

International Journal of Population Data Science

Journal Website: www.ijpds.org



Estimates of drug treated diabetes incidence and prevalence using Australian administrative pharmaceutical data

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

Submission History

Submitted:	14/09/2020
Accepted:	04/03/2021
Published:	10/05/2021

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

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

Abstract

Introduction

The incidence and prevalence of diabetes within a population are important public health metrics. Pharmaceutical administrative data may offer a resource that can contribute to quantifying these measures using the recorded signals derived from the drugs used to treat people with diabetes.

Objective

To estimate the longitudinal incidence and prevalence of drug treated (DT) diabetes in Australia utilising an Australian Pharmaceutical Benefits Scheme (PBS) dataset and compare estimates with community survey data for all diabetes reported in the Australian National Health Survey (NHS).

Methods

Persons with DT diabetes were identified within the PBS dataset using assigned Anatomic Therapeutic Chemical codes for 'Drugs used in diabetes'. Prevalent persons with DT diabetes were determined by a single annual treatment, and incident cases from the earliest treatment with diabetes medications. Counts were aggregated by age group and utilised Australian national census data as a denominator to calculate diabetes disease frequencies for the period 2004–14. Comparison of PBS prevalence data was made with NHS surveys over equivalent years.

Results

The age adjusted incidence of DT diabetes was 3.4/1000 in 2006 and increased to 3.8/1000 in 2011 and 5.1/1000 in 2014. Age adjusted prevalence of DT diabetes in Australia also rose from 26.7/1000 in 2006 to 32.1/1000 in 2011 and 42.1/1000 in 2014. DT diabetes prevalence estimates correlated with NHS estimates of self-reported diabetes prevalence across age groups and in 2014 was $r = 0.987$. However, PBS estimates of DT diabetes prevalence generally underestimated NHS values of self-reported diabetes in older age groups with mean percentage differences of -22% to -3% . In contrast, PBS data captured more younger persons with diabetes in comparison to NHS data. These differences were then used to adjust DT diabetes incidence rates to provide age specific estimates that could potentially reflect diabetes incidence estimates acquired by community survey.

Conclusions

PBS data representing dispensed medications prescribed to persons with diabetes offers a perspective for the assessment of diabetes incidence and prevalence. PBS derived DT diabetes prevalence estimates correlate well with community survey estimates of self-reported diabetes, but underestimate NHS data in older age groups. Calibrated DT incidence estimates may potentially reflect community survey derived diabetes incidence estimates and may offer a method for longitudinal monitoring.

Keywords

diabetes; Australia; incidence; pharmaceutical data; administrative; prevalence

*Corresponding Author:

Email Address: s.purkiss@latrobe.edu.au (Shaun Purkiss)

Introduction

The rising prevalence of diabetes is a health problem of worldwide significance [1, 2]. Diabetes contributes significantly to the burden of disease for many communities bringing poorer health to individuals and financial costs to societies [2]. To adequately quantify the community, individual, and economic impact of diabetes, accurate and reliable measures of diabetes incidence and prevalence are required [2]. If possible, these estimates of disease frequency should be obtained with consistent methods so that the information concerning the longitudinal profile of diabetes is reliably defined [1–3].

Concordant with world trends, the number of people with diabetes in Australia is increasing [1]. At present, there are approximately 1 million persons with diabetes, and 100,000 persons annually developing new onset diabetes in Australia [3]. Poor diets and sedentary lifestyles are well recognised behavioural associations of the increasing risk of developing diabetes within Australian communities [4]. However, demographic changes such as increasing life expectancy and ethnic diversity are other factors also likely to influence disease frequency [3]. As these population characteristics are changing over time in a rapidly growing and evolving Australian society, the incidence and prevalence of diabetes are also likely to become changeable and unpredictable [4]. Diabetes prevalence rates in Australia have been estimated using a variety of techniques including community surveys such as the National Health Survey (NHS) and the Australian Diabetes and Obesity Lifestyle study [3, 5]. These are very comprehensive instruments which provide a detailed perspective of diabetic population health and can be cross referenced and validated with other Australian data sources such as the National Diabetes Services Scheme [6]. An assessment of the strengths and weaknesses of the various methods that can be used to estimate diabetes prevalence has been communicated [3]. However, the approaches that may be utilised to assess diabetes incidence are less well represented.

