

Categorising major cardiovascular disease hospitalisations from routinely collected data

Grace Joshy^{a,d}, Rosemary J Korda^a, Walter P Abhayaratna^b, Kay Soga^a and Emily Banks^{a,c}

^a National Centre for Epidemiology and Population Health, Australian National University, Canberra, ACT, Australia

^b Medical School, College of Medicine, Biology and the Environment, Australian National University, Canberra, ACT, Australia

^c The Sax Institute, Sydney, NSW, Australia

^d Corresponding author: grace.joshy@anu.edu.au

Article history

Publication date: July 2015

Citation: Joshy G, Korda RJ,

Abhayaratna WP, Soga K, Banks E.

Categorising major cardiovascular disease hospitalisations from routinely collected data. Public Health Res Pract. 2015;25(3):e2531532. doi: <http://dx.doi.org/10.17061/phrp2531532>

Key points

- International Classification of Diseases (ICD) codes under 'Diseases of the circulatory system' (I00–I99) are often grouped together, combining common severe atherosclerotic and thrombotic cardiovascular disease (CVD) with common, less severe and pathologically dissimilar conditions
- We propose a pragmatic approach to using ICD diagnosis and procedure codes to capture major atherosclerotic and arteriovenous thromboembolic and related CVD
- The implications of the proposed approach for classifying CVD are investigated using linked data

Abstract

Objectives and importance of the study: Routine hospital administrative data provide an important source of information about cardiovascular disease (CVD) for health statistics reporting, health services and research. How such conditions are classified and grouped has implications for the use and interpretation of these data. International Classification of Diseases (ICD) diagnosis codes from hospital data collections are often used to classify CVD, but there is little published evidence on the most appropriate ways to use these codes to categorise CVD in a way that maximises the usefulness of hospital data for reporting and research. In particular, ICD codes under 'Diseases of the circulatory system' (I00–I99) are often grouped together into a general CVD category. However, this category is heterogeneous and combines common severe atherosclerotic and thrombotic CVDs (such as myocardial infarction and pulmonary embolism) with common, less severe and pathologically dissimilar conditions (such as varicose veins and haemorrhoids). In addition, hospital data collections contain a range of data fields, including those relating to primary and additional diagnoses and those relating to procedures. All of these have the potential to contribute valuable information on CVD. This paper proposes a pragmatic approach to using ICD diagnosis codes and procedure codes to capture major atherosclerotic and arteriovenous thromboembolic and related CVD.

Methods: We reviewed the ICD diagnosis codes and procedure codes and developed an algorithm for classifying and categorising major CVD diagnoses. This approach was then applied to linked hospitalisation data from individuals participating in the 45 and Up Study, a cohort study of 267 153 New South Wales residents aged 45 and over, to investigate the implications of the proposed approach for quantifying CVD.

Results: Large differences were observed in the numbers of events in grouped CVD outcomes, depending on the methods used.

Conclusions: In cases where the reporting and research interest relates to incident disease, it may be appropriate to prioritise specific disease categories and pathological homogeneity.

Introduction

Routine administrative hospital data are an important source of information for population health, health services and research. They have been particularly important for cardiovascular disease (CVD) statistics and research because there are few disease registers for CVD, and reporting of CVD is not mandatory, unlike most cancer types and certain infectious diseases.

