22 23 **Study population**

24
25 The study population comprised patients aged 45 years and above who participated in the 45
26 and Up Study and who had ~~an overnight hospitalisation~~ [hospitalization lasting at least one](#)
27 [night in the period](#) up to 365 days prior to filling out the baseline 45 and Up Study survey.
28
29 [Day stay patients were excluded from the analysis to make the study more robust and](#)
30 [generalizable beyond NSW and Australia, as there are differences in admission practices for](#)
31 [same day patients between Australia and most other comparable countries.](#)[23] NSW is home
32
33
34
35
36 to 7.4 million people or one-third of the population of Australia.
37
38
39

40 41 42 **Measuring morbidity**

43
44 We examined four health conditions (diabetes, heart disease, hypertension and stroke) and
45 two health risk factors (obesity and smoking), referred to hereafter collectively as
46 "morbidity". For each participant, these health conditions were measured using self-report
47 and administrative hospital data.
48
49

50
51 Self-reported morbidities were ascertained on the basis of responses to questions in the
52 baseline 45 and Up Study survey. Diabetes, hypertension, stroke and heart disease were
53
54
55

1
2
3
4
5
6 identified using the question “Has a doctor ever told you that you have [name of condition]?”.

7
8 Participants who did not answer the question were excluded from analyses (n=1,242).

9
10 Smoking was classified on the basis of answering “yes” to both of the questions “Have you
11 ever been a regular smoker?” and “Are you a regular smoker now?”. Participants’ responses
12 to the questions “How tall are you without shoes?” and “About how much do you weigh?”
13 were used to derive body mass index (BMI), defined as body weight divided by height
14 squared (kg/m²). The World Health Organization’s [24] classification system was used to
15 categorize individuals as obese (BMI ≥ 30kg/m²).
16
17
18
19
20
21

22 Morbidity information in administrative hospital data was ascertained using all 55 diagnosis
23 codes in the APDC records (ICD-10-AM: E10-~~E16~~-E14 for diabetes, I20-I52 for heart
24 disease, I60-I69, G45, G46 for stroke, I10-I15 and R03.0 for hypertension, F17.2 or Z72.0 for
25 smoking and E66 for obesity). The inclusion of broader ICD-10-AM codes for heart disease
26 and stroke was chosen because of the broad definition of disease type in the self-reported
27 data. Thus, heart disease codes were inclusive of coronary heart disease, pulmonary heart
28 disease, and other forms of heart diseases including heart failure and arrhythmias. Stroke
29 codes included cerebrovascular diseases without infarction among others.
30
31
32
33
34
35
36
37
38
39
40

41 **Predictors of agreement**

42
43 We explored both patient- and hospital-level factors as predictors of agreement between the
44 two data sources.
45

46
47 Patient-level factors were self-reported in the 45 and Up Study baseline survey and included
48 age, sex, education, country of birth, income and functional limitation. Functional limitation
49 was measured using the Medical Outcomes Study – Physical Functioning scale, [25] and
50
51
52
53
54
55

1
2
3
4
5
6
7 classified into 5 groups: no limitation (score of 100), minor limitation (score 95-99), mild
8 limitation (score 85-94), moderate limitation (60-84) and severe limitation (score 0-59).

9
10 Facility-level factors were type of hospital (public/private), hospital peer group (akin to
11 hospital size defined by number of case-mix weighted separations, [26] which includes
12 hospital remoteness in the classification), remoteness of hospital and depth of coding.

13
14
15
16 Remoteness of the Statistical Local Area in which the hospital was located was classified
17 according to the Accessibility/Remoteness Index of Australia (ARIA+), grouped into four
18 categories (major city, inner regional, outer regional, remote/very remote). [27] Depth of

19
20
21
22 hospital coding was the mean number of additional diagnoses coded per episode of care for
23 each hospital, calculated using all overnight hospitalizations for the full 45 and Up Study
24 cohort from 2000 to 2010, and divided into four groups at the 25th, 50th and 75th percentile.

25
26
27
28 Hospital peer groups were divided into 5 categories: principal referral ($\geq 25,000$ separations
29 per year), major (10,000-24,999 separations per year), district (2,000-9,999 separations per
30 year), community (up to 2,000 separations per year) and other (non-acute, un-peered
31 hospitals). Missing information was treated as a separate category for any variables with
32 missing data.
33
34
35
36
37
38
39
40

41 **Statistical methods**

42
43 We examined patient-level agreement between data sources for each of the six morbidities
44 individually, as well as for their 15 two-way combinations. We compared the self-reported
45 responses (yes/no) with all the diagnoses provided in the hospital records both for 'index'
46 admissions and for the 'lookback' period admissions. [28] The 'index' admission was the
47 overnight hospital stay with admission date closest to the survey completion date and no
48 longer than a year prior. Morbidity was coded as 'yes' if any of the diagnoses during that stay
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 contained a mention of that morbidity. The ‘lookback’ admissions included all overnight
8 stays in the 365-day period that preceded and included the ‘index’ admission. Morbidity was
9 coded as ‘yes’ if any of the diagnoses from any lookback admissions contained a mention of
10 that morbidity.
11
12

13
14 Agreement between the two data sources (yes/no) was measured using Cohen’s kappa
15 statistic (κ). Kappa values above 0.75 denote excellent agreement, 0.40 to 0.75 fair to good
16 agreement and below 0.45 poor agreement. [29] Agreement was computed for all 313
17 hospitals in the state, regardless of size, as well as for the 82 largest public hospitals, for
18 which performance metrics are publicly reported.
19
20

21
22 Multilevel logistic regression was used to estimate odds ratios (OR) with 95% confidence
23 intervals (CI) for patient- and hospital-level factors that predicted positive agreement between
24 the two data sources. Multilevel models were chosen because of the clustering of patients
25 within hospitals. Models were run for each of the six morbidities separately. These analyses
26 were constrained to only those participants who self-reported the morbidity of interest, and
27 the outcome was whether the index hospital record contained a mention of the morbidity or
28 not. Addition of the hospital-level characteristics was done one at a time, due to the
29 collinearity between variables. All ORs presented are adjusted for all other demographic
30 variables in the model.
31
32

33
34 Variation at the hospital level was expressed as a median odds ratio (MOR), which is the
35 median of the odds ratios of pair-wise comparisons of patients taken from randomly chosen
36 hospitals, calculated as $exp^{0.954 \times \sqrt{\text{variance}}}$; [30] and the intraclass correlation coefficient (ICC),
37 which is the percentage of the total variance attributable to the hospital level. [31] Large ICCs
38 indicate that differences among hospitals account for a considerable part of the variation in
39 the outcome, whereas a small ICC means that the hospital effect on the overall variation is
40 minimal. The relative influence of the hospital on reporting of morbidity was calculated using
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

1
2
3
4
5
6
7 a variance partitioning coefficient expressed as a percentage of the total variance using the
8 Snijders and Bosker latent variable approach. [31]
9

10 All data management was done using SAS 9.2 [32] and multilevel modeling using MLwiN
11 2.24. [33]
12

13
14 The conduct of the 45 and Up Study was approved by the University of New South Wales
15 Human Research Ethics Committee (HREC), while ethical approval for this particular study
16 was provided by the NSW Population and Health Services Research Ethics Committee and
17 the University of Western Sydney HREC.
18
19
20
21
22
23

24 25 **Results**

26 *Descriptive characteristics*

27
28 A total of 32,832 study participants were admitted to 313 hospitals up to a year prior to
29 completing the 45 and Up Study baseline survey. Just over half of the index admissions
30 (53%) were planned stays in hospital, and 57% were to a public hospital. Around one-third of
31 the index admissions occurred within the three months before study entry, and the mean
32 length of stay was 4.8 days (median = 3 days). Just under half of the sample (47%) reported
33 having hypertension, with heart disease or obesity reported by 25%, and current smoking by
34 6.1% of the sample. One-third (34%) of participants had two or more morbidities (data not
35 shown). Other characteristics of the sample at their index admission are shown in Table 1.
36
37
38
39
40
41
42
43
44

45 [Characteristics of hospitals are summarized in Table 2.](#)
46
47
48
49

50 Table 1 about here
51
52
53
54
55

[Table 2 about here](#)

Concordance between self-report and hospital records

Overall, reporting of morbidity differed between the two data sources with 23,257 (71%) participants having at least one of the six self-reported morbidities, and 11,977 (36.5%) and 14,335 (43.7%) of the sample having at least one morbidity recorded on their index or lookback hospital admissions, respectively.

