

A comparison of Australian chronic disease prevalence estimates using administrative pharmaceutical dispensing data with international and community survey data

Shaun Francis Purkiss^{1*}, Tessa Keegel^{1,2}, Hassan Vally¹, and Dennis Wollersheim¹

Submission History

Submitted:	15/02/2020
Accepted:	16/09/2020
Published:	xx/xx/20xx

¹Department of Public Health, La Trobe University, Bundoora, Victoria, Australia

²Monash Centre for Occupational and Environmental Health, Monash University, Victoria, Australia

Abstract

Introduction

Chronic disease (CD) is a leading cause of population mortality, illness and disability. Identification of CD using administrative data is increasingly used and may have utility in monitoring population health. Pharmaceutical administrative data using World Health Organization, Anatomic Therapeutic Chemical Codification (ATC) assigned to prescribed medicines may offer an improved method to define persons with certain CD and enable the calculation of population prevalence.

Objective

To assess the feasibility of Australian Pharmaceutical Benefits Scheme (PBS) dispensing data, to provide realistic measures of chronic disease prevalence using ATC codification, and compare values with international data using similar ATC methods and Australian community surveys.

Methods

Twenty-two chronic diseases were identified using World Health Organization (WHO) formulated ATC codes assigned to treatments received and recorded in a PBS database. Distinct treatment episodes prescribed to individuals were counted annually for prevalence estimates. Comparisons were then made with estimates from international studies using pharmaceutical data and published Australian community surveys.

Results

PBS prevalence estimates for a range of chronic diseases listed in European studies and Australian community surveys demonstrated good correlation. PBS estimates of the prevalence of diabetes, cardiovascular disease and hypertension, dyslipidemia, and respiratory disease with comparable Australian National Health Survey in older adults showed correlations of between ($r = 0.82 - 0.99$) and a range of percentage error of -11% to 59%. However, other conditions such as psychological disease and migraine showed greater disparity and correlated less well.

Conclusions

Although not without limitations, Australian administrative pharmaceutical dispensing data may provide an alternative perspective on population health and a useful resource to estimate the prevalence of a number of chronic diseases within the Australian population.

Keywords

chronic disease; prevalence; pharmaceutical data; administrative data

Introduction

Chronic disease is a major determinant of patient and population health and its extent has a significant effect on

the welfare of any society [1, 2]. As chronic conditions are in general related to age; their effects will increasingly impact on the health of many communities as life expectancies improve [3]. Monitoring chronic disease within a population with accurate prevalence estimates is therefore important to provide

*Corresponding Author:

Email Address: S.Purkiss@latrobe.edu.au (Shaun Francis Purkiss)

background for health policy formulation and the prediction of health care expenditure [4].

Currently, in the Australian setting, the Australian Bureau of Statistics provide chronic disease prevalence data every three years [5]. These National Health Surveys (NHS) are based on self-reported responses to structured questionnaires and rely on participants providing correct information [6]. Suboptimal respondent understanding of chronic disease conditions and unreliability of recall can lead to imprecise diagnostic categorization [7]. The National Health Survey is also community based and acknowledged to have poor coverage of remote locations and indigenous people in these areas [5]. Given these weaknesses in the current system, alternative data sources may be studied in an attempt to offer new perspectives of chronic disease and improve population prevalence estimates [8].

Pharmacy prescription data may be used to calculate chronic disease prevalence [8–10]. Prescriptions represent a treatment decision by a health professional and provide units of information that include the dispensed drugs supplied, their dose and quantity [9, 10]. This data may also be used to construct a medical diagnosis from the patterns of supplied medications [8]. The number of patients with a drug defined proxy diagnosis can be subsequently counted to estimate population disease prevalence [8–11].

In Australia, pharmacy claims data are collected as part of the national Pharmaceutical Benefits Scheme (PBS). This program is central to the Australian government's universal health care system that subsidises most medicines to all residents who hold a Medicare card [12]. The PBS scheme produces data with extensive coverage of the Australian population including indigenous communities and lower decile groups. However, a small number of persons obtain medicines privately [12]. In international settings, similar data sets have been used for a variety of purposes such as studies of disease prevalence, illness severity and insurance purposes [8–10, 13]. In a Swiss study of insured patients, Huber, Szucs, Rapold and Reich employed administrative pharmaceutical data to measure chronic disease prevalence based on an updated allocation of Anatomical Therapeutic Chemical (ATC) classifications to the medications prescribed [8]. ATC classifications are alphanumeric codes designated by the World Health Organization (WHO) which can allow international comparison of studies with different pharmacopeia [14].

The aim of this study was to assess the feasibility of the PBS database and the ATC classification system to calculate chronic disease prevalence estimates for the Australian population. Our study also compared the prevalence estimates derived from PBS data with results from international studies using similar pharmaceutical data sources and ATC methodologies. Finally, the study will draw comparisons between PBS disease prevalence estimates with similar prevalence estimates reported in Australian community data such as the National Health Survey [4, 5, 16].

Methods

Study design

This was a retrospective cross-sectional study using a published 10% sample of the PBS administrative database

that details the exchanges of dispensed prescribed medications occurring at Australian pharmacies [17, 18]. The medications recorded in the prescriptions and the PBS database were allocated ATC codes according to World Health Organization criteria [8, 14]. ATC codes of treatments prescribed to individuals were then used to define persons with particular chronic diseases. The number of persons with these specific conditions were counted and prevalence estimates calculated using population data from the Australian Bureau of Statistics (ABS). The prevalence estimates for individual chronic diseases were then compared with European data using similar methodology and Australian data from community surveys.

Data source and setting

A PBS claims database published by the Australian government was used for this study. [7, 18]. The data represents the complete dispensed prescription information of 10% of the overall population of Australian persons that utilised the PBS scheme during the period 2003–2014. The PBS data are de-identified but include an assigned numeric unique identifier (UI), person's gender, state of residence and year of birth. The ten percent data set were selected randomly by government statisticians from the entire current and historical PBS data prior to publication. The pharmaceutical information listed for each individual patient includes the prescribed drug name and its date of supply, its form, and strength [17, 18].