Hospital data can be coded using a variety of systems. In practice, the International Classification of Diseases (ICD) has become the international standard diagnostic classification for many general epidemiological and health management purposes.¹ These include analysing the general health status of population groups, and monitoring the incidence and prevalence of diseases and other health problems in relation to other variables such as the characteristics and circumstances of the individuals affected. Generally, the ICD orients diagnoses around the organ system affected and the pathological or clinical diagnosis within each organ system, with additional broader and less specific codes where this is not possible or appropriate. The ICD allows ascertainment of specific CVD endpoints (e.g. I21: acute myocardial infarction), homogeneous 'blocks' of three-character categories (e.g. I20–I25: ischaemic heart diseases), through to an entire 'chapter' (Chapter IX, 'Diseases of the circulatory system': ICD-10-AM codes I00–I99), although hospital data vary in their validity for this purpose.^{2–6} However, methods of ascertaining and categorising CVD using hospital data vary considerably^{7–10}, often with little explicit justification for why a particular method has been applied. In particular, it is common for all codes at the chapter level to be used to identify 'All CVD' or 'All circulatory disease' hospitalisations.^{10–12} However, this results in a heterogeneous category that combines common severe atherosclerotic and thrombotic CVD (caused by plaque formation or blood clots in blood vessels), such as myocardial infarction and pulmonary embolism, with common, less severe and pathologically dissimilar conditions such as varicose veins and haemorrhoids. In addition, hospital data collections contain a range of data fields, including those relating to primary and additional diagnoses, and those relating to procedures. All of these have the potential to contribute valuable information on CVD. Depending on how CVD is defined, the reporting of the total burden of CVD in the population can vary considerably – for example, if CVD is defined as 'heart, stroke and vascular diseases', 1 million Australians were affected in 2011–12, but if the definition includes all circulatory diseases, 3.7 million Australians were affected.^{10,13}

This paper proposes a pragmatic approach to using primary and additional ICD diagnosis codes and procedure codes to capture major atherosclerotic, arteriovenous thromboembolic and related CVD. The implications of differences in CVD ascertainment methods are then explored using a large-scale cohort study and its linked datasets.

Methods

Classifying and categorising major CVD diagnoses using hospital data

Context

Australia uses the International Statistical Classification of Diseases and Related Health Conditions, to code and classify a range of health data, including hospital diagnoses and causes of death, with the Tenth Revision Australian Modification (ICD-10-AM) in use since 1997–98.¹⁴ Procedures are coded using the Australian Classification of Health Interventions, which is used in conjunction with ICD-10-AM.¹⁵

The New South Wales (NSW) Admitted Patient Data Collection (APDC) is a complete census of all public and private hospital admissions in NSW. In this dataset, a 'hospitalisation' refers to an episode of admitted care, which can be a total hospital stay (from admission to discharge, transfer or death) or a portion of a hospital stay beginning or ending in a change of type of care (e.g. from acute care to rehabilitation). A person can have multiple 'separations' within the same hospitalisation episode (e.g. transfers between hospitals or wards). Data linkage – either internal or to an external cohort – is required to track individuals to allow the number and proportion of people admitted to hospital with CVD to be estimated.

Diagnoses are recorded as principal and additional in the APDC.¹⁰ The principal diagnosis is the diagnosis established after an episode of admitted patient care that is considered to be mainly responsible for the care episode. The principal diagnosis is usually a disease, but can also be injury or poisoning, a specific care or service provided for a current condition (e.g. dialysis for kidney disease), or other reasons for hospitalisation.¹⁰ An additional diagnosis is a condition or complaint that affects patient management and either coexists with the principal diagnosis or arises during the episode of care. Up to 54 additional diagnoses may be recorded.

Outcome and grouping definitions

The aim of this study was to develop a pragmatic and inclusive definition of major atherosclerotic and arteriovenous thromboembolic CVD for use in reporting and research. We were not aiming to identify clinically or pathologically 'pure' atherosclerosis or arteriovenous thromboembolic disease (such a grouping is not necessarily feasible using routine data), but to focus on serious disease that has a significant atherosclerotic or thrombotic component.

The authors (GJ, EB, RK, KS) reviewed the ICD-10-AM three-character codes (one letter and two digits) for all circulatory diseases, including 'Diseases of the circulatory system' (I00–99), as well as G45 ('Transient cerebral ischaemic attacks and related syndromes') and G46 ('Vascular syndromes of brain in cerebrovascular

diseases'). From these, we selected conditions that were considered to have a significant atherosclerotic or arteriovenous thromboembolic contribution, and placed them in a group that we termed 'major CVD'. Where aetiology was likely to be mixed, we erred on the side of inclusiveness. We excluded conditions that chiefly and primarily affect the lymphatic system or are caused by other factors, such as CVD caused by infections; nonthrombotic diseases of veins, lymphatic vessels and lymph nodes; and other or unspecified disorders of the circulatory system. We excluded essential hypertension on the grounds that it was a risk factor rather than a disease outcome. The excluded conditions were grouped as 'other circulatory diseases' for the purposes of this paper. We also considered procedure codes related to CVD interventions¹⁰ and identified those likely to relate to those conditions considered 'major CVD'. Groupings were refined further in consultation with an expert cardiologist (author WA).