Community surveys such as the Australian National Health Surveys (NHS) have the acknowledged potential for bias, relating to factors such as self-reporting, participation, and geographic and demographic coverage [5]. As a consequence, the results may not fully represent the profile of diabetes prevalence in Australia. In addition, survey methods will have a lag time from the collection of data to the processing and publication of results [5]. NHS estimates of prevalence for chronic conditions such as diabetes report on different populations for each survey. The sample size is used as a denominator and national estimates obtained by extrapolation. In contrast, incidence estimates have logistical difficulties as sampling of defined populations over time are required. Given that the incidence of diabetes in Australia is about 0.1% per annum, a large sample population is required to accurately assess disease frequency and change [6].

There is scope for new methods to estimate both diabetes incidence and prevalence. These approaches should be of sufficient flexibility to keep pace with societal demographic changes and would be a useful adjunct to assess the population profile of diabetes [3]. Collectively, they could help inform health policy and predict future health requirements [2, 3].

Administrative pharmaceutical data is a potential data source [7]. The information from these and similar repositories, when combined with appropriate methodology, has demonstrated utility for estimates of chronic disease frequency [7, 8]. The drug dispensed in a prescription can act as a proxy diagnosis and be used for case definition of persons with diabetes [7, 8]. In Australia, these data have been collected and retained by the Pharmaceutical Benefits Scheme (PBS), and a ten per-cent sample of the complete data base was made available for public use [9]. Similar data sources from Australia and Europe have already been successfully applied to quantify the prevalence of a range of chronic diseases including diabetes [7–11]. Administrative pharmaceutical data representing the quantity and diversity of prescriptions exchanged at pharmacies can provide a viewpoint of population health [7, 8, 11]. PBS data as collected in Australia also represent the extensive geographic coverage provided by pharmacies in remote and metropolitan areas, and details services provided to patients utilising a universal system of health care available to most Australians [9–13].

In this study, we consider the utility of Australian PBS data as a source to estimate drug treatment (DT) defined diabetes incidence and prevalence. We compare these estimates of disease frequency with community survey data of self-reported responses for all forms of diabetes for the Australian population. In addition, we evaluate the data's ability to provide longitudinal information about incidence and prevalence of DT diabetes for the period 2006–2014.

Methods

Study design

This was a retrospective cohort study using a sample of the PBS administrative database that provides details of the prescription exchanges occurring at Australian pharmacies [9].

Dataset and setting

The PBS dataset employed was published by the Australian Department of Health under a Creative Commons licence. This dataset has been used in previous communications and its strengths and weakness evaluated [9–11]. The data is a 10% sample of all persons that utilised the PBS for their prescriptions and includes complete pharmaceutical data for a random 2.2 million Australians from all states and territories covering the period 2003–2014. The information represents a perspective of Australia's universal health care system which is available to all residents holding a Medicare card [13]. The system allows people access to a range of medical services and prescriptions at lower cost [13].

Persons were de-identified within the data set as part of the methodology employed by the Australian Department of Health to ensure privacy. A unique identifier assigned to individuals within the database enabled continuity of referencing over time and facilitated cross linkage with other data bases [9]. For all individuals listed in the database a comprehensive longitudinal profile of PBS activity is recorded. The ten percent data set were selected randomly by government statisticians from the entire current and

historical PBS data prior to publication. The data includes the administrative details of the date of supply for medications prescribed, the drug's PBS code, its form and strength. We applied World Health Organization Anatomic Therapeutic Codes (ATC), a system for the classification of medications, to the pharmaceutical items listed in the data [7, 8, 11, 14]. Approval for the study was obtained from La Trobe University ethics committee (Approval number S17-177).