Identification of chronic disease and prevalence calculation

ATC codes were employed to identify persons with specific chronic conditions using an updated system described in detail by Huber et al in 2013 (Appendix 1) [8]. Of the several hierarchies in the ATC codification, the first three levels are generally sufficient for case definition of certain chronic conditions [8, 14, 15]. The PBS have also re-assigned a small number of drugs with ATC codes that align more with approved Australian indications [15, 17]. We allocated these ATC codes to the medications listed and where appropriate, the proxy chronic conditions defined by these treatments as separate fields into the dataset. The number of distinct persons exchanging one or more prescriptions with an ATC assigned chronic condition was counted for the year 2013 to calculate the numerator of prevalence estimates. Measures of the Australian population were obtained from published Australian Bureau of Statistics data and used as the denominator values for prevalence [5]. This use of the ABS data and the population for denominator values was employed as not all Australians will be represented within the PBS data-set. Healthier Australians not requiring dispensed treatments are not identified in PBS data which may lead to prevalence overestimates.

Comparison analysis

Comparisons of chronic disease prevalence estimates derived using Australian PBS data for 2013 were made with published international studies using similar pharmaceutical data and methodology. The study by Huber et al was the principal standard for these comparisons [8]. Potential confounders such

as control of treatments for other conditions, off label use, and treatment threshold considerations as in other studies were not controlled in this analysis [8–10, 15, 17–19].

PBS chronic disease prevalence estimates were also compared with prevalence estimates from Australian community surveys [8–10]. The chronic conditions identified in the NHS, surveying the self-reporting of patients and the Bettering the Evaluation of Care of Health (BEACH) studies, examining the responses of primary physicians were used predominantly for comparisons [4, 5]. The chronic disease domains chosen for comparison analysis from community surveys were those that were considered subjectively the best fit to the PBS ATC coded domains. This on occasion produced recognised broad categories of PBS case definition such as 'Cardiovascular disease and hypertension' [8]. This condition for example was compared with 'Any Cardiovascular' in the Beach study and 'Total diseases of the circulatory system' in the NHS study [4, 5]. However, other conditions such as Tuberculosis and Cancer prevalence estimates were more readily compared between PBS and community surveys. The availability of NHS chronic disease prevalence data by age groupings enabled evaluations with similar PBS chronic disease prevalence estimates at this level of detail. The percentage error between PBS prevalence estimates and NHS studies performed between 2003 and 2014 were compared for several chronic diseases in older adults. This cohort were considered the most likely to suffer these conditions.

Other PBS ATC defined chronic conditions not listed in the Beach and NHS studies were compared with other Australian community survey reports. These studies were identified from a literature search using the index 'chronic condition', 'prevalence', 'Australia' and a date range between 2008–2017 as key terms. This enabled PBS chronic disease prevalence estimates to be compared with similar contemporaneous data from community surveys. Australian PBS derived chronic disease prevalence's were compared for homogeneity with values derived from the Australian community surveys [4, 5, 16].

Statistical analysis

Pearson's correlation was used to compare PBS data derived prevalence estimates with international studies using pharmaceutical data and Australian community survey data over the range of chronic diseases studied. Percentage errors were employed for individual chronic disease comparisons between NHS and other community surveys, and PBS prevalence estimates [20]. Confidence intervals when used are at the 95% level. All other computations were performed using SPSS (Version 23, 2013, IBM SPSS Statistics, New York, NY).

Ethics

Ethics approval for the study was obtained from the La Trobe University ethics committee (S17–198). The construction of this manuscript followed the guidelines recommended in the Reporting of studies conducted using observational routinely-collected health data (RECORD) statement [21].

Results

Data demographics

The total number of individuals in the PBS data available for analysis during 2013 for this study was 1,580,928 (Table 1).

Australian chronic disease prevalence estimates: Comparison with published studies

The Australian prevalence of twenty-two chronic diseases using PBS data for estimates in 2013 are listed in table 2. The conditions with the highest prevalence identified using PBS data in Australia include cardiovascular disorders (18%), pain (15%) and psychological disease (14%). Comparative data reported by Huber et al from a Swiss population is shown in table 2 [8]. The profile of prevalences show a correlation ($r = .89$) over the range of conditions but many chronic diseases demonstrate higher prevalence within the Swiss population as compared to the figures derived from the Australian PBS data (mean percentage error = 40.5%, ($SD = 30.6$)). The conditions included rheumatological disorders at 36% in the Swiss cohort and 10% using Australian PBS data. Similar higher prevalences occurred with cardiovascular disease and hypertension (29% Swiss, 18.1% PBS), pain (28%, 16%), psychological disease (21%, 15%), iron deficiency treatments (4%, 0.3%) and acid disorders (20%, 12%). The prevalence of chronic respiratory conditions in contrast was higher in Australian PBS data (10%) as compared to the Swiss study (7%). The prevalence for dyslipidemia was similar in both populations. The average age of the Australian population (56.6 years ($SD = 21.6$)) was higher than that recorded in the insured Swiss population ($M = 51.4$, $SD = 19.2$), ($p < .001$).

A comparison of PBS derived prevalence rates of chronic conditions with data available from the Swiss and two Italian studies are also presented in table 2 [8–10]. The studies from Italy do not report on all the conditions listed in this PBS derived study, limiting comparison to 17 and 19 chronic conditions [9, 10]. The data from Chini, Pezzotti, Orzella, Borgia, and Guasticchi and Maio et al correlate with PBS data ($r = .83$) [9, 10]. Where comparisons could be made between PBS data and all three European studies PBS prevalence estimates were similar for diabetes, cardiovascular disease, epilepsy, respiratory disease, hyperlipidemia, and thyroid disorders.