Applying the approach to a cohort study with linked hospitalisation data

Data sources

The Sax Institute's 45 and Up Study is a large-scale Australian cohort study of 267 153 men and women aged 45 years and over who were randomly sampled from the general population of NSW, Australia. Individuals joined the study by completing a baseline postal questionnaire (distributed from 1 January 2006 to 31 December 2008) and giving informed consent for follow-up through repeated data collection and linkage of their data to population health databases. The study methods are described in detail elsewhere.¹⁶

Questionnaire data from study participants have been linked probabilistically to the NSW APDC by the Centre for Health Record Linkage (www.cherel.org.au). The linked data contain details of admissions in participants from 1 July 2000 to 31 December 2011, including dates of admission and discharge, the primary reason for admission using ICD-10-AM, up to 54 additional clinical diagnoses and up to 50 operations or procedures, coded using the Australian Classification of Health Interventions.¹⁵ Dates of death were ascertained from the date of recruitment up to 31 December 2011, using linkage to the NSW Register of Births, Deaths and Marriages. Death registrations capture all deaths in NSW.

Ethical approval for the study was obtained from the NSW Population and Health Services Research Ethics Committee and the Australian National University Human Research Ethics Committee.

Analysis

The 45 and Up Study linked data were used to follow participants from baseline until 31 December 2011 or the date of death, whichever occurred first. We ascertained the total number of hospitalisations attributable to the

CVD codes under investigation and the total number of participants hospitalised with these codes over the follow-up period. Participants with invalid age or date of recruitment and those with linkage errors ($n = 428$, 0.16%) were excluded.

The numbers of CVD hospitalisations were identified using two different definitions: the proposed definition of major CVD and the definition using codes for all circulatory diseases (I00–99, plus G45, G46 and relevant procedures). Using both definitions, the numbers of cases were ascertained using primary diagnosis codes only, as well as using both primary and additional diagnosis codes. We also investigated the number of additional cases ascertained when using procedure codes with and without paired diagnosis codes (specific to each procedure).

Results

The Australian Classification of Health Interventions procedure codes and ICD-10-AM diagnosis codes categorised as major CVD are outlined in Tables 1 and 2, respectively, and the other circulatory disease codes excluded by the review process (categorised as 'other circulatory disease') are shown in Table 3.

Study participants ($n = 266\,725$) were followed for a median of 3.42 years (range 0.005–6.33 years). Overall, there were 39 459 hospital admissions and 24 543 participants admitted for major CVD as a primary diagnosis; the corresponding figures for all diagnosis codes were 65 340 and 33 242 (Table 2). Considering primary diagnoses only, the greatest contributions to this group were from angina pectoris (I20: 6725 hospitalisations), atrial fibrillation and flutter (I48: 6270 hospitalisations) and chronic ischaemic heart disease (I25: 4845 hospitalisations). Including secondary diagnoses substantially increased the number of hospitalisations and the number of participants hospitalised (by 66% and 35%, respectively), but the proportions of hospitalisations attributed to the individual codes were broadly similar to those for primary diagnosis codes alone. Procedures (coded with a relevant diagnosis code) contributed an additional 1.1% of major CVD admissions (increasing from 39 459 to 39 899); dropping the requirement for a procedure code to include a relevant diagnosis code made little difference to the numbers ascertained (1.6% increase, from 39 459 to 40 100).