Table [2-3](#) gives the summary concordance measures for each morbidity and two-way morbidity combination. For the index admission, good agreement was found for diabetes ($\kappa=0.79$), moderate agreement for smoking ($\kappa=0.59$), fair agreement for heart disease ($\kappa=0.40$), and poor agreement for stroke ($\kappa=0.30$), hypertension ($\kappa=0.24$) and obesity ($\kappa=0.09$). In two-way combinations, moderate levels of agreement were found only for diabetes combinations (with smoking, hypertension and heart disease).

Table [2-3](#) about here

Incorporating a one-year lookback period increased the numbers of participants with a morbidity recorded in a hospital record, with average relative increases in the kappa values of 20% (ranging from 2% increase for smoking, to 41% increase for obesity). Good to excellent level of agreements were still found only for diabetes ($\kappa =0.83$) and smoking ($\kappa =0.60$).

Agreement was only slightly higher among the 82 large public hospitals (see Supplementary Table 1) with relative kappa values higher by 4%, on average.

Patient- and hospital-level predictors of positive agreement

The patient factors which predicted positive agreement between the two data sources differed between morbidities (Table 34). Male sex was associated with better agreement for diabetes (OR=1.37, 95% CI 1.19 – 1.58), heart disease (OR=1.30, 95% CI 1.17 – 1.44) and hypertension (OR=1.28, 95% CI 1.18 – 1.38) (Supplementary Table 2).

Older patients were significantly less likely to have smoking (80+ years OR=0.48, 95% CI 0.31 – 0.74) and obesity (OR=0.14, 95% CI 0.08 – 0.26) recorded in their hospital records, and significantly more likely to have hypertension recorded (OR=1.32, 95% CI 1.16 – 1.49), compared to younger patients (45 – 59 years). People with higher levels of functional limitation were significantly more likely to have hypertension, diabetes and obesity recorded on their most recent hospital stay. Planned admissions to hospital had lower odds of having any of the six conditions recorded, as did medical admissions (for diabetes, smoking and obesity only). Agreement did not vary significantly for any other patient factors.

Table 34 about here

The four hospital-level covariates (hospital type, hospital peer group, hospital remoteness and depth of coding) were added to multilevel models (including a random intercept for hospital) one at a time, separately. Positive agreement between self-report and hospital records was significantly lower for hospitals with lower depth of coding across all morbidities. The odds of recording were also lower among private hospitals for all six morbidities, with this difference being statistically significant for hypertension, heart disease and stroke only. Records from smaller hospitals (district and community peer groups) were significantly less likely to agree with self-reported data on hypertension, diabetes and heart disease. Positive

1
2
3
4
5
6
7 agreement did not vary significantly with remoteness of hospital, with the exceptions of
8 diabetes (lower agreement for outer regional, remote and very remote hospitals) and smoking
9 (lower agreement for remote and very remote hospitals) (Supplementary Table 3).
10
11

12 13 14 15 *Quantifying variation between hospitals*

16
17 Before any hospital-level variables were added into the multilevel model, the intraclass
18 correlation coefficient indicated that between 8% (diabetes) and 22% (heart disease) of the
19 residual (unexplained) variation in agreement was attributable to the hospital, after
20 adjustment for the patient-level factors (Table 45). This equated to median odds ratios
21 (MORs) of 1.64 and 2.48, respectively, indicating that a patient in one hospital had an
22 average of between 64% and 148% higher odds of having a particular morbidity recorded
23 than a patient in a hospital with lower levels of recording. Less variation at the hospital level
24 was found for the recording of diabetes, smoking and stroke, while more variation at the
25 hospital level was found for the recording of hypertension, heart disease and obesity. When
26 the analyses were restricted to 82 large public hospitals only, the between-hospital variation
27 decreased to between 2% (stroke) and 13% (hypertension), or MOR of 1.24 and 1.94 (Figure
28 1). This between-hospital variation was still significant for all morbidities except for stroke.
29 Between-hospital variation was further reduced once lookback admissions were used to
30 identify morbidities.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 Figure 1 about here

50
51
52 The addition of hospital-level variables to multilevel models, one at the time, separately,
53 helped ascertain which factors explained the variation between hospitals (Table 45). The
54
55

1
2
3
4
5
6
7 addition of ~~at least one of the four~~ hospital-level factors contributed to explaining (*i.e.*
8 ~~decreasing~~) the residual variation ~~between hospitals~~ for all conditions, except obesity. For the
9 other morbidities, differences in the depth of coding explained from 16% (smoking) to 42%
10 (hypertension) of residual variation between hospitals, while hospital type (public/private)
11 explained from 0% (smoking) to 59% (stroke), and hospital peer group explained from 10%
12 (hypertension) to 27% (diabetes) residual variation between hospitals.
13
14
15
16
17
18
19
20

21 Table 4-5 about here

22 Discussion

23
24
25
26
27
28 Our study found that the concordance of administrative hospital and self-reported data varied
29 between the six morbidities examined, with agreement ranging from good for diabetes,
30 moderate for smoking, through to fair for heart disease, and poor for hypertension, stroke and
31 obesity. We demonstrated considerable between-hospital variation in the recording of these
32 common health conditions. Smaller, but still significant, between-hospital variation was
33 found when restricting the analyses to the largest public hospitals in the state.
34
35
36
37
38
39

40 Previous studies have validated information recorded in NSW administrative hospital data for
41 demographic factors, [34 35] and recording of perinatal conditions, [36-39] but there have
42 been limited studies of the accuracy of the recording of health conditions commonly used for
43 case-mix or risk-adjustment. Our findings regarding agreement for the recording of diabetes
44 ($\kappa=0.83$) were similar to previous Australian studies [3 10], while agreement for hypertension
45 ($\kappa=0.30$) and heart disease ($\kappa=0.47$) was considerably lower in our study. These differences
46 may be due to the fact that both previous studies used medical records as a 'gold standard',
47 while we used self-report. Lower agreement rates for heart disease could be due to the
48
49
50
51
52
53
54
55

1
2
3
4
5
6
7 broader range of heart disease types included in our study, with known lower levels of
8 agreement for heart failure compared to myocardial infarction.[9],[40] ~~[3]~~-Higher
9
10 sensitivities reported in a study from the state of Victoria [10] could also be attributable to the
11 differences in public hospital funding models between the two states. Specifically, Victoria
12 has used activity- based funding since 1993, while this method of funding was introduced in
13 NSW and other Australian states only subsequent to our study period. [41] Introduction of
14 activity-based funding has been shown to increase recording of additional diagnoses and
15 procedures in Europe. [42]

16
17
18
19
20
21
22 Some of the apparent discrepancies in the levels of coding between conditions can be
23 attributed to the coding rules that govern whether or not a diagnosis is recorded in
24 administrative hospital data. Additional diagnoses, recorded on administrative hospital data,
25
26 are coded only if they affect the patient's treatments received, investigations required and/or
27 resources used during the hospital stay. Thus, diagnoses that relate to an earlier episode, and
28 which have no bearing on the current hospital stay, are ~~excluded~~not coded for that particular
29 stay. Therefore, it is not surprising that (managed) hypertension, in particular, might not be
30 recorded in hospital data relating to, for example, elective surgery. On the other hand, we
31 found that diabetes is well recorded, suggesting that it is considered to affect patient
32 management in most hospital stays, and possibly reflecting the impact of changes to the
33 Australian Coding Standards for diabetes such that between 2008 and 2010 diabetes with
34 complications could be coded even where there was no established cause and effect
35 relationship between diabetes and the complication. -[43] It is for these reasons that
36 researchers using administrative datasets are encouraged to incorporate information from
37 previous hospitalizations, to increase the likelihood of capturing morbidity, as demonstrated
38 in this as well as other Australian studies.[44]