Comparison with Australian community data

Comparison of PBS disease prevalence rates with an amalgamation of Australian community surveys including the National Health Survey demonstrated population homogeneity and were not statistically different. These data are shown in table 3. Variations in prevalence estimates were noted with a higher prevalence of rheumatological conditions (15%) and cancer (2%) reported in the National Health Survey. In contrast, the prevalence estimate for dyslipidemia is higher in the PBS data set (12%). Further comparison of the PBS prevalence estimates with Australian NHS survey data shows areas of similarity for diabetes,

Table 1: The number of PBS data study participants by age grouping and gender

Age range	Female	Percentage	Male	Percentage
[0,15)	111740	7.1	121674	7.7
[15,25)	101075	6.4	74164	4.7
[25,35)	117243	7.4	81417	5.1
[35,45)	119330	7.5	94662	6.0
[45,55)	119500	7.6	104570	6.6
[55,65)	115477	7.3	104815	6.6
[65,75)	89394	5.7	84143	5.3
[75,85)	54978	3.5	46424	2.9
[85,95)	24068	1.5	13365	0.8
[95,105)	2286	0.1	603	0.0
Totals	855091	54.1	725837	45.9

Table 2: Comparisons of Australian population chronic disease prevalence estimates with international data using administrative pharmaceutical data and ATC methodology for case definition

Chronic condition	Population prevalence estimate (%) ^f			
	Australian PBS data (2013)	Swiss data Huber et al ^a (2013)	Italian data Maio et al [10]. (2005)	Italian data Chini et al [9]. (2011)
Acid disorders	12.4	19.6	5.1	
Bone disease	1.8	2.4 ^b		1.8
Cancer	0.8	1.4	1.1	1.5
CVS ^c	18.1	29.0	22.9	22.6
Dementia	0.2	1.8	0.02	0.2
Diabetes	4.5	5.4	3.6	4.4
Epilepsy	1.5	2.9	1.2	1.9
Glaucoma	1.4	3.6	1.9	1.7
Gout	1.8	1.8	0.9	0.1
HIV	0.1	0.2	0.01	
Hyperlipidaemia	12.0	12.3	4.0	
IBD ^d	0.4	0.5	0.3	0.4
Iron deficiency	0.3	4.1		
Migraine	0.9	1.2	0.3	0.4
Pain	15.5	28.0	0.1 ^e	
Rheumatological conditions	10.2	36.1	6.1	8.7
Parkinson's disease	0.5	1.2	0.5	0.5
Psychological disease	14.7	21.2	4.4	3.8
Psychoses	1.8	3.2		
COPD/Asthma	9.5	7.4	4.5	4.7
Thyroid disorders	3.1	3.8	3.2	4.4
Tuberculosis	0.01	0.1	0.03	0.04

a, Prevalence in an insured Swiss population of 936612 persons [8].

b, 23.8 in the Huber et al paper and confirmed as an editorial error [8].

c, Cardiovascular system including hypertension

d, IBD = Inflammatory Bowel Disease

e, A low estimate for pain prevalence in this study is noted with caution

f, 95% Confidence intervals not included in prevalence data as all less than 0.01%.

psychological disease, cardiovascular disease, respiratory disease and thyroid disorders with percentage errors of less than 25%.

Comparison of PBS data with chronic disease categories in the BEACH and NHS studies demonstrated correlations of ($r = .94$) and ($r = .93$) respectively. Comparison of PBS data with NHS data also showed good correlation over a range

of ten interval age groupings for the prevalence of diabetes ($r = .99$) (figure 1), cardiovascular disorders ($r = .99$), dyslipidemia ($r = .99$), respiratory disease ($r = .89$) and glaucoma ($r = .85$). However, correlations between NHS and PBS prevalence estimates were much lower for psychological disease ($r = .14$) and migraine ($r = .33$) over similar age groupings.

Table 3: Comparisons of Australian population chronic disease prevalence estimates using Pharmaceutical Benefits Scheme data with Australian community surveys

Australian chronic disease population prevalence estimates (% , CI) ^a					
Chronic condition	PBS Prevalence ^a (2013)	NHS [5]. (2014)	PBS, NHS PE % ^b	Other Australian studies ^c	PE% ^g
Acid disorders	12.4			4.9(4.6, 5.2) ^f 12.0(9.0, 14.0)	47 ^g
Bone disease	1.8			4.6(na)	-61
Cancer	0.8	1.8(1.6, 2.0)	-56	2.8(2.6, 3.0) ^f	-71
CVS ^d	18.1	18.3(17.7, 18.9)	-1	15(14.3, 15.6) ^f	21
Dementia	0.2			1.0(na)	-80
Diabetes	4.5	5.1(4.8, 5.4)	-12	4.6(4.4, 4.9) ^f	-2
Epilepsy	1.5	0.6(0.5, 0.7)	150		
Glaucoma	1.4	0.8(0.6, 1.0)	75	2.8(1.2, 3.8)	-50
Gout	1.8			1.54(1.52, 1.56)	17
HIV	0.1			0.06(na)	67
Hyperlipidaemia	12	7.1(6.7, 7.5)	69	8.2(7.7, 8.6) ^f	46
IBD ^e	0.4			0.33(0.31, 0.38)	21
Iron deficiency	0.3			17(na)	-98
Migraine	0.9	6.2(5.7, 6.7)	-85		
Pain	15.5	12.7(na)	22	19.2(17.4, 21.0) ^f	-19
Rheumatological conditions	10.2	15.3(14.8, 15.8)	-33		
Parkinson's disease	0.5			0.6(0.2, 1.3)	-17
Psychological disease	14.7	17.5(16.8, 18.2)	-16	13.7(13.1, 14.2) ^f	7
Psychoses	1.8			0.35(0.33, 0.36)	414
COPD/asthma	9.5	10.8(10.2, 11.3)	-12	6.8(6.5, 7.1) ^f	40
Thyroid disorders	3.1	3.5(3.2, 3.8)	-11		
Tuberculosis	0.001			0.001(na)	0

a, 95% Confidence intervals not included in PBS prevalence data as all less than 0.01%

b, Percentage error between NHS and PBS estimates

c, References [25–36]. and includes the results of the BEACH survey

d, Cardiovascular system including hypertension

e, IBD = Intestinal Inflammatory Bowel Disease

f, Data derived from the Beach study

g, Percentage error calculated from average of other community survey prevalence estimates