Of the other circulatory diseases, the largest contributors based on primary diagnosis codes were haemorrhoids (I84: 3810 hospitalisations), varicose veins (I83: 1281 hospitalisations) and hypotension (I95: 1002 hospitalisations) (Table 3). When both primary and additional diagnosis codes were included, 59% of all hospital admissions with these codes were for essential hypertension (I10: 39 820 hospitalisations). The grouping of 'all circulatory disease' (I00–99, G45 and G46)

Table 1. Australian Classification of Health Interventions codes for CVD-related intervention procedures, used to define major CVD

Procedures	Australian Classification of Health Interventions codes
Percutaneous coronary interventions	35304-00, 35305-00, 35304-01, 35305-01, <i>38300-00, 38303-00</i> (block: 670) 35310-00, 35310-01, 35310-02, 35310-03, 35310-04, 35310-05, <i>38306-00, 38306-01, 38306-02, 38306-03, 38306-05</i> (block: 671) and diagnosis codes I20–I25
Coronary artery bypass grafting	38497-00 to 38497-07, 38500-00 to 38500-04, 38503-00 to 38503-04, 90201-00 to 90201-03, <i>38500-05, 38503-05</i> (blocks: 672–679) and diagnosis codes I20–I25, I34, I35
Heart transplant	90205-00, 90205-01 (block: 660)
Cardiac defibrillator implants	38524-00, 38521-01, 38521-02, 38521-03, <i>38393-00</i> (block:653) and diagnosis codes I20, I21,I25, I42–I51
Valve replacement, repair or reconstruction	Aortic valve (blocks 621, 622, 623, 624) 38456-10, 38483-00, 38270-01, 38480-00, 38481-00, 38488-00, 38488-01, 38489-00, 38489-01, 38456-15, 38653-04, 38475-02, 38477-02. Mitral valve (blocks 625, 626, 627, 628, 629, 630) 38487-00, 38485-01, 38270-02, 38480-01, 38481-01, 38475-00, 38477-00, 38488-02, 38488-03, 38489-02, 38485-00, 38456-16, 38653-05. Tricuspid valve (blocks 631, 632, 633, 634, 635) 38456-11, 38480-02, 38481-02, 38475-01, 38477-01, 38488-04, 38488-05, 38489-03, 38456-17, 38653-06. Pulmonary valve (blocks 636, 637, 638) 38456-01, 38270-03, 38488-06, 38488-07, 38489-04, 38489-05, 38456-18, 38653-07. and diagnosis codes I01–I09, I20–I25, I33–I39, I42–I52, Q20–Q25, T82
Pacemaker insertion	38281-00, 38281-01, 38281-02, 38281-03, 38281-04, 38281-05, 38281-06, 38281-07, 38281-08, 38281-09, 38281-10, 38281-11, 38281-12, 38281-13, <i>38353-00</i> (blocks: 650, 651, 652) and diagnosis codes I08, I20, I21, I25, I34–I39, I42–I52
Carotid endarterectomy	33500-00 (block: 700) and diagnosis codes G45, I63–I67

CVD = cardiovascular disease

Notes: Codes in italics are additional codes that we have included to reflect coding changes during the follow-up period.

Computerised tomographic brain scans (56001-00, 56007-00, 56010-02, 56010-03 and diagnosis codes G45, I60–I64), echocardiography (55112-00, 55118-00, 55130-00 and diagnosis codes I00–I52), magnetic resonance imaging brain scans (90901-00 and diagnosis codes G45, I60–I64) and coronary angiography (38215-00, 38218-00, 38218-01, 38218-02, and diagnosis codes I20–I25, I30–I52) are CVD-related diagnostic procedures and do not amount to diagnosis of CVD. These procedure codes are not used to define major CVD. Source for the full list of procedure codes: Australian Institute of Health and Welfare.¹⁰

included 47 678 hospitalisations and 30 389 participants using primary diagnosis codes, increasing to 105 759 and 55 036, respectively, when all diagnosis codes were used.

Discussion

To maximise the usefulness of routine administrative hospital data for CVD reporting and research, such data must be classified and categorised judiciously. The findings presented here demonstrate large differences in the composition and numbers of events in grouped CVD outcomes depending on the methods used. In particular, use of the ICD-10-AM 'all circulatory disease' (I00–I99, G45, G46) category results in a large, heterogeneous group; inclusion of both primary and secondary diagnoses in this category leads to a substantial contribution from common comorbidities such as hypertension.