Participants

Participants with drug treated (DT) diabetes were characterised as those who were prescribed and dispensed a drug categorised by an ATC code starting with A10 [11]. This case definition, enabled the capture of persons receiving injectable and oral treatments for diabetes [7]. It is recognised that certain drugs used to treat diabetes can be also be used for other conditions and this is accepted as a confounding variable in our analyses [15, 16]. The available demographic data retrieved for persons identified with diabetes in this way included age, gender and Australian state of origin. The prevalence of persons with DT diabetes was calculated from the number of individuals who had received at least one prescription containing an A10 ATC code within each calendar year over the period of the study. Incident cases were similarly identified by their earliest recorded A10 ATC coded prescription, but only classified with new-onset diabetes if they had not received any diabetes medication in the previous two years. Accordingly, DT diabetes incidence estimates could only be calculated for persons after 2005. These annual distinct and earliest calendar counts of A10 prescriptions were used as numerator values for estimates of diabetes prevalence and incidence. The denominator values used to calculate incidence and prevalence utilised the Australian national populations over these years and were obtained from Australian Bureau of Statistics (ABS) census data and modified to 10% to correspond with PBS data [17]. Age adjusted standardisation of longitudinal estimates of disease frequency used the direct method of standardisation and the 2001 ABS census data as the reference population. Comparison data on Australian diabetes prevalence were obtained from the NHS [6].

Statistical analyses

Age groups of 10 years from the ages 15–85 were used to aggregate measures of disease frequency, and for comparison

with similar NHS community survey data. We used the published margins of errors of NHS diabetes prevalence estimates to calculate 95% confidence intervals for comparison with PBS data [5]. Assessment of the accuracy of PBS estimates of DT diabetes prevalence over the period 2003–2014 used the mean percentage error for comparison with published NHS data, considered as the reference population [18]. Longitudinal estimates of diabetes incidence were calculated using adjusted DT diabetes incidence estimates. The factors used for modification of DT diabetes incidence were derived from the mean percentage error between PBS DT diabetes and NHS diabetes prevalence estimates. These values were calculated by age groups and used the most contemporaneous NHS survey data relative to the year of the PBS incidence estimate.

Graphical representation and statistical analysis were performed using a combination of Excel, and SPSS (IBM SPSS Statistics, New York, NY, USA Version 23 2013) software packages. Independent samples t-test was used to compare mean ages. Chi-squared analysis was used for the assessment of proportions and Pearson's correlation utilised as a measure of association. Statistical significance was set at $p < .05$, and 95% confidence intervals are presented. This manuscript followed the guidelines recommended in the REporting of studies Conducted using Observational Routinely collected health Data (RECORD) statement [19].

Results

A total of 158,798 persons with DT defined diabetes were identified within the PBS database for the period 2003–2014, using the exchange of a prescription with ATC coded A10 medications as a case definition. The average ages (M , SD) of prevalent males and females with DT diabetes in 2014 were 59.0 (15.1) and 55.6 (18.6). For persons with incident DT diabetes in 2014, the average ages were 56.1 (15.3) for males and 47.5 (18.5) for females (Table 1).

The mean age (SD) of prevalent persons with diabetes gradually rose from 62.5 (16.0) to 63.5 (15.9) years ($p < .001$). The mean age of incident persons with diabetes in contrast reduced modestly from 57.8 (16.9) to 57.0 (17.2) over the same period ($p < .001$).

Table 1: Person demographics of incident and prevalent Australian Drug Treated diabetes for the period 2006–2014.