The percentage error between PBS prevalence estimates and NHS results across age groupings (0–100 years) were -8.4%, ($SD = 24.5\%$) for diabetes (figure 1), cardiovascular disease 19.7% ($SD = 23.5\%$), dyslipidemia 37%, ($SD = 31.4\%$), and for respiratory disease 46.5%, ($SD = 38.5\%$). The mean percentage error (MPE) of PBS prevalence estimates when compared to NHS data for glaucoma ($MPE = 41.6\%$, $SD = 41.1\%$) and psychological disease ($MPE = 14.8\%$, $SD = 31.8\%$) demonstrated inconsistency with wide variation over age groupings. Table 4 shows the percentage error of PBS prevalence estimates of selected chronic conditions when compared with NHS survey data for persons aged more than 45 years. These data show wide ranging variations in percentage error between PBS and NHS prevalence estimates for various chronic diseases in older adults likely to suffer these conditions. However, PBS diabetes and all cardiovascular disease prevalence estimates demonstrated the closest comparison with NHS data over this age range.

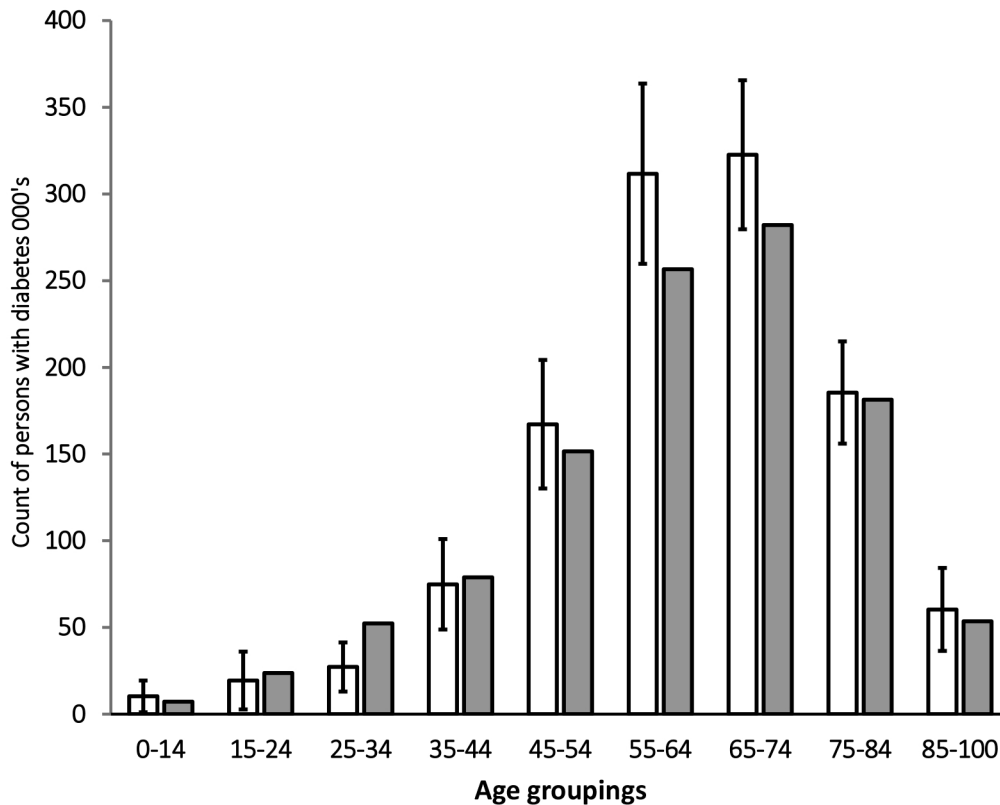
Assessment of PBS prevalence estimates for eight chronic diseases with similar case definitions from the BEACH, NHS studies are shown in Figure 2 [4, 5]. PBS prevalence estimates

are similar to NHS and BEACH estimates for cardiovascular disease, psychological disease, asthma and diabetes [4, 5]. However, PBS estimates show higher prevalence estimates for hyperlipidemia and GORD and underestimates the prevalence of malignant neoplasms when compared to BEACH and NHS studies [4, 5]. The PBS estimate of the prevalence of any arthritis was similar to that reported in the BEACH study but were below that reported in the NHS [4, 5].

Discussion

This study supports the feasibility of Australian administrative pharmaceutical data as a resource for particular chronic disease prevalence estimates from within the Australian population. It builds on the methods introduced by Huber et al in an insured Swiss population, but utilises the advantages of PBS data which represents Australia's universal health care system and well distributed pharmacy access [8, 22]. The extensive coverage of the PBS scheme across Australia may therefore offer a more representative perspective of disease prevalence

Figure 1: Comparison of National Health Survey reported diabetic population sizes (CI) by age group with Australian PBS data for 2013–14 (correlation $r = .99$)



than results obtained from patients who subscribe to insurance schemes as in the European studies presented [8–10, 15, 22]. A comparison of the chronic disease prevalence estimates obtained with this study and Australian community survey data support the use of PBS dispensing data as well as the methodology to define certain chronic diseases by proxy. The data contained within the PBS database also enables repeated calculation of chronic disease prevalence, providing useful comparative trends over time. As a consequence, PBS derived prevalence estimates may in specific circumstances act as a proxy to calculate population prevalence for certain chronic diseases. This study suggests that the conditions dyslipidemia, all diabetes, and asthma may be amenable for population prevalence estimates using this method. Other conditions such as iron deficiency anemia, migraine and cancer correlate poorly with community survey data and would be less useful conditions to examine. Furthermore, certain ATC classified chronic conditions such as 'all cardiovascular disease' and 'all psychological disease' are very broad groupings. The categories correlate well with similar community survey data but offer only rudimentary chronic disease case definition and may have limited use.