Where the reporting and research interest relates to incident disease, it may be appropriate to

prioritise specific disease categories and pathological homogeneity. Where grouping is necessary, it is important that the goals of and underlying rationale for such a grouping is clear. In defining the category for 'major CVD', we included conditions with a substantial contribution from atherosclerotic and arteriovenous thromboembolic disease. Along with its specific subcomponents, this category is particularly suited for research aimed at understanding conditions with this aetiology. We excluded conditions known to be secondary to nonatherosclerotic causes, such as those caused by infectious and infiltrative disease (e.g. rheumatic heart disease, infectious endocarditis and portal vein thrombosis). We aimed for overall pathological homogeneity in outcome ascertainment while maintaining inclusiveness. We also wanted to minimise assumptions about underlying relationships. For example, excluding venous thrombosis from the definition of major CVD could be seen as assuming that risks for arterial and venous thrombosis differ fundamentally for all exposures.

Table 2. 'Major CVD' codes, number of hospitalisations and number of participants hospitalised during follow-up

ICD-10-AM code	Using primary diagnosis code only		Using all 55 diagnosis codes	
	No. of hospitalisations	No. of participants	No. of hospitalisations	No. of participants
I11-I13 (selected hypertensive diseases)	57	51	426	214
I11 Hypertensive heart disease	25	22	87	78
I12 Hypertensive kidney disease	22	20	314	116
I13 Hypertensive heart and kidney disease	10	10	25	22
I20-I25 Ischaemic heart disease	14 846	10 663	24 154	15 012
I20 Angina pectoris	6 725	5 203	8 124	6 046
I21 Acute myocardial infarction	4 057	3 535	5 740	4 692
I22 Subsequent myocardial infarction	17	17	36	36
I23 Certain current complications following acute myocardial infarction	21	21	45	43
I24 Other acute ischaemic heart diseases	45	43	197	174
I25 Chronic ischaemic heart disease	4 845	4 059	18 517	11 861
I26-I28 Pulmonary heart disease and diseases of pulmonary circulation	1 154	1 029	3 247	2 613
I26 Pulmonary embolism	1 002	921	1 826	1 600
I27 Other pulmonary heart diseases	150	111	1 465	1 085
I28 Other diseases of pulmonary vessels	2	2	14	14
I34-36, I42, I44, I46-I51 (selected other forms of heart disease)	14 651	9506	33 796	18 857
I34 Nonrheumatic mitral valve disorders	311	236	1 115	882
I35 Nonrheumatic aortic valve disorders	855	627	2 261	1 444
I36 Nonrheumatic tricuspid valve disorders	3	3	27	24
I42 Cardiomyopathy	328	283	1 425	965
I44 Atrioventricular and left bundle-branch block	597	569	2 039	1 796
I46 Cardiac arrest	134	134	658	647
I47 Paroxysmal tachycardia	1 051	836	2 126	1 727
I48 Atrial fibrillation and flutter	6 270	4 320	19 425	11 701
I49 Other cardiac arrhythmias	883	824	1 788	1 610
I50 Heart failure	4 376	2 977	10 883	6 347
I51 Complications and ill-defined descriptions of heart disease	98	96	1 397	1 226
I61-I67, I69 (selected cerebrovascular disease)	3 522	3 139	8 441	4 587
I61 Intracerebral haemorrhage	422	398	791	523
I62 Other nontraumatic intracranial haemorrhage	244	227	398	323
I63 Cerebral infarction	1 501	1 432	3 008	1 754
I64 Stroke, not specified as haemorrhage or infarction	812	779	1 663	1 041
I65 Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction	425	362	968	779
I66 Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction	78	76	144	132
I67 Other cerebrovascular diseases	193	160	803	508
I69 Sequelae of cerebrovascular disease	5	5	1 633	774

(continued)

Table 2. 'Major CVD' codes, number of hospitalisations and number of participants hospitalised during follow-up (continued)