	Incident persons with diabetes mean age (Years, SD)					Prevalent persons with diabetes mean age (Years, SD)				
	2006	2008	2010	2012	2014	2006	2008	2010	2012	2014
Female	55.3 (18.7)	55.6 (18.4)	55.1 (18.9)	49.9 (16.1)	47.5 (18.5)	63.0 (17.0)	63.3 (16.7)	63.4 (16.6)	60.9 (17.1)	60.2 (17.8)
Male	59.4 (15.6)	59.9 (16.1)	58.9 (16.1)	56.1 (12.9)	56.1 (14.9)	62.5 (14.5)	63.2 (15.8)	63.5 (15.5)	62.3 (16.7)	62.6 (14.4)

Table 2: Age and sex specific, and age adjusted longitudinal estimates of incidence and prevalence (n/1000), and percentage change for persons with Drug Treated diabetes in Australia from 2006 to 2011

Age specific incidence						Change (%)	Age specific prevalence					Change (%)
Female	2006	2008	2010	<u>2012^a</u>	2014	2006–11	2006	2008	2010	<u>2012^a</u>	2014	2006–11
0–14	0.3	0.4	0.3	<u>0.4</u>	0.5	21	1.4	1.6	1.6	1.7	1.7	14
15–24	1.2	1.2	1.3	<u>3.6</u>	3.5	25	5.1	5.4	5.6	8.3	10.2	17
25–34	2.6	2.5	2.5	<u>8.6</u>	8.0	6	7.3	7.5	8.1	15.8	19.9	21
35–44	2.9	2.9	3.3	<u>9.5</u>	6.8	11	11.5	12.3	13.8	22.4	26.9	25
45–54	3.5	3.6	4.1	<u>13.2</u>	6.8	21	22.6	23.5	25.7	39.2	42.7	19
55–64	7.4	7.0	7.7	<u>20.3</u>	8.7	7	52.4	54.8	57.8	77.7	81.6	15
65–74	10.2	10.7	10.0	<u>12.6</u>	8.5	10	102.4	109.1	112.2	120.5	122.1	13
75–84	8.6	8.8	8.7	<u>7.9</u>	7.9	5	108.4	119.8	131.2	142.3	147.9	25
85–100	5.5	6.2	5.7	<u>5.9</u>	4.8	–5	82	90.3	95.1	103	107.6	20
Male	2006	2008	2010	<u>2012^a</u>	2014	2006–11	2006	2008	2010	<u>2012^a</u>	2014	2006–11
0–14	0.9	0.8	0.8	<u>1.1</u>	0.9	6	1.5	1.5	1.5	1.6	1.5	1
15–24	0.5	0.4	0.6	<u>0.8</u>	0.7	–4	4.2	4.5	4.7	5.0	5.2	16
25–34	0.6	0.8	1.0	<u>2.2</u>	1.5	47	5.0	5.5	6.1	7.9	8.4	24
35–44	1.6	2.0	2.3	<u>7.8</u>	4.3	50	10.6	11.1	12.5	19.8	22.1	24
45–54	4.5	4.9	5.2	<u>19.9</u>	7.9	16	26.9	28.2	31.1	51.7	55.2	19
55–64	8.7	9.0	9.4	<u>32.8</u>	11.1	14	65.7	68.1	71.9	106.0	109.8	12
65–74	14.3	15.1	13.9	<u>22.3</u>	11.6	2	132.3	143.5	147.8	164.3	169.9	15
75–84	10.1	12.7	11.9	<u>12.6</u>	9.6	16	128.7	154.4	176.2	193.5	197.9	44
85–100	7.1	7.3	6.7	<u>7.9</u>	6.4	12	78.5	84.3	94.8	123.2	144.7	35
Persons	2006	2008	2010	<u>2012^a</u>	2014	2006–11	2006	2008	2010	<u>2012^a</u>	2014	2006–11
0–14	0.4	0.4	0.4	<u>0.5</u>	0.5	14	1.5	1.5	1.5	1.6	1.6	7
15–24	0.8	0.8	0.9	<u>2.2</u>	2.0	17	4.6	4.9	5.1	6.6	7.6	16
25–34	1.7	1.7	1.9	<u>5.8</u>	5.2	13	6.5	6.9	7.5	12.6	15.2	22
35–44	2.3	2.5	2.8	<u>8.6</u>	5.6	25	11.0	11.7	13.2	21.1	24.5	24
45–54	4.0	4.3	4.6	<u>16.5</u>	7.3	18	24.7	25.8	28.4	45.4	48.9	19
55–64	8.1	8.0	8.6	<u>26.5</u>	9.9	11	59.1	61.5	64.8	91.7	95.5	13
65–74	12.2	12.9	11.9	<u>17.4</u>	10.0	6	117.0	126.0	129.8	142.2	145.7	14
75–84	9.3	10.5	10.2	<u>10.0</u>	8.7	11	117.3	135.1	151.3	165.4	170.8	35
85–100	6.0	6.2	6.0	<u>6.5</u>	5.2	–3	78.8	86.3	93.4	107.5	117.9	25
Age adjusted incidence						Change (%)	Age adjusted prevalence					Change (%)
	2006	2008	2010	<u>2012^a</u>	2014	2006–11	2006	2008	2010	<u>2012^a</u>	2014	2006–11
	3.42	3.57	3.69	<u>9.05</u>	5.06	12	26.73	28.79	30.86	39.59	42.08	20