1
2
3
4
5
6
7 As well as looking at single morbidities, ours is the first study to our knowledge to explore
8 the variations of recording of multiple conditions in hospital data. Concordance of two-way
9 condition combinations was very low, with best results found for combinations of diseases
10 involving diabetes, which had the highest single-condition level of agreement with self-
11 reported data ($\kappa = 0.83$). Agreement measures for two-way combinations were found to be fair
12 to good at best, with agreement on three-way condition combinations (not investigated here)
13 expected to be even lower. These findings have implications for research into multimorbidity
14 (the co-occurrence of multiple chronic or acute diseases and medical conditions within one
15 person [45]). We suggest that researchers who use administrative data for research into
16 multimorbidity should use linked data to increase ascertainment, and, if possible, supplement
17 this information from other data sources, such as physician claims data or self-reported data.

18
19
20
21
22
23
24
25
26
27
28 We identified considerable between-hospital variability in the levels of recording of common
29 health conditions, with between 8% and 22% of the variation attributable to hospital-level
30 factors, after adjustment for patient factors. This was similar in magnitude to the variability
31 previously reported for performance measures (varying from patient satisfaction, mortality,
32 length of stay to quality of care) clustered at the facility level (0-51%) [46] and hospital-level
33 variations in the use of services. [47-49] Significant between-hospital variation was still
34 present after constraining the analyses to the 82 largest public hospitals in the state.

35
36
37
38
39
40
41
42 The recording of hypertension and heart disease was particularly variable between hospitals,
43 those with better reporting having on average 2.3 and 2.5 times, respectively, the odds of
44 recording these conditions than those with lower levels of reporting. The corresponding
45 figures were 1.9 and 1.6 times for the 82 largest hospitals in the state. These findings indicate
46 the potential for reporting bias to influence comparisons of health performance indicators
47 between hospitals, especially for indicators that use conditions such as heart disease or
48 hypertension for case-mix adjustment. To our knowledge, no previous studies have provided
49
50
51
52
53
54

1
2
3
4
5
6
7 detailed information about how the validity of morbidity reporting varies among hospitals
8 after accounting for patient factors.
9

10 Further, we have shown that variations in the accuracy of morbidity reporting between
11 hospitals are predominantly driven by the hospital's depth of coding – concordance between
12 self-reported and hospital data is lower in hospitals with a lower average number of
13 additional diagnoses recorded. Up to 42% of the variation in recording at the hospital level
14 could be attributed to differences in hospital depth of coding. Even though the measure of
15 depth of coding we used was crude, and related to hospital size, it still helps in highlighting
16 the impact of coding practices on variations among hospitals. Other research using the same
17 depth of coding measure has shown that the lower depth of coding can disproportionately
18 disadvantage hospitals' standardised mortality ratios, one of the commonly reported measures
19 of hospital performance. [2] It will be important to track changes in the levels of the depth of
20 coding across Australian states, and to consider the implications of these for state-based
21 performance comparisons, following the national rollout of activity-based funding and
22 comparative performance reporting.
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Several factors might explain variation in depth of coding between hospitals. Clinical coders
37 can code only information that has been recorded in the patient's medical record, so varying
38 level of details recorded by clinicians will influence what gets coded. The training and
39 professional development opportunities for coding staff might also influence the depth of
40 coding. Also, casemix funding systems, such as the Diagnosis Related Group (DRG)
41 classification, are prone to 'upcoding' in order for services to receive higher reimbursement
42 costs. [14]
43
44
45
46
47
48
49

50 We found that the reporting of conditions varied with hospital size, larger metropolitan
51 hospitals having higher concordance, with kappa values higher by 7% on average when
52 comparing large tertiary with smaller urban hospitals. This finding echoes those of Powell et
53
54
55

1
2
3
4
5
6
7 al. [3] in NSW, Australia during 1996 – 1998 and Rangachari et al. [16] in the US, during
8
9 2000 – 2004. Our study showed that large tertiary hospitals had better concordance for the
10
11 recording of hypertension and heart disease than smaller urban hospitals, but the reverse was
12
13 true for stroke and smoking. Our finding that between-hospital variation in the recording of
14
15 morbidities was up to two times higher when all hospitals, rather than just the largest ones,
16
17 were included has implications for further research using data from smaller hospitals. This
18
19 high variability in concordance among smaller hospitals may mean that morbidity-adjusted
20
21 comparisons are not as valid as for larger hospitals. Researchers using information from these
22
23 hospitals are encouraged to supplement their data with either self-report information and/or
24
25 data linkage. The value-add of incorporating previous hospitalizations was also highlighted in
26
27 our results for stroke and obesity, with 43% – 47% more patients identified using lookback
28
29 admissions than from a single admission only.

30
31 A particular strength of our study lies in the use of linked data from a large-scale cohort study
32
33 to comprehensively evaluate the recording of common conditions in hospital data, and
34
35 explore the variation in recording among hospitals. The 45 and Up Study contains records for
36
37 one in every 10 persons aged 45 and over in NSW, so it provides a rich resource to answer
38
39 research questions. Additionally, we used advanced multilevel modeling methods to quantify
40
41 the amount of between-hospital variation in the level of recording of common health
42
43 conditions, a finding which is of importance for both research and policy paradigms due to its
44
45 impact on adjusted comparisons among hospitals and the highlighted need to improve
46
47 consistency of recording in hospitals across the State. To date, hospital-level variation has
48
49 only been explored with a set outcome (e.g. mortality, readmission) in mind.

50
51 A potential limitation of our study was its use of self-reported information to explore
52
53 concordance, in the absence of another ‘gold standard’, such as medical records. Access to
54
55 medical records was not possible given the de-identified nature of our data, and the large

1
2
3
4
5
6
7 number of records in the dataset. Moreover, studies that have examined accuracy of self-
8 reported conditions against medical records have found high levels of agreement, ranging
9 from 81% [50] to 87% [51] for hypertension, 66% [40] to 96% [50 51] for diabetes and 60%
10 [50] to 98% [52] for acute myocardial infarction. Validation studies in the 45 and Up Study
11 cohort have reported strong correlations and excellent levels of agreement between self-
12 reported and measured height and weight, and derived BMI [53] as well as self-reported
13 diabetes. [54] Although the 45 and Up Study had a response rate of 18%, the study sample is
14 very large and has excellent heterogeneity. Furthermore, exposure-outcome relationships
15 estimated from the 45 and Up Study data have been shown to be consistent with a large
16 'representative' population survey of the same population.[55]
17
18
19
20
21
22
23
24
25
26
27
28

29 **Conclusion**

30 The recording of common comorbid conditions in routine hospital data is highly variable,
31 and, for some conditions, very poor. Recording varies considerably among hospitals,
32 presenting the potential to introduce bias into risk-adjusted comparisons of hospital
33 performance, especially for indicators that use heart disease or hypertension for risk
34 adjustment. Furthermore, between-hospital variation is amplified when smaller and private
35 hospitals are included in the analyses. Stratification of analyses according to factors that
36 predict the completeness of recording, including hospital depth of coding and hospital type
37 and size, supplementing morbidity information with linked data from previous
38 hospitalizations and increases in efforts to standardize recording across hospitals, all offer
39 potential for increasing the validity of risk-adjusted comparisons.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgement

~~This work was supported by the National Health and Medical Research Council [grant number 1036858].~~

~~We thank all of the men and women who participated in the 45 and Up Study. This research was completed using data collected through the 45 and Up Study (www.saxinstitute.org.au).~~

The 45 and Up Study is managed by the Sax Institute (www.saxinstitute.org.au) in collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; *beyondblue*; Ageing, Disability and Home Care, Department of Family and Community Services; the Australian Red Cross Blood Service; and UnitingCare Ageing. ~~We thank the many thousands of people participating in the 45 and Up Study.~~

We would like to acknowledge the [Sax Institute and the](#) NSW Ministry of Health for allowing access to the data, and the Centre for Health Record Linkage for conducting the probabilistic linkage of records.