ICD-10-AM code	Using primary diagnosis code only		Using all 55 diagnosis codes	
	No. of hospitalisations	No. of participants	No. of hospitalisations	No. of participants
I70-I77 (selected diseases of arteries, arterioles and capillaries)	3 183	2 088	6 268	3 657
I70 Atherosclerosis	2 046	1 302	4 180	2 366
I71 Aortic aneurysm and dissection	634	501	1 121	799
I72 Other aneurysm and dissection	203	153	522	353
I73 Other peripheral vascular diseases	38	34	179	135
I74 Arterial embolism and thrombosis	157	145	372	329
I77 Other disorders of arteries and arterioles	122	109	330	281
I80 (selected diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified)				
I80 Phlebitis and thrombophlebitis	689	643	2 334	1 882
G45,G46 (selected episodic and paroxysmal disorders)				
G45 Transient cerebral ischaemic attacks and related syndromes	1 946	1 775	2 371	2 078
G46 Vascular syndromes of brain in cerebrovascular diseases	0	0	23	20
a. Major CVD diagnosis (all of the above)	39 459	24 543	65 340	33 242
b. CVD-related intervention procedures* (including diagnosis codes in the definition)	6 998	6 179	6 998	6 179
c. CVD-related intervention procedures* (excluding diagnosis codes in the definition)	7 222	6 373	7 222	6 373
d. Major CVD outcome (a or b)	39 899	24 755	65 410	33 273
e. Major CVD outcome (a or c)	40 100	24 866	65 602	33 361

ICD-10-AM = International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification.

* Procedure codes in any of the 50 code fields; see Table 1 for the full list of procedure codes.

Notes: Episodes of hospitalisation with changes in type of care (e.g. transfer from acute to rehabilitation) could contribute to more than one ICD-10-AM three-character code; as a result, numbers may not add up precisely.

Codes I70–I74 are often grouped together to identify hospitalisations for peripheral vascular disease.

Codes I60–I64 are often grouped together to identify hospitalisations for stroke. Although I60 (subarachnoid haemorrhage) is a major illness, it has a different aetiology from ischaemic stroke and is neither atherosclerotic nor arterothrombotic. We have excluded I60 from the definition of major CVD.

How the proposed groupings are implemented depends on the outcome and research questions of interest. For example, where the interest is in incident hospitalisation attributable to major atherosclerotic and related CVD, then the primary diagnosis codes for these conditions may be most relevant. When evidence of a history of CVD is sought, the use of 'major CVD' diagnoses, including primary and secondary diagnoses codes and CVD-related intervention procedures may be more useful. Because procedure codes introduce only a marginal gain in case identification, these may be dropped from the algorithm. In addition, because linked hospitalisation data may cover only a limited time period before baseline, self-reported survey data may be used as a complementary source of information about history of CVD. Where there is interest in comorbidity such as hypertension, and other elements of health status and health services use, broader groupings and those that

include primary and secondary diagnoses may be more appropriate, but the validity of any included diagnoses should be borne in mind.^{2–6} Investigations involving CVD mortality are also likely to benefit from the use of secondary diagnosis codes. For example, a significant number of fatal cases of myocardial infarction are coded in secondary diagnosis fields, with the principal diagnosis being a code for arrhythmia.

The classification method and grouping outlined here acknowledges that hospital administrative diagnoses are imperfect and variable⁶, and we chose not to go beyond the three-character ICD-10-AM codes. Reporting patterns are also likely to change with the recent introduction of activity-based funding in many Australian states.¹⁷ The pragmatic method outlined here is open to being refined and may benefit from adaptation to address specific research questions, or according to the preferences of other research groups.

Table 3. Codes not included in ‘major CVD’, number of hospitalisations and number of participants hospitalised during follow-up