^a The abrupt temporary rise in age adjusted and specific incidences underlined in 2012 is likely associated with changes in Australian PBS legislation of that year, resulting in the capture of more persons with DT diabetes that were prevalent in preceding years.

Incidence

A sudden increase in the monthly counts of persons exchanging prescriptions for diabetes medications occurred during 2012. This finding, observed and documented in other reports, was considered a confounding factor and likely related to the introduction of the *National Health Amendment (PBS) 2010 Act* during 2012 [9]. This had consequential effects upon the longitudinal trends in prevalence and incidence. Accordingly, longitudinal analyses were presented for the period 2005–11 before the introduction of the act.

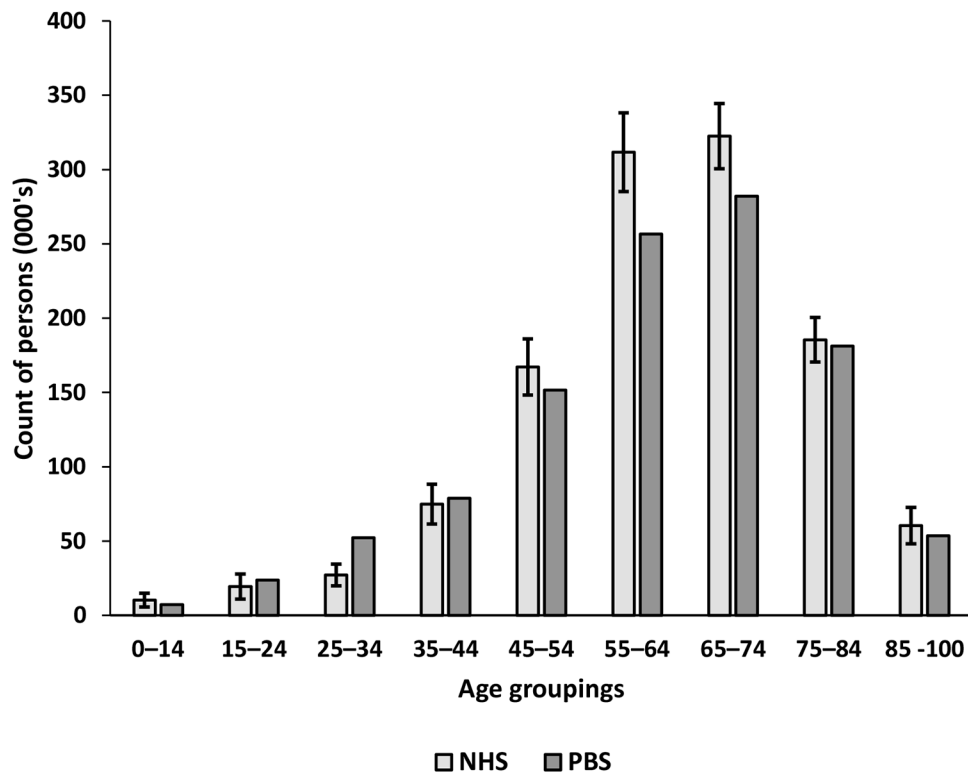
Table 2 illustrates the age specific incidence estimates of persons with DT diabetes by gender and age group for the period 2006–2014. Age adjusted incidences for persons with DT diabetes over the same time period are also included. The