This lack of diagnostic detail provided by the ATC methodology used in this study to identify conditions is another major limitation. A chronic disease is identified by a single prescription of the associated ATC class medication dispensed during the year. As a consequence, for individuals with pain this approach can neither distinguish between acute and chronic pain nor define the anatomical site of the pain. Furthermore, the simplicity of the methodology used

when applied across the range of conditions will vary in sensitivity and specificity between conditions. Therefore, the PBS prevalence estimates of certain chronic diseases using the method employed will have less utility and the results need to be interpreted with caution.

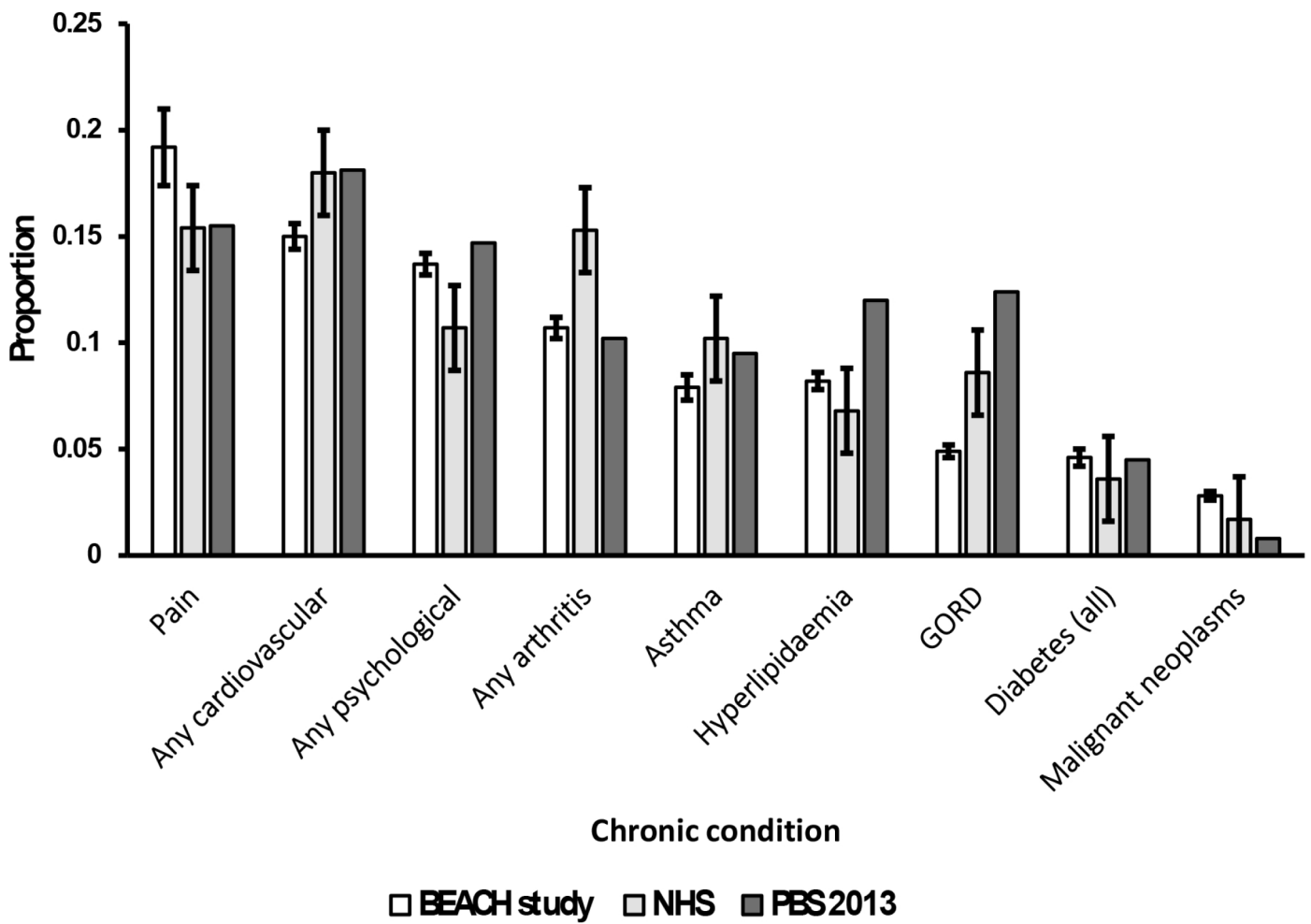
The prevalence of chronic conditions calculated for the Australian population in this study display good correlation with the estimates using pharmaceutical data from European populations [8–10]. The earlier studies by Chini et al and Maio et al from Italy represent an initial use of ATC groupings for case definition of chronic conditions [8–10]. The differences between the prevalence estimates from PBS data and these European studies will therefore reflect both population and methodological differences. A comparison of PBS data with the Huber et al study from Switzerland which used very similar ATC methodology to that employed in this study showed better correlation between the two populations as compared to the Italian studies [8–10]. However, the prevalence's for rheumatological, cardiovascular disease and hypertension, pain, psychological disease and acid disorders are all higher than Australian PBS derived estimates. The Swiss study used approximately 0.9 million insured persons as its denominator cohort which represents 12% of the total population, whereas this study is a ten per cent sample of the full PBS data set representing more than 1.5 million Australians.

These cohort differences likely reflect the societal variations separating the two populations and their susceptibility to chronic diseases. Furthermore, there are methodological differences between the PBS prevalence estimates and those reported in the study by Huber et al [8]. The denominators

Table 4: Comparison of mean percentage error between PBS prevalence estimates and National Health Survey data for persons age >45 years and selected chronic diseases

Chronic condition	Mean percentage error (95%CI)
Glaucoma	50 (-10, 110)
COPD	26 (-2, 54)
Dyslipidaemia	59 (41, 77)
Cardiovascular disease	-4 (-13, 5)
Diabetes	-11 (-16, 6)

Figure 2: Comparison of chronic condition prevalence estimates (mean proportion, 95% CI) from the BEACH, NHS and PBS surveys



used to calculate our estimates are derived from national ABS statistics. A proportion of these individuals will be healthier people not requiring dispensed medications and therefore not represented within the PBS data. This will lower the prevalence estimates generally as compared to denominator values using PBS data only. It is also recognised that the PBS assign slightly different ATC codes as compared to the WHO in several areas. This is to align the PBS data more closely with the approved Australian indications for treatments of certain diseases [17, 18]. These are further methodological differences

but may offer an improved local perspective for case definition of certain conditions found in Australia.