Disease type	ICD-10-AM code	Using primary diagnosis code only		Using all 55 diagnosis codes	
		No. of hospitalisations	No. of participants	No. of hospitalisations	No. of participants
Acute rheumatic fever	I00 Rheumatic fever without mention of heart involvement	1	1	1	1
	I01 Rheumatic fever with heart involvement	0	0	3	3
	I02 Rheumatic chorea	1	1	1	1
Chronic rheumatic heart diseases	I05 Rheumatic mitral valve diseases	56	51	169	136
	I06 Rheumatic aortic valve diseases	6	6	8	8
	I07 Rheumatic tricuspid valve diseases	18	15	195	172
	I08 Multiple valve diseases	93	85	654	522
	I09 Other rheumatic heart diseases	0	0	16	15
Selected hypertensive diseases	I10 Essential (primary) hypertension	620	556	39 820	23 724
	I15 Secondary hypertension	10	9	37	32
Selected other forms of heart disease	I30 Acute pericarditis	82	75	119	111
	I31 Other diseases of pericardium	202	181	569	500
	I32 ^a Pericarditis in diseases classified elsewhere	0	0	3	3
	I33 Acute and subacute endocarditis	121	77	178	127
	I38 Endocarditis, valve unspecified	24	23	105	81
	I39 ^a Endocarditis and heart valve disorders in diseases classified elsewhere	0	0	1	1
	I40 Acute myocarditis	2	2	7	7
	I41 ^a Myocarditis in diseases classified elsewhere	0	0	0	0
	I43 ^a Cardiomyopathy in diseases classified elsewhere	1	1	36	22
	I45 Other conduction disorders	232	223	885	790
Selected cerebrovascular diseases	I60 Subarachnoid haemorrhage	144	135	353	175
	I68 ^a Cerebrovascular disorders in diseases classified elsewhere	1	1	30	19
Selected diseases of arteries, arterioles and capillaries	I78 Diseases of capillaries	58	31	123	82
	I79 ^a Disorders of arteries, arterioles and capillaries in diseases classified elsewhere	1	1	4	3
Selected diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified	I81 Portal vein thrombosis	19	13	63	47
	I82 Other venous embolism and thrombosis	68	66	390	358
	I83 Varicose veins of lower extremities	1 281	1 131	1 532	1 310
	I84 Haemorrhoids	3 810	3 595	13 419	12 244
	I85 Oesophageal varices	139	75	239	139
	I86 Varicose veins of other sites	42	39	120	101
	I87 Other disorders of veins	97	87	342	289
I88 Nonspecific lymphadenitis	23	23	42	42	

(continued)

Table 3. Codes not included in ‘major CVD’, number of hospitalisations and number of participants hospitalised during follow-up (*continued*)

Disease type	ICD-10-AM code	Using primary diagnosis code only		Using all 55 diagnosis codes	
		No. of hospitalisations	No. of participants	No. of hospitalisations	No. of participants
Selected other and unspecified disorders of the circulatory system	I89 Other noninfective disorders of lymphatic vessels and lymph nodes	77	52	1 057	329
	I95 Hypotension	1 002	943	10 354	8 642
	I97 Postprocedural disorders of circulatory system, not elsewhere classified	238	231	2 324	1 840
	I98 ^a Other disorders of circulatory system in diseases classified elsewhere	0	0	163	103
	I99 Other and unspecified disorders of circulatory system	13	12	76	73
Codes not included in the definition of Major CVD (all of the above)		8 454	7 523	67 359	41 979
All circulatory diseases (I00–I99, G45, G46)		47 678	30 389	105 759	55 036

ICD-10-AM = International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification a Categories for ICD-10-AM Chapter IX (Diseases of the circulatory system), which implies that they are optional additional codes for the manifestation.

Conclusion

In investigations of CVD outcomes using routine administrative hospital data, where classification and grouping is necessary, it is important to clearly specify the goal and rationale for the grouping method employed. Prioritisation of specific disease categories and pathological homogeneity may be appropriate, particularly where the reporting and research interest relates to incident disease.

Acknowledgements

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; NSW Government Family and Community Services – Carers, Ageing and Disability Inclusion; and the Australian Red Cross Blood Service. We thank the many thousands of people participating in the 45 and Up Study.

This specific study was supported by a National Health and Medical Research Council (NHMRC) of Australia project grant (reference: 1024450), a NHMRC partnership grant (reference: 1092674), a NHMRC Centre for Research Excellence grant and a Cardiovascular Disease Network Development Grant from the Heart Foundation of Australia. Emily Banks is supported by an NHMRC Fellowship (no. 042717).