overall age adjusted incidence of DT diabetes was 3.4/1000 in 2006 and rose to 3.8/1000 in 2011 and 5.1/1000 in 2014 (χ^2 , $p < .001$).

The highest values for age-specific incidences in 2014 occurred in the age-groups 55–64 and 65–74 with 9.9/100 and 10.0/1000 persons with new-onset DT diabetes respectively. These age-groups cohorts were observed to have higher incidences for both genders over the period of the study. Towards the end of the study period there was a tendency for these older groups to have a slightly reduced incidence of DT diabetes compared to previous years.

The greatest changes of incidence for persons with DT diabetes occurred in the younger age groups 0–14, 15–24, and 35–44 with average annual percentage rises of 3.1%, 3.4% and 4.7% respectively for the period 2006–12. The greatest

Figure 1: Comparison of NHS counts of persons with self-reported diabetes (95% CI) and PBS derived estimates of persons with DT diabetes by age-group for the year 2014 (correlation $r = 0.99$)



changes occurred in males aged 35–44 (50%) and 25–34 (47%) and females aged 0–14 (21%) and 15–24 (25%) in the period 2006–12.

Prevalence

The age adjusted prevalence of DT diabetes was 26.7/1000 in 2006 and 42.1/1000 in 2014 (χ^2 , $p > 0.001$). The age-specific prevalence's of all age-groups for the period 2006–2014 are shown in Table 2. For all age-groups there is an increasing prevalence over this period. The highest prevalence for persons with DT diabetes occurs in the age groups over 55 in both genders. Elderly males however, have in general the highest prevalence in all age groups above age 45 years. The greatest change of DT diabetic prevalence for the period 2006–11 occurred in males aged 75–84 (44%).

A comparison of PBS data for Australian age-specific estimates of the number of persons with DT diabetes with similar NHS data on self-reported diabetes is shown in Figure 1. PBS estimates of DT diabetes generally fall within the calculated confidence intervals of NHS data for self-reported diabetes for the different age groups. However, measures of DT diabetes person counts from PBS data underestimate NHS values for the person age groups 55–74, and overestimates values for the person age group 25–34. This overestimate in young adults was noticed to be predominantly associated with a higher PBS estimates of DT diabetes counts of females in this age group.

The correlation between DT diabetes prevalence estimates and NHS survey data for persons with self-reported diabetes across age groups was high for reports in 2004, 2008 and 2011 ($r > 0.9$). These associations improved over

time with the data in 2014 attaining a correlation of ($r = 0.987$). The accuracy of DT diabetes prevalence estimates compared to NHS survey results of self-reported diabetes prevalence for comparable years using the mean percentage errors is demonstrated in Figure 2. The results show that, in general, PBS data overestimate the values in NHS data for age-groups less than ages 34. The percentage error of PBS measures of DT diabetes prevalence with NHS estimates of self-reported diabetes in these age groups vary between 20% and 35% with wide confidence intervals. In the more elderly age groups PBS values for DT diabetes are in general underestimates between 30% and 5% of NHS data. However, the difference in these age cohorts is more systematic with more compact confidence intervals (Figure 2).

Modification of DT diabetes incidence estimates to reflect community survey derived diabetes incidence

The calculated differences between age specific NHS community survey and PBS DT estimates for prevalent diabetes as shown in Figure 2 and Table 2 were used to calibrate PBS derived age specific DT incidence estimates. These adjustments of DT incidence changed values to represent hypothetical NHS (community survey) diabetes incidence. The adjustments of PBS derived DT incidence utilised contemporaneous NHS survey data for calibration. The values of diabetes incidence for persons calculated in this manner are shown in Table 3 and Supplementary Appendix 1 for the years 2006–2014.