We are grateful to Dr Fiona Blyth and Dr Kris Rogers on their advice at the early stages of the project.

Funding statement

The Assessing Preventable Hospitalisation InDicators (APHID) study is funded by a National Health and Medical Research Council Partnership Project Grant (#1036858) and by partner agencies the Australian Commission on Safety and Quality in Health Care, the Agency for Clinical Innovation and the NSW Bureau of Health Information.

Contributorship

SL had overall responsibility for the design of this study, data management, statistical analysis and drafting this paper. DW and LJ contributed to the conception and design of the study. LJ helped with data acquisition, and provided oversight for all analyses. DR and JS provided oversight and advice for the design and interpretation of the statistical analyses. All authors contributed to the interpretation of the findings, the writing of the paper and approved the final draft.

Competing interests: None

Data sharing statement

This study used the data from the Assessing Preventable Hospitalisation InDicators (APHID) project. The data has been constructed with the permission of each of the custodians of the respective source datasets and with specific ethical approval. The data are available to researchers on request and subject to approval from the relevant data custodians and ethics committees. More information about these approvals is available from the authors on request.

No additional data are available.

References

1. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004;**363**(9415):1147-54
2. Mohammed MA, Deeks JJ, Girling A, et al. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* 2009;**338**
3. Powell H, Lim LL, Heller RF. Accuracy of administrative data to assess comorbidity in patients with heart disease: an Australian perspective. *J Clin Epidemiol* 2001;**54**(7):687-93
4. Pine M, Jordan HS, Elixhauser A, et al. Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA* 2007;**297**(1):71-76
5. Pine M, Norusis M, Jones B, et al. Predictions of Hospital Mortality Rates: A Comparison of Data Sources. *Ann Intern Med* 1997;**126**(5):347-54
6. Hawker GA, Coyte PC, Wright JG, et al. Accuracy of administrative data for assessing outcomes after knee replacement surgery. *J Clin Epidemiol* 1997;**50**(3):265-73
7. Kieszak SM, Flanders WD, Kosinski AS, et al. A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *J Clin Epidemiol* 1999;**52**(2):137-42
8. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Med Care* 2002;**40**(8):675-85
9. Preen DB, Holman CAJ, Lawrence DM, et al. Hospital chart review provided more accurate comorbidity information than data from a general practitioner survey or an administrative database. *J Clin Epidemiol* 2004;**57**(12):1295-304
10. Henderson T, Shephard J, Sundararajan V. Quality of diagnosis and procedure coding in ICD-10 administrative data. *Med Care* 2006;**44**(11):1011-9
11. Quach S, Blais C, Quan H. Administrative data have high variation in validity for recording heart failure. *Can J Cardiol* 2010;**26**(8):e306-e12
12. Leal J, Laupland K. Validity of ascertainment of co-morbid illness using administrative databases: a systematic review. *Clin Microbiol Infect* 2010;**16**(6):715-21
13. Chong WF, Ding YY, Heng BH. A comparison of comorbidities obtained from hospital administrative data and medical charts in older patients with pneumonia. *BMC Health Serv Res* 2011;**11**(1):105
14. Steinbusch PJ, Oostenbrink JB, Zuurbier JJ, et al. The risk of upcoding in casemix systems: A comparative study. *Health Policy* 2007;**81**(2):289-99
15. Lorence DP, Ibrahim IA. Benchmarking Variation in Coding Accuracy Across the United States. *J Health Care Finance* 2003;**29**(4):29
16. Rangachari P. Coding for quality measurement: the relationship between hospital structural characteristics and coding accuracy from the perspective of quality measurement. *Perspectives in health information management/AHIMA, American Health Information Management Association* 2007;**4**
17. Coory M, Cornes S. Interstate comparisons of public hospital outputs using DRGs: Are they fair? *Aust N Z J Public Health* 2005;**29**(2):143-48
18. National Health Performance Authority. Healthy Communities: Overweight and obesity rates across Australia, 2011–12 (In Focus). 2013.

19. Banks E, Redman S, Jorm L, et al. Cohort profile: the 45 and up study. *Int J Epidemiol* 2008;**37**(5):941
20. National Centre for Classification in Health. International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification (ICD-10-AM), Australian Classification of Health Interventions (ACHI). Sydney: National Centre for Classification in Health, 2006.
21. Australian Institute of Health and Welfare. National Health Data Dictionary. Version 16. Canberra: AIHW, 2012.
22. Kelman CW, Bass AJ, Holman CDJ. Research use of linked health data — a best practice protocol. *Aust N Z J Public Health* 2002;**26**(3):251-55
23. Eagar K. Counting acute inpatient care. ABF Information Series No 5 2010. <https://ahsri.uow.edu.au/content/groups/public/@web/@chsd/documents/doc/uow082637.pdf> (accessed July 2014).
24. World Health Organization. *Obesity: preventing and managing the global epidemic*: World Health Organization, 2000.
25. Stewart AL, Ware JE. *Measuring functioning and well-being: The medical outcomes study approach*: Duke University Press, 1992.
26. NSW Health. NSW Health Services Comparison Data Book 2008/09. 2010. http://www0.health.nsw.gov.au/pubs/2010/pdf/yellowbook_09.pdf (accessed April 2014).
27. Australian Institute of Health and Welfare. Regional and remote health: A guide to remoteness classifications Canberra: AIHW, 2004.
28. Zhang JX, Iwashyna TJ, Christakis NA. The performance of different lookback periods and sources of information for Charlson comorbidity adjustment in Medicare claims. *Med Care* 1999;**37**(11):1128-39
29. Fleiss JL, Levin B, Paik MC. The measurement of interrater agreement. *Statistical methods for rates and proportions* 1981;**2**:212-36
30. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health* 2006;**60**(4):290-97
31. Snijders T, Bosker R. *Multilevel analysis*. 2nd ed. London: Sage, 2012.
32. SAS Version 9.3 [software] [program]. Cary, North Carolina, 2010.
33. MLwiN Version 2.25 [software] [program]. Centre for Multilevel Modelling, University of Bristol, Bristol., 2012.
34. Bentley JP, Taylor LK, Brandt PG. Reporting of Aboriginal and Torres Strait Islander peoples on the NSW Admitted Patient Data Collection: the 2010 data quality survey. *N S W Public Health Bull* 2012;**23**(1-2):17-20
35. Tran DT, Jorm L, Lujic S, et al. Country of birth recording in Australian hospital morbidity data: accuracy and predictors. *Aust N Z J Public Health* 2012;**36**(4):310-16
36. Taylor LK, Travis S, Pym M, et al. How useful are hospital morbidity data for monitoring conditions occurring in the perinatal period? *Aust N Z J Obstet Gynaecol* 2005;**45**(1):36-41
37. Roberts CL, Bell JC, Ford JB, et al. The accuracy of reporting of the hypertensive disorders of pregnancy in population health data. *Hypertens Pregnancy* 2008;**27**(3):285-97
38. Hadfield RM, Lain SJ, Cameron CA, et al. The prevalence of maternal medical conditions during pregnancy and a validation of their reporting in hospital discharge data. *Aust N Z J Obstet Gynaecol* 2008;**48**(1):78-82