A comparison of PBS and Australian community survey derived prevalence estimates showed similarities with a number of conditions. Correlations between PBS data and NHS data for diabetes, dyslipidemia, cardiovascular disease and obstructive airways diseases over age groupings are particularly strong. However, there are areas of divergence for particular conditions. These discrepancies are likely to represent the differences of chronic disease case definition within each type of study. As an example, acid disorders are categorised in our

study using prescribed medications ATC coded A02, 'drugs for acid related disorders' as a proxy. This definition will include patients with the clinical diagnoses of peptic ulcer disease and oesophageal reflux. The ATC system cannot discriminate between these diagnoses, whereas a respondent in a questionnaire study might be able to separate the conditions. These responses are however subjective, which may lead to a lack of specificity for the prevalence estimates obtained using survey data. Conversely, PBS prescriptions will not identify persons who acquire treatments privately, off prescription or using alternative insurance schemes. This may lead to an overall underestimate of the prevalence of dyspeptic disorders and may explain the disparity between PBS data and community survey analyses.

The percentage error of PBS chronic disease estimates when compared to NHS estimates over a range of conditions demonstrated variations in both polarity and magnitude. This variation between NHS and PBS prevalence estimates was also noted when age groupings within specific chronic diseases such as diabetes are compared. The systematic errors between PBS prevalence estimates and NHS values, was observed to be inconsistent and random over an array of ages and as a consequence, represents a major inconsistency. These variances between NHS and PBS prevalence estimates across the range of age groupings may reflect different propensities to use certain types of medicines among different patients. Older persons for example may be less likely to receive anti-dementia medicines than younger people leading to a differential identification of prevalence among different population groups. This might explain the lower prevalence of dementia using PBS data (0.2%) in contrast to a prevalence of 1.3% in other Australian community surveys. Similarly, diabetes may be treated more conservatively without medication in older persons as compared to younger people. In addition, a PBS estimate for the prevalence of cancer is likely to be an underestimate as certain cancers may be treated by surgery alone and not have a distinguishing drug signal within PBS data. Both PBS data and community survey therefore have challenges in their ability to accurately measure the true prevalence of chronic disease within populations.

The prevalence of diabetes reported in this study represent patients with overt disease and a level of severity that requires medication. Patients managed by non-pharmacological interventions such as diet and exercise will not be included in the prevalence calculation. This will introduce an overall systematic bias to our estimates of diabetes prevalence. We consider that our overall PBS derived measures will be an underestimate of the actual prevalence of "real diabetes". This hypothesis is supported by the strong correlation between NHS and PBS data for estimates of diabetes prevalence and an average percentage error of 8%, representing a small overall underestimation. This association has been confirmed in other Australian pharmaceutical data and supports the use of PBS data for diabetes prevalence estimates [15, 23]. The prevalence estimates of people with drug-treated diabetes as defined in this study would also have application for studies of disease burden and health economic enquiry as these patients are actively consuming medications as part of their care.

The comparison of PBS prevalence data with NHS and BEACH surveys offers the potential for triangulation of

chronic disease prevalence estimates [24]. It is likely that none of these estimates represent a true population prevalence estimate for the conditions shown but values may represent the different perspectives of each survey. NHS self-reported data representing the conditions from which individuals consider they are suffering and the BEACH data representing the view point of primary physicians and the conditions they consider they are treating. PBS data may represent an intermediary position between these two perspectives and denote the outcome of medical consultations. The prescription output recorded in the administrative pharmaceutical data representing a condition that both doctor and patient consider of sufficient severity to require drug treatment.

Conclusions

The estimates of chronic disease prevalence using Australian pharmaceutical data as presented in this study offer the advantages of figures derived from an extremely large and representative dataset. The techniques employed in our study using proxy definitions of chronic conditions provide an alternative perspective of population health as compared to community interview survey data. These include the contextual domains of societal disease burden related to disease severity that is defined by the need for prescribed treatments, and an economic view point from the cost of medications used. The results from PBS data can be interpreted either alone or in connection by triangulation with other methods of enquiry. The utility of Australian PBS data as an instrument to estimate chronic disease prevalence comes from improved precision and objectivity and a potentially wider demographic and geographic population coverage when compared with questionnaire studies. PBS data are also continuously collected, enabling repeated surveillance of chronic disease over time that may act as a reference to calibrate other forms of prevalence inquiry and identify disease trends [18, 24]. Accordingly, despite the discussed limitations we consider pharmaceutical administrative (PBS) data is a useful resource for a perspective on the prevalence of certain chronic diseases within Australia.

Acknowledgements

We are grateful for the authorisation of the study by the La Trobe research ethics committee. (Approval number S17-177). Supplementary Appendix 1 is reproduced from the article by Huber et al [8]. This is an open access article distributed under the terms of the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium.

Statements of conflicts of interest

The authors have no conflicts of interest to declare.