Figure 2: The mean percentage error (95% CI) of PBS estimates of DT diabetes prevalence as compared to estimates from NHS self-reported data for the period 2004–2014

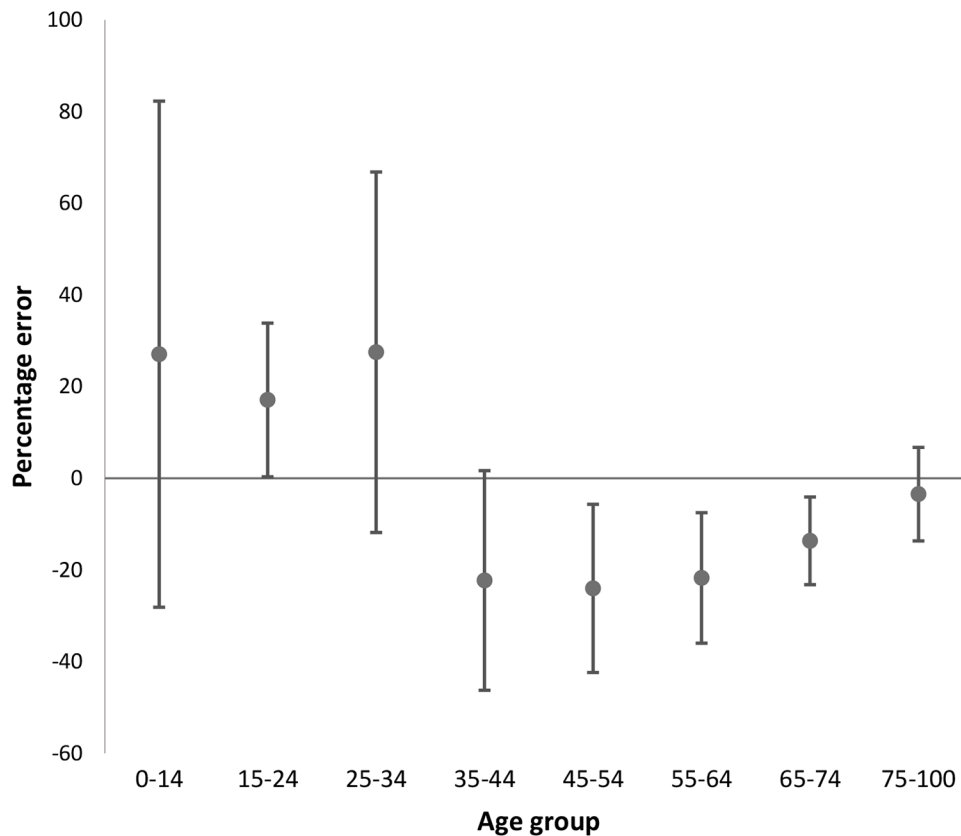


Table 3: PBS estimates of calibrated age specific diabetes incidence (n/1000, CI) for persons in Australia from 2006 to 2014