- 1
2
3
4
5
6
7 39. Lain SJ, Hadfield RM, Raynes-Greenow CH, et al. Quality of data in perinatal population health
8 databases: a systematic review. *Med Care* 2012;**50**(4):e7-20
- 9 40. Okura Y, Urban LH, Mahoney DW, et al. Agreement between self-report questionnaires and
10 medical record data was substantial for diabetes, hypertension, myocardial infarction and
11 stroke but not for heart failure. *J Clin Epidemiol* 2004;**57**(10):1096
- 12 41. Eagar KM. What is activity-based funding? ABF Information Series No. 1. 2011.
13 <http://ro.uow.edu.au/gsbpapers/47/> (accessed April 2014).
- 14 42. O'Reilly J, Busse R, Häkkinen U, et al. Paying for hospital care: the experience with
15 implementing activity-based funding in five European countries. *Health Economics Policy*
16 *and Law* 2012;**7**(1):73
- 17 43. Knight L, Halech R, Martin C, et al. Impact of changes in diabetes coding on Queensland hospital
18 principal diagnosis morbidity data. 2011.
19 http://www.health.qld.gov.au/hsu/tech_report/techreport_9.pdf (accessed July 2014).
- 20 44. Preen DB, Holman CDAJ, Spilsbury K, et al. Length of comorbidity lookback period affected
21 regression model performance of administrative health data. *J Clin Epidemiol*
22 2006;**59**(9):940-46
- 23 45. van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity: what's in a name? A
24 review of literature. *Eur J Gen Pract* 1996;**2**(2):65-70
- 25 46. Fung V, Schmittiel JA, Fireman B, et al. Meaningful variation in performance: a systematic
26 literature review. *Med Care* 2010;**48**(2):140-48
- 27 47. Bennett-Guerrero E, Zhao Y, O'Brien SM, et al. Variation in use of blood transfusion in coronary
28 artery bypass graft surgery. *JAMA* 2010;**304**(14):1568-75
- 29 48. Seymour CW, Iwashyna TJ, Ehlenbach WJ, et al. Hospital-Level Variation in the Use of Intensive
30 Care. *Health Serv Res* 2012;**47**(5):2060-80
- 31 49. Randall DA, Jorm LR, Lujic S, et al. Disparities in Revascularization Rates After Acute
32 Myocardial Infarction Between Aboriginal and Non-Aboriginal People in Australia Clinical
33 Perspective. *Circulation* 2013;**127**(7):811-19
- 34 50. Merkin S, Cavanaugh K, Longenecker JC, et al. Agreement of self-reported comorbid conditions
35 with medical and physician reports varied by disease among end-stage renal disease patients.
36 *J Clin Epidemiol* 2007;**60**(6):634
- 37 51. Barber J, Muller S, Whitehurst T, et al. Measuring morbidity: self-report or health care records?
38 *Fam Pract* 2010;**27**(1):25-30
- 39 52. Machón M, Arriola L, Larrañaga N, et al. Validity of self-reported prevalent cases of stroke and
40 acute myocardial infarction in the Spanish cohort of the EPIC study. *J Epidemiol Community*
41 *Health* 2013;**67**(1):71-75
- 42 53. Ng SP, Korda R, Clements M, et al. Validity of self-reported height and weight and derived body
43 mass index in middle-aged and elderly individuals in Australia. *Aust N Z J Public Health*
44 2011;**35**(6):557-63
- 45 54. Comino EJ, Tran DT, Haas M, et al. Validating self-report of diabetes use by participants in the 45
46 and up study: a record linkage study. *BMC Health Serv Res* 2013;**13**(1):481
- 47 55. Mealing NM, Banks E, Jorm LR, et al. Investigation of relative risk estimates from studies of the
48 same population with contrasting response rates and designs. *BMC Med Res Methodol*
49 2010;**10**:26
- 50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 **List of tables and figures**

8 **Table 1.** Characteristics of the study sample at their index admission

9
10 **Table 2.** Characteristics of the hospital of admission

11
12 **Table 23.** Agreement measures between self-report and hospital data, index and lookback admissions, all public and private hospitals in New South Wales, Australia (n=313)

13
14
15 **Table 34.** Factors that predict positive agreement between self-report and hospital data, using multilevel modeling, all public and private hospitals in New South Wales, Australia (n=313)

16
17
18 **Table 45.** Variance and intraclass correlation coefficient for hospital-level random effects from multilevel logistic regression, all public and private hospitals in New South Wales, Australia (n=313)

19
20
21
22 **Figure 1 Title.** Variance for hospital-level random effects from multilevel logistic regression, for index and lookback admissions, by hospital size

23
24
25 **Figure 1 Legend:**

26 * Significantly different from 0 at 5% level

27 ** Significantly different from 0 at 1% level

Table 1. Characteristics of the study sample at their index admission

Characteristics	All participants	
	(N = 32,832)	
	N	%
Demographic characteristics		
Sex		
Male	16,812	51.2
Female	16,020	48.8
Age		
45-59	9,666	29.4
60-79	16,624	50.6
80+	6,540	19.9
Country of birth		
Australia	25,001	76.2
Other	7,448	22.7
Unknown	383	1.2
Highest education level		
No school	5,196	15.8
Year 10 or equivalent	7,894	24.0
Year 12 or equivalent	2,975	9.1
Trade	4,270	13.0
Certificate	6,109	18.6
University degree	5,662	17.3
Unknown	726	2.2
Household income (\$, per annum)		
<20,000	9,077	27.7
20,000 - <50,000	8,223	25.1
50,000 - <70,000	2,560	7.8
70,000+	5,042	15.4
Not disclosed	6,003	18.3
Missing	1,927	5.9
Functional status		
No limitation	4,915	15.0
Mild limitation	6,011	18.3
Moderate limitation	8,701	26.5
Severe limitation	10,121	30.8
Missing	3,084	9.4
Admission characteristics		
Admission type		
Surgical	15,464	47.1
Other	1,439	4.4
Medical	15,929	48.5
Emergency status		
Emergency	13,484	41.1

Formatted Table

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Planned	17,544	53.4
Other	1,803	5.5
<i>Hospital characteristics of admission</i>		
Hospital type		
Public	18,734	57.1
Private	14,096	42.9
Hospital remoteness		
Major city	19,754	60.2
Inner regional	8,424	25.7
Outer regional	4,137	12.6
Remote/very remote	363	1.1
Hospital depth of coding		
1 - least comprehensive	1,629	5.0
2	8,803	26.8
3	11,543	35.2
4 - most comprehensive	10,857	33.1
Hospital peer group		
Principal referral	6,329	19.3
Major	11,052	33.7
District	6,862	20.8
Community	7,018	21.4
Other	1,571	4.8

*for comparisons of hospital characteristics

Table 2. Characteristics of the hospital of admission

	All hospitals^d	
	(N = 313)	
	N	%
Hospital type		
Public	224	71.6
Private	88	28.1
Hospital remoteness		
Major city	124	39.6
Inner regional	72	23.0
Outer regional	94	30.0
Remote/very remote	20	6.4
Hospital depth of coding		
1 - least comprehensive	48	15.3
2	91	29.1
3	89	28.4
4 - most comprehensive	85	27.2
Hospital peer group		
Principal referral	14	4.5
Major	33	10.5
District	51	16.3
Community	121	38.7
Other	94	30.0

Formatted Table

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Table 23. Agreement measures between self-report and hospital data, index and lookback admissions, all public and private hospitals in New South Wales, Australia (n=313)