Competing interests

None declared

Author contributions

GJ, EB and RK conceived of the original idea for the paper. GJ and KS conducted the data analyses. GJ and EB produced the initial draft of the paper. WA provided clinical and epidemiological expertise. All authors were involved in the interpretation of the analyses and writing the paper. All authors approved the final version of the paper and take responsibility for its content.

References

1. World Health Organization. International statistical classification of diseases and related health problems 10th revision. Geneva: World Health Organization; 2010 [cited 2015 May 20]. Available from: apps.who.int/classifications/icd10/browse/2015/en
2. Boyle CA, Dobson AJ. The accuracy of hospital records and death certificates for acute myocardial infarction. *Aust N Z J Med.* 1995;25(4):316–23.
3. Tirschwell DL, Longstreth WT Jr. Validating administrative data in stroke research. *Stroke.* 2002;33(10):2465–70.
4. Teng TH, Finn J, Hung J, Geelhoed E, Hobbs M. A validation study: how effective is the Hospital Morbidity Data as a surveillance tool for heart failure in Western Australia? *Aust N Z J Public Health.* 2008;32(5):405–7.

5. Ruigomez A, Martin-Merino E, Rodriguez LA. Validation of ischemic cerebrovascular diagnoses in the health improvement network (THIN). *Pharmacoepidemiol Drug Saf.* 2010;19(6):579–85.
6. Lujic S, Watson D, Randall DA, Simpson JM, Jorm L. Variation in the recording of common health conditions in routine hospital data: study using linked survey and administrative data in New South Wales, Australia. *BMJ Open.* 2014;4(9):e005768.
7. Wilson PW, D'Agostino RB, Sullivan L, Parise H, Kannel WB. Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. *Arch Intern Med.* 2002;162(16):1867–72.
8. Guh D, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis A. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC Public Health.* 2009;9(1):88.
9. Population Health Division. The health of the people of New South Wales – report of the Chief Health Officer. Sydney: NSW Department of Health; 2010.
10. Australian Institute of Health and Welfare. Cardiovascular disease: Australian facts 2011. Canberra: Australian Institute of Health and Welfare; 2011 [cited 2015 May 20]. Available from: www.aihw.gov.au/publication-detail/?id=10737418510
11. Hall MJ, Williams SN, Golosinskiy A, Schwartzman A. National health statistics reports: national hospital discharge survey: 2007 summary. Washington, DC: US Department of Health and Human Services; 2010 [cited 2014 Nov 20]. Available from: www.cdc.gov/nchs/data/nhsr/nhsr029.pdf
12. Banks E, Joshy G, Abhayaratna WP, Kritharides L, Macdonald PS, Korda RJ, et al. Erectile dysfunction severity as a risk marker for cardiovascular disease hospitalisation and all-cause mortality: a prospective cohort study. *PLoS Med.* 2013;10(1):e1001372.
13. Australian Bureau of Statistics. Australian health survey: first results, 2011–12. Canberra: Australian Bureau of Statistics; 2012 [cited 2015 May 20]. Available from: www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012011-12?OpenDocument
14. National Centre for Classification in Health. International statistical classification of diseases and related health problems, 10th revision, Australian modification 5th edition. Sydney: National Centre for Classification in Health; 2006.
15. National Centre for Classification in Health. The Australian classification of health interventions (ACHI) 6th edition. Tabular list of interventions and alphabetic index of interventions. Sydney: National Centre for Classification in Health; 2007.
16. 45 and Up Study Collaborators, Banks E, Redman S, Jorm L, Armstrong B, Bauman A, et al. Cohort profile: the 45 and up study. *Int J Epidemiol.* 2008;37(5):941–7.
17. O'Reilly J, Busse R, Hakkinen U, Or Z, Street A, Wiley M. Paying for hospital care: the experience with implementing activity-based funding in five European countries. *Health Econ, Policy, Law.* 2012;7(1):73–101.

Copyright: 

© 2015 Joshy et al. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence, which allows others to redistribute, adapt and share this work non-commercially provided they attribute the work and any adapted version of it is distributed under the same Creative Commons licence terms. See: www.creativecommons.org/licenses/by-nc-sa/4.0/