References

1. Bauer UE, Briss PA, Goodman RA, Bowman BA. Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. *The Lancet*. 2014 Jul 5;384(9937):45–52. [https://doi.org/10.1016/S0140-6736\(14\)60648-6](https://doi.org/10.1016/S0140-6736(14)60648-6)
2. Gunn JM, Ayton DR, Densley K, Pallant JF, Chondros P, Herrman HE, Dowrick CF. The association between chronic illness, multimorbidity and depressive symptoms in an Australian primary care cohort. *Social psychiatry and psychiatric epidemiology*. 2012 Feb 1;47(2):175–84. <https://doi.org/10.1007/s00127-010-0330-z>
3. Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Archives of internal medicine*. 2002 Nov 11;162(20):2269–76. <https://doi.org/10.1001/archinte.162.20.2269>
4. Harrison C, Britt H, Miller G, Henderson J. Prevalence of chronic conditions in Australia. *PLoS one*. 2013 Jul 23;8(7):e67494. <https://doi.org/10.1371/journal.pone.0067494>
5. Australian Bureau of Statistics. National Health Survey: First Results, 2014–15 Available from <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0-5.0012014-15?OpenDocument> last accessed 07/09/2020
6. Moore Jr JR. Accuracy of a health interview survey in measuring chronic illness prevalence. *Health services research*. 1975;10(2):162.
7. West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. *American journal of epidemiology*. 1995 Nov 15;142(10):1103–12. <https://doi.org/10.1093/oxfordjournals.aje.a117563>
8. Huber CA, Szucs TD, Rapold R, Reich O. Identifying patients with chronic conditions using pharmacy data in Switzerland: an updated mapping approach to the classification of medications. *BMC public health*. 2013 Dec;13(1):1030. <https://doi.org/10.1186/1471-2458-13-1030>
9. Chini F, Pezzotti P, Orzella L, Borgia P, Guasticchi G. Can we use the pharmacy data to estimate the prevalence of chronic conditions? a comparison of multiple data sources. *BMC Public Health*. 2011 Dec;11(1):688. <https://doi.org/10.1186/1471-2458-11-688>
10. Maio V, Yuen E, Rabinowitz C, Louis D, Jimbo M, Donatini A, Mall S, Taroni F. Using pharmacy data to identify those with chronic conditions in Emilia Romagna, Italy. *Journal of health services research & policy*. 2005 Oct 1;10(4):232–8. <https://doi.org/10.1258/135581905774414259>
11. Herings RM, Bakker A, Stricker BH, Nap G. Pharmaco-morbidity linkage: a feasibility study comparing morbidity in two pharmacy based exposure cohorts. *Journal of Epidemiology & Community Health*. 1992 Apr 1;46(2):136–40. <https://doi.org/10.1136/jech.46.2.136>
12. Henry DA, Hill SR, Harris A. Drug prices and value for money: the Australian Pharmaceutical Benefits Scheme. *Jama*. 2005 Nov 23;294(20):2630–2. <https://doi.org/10.1001/jama.294.20.2630>
13. Lamers LM, Vliet RC. Health-based risk adjustment Improving the pharmacy-based cost group. *The European Journal of Health Economics*. 2003 Jun 1;4(2):107–14. <https://doi.org/10.1007/s10198-002-0159-9>
14. World Health Organization. WHO Collaborating Centre for Drug Statistics Methodology: ATC classification index with DDDs and Guidelines for ATC classification and DDD assignment. Oslo, Norway: Norwegian Institute of Public Health. 2006.
15. Pratt NL, Kerr M, Barratt JD, Kemp-Casey A, Ellett LM, Ramsay E, Roughead EE. The validity of the Rx-Risk Comorbidity Index using medicines mapped to the Anatomical Therapeutic Chemical (ATC) classification system. *BMJ open*. 2018 Apr 1;8(4):e021122. <https://doi.org/10.1136/bmjopen-2017-021122>
16. Knox SA, Harrison CM, Britt HC, Henderson JV. Estimating prevalence of common chronic morbidities in Australia. *Medical Journal of Australia*. 2008 Jul;189(2):66–70. <https://doi.org/10.5694/j.1326-5377.2008.tb01918.x>
17. Paige E, Kemp-Casey A, Korda R, Banks E. Using Australian Pharmaceutical Benefits Scheme data for pharmacoepidemiological research: challenges and approaches. <https://doi.org/10.17061/phrp2541546>
18. Mellish L, Karanges EA, Litchfield MJ, Schaffer AL, Blanch B, Daniels BJ, Segrave A, Pearson SA. The Australian Pharmaceutical Benefits Scheme data collection: a practical guide for researchers. *BMC research notes*. 2015 Dec;8(1):634. <https://doi.org/10.1186/s13104-015-1616-8>
19. Brookhart MA, Stürmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. *Medical care*. 2010 Jun;48(6 0):S114. <https://doi.org/10.1097/MLR.0b013e3181d8be3>
20. Furey E. Percentage Error Calculator. [online]. Calculator soup. 2020. Available at: <https://www.calculatorsoup.com>. Last accessed 07/09/2020
21. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS medicine*. 2015 Oct 6;12(10):e1001885. <https://doi.org/10.1371/journal.pmed.1001885>