Morbidity ^a	Index admission					Lookback admissions						
	45 and Up Yes:		45 and Up No:		Kappa %	95% CI	45 and Up Yes:		45 and Up No:		Kappa %	95% CI
	APDC yes	APDC no	APDC yes	APDC no			APDC yes	APDC no	APDC yes	APDC no		
Hypertension	4,767	10,512	1,434	16,119	24.0	(22.9-25.0)	6,260	9,019	2,051	15,502	30.2	(29.1-31.2)
Heart disease	3,639	4,668	1,942	22,583	40.3	(39.0-41.5)	4,673	3,634	2,697	21,828	47.0	(45.8-48.2)
Diabetes	3,560	1,234	347	27,691	79.1	(78.1-80.1)	3,928	866	479	27,559	83.0	(82.1-83.9)
Stroke	541	1,939	306	30,046	29.8	(27.0-32.6)	776	1,704	488	29,864	38.3	(35.8-40.8)
Smoking	1,205	804	727	30,096	58.7	(56.7-60.7)	1,411	598	1,076	29,747	60.1	(58.2-61.9)
Obesity	551	7,611	114	24,556	9.1	(7.3-10.9)	810	7,352	209	24,461	12.8	(11.1-14.6)
Hypertension + heart disease	1,172	3,481	1,270	26,909	25.8	(23.8-27.7)	1,807	2,846	2,008	26,171	34.3	(32.6-36.0)
Hypertension + diabetes	1,819	1,238	759	29,016	61.3	(59.6-62.9)	2,186	871	1,021	28,754	66.6	(65.2-68.1)
Hypertension + stroke	203	1,317	189	31,123	19.7	(15.7-23.7)	329	1,191	340	30,972	28.0	(24.5-31.5)
Hypertension + smoking	133	598	180	31,921	24.5	(19.2-29.7)	199	532	319	31,782	30.6	(26.0-35.2)
Hypertension + obesity	234	4,574	93	27,931	7.4	(4.9-9.8)	383	4,425	183	27,841	11.5	(9.2-13.9)
Heart disease + diabetes	646	1,154	404	30,628	43.0	(40.3-45.8)	904	896	661	30,371	51.2	(48.9-53.6)
Heart disease + stroke	76	973	126	31,657	11.2	(6.1-16.4)	149	900	261	31,522	19.0	(14.4-23.5)
Heart disease + smoking	76	294	222	32,240	22.0	(15.3-28.6)	118	252	373	32,089	26.5	(20.8-32.2)
Heart disease + obesity	79	1,938	79	30,736	6.4	(2.5-10.4)	151	1,866	169	30,646	11.4	(7.7-15.2)
Diabetes + stroke	85	555	58	32,134	21.1	(15.0-27.3)	140	500	119	32,073	30.4	(24.9-35.8)
Diabetes + smoking	143	161	108	32,420	51.1	(45.3-56.9)	171	133	176	32,352	52.1	(46.7-57.4)
Diabetes + obesity	232	1,701	65	30,834	19.5	(15.9-23.2)	351	1,582	120	30,779	27.5	(24.2-30.9)
Stroke + smoking	13	142	28	32,649	13.1	(0.1-26.1)	23	132	57	32,620	19.3	(7.8-30.8)
Stroke + obesity	6	558	9	32,259	2.0	(0.0-10.0)	13	551	21	32,247	4.2	(0.0-11.9)
Smoking + obesity	27	447	29	32,329	9.9	(1.9-17.9)	38	436	47	32,311	13.2	(5.5-20.9)

^a ICD-10-AM codes: hypertension (I10-I15, R03.0), heart disease (I20-I25, I26-I28, I30-I52, I20-I52), diabetes (E10-E14), stroke (I60-I69, G45, G46), smoking (F17.2, Z72.0), obesity (E66)

Table 34. Factors that predict positive agreement between self-report and hospital data, using multilevel modelling, all public and private hospitals in New South Wales, Australia (n=313)

	Hypertension (N = 15,279)	Diabetes (N = 4,794)	Heart disease (N = 8,307)	Stroke (N = 2,480)	Smoking (N = 2,099)	Obesity (N = 8,162)
Person-level variables						
Sex ¹	**	**	**			
Age ¹	**				**	**
Education ¹			*	**	**	
Country of birth ¹						
Functional limitation ¹	**	**				**
Income ¹						
Admission type ²	**	**	**	**	**	**
Emergency status ²	**	**	**	**		**
Hospital-level variables						
Hospital type (public/private) ³	**		**	**		
Hospital remoteness ³					*	
Hospital depth of coding ³	**	**	**	**	**	**
Hospital peer group ³	**	**	**		**	

* Significant at 5% level
 ** Significant at 1% level

1 – Model 0: adjusted for demographic factors + random intercept for hospital
 2 – Model 0 + admission type + emergency status
 3 – Model 0 + hospital-level variables (entered one at a time)

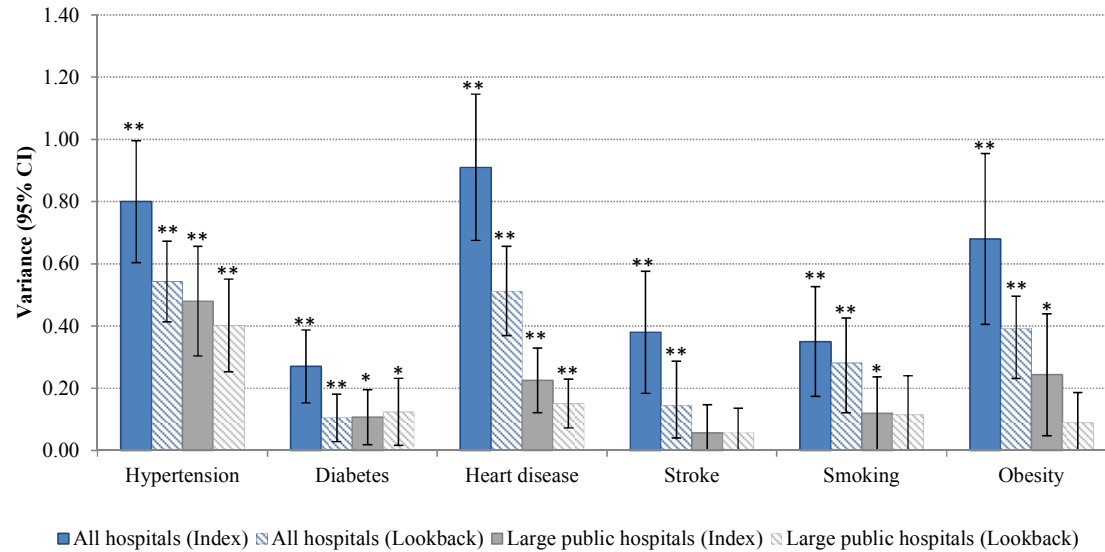
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Table 45. Variance and intraclass correlation coefficient for hospital-level random effects from multilevel logistic regression, all public and private hospitals in New South Wales, Australia (n=313)

		Hypertension (N [†] = 15,279)	Diabetes (N [†] = 4,794)	Heart disease (N [†] = 8,307)	Stroke (N [†] = 2,480)	Smoking (N [†] = 2,099)	Obesity (N [†] = 8,162)
Hospital-level variance (SE)*							
Model 0.	Patient factors	0.80 (0.10)	0.27 (0.06)	0.91 (0.12)	0.38 (0.10)	0.35 (0.09)	0.68 (0.14)
Model 1.	Model 0 + hospital type (public/private)	0.65 (0.08)	0.27 (0.06)	0.71 (0.10)	0.16 (0.06)	0.35 (0.09)	0.69 (0.14)
Model 2.	Model 0 + hospital remoteness	0.77 (0.09)	0.25 (0.05)	0.92 (0.12)	0.37 (0.10)	0.33 (0.08)	0.68 (0.14)
Model 3.	Model 0 + hospital depth of coding	0.46 (0.06)	0.20 (0.05)	0.56 (0.08)	0.26 (0.08)	0.29 (0.08)	0.68 (0.14)
Model 4.	Model 0 + hospital peer group	0.72 (0.09)	0.21 (0.05)	0.75 (0.10)	0.34 (0.09)	0.31 (0.08)	0.67 (0.14)
Intraclass correlation coefficient (ICC)**		19.5%	7.6%	21.6%	10.4%	9.6%	17.1%
Median odds ratio (MOR)**		2.34	1.64	2.48	1.80	1.76	2.19

[†] N = number of patients who self-reported condition
 * Patient-level variance in a logistic regression is set at $\pi^2/3=3.29$ [31]
 ** ICC and MOR calculated from Model 0 [ICC = hospital-level variance divided by total variance (hospital-level + patient-level); MOR is calculated as $exp^{0.954 \times \sqrt{variance}}$] [30]

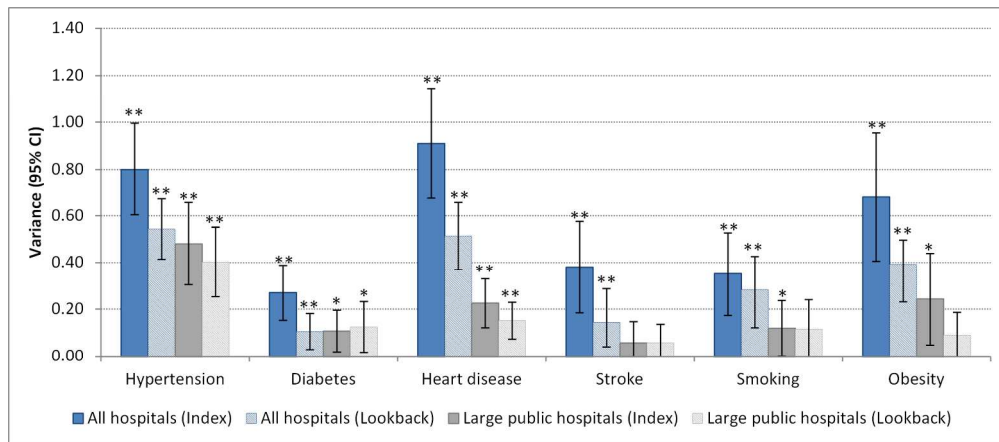
Figure 1. Variance for hospital-level random effects from multilevel logistic regression, for index and lookback admissions, by hospital size



* Significantly different from 0 at 5% level
 ** Significantly different from 0 at 1% level

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

For peer review only



* Significantly different from 0 at 5% level
 ** Significantly different from 0 at 1% level

193x84mm (300 x 300 DPI)

er review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Supplementary Table 1.** Agreement measures between self-report and hospital data, index and lookback admissions, large public
4 hospitals in New South Wales, Australia (n=82)
5
6

7 **Supplementary Table 2.** Adjusted ORs for patient-level variables from the multilevel logistic regression with random intercept for
8 hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)
9

10 **Supplementary Table 3.** Adjusted ORs for hospital-level variables from the multilevel logistic regression with random intercept for
11 hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Supplementary Table 1. Agreement measures between self-report and hospital data, index and lookback admissions, large public hospitals in New South Wales, Australia (n=82)

Morbidities ^a	<u>Index admission</u>						<u>Lookback admissions</u>					
	45 and Up Yes:		45 and Up No:		Kappa %	95% CI	45 and Up Yes:		45 and Up No:		Kappa %	95% CI
	APDC yes	APDC no	APDC yes	APDC no			APDC yes	APDC no	APDC yes	APDC no		
Hypertension	3,061	4,803	983	7,634	28.1	(26.6-29.6)	3,829	4,035	1,339	7,278	33.7	(32.2-35.1)
Heart disease	2,306	2,455	1,309	10,411	40.1	(38.5-41.8)	2,910	1,851	1,710	10,010	46.9	(45.4-48.5)
Diabetes	2,168	661	214	13,438	80.1	(78.8-81.4)	2,355	474	289	13,363	83.3	(82.1-84.4)
Stroke	414	1,210	213	14,644	33.1	(29.8-36.4)	563	1,061	311	14,546	41	(38.0-44.0)
Smoking	820	507	468	14,686	59.5	(57.0-62.0)	948	379	692	14,462	60.4	(58.1-62.7)
Obesity	265	3,857	61	12,298	8.6	(6.1-11.1)	414	3,708	114	12,245	12.9	(10.4-15.3)
Hypertension + heart disease	799	1,878	893	12,911	27.4	(25.0-29.9)	1,159	1,518	1,327	12,477	34.7	(32.5-36.9)
Hypertension + diabetes	1,129	670	518	14,164	61.5	(59.4-63.6)	1,317	482	662	14,020	65.8	(63.9-67.7)
Hypertension + stroke	160	825	145	15,351	22.6	(17.9-27.3)	238	747	237	15,259	29.9	(25.6-34.1)
Hypertension + smoking	106	399	135	15,841	27	(20.9-33.1)	154	351	237	15,739	32.6	(27.2-37.9)
Hypertension + obesity	157	2,291	62	13,971	9.6	(6.2-13.0)	251	2,197	117	13,916	14.5	(11.3-17.7)
Heart disease + diabetes	452	686	293	15,050	45	(41.7-48.3)	620	518	442	14,901	53.2	(50.4-56.1)
Heart disease + stroke	61	641	93	15,686	12.9	(6.8-19.1)	107	595	171	15,608	19.9	(14.4-25.4)
Heart disease + smoking	65	209	163	16,044	24.8	(17.2-32.3)	98	176	280	15,927	28.7	(22.2-35.1)
Heart disease + obesity	63	1,145	62	15,211	8.2	(3.2-13.2)	117	1,091	119	15,154	14.1	(9.5-18.8)
Diabetes + stroke	66	374	38	16,003	23.5	(16.2-30.8)	110	330	69	15,972	34.5	(28.2-40.9)
Diabetes + smoking	107	112	83	16,179	51.7	(45.0-58.5)	130	89	132	16,130	53.4	(47.3-59.5)
Diabetes + obesity	150	994	39	15,298	20.9	(16.3-25.6)	229	915	76	15,261	29.5	(25.3-33.8)
Stroke + smoking	13	106	23	16,339	16.5	(2.1-30.8)	19	100	41	16,321	20.8	(7.8-33.9)
Stroke + obesity	4	368	7	16,102	2	(0.0-11.8)	9	363	13	16,096	4.3	(0.0-13.9)
Smoking + obesity	17	292	18	16,154	9.5	(0.0-19.5)	25	284	29	16,143	13.3	(3.8-22.8)

^a ICD-10-AM codes: hypertension (I10-I15, R03.0), heart disease (I20-I25, I26-I28, I30-I52), diabetes (E10-E14), smoking (F17.2, Z72.0), obesity (E66)

Supplementary Table 2. Adjusted ORs^a for patient-level variables from the multilevel logistic regression with random intercept for hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)