22. Benrimoj SI, Frommer MS. Community pharmacy in Australia. *Australian Health Review*. 2004;28(2):238–46. <https://doi.org/10.1071/AH040238>
23. Lujic S, Simpson JM, Zwar N, Hosseinzadeh H, Jorm L. Multimorbidity in Australia: Comparing estimates derived using administrative data sources and survey data. *PloS one*. 2017 Aug 29;12(8):e0183817. <https://doi.org/10.1371/journal.pone.0183817>
24. Morse JM. Approaches to qualitative-quantitative methodological triangulation. *Nursing research*. 1991 Mar 1;40(2):120–3. <https://doi.org/10.1097/00006199-199103000-00014>
25. El-Serag HB, Sweet S, Winchester CC, Dent J. Update on the epidemiology of gastro-oesophageal reflux disease: a systematic review. *Gut*. 2014 Jun 1;63(6):871–80. <https://doi.org/10.1136/gutjnl-2012-304269>
26. Australian Institute of Health and Welfare 2014. Estimating the prevalence of osteoporosis. Cat. no. PHE 178. Canberra: AIHW.
27. Mitchell P, Smith W, Chey T, Healey PR. Open-angle glaucoma and diabetes: the Blue Mountains eye study, Australia. *Ophthalmology*. 1997 Apr 1;104(4):712–8. [https://doi.org/10.1016/S0161-6420\(97\)30247-4](https://doi.org/10.1016/S0161-6420(97)30247-4)
28. Robinson PC, Taylor WJ, Dalbeth N. An observational study of gout prevalence and quality of care in a national Australian general practice population. *The Journal of rheumatology*. 2015 Sep 1;42(9):1702–7. <https://doi.org/10.3899/jrheum.150310>
29. Falster K, Gelgor L, Shaik A, Zablotska I, Prestage G, Grierson J, Thorpe R, Pitts M, Anderson J, Chuah J, Mulhall B. Trends in antiretroviral treatment use and treatment response in three Australian states in the first decade of combination antiretroviral treatment. *Sexual health*. 2008 Jun 23;5(2):141–54. <https://doi.org/10.1071/SH07082>
30. Wilson J, Hair C, Knight R, Catto-Smith A, Bell S, Kamm M, Desmond P, McNeil J, Connell W. High incidence of inflammatory bowel disease in Australia: a prospective population-based Australian incidence study. *Inflammatory bowel diseases*. 2010 Apr 12;16(9):1550–6. <https://doi.org/10.1002/ibd.21209>
31. Australian Bureau of Statistics 4364.0.55.008 – Australian Health Survey: Usual Nutrients Intakes, 2011–12. <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4364.0.55.008~2011-12~Main%20Features~Iron~402>. Last accessed 04/24/2018
32. Vickland V, McDonnell G, Werner J, Draper B, Low LF, Brodaty H. A computer model of dementia prevalence in Australia: foreseeing outcomes of delaying dementia onset, slowing disease progression, and eradicating dementia types. *Dementia and Geriatric cognitive disorders*. 2010;29(2):123–30. <https://doi.org/10.1159/000272436>
33. Pringsheim T, Jette N, Frolkis A, Steeves TD. The prevalence of Parkinson's disease: A systematic review and meta-analysis. *Movement disorders*. 2014 Nov;29(13):1583–90. <https://doi.org/10.1002/mds.25945>
34. Morgan VA, McGrath JJ, Jablensky A, Badcock JC, Waterreus A, Bush R, Carr V, Castle D, Cohen M, Galletly C, Harvey C. Psychosis prevalence and physical, metabolic and cognitive co-morbidity: data from the second Australian national survey of psychosis. *Psychological medicine*. 2014 Jul;44(10):2163–76. <https://doi.org/10.1017/S0033291713002973>
35. Toms C, Stapledon R, Waring J, Douglas P. Tuberculosis notifications in Australia, 2012 and 2013. *Communicable diseases intelligence quarterly report*. 2015 Jun;39(2):E217–35.
36. Miller A, Sanderson K, Bruno R, Breslin M, Neil AL. The prevalence of pain and analgesia use in the Australian population: Findings from the 2011 to 2012 Australian National Health Survey. *Pharmacoepidemiology and Drug Safety*. 2017 Nov;26(11):1403–10. <https://doi.org/10.1002/pds.4301>
37. Henderson JV, Harrison CM, Britt HC, Bayram CF, Miller GC. Prevalence, causes, severity, impact, and management of chronic pain in Australian general practice patients. *Pain Medicine*. 2013 Sep 1;14(9):1346–61. <https://doi.org/10.1111/pme.12195>

Abbreviations

ATC:	Anatomic Therapeutic Chemical Codes
BEACH:	Bettering the Evaluation of Care of Health study
COPD:	Chronic Obstructive Pulmonary Disease
CVS:	Cardiovascular System
GORD:	Gastro Oesophageal Reflux Disease
HIV:	Human Immunodeficiency Disease
IBD:	Intestinal Inflammatory Bowel Disease
MPE:	Mean Percentage Error
NHS:	National Health Survey
PBS:	Pharmaceutical Benefits Scheme
PE:	Percentage Error
SPSS:	Statistical Package for the Social Sciences
WHO:	World Health Organization

Supplementary Appendix 1: Chronic diseases and updated assigned ATC-codes and medication classes (Huber et al [8])

Chronic condition	ATC classification	Medication class
Acid related disorders	A02	Antacids Drugs for peptic ulcer and gastroesophageal reflux disease (GERD) Other drugs for acid related disorders
Bone diseases (osteoporosis)	M05	Drugs for treatment of bone diseases
Cancer	L01	Antineoplastic agents
Cardiovascular diseases (incl. hypertension)	B01AA, B01AC, C01, C04A, C02, C07, C08, C09	Cardiac agents (excl. ACE inhibitors) Anti-hypertensives Peripheral vasodilators Beta blocking agents Calcium channel blockers Agents acting on the renin-angiotensin system Vitamin K antagonists Platelet aggregation inhibitors (excl. heparin)
Dementia	N06D	Anti-dementia drugs
Diabetes mellitus	A10A, A10B, A10X	Insulins and analogues Blood glucose lowering drugs (excl. insulins) Other drugs used in diabetes
Epilepsy	N03	Anti-epileptics
Glaucoma	S01E	Anti-glaucoma preparations and miotics
Gout, Hyperuricemia	M04	Anti-gout preparations
HIV	J05AE, J05AG, J05AR	Protease inhibitors Non-nucleoside reverse transcriptase inhibitors Antivirals for treatment of HIV infections, combinations
Hyperlipidaemia	C10	Lipid modifying agents
Intestinal inflammatory diseases	A07EA, A07EC	Corticosteroids acting locally Amino salicylic acid and similar agents
Iron deficiency anaemia	B03AA, B03AB, B03AC	Iron bivalent, oral preparations Iron trivalent, oral preparations Iron trivalent, parenteral preparations
Migraines	N02C	Antimigraine preparations
Pain	N02A, N02B	Opioids Other analgesics
Parkinson's disease	N04	Anti-Parkinson drugs
Psychological disorders (sleep disorder, depression)	N05B, N05C, N06A	Anxiolytics Hypnotics and sedatives Antidepressants
Psychoses	N05A	Antipsychotics
Respiratory illness (asthma, COPD)	R03	Drugs for obstructive airway diseases
Rheumatologic conditions	M01, M02 L04AA, L04AB	Anti-inflammatory and anti-rheumatic products Topical products for joint and muscular pain Selective immunosuppressants TNF-alpha inhibitors
Thyroid disorders	H03	Drugs for thyroid therapy
Tuberculosis	J04A	Drugs for treatment of tuberculosis