Patient characteristics	Hypertension (N = 15,279)		Diabetes (N = 4,794)		Heart disease (N = 8,307)		Stroke (N = 2,480)		Smoking (N = 2,099)		Obesity (N = 8,162)	
	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b
Sex												
Female	1		1		1		1		1		1	
Male	1.28	(1.18,1.38)	1.37	(1.19,1.58)	1.30	(1.17,1.44)	1.13	(0.91,1.40)	1.14	(0.94,1.40)	0.85	(0.70,1.04)
Age												
45-59	1		1		1		1		1		1	
60-79	1.27	(1.15,1.41)	0.97	(0.81,1.16)	0.94	(0.82,1.09)	1.08	(0.78,1.52)	0.80	(0.65,0.99)	0.57	(0.47,0.70)
80+	1.32	(1.16,1.49)	1.00	(0.80,1.25)	1.01	(0.86,1.19)	1.02	(0.72,1.46)	0.48	(0.31,0.74)	0.14	(0.08,0.26)
Education												
None	1		1		1		1		1		1	
Trade	0.90	(0.79,1.03)	0.93	(0.73,1.18)	0.80	(0.68,0.94)	1.08	(0.76,1.54)	0.67	(0.48,0.94)	1.42	(1.01,2.02)
School certificate	0.96	(0.86,1.07)	1.00	(0.82,1.23)	0.90	(0.78,1.05)	1.22	(0.90,1.66)	0.87	(0.66,1.16)	1.04	(0.77,1.41)
HSC	0.99	(0.85,1.15)	0.89	(0.68,1.17)	0.91	(0.75,1.11)	2.23	(1.51,3.30)	0.53	(0.37,0.76)	1.24	(0.84,1.83)
Diploma	0.96	(0.84,1.09)	0.90	(0.72,1.14)	0.87	(0.74,1.03)	1.08	(0.75,1.56)	1.00	(0.73,1.37)	1.15	(0.83,1.59)
University	0.85	(0.74,0.98)	0.90	(0.70,1.16)	0.72	(0.60,0.86)	1.23	(0.83,1.81)	0.54	(0.37,0.80)	1.25	(0.88,1.79)
County of birth												
Australia	1		1		1		1		1		1	
Overseas	1.00	(0.91,1.09)	0.95	(0.81,1.11)	1.10	(0.98,1.23)	1.19	(0.94,1.51)	1.17	(0.92,1.48)	0.89	(0.69,1.14)
Functional limitation												
No limitation	1		1		1		1		1		1	
Mild	1.07	(0.91,1.25)	0.91	(0.68,1.23)	1.02	(0.82,1.28)	0.82	(0.48,1.42)	0.92	(0.65,1.30)	1.07	(0.72,1.60)
Moderate	1.23	(1.07,1.42)	1.14	(0.87,1.50)	0.92	(0.75,1.13)	0.68	(0.42,1.11)	0.79	(0.57,1.09)	1.06	(0.73,1.53)
Severe	1.53	(1.33,1.76)	1.54	(1.18,2.01)	0.97	(0.79,1.19)	0.84	(0.53,1.33)	0.82	(0.60,1.12)	2.27	(1.59,3.24)
Income												
<20,000	1		1		1		1		1		1	
20-50,000	0.89	(0.81,0.99)	0.95	(0.79,1.14)	1.03	(0.91,1.17)	1.14	(0.87,1.49)	1.17	(0.90,1.53)	0.76	(0.57,1.00)

1													
2													
3													
4	50-70,000	0.89	(0.75,1.05)	0.87	(0.64,1.19)	1.11	(0.89,1.38)	1.16	(0.68,1.99)	1.37	(0.93,2.02)	0.89	(0.60,1.30)
5	>70,000	0.86	(0.74,1.00)	1.03	(0.77,1.38)	1.24	(1.02,1.50)	1.07	(0.63,1.82)	0.95	(0.66,1.36)	1.15	(0.83,1.59)
6	Not disclosed	1.00	(0.90,1.12)	1.04	(0.86,1.27)	1.14	(0.99,1.31)	1.18	(0.89,1.56)	1.36	(1.02,1.80)	1.07	(0.81,1.41)
7													
8	Admission type^c												
9	Surgical	1		1		1		1		1		1	
10	Other	1.45	(1.23,1.72)	1.01	(0.72,1.42)	2.34	(1.91,2.87)	0.47	(0.16,1.37)	0.69	(0.41,1.14)	0.62	(0.36,1.09)
11	Medical	1.14	(1.03,1.27)	0.66	(0.55,0.80)	0.97	(0.84,1.11)	4.36	(3.02,6.29)	0.50	(0.38,0.65)	0.64	(0.50,0.84)
12													
13	Emergency status^c												
14	Emergency	1		1		1		1		1		1	
15	Planned	0.63	(0.56,0.71)	0.64	(0.52,0.77)	0.42	(0.36,0.49)	0.65	(0.48,0.88)	0.86	(0.65,1.13)	0.58	(0.44,0.78)
16	Other	0.96	(0.80,1.15)	0.98	(0.72,1.33)	1.02	(0.81,1.28)	1.19	(0.80,1.76)	1.03	(0.65,1.62)	0.80	(0.50,1.28)
17													

^a Odds ratio of a hospital record of a condition, among those that self-reported having a condition. Adjusted for age, sex, income, education, country of birth and functional limitation

^b Confidence interval

^c Model included both admission type and emergency status together with other listed patient characteristics

Supplementary Table 3. Adjusted ORs^a for hospital-level variables from the multilevel logistic regression with random intercept for hospital of admission, all public and private hospitals in New South Wales, Australia (n=313)

Hospital characteristics	Hypertension (N = 15,279)		Diabetes (N = 4,794)		Heart disease (N = 8,307)		Stroke (N = 2,480)		Smoking (N = 2,099)		Obesity (N = 8,162)	
	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b	OR ^a	(95%CI) ^b
Hospital type^c												
Public	1		1		1		1		1		1	
Private	0.49	(0.38,0.63)	0.98	(0.78,1.23)	0.35	(0.26,0.47)	0.31	(0.22,0.43)	0.99	(0.72,1.35)	0.91	(0.64,1.31)
Hospital remoteness^c												
Major city	1		1		1		1		1		1	
Inner regional	0.89	(0.64,1.23)	0.86	(0.67,1.11)	1.01	(0.70,1.47)	1.29	(0.91,1.83)	1.04	(0.75,1.45)	0.91	(0.60,1.38)
Outer regional	0.75	(0.55,1.02)	0.69	(0.52,0.91)	0.97	(0.67,1.41)	1.09	(0.72,1.67)	0.82	(0.57,1.18)	0.91	(0.58,1.44)
Remote/very remote	1.05	(0.57,1.94)	0.53	(0.28,1.00)	1.70	(0.81,3.59)	0.66	(0.22,1.98)	0.33	(0.15,0.71)	0.52	(0.16,1.68)
Hospital size^c												
Principal referral	1		1		1		1		1		1	
Major	0.59	(0.34,1.01)	0.89	(0.62,1.27)	0.76	(0.44,1.34)	0.93	(0.58,1.47)	1.03	(0.66,1.61)	1.10	(0.59,2.05)
District	0.45	(0.27,0.76)	0.83	(0.58,1.19)	0.45	(0.26,0.78)	0.97	(0.60,1.55)	0.73	(0.47,1.15)	1.02	(0.55,1.91)
Community	0.41	(0.25,0.68)	0.61	(0.43,0.87)	0.35	(0.20,0.59)	0.57	(0.35,0.94)	0.89	(0.56,1.39)	0.88	(0.47,1.62)
Other	0.52	(0.30,0.89)	0.44	(0.29,0.68)	0.35	(0.19,0.65)	1.19	(0.66,2.14)	0.39	(0.22,0.68)	1.22	(0.59,2.53)
Depth of coding^c												
1 - least comprehensive	0.17	(0.11,0.27)	0.26	(0.17,0.40)	0.09	(0.04,0.17)	0.38	(0.17,0.82)	0.22	(0.12,0.42)	0.28	(0.12,0.65)
2	0.29	(0.22,0.38)	0.66	(0.52,0.85)	0.41	(0.29,0.56)	0.31	(0.21,0.48)	0.74	(0.52,1.06)	0.59	(0.38,0.92)
3	0.58	(0.45,0.76)	0.85	(0.66,1.08)	0.75	(0.55,1.02)	0.66	(0.48,0.91)	0.89	(0.65,1.24)	0.65	(0.43,0.99)
4 - most comprehensive	1		1		1		1		1		1	

^a Odds ratio of a hospital record of a condition, among those that self-reported having a condition. Adjusted for age, sex, income, education, country of birth and functional limitation

^b Confidence interval

^c Hospital-level covariates added one at a time, separately

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8
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	6-8
Bias	9	Describe any efforts to address potential sources of bias	9
Study size	10	Explain how the study size was arrived at	7, 10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	9-10
		(b) Describe any methods used to examine subgroups and interactions	9
		(c) Explain how missing data were addressed	8
		(d) If applicable, describe analytical methods taking account of sampling strategy	n/a
		(e) Describe any sensitivity analyses	9
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10-11, 25-26
		(b) Indicate number of participants with missing data for each variable of interest	25-26
Outcome data	15*	Report numbers of outcome events or summary measures	11, 27,28,29,31-34
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10, 12-13, 32-34
		(b) Report category boundaries when continuous variables were categorized	30,32-34
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-13,30
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
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	18-19
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-19
Generalisability	21	Discuss the generalisability (external validity) of the study results	18-19
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	20

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.