Persons	2006	2008	2010	2012 ^a	2014
0–14	0.4 (0.2, 0.6)	0.4 (0.1, 0.5)	0.4 (0.1, 0.5)	0.4 (0.6, 1)	0.4 (0.5, 0.9)
15–24	0.7 (0.1, 1.3)	0.6 (0, 1.2)	0.6 (0, 1.2)	2.7 (2.1, 3.3)	1.6 (1, 2.2)
25–34	1.7 (1.3, 2.1)	1.2 (0.8, 2.1)	1.4 (1, 2.3)	6.8 (6.4, 6.2)	2.7 (2.3, 5.6)
35–44	4.1 (3.3, 4.9)	4.6 (3.8, 3.3)	5.2 (4.4, 3.6)	15.8 (15, 9.4)	5.3 (4.5, 6.4)
45–54	6.4 (5.2, 7.6)	7.4 (6.2, 5.5)	7.9 (6.7, 5.8)	32.5 (31.3, 17.7)	8.1 (6.9, 8.5)
55–64	11.5 (9.6, 13.4)	12.8 (10.9, 9.9)	13.8 (11.9, 10.5)	51.4 (49.5, 28.4)	12.0 (10.1, 11.8)
65–74	15.8 (13.6, 18)	14.4 (12.2, 15.1)	13.2 (11, 14.1)	27 (24.8, 19.6)	11.4 (9.2, 12.2)
75–84	11.7 (9, 14.4)	10.1 (7.4, 13.2)	9.8 (7.1, 12.9)	13.5 (10.8, 12.7)	9.2 (6.5, 11.4)
85–100	7.5 (3.5, 11.5)	5.9 (1.9, 10.2)	5.8 (1.8, 10.0)	8.7 (4.7, 10.5)	5.5 (1.5, 9.2)

^aThe spike in age specific incidences in 2012 noted in table 2 is again shown in these data of adjusted diabetes incidence estimates and is related to changes in the PBS legislation improving capture of persons with prevalent diabetes during this year.

Discussion

This study has demonstrated that Australian pharmaceutical data may be a useful resource for the measurement of drug treated diabetes prevalence and incidence. The data may also have a contributing role in estimates of overall diabetes incidence and prevalence. The context of this premise is that diabetes treatments authorised by a doctor and recorded by prescription exchange can act as a proxy diagnosis for an individual with diabetes [7–9]. However, it is acknowledged that the estimates presented here with PBS data are more accurately described as the incidence

and prevalence of medications dispensed for the treatment of diabetes [20]. Furthermore, it is also recognised that certain ATC categorised ‘Drugs used in diabetes’ can be prescribed for other non-diabetes conditions such as polycystic ovarian syndrome [15, 16].

The methodology and the PBS data utilised in this study has yielded estimated measures of incidence and prevalence of DT diabetes in Australia and enabled subgroup comparisons by age and gender. The prevalence estimates of DT diabetes have also demonstrated significant correlation and reasonable accuracy with National Health Survey data describing Australian self-reported diabetes prevalence. The

similarity of case definition using ATC codes of diabetes treatments received in prevalence data has been extrapolated to an assumption that incidence estimates may also have comparable correlative value to the Australian population. Accordingly, the potential to utilise the differences between NHS community survey and PBS data allowed age group calibrated diabetes incidence rates to be calculated and trends demonstrated over time.

Administrative data as used in this study is recognised as having limitations in the granularity of data and the number of available co-variables to observe [21–23]. These limits can be applied to PBS data where person details were restricted to gender, year of birth, and state of origin. Several other limitations of the data source and the methodology also require consideration. First, the coverage of the PBS data source changes in 2012 as a result of the Australian government's *National Health Amendment (PBS) Act 2010*, and is considered a confounding factor [9]. The implementation of this act resulted in the amalgamation of under co-payment data into the PBS data collection set during 2012 [10]. The increase in the overall numbers recorded into the PBS data set altered the relationship between PBS data and the Australian population. This change consequently affected the estimates of DT diabetes disease frequency that span these years and produced an artefactual temporary rise in incidence in 2012. This represents a major limitation of the data source and accordingly, time trend analyses can only be appropriately performed for the time periods either before or after 2012.

Second, using this data for longitudinal analysis requires the assumption that the incidence and prevalence of diabetes treatments have a stable correlation with diabetes disease frequencies over time. It is recognised however that treatment strategies change, and administrative/legal frameworks change, and these are likely to impact upon the observed trends over time [9, 11]. This may explain some of the trends noticed in the data. The possible overestimates of PBS measures of diabetes frequency in younger adult females for example could reflect treatments for other conditions such as polycystic ovarian syndrome, weight loss and prediabetes [14, 15]. The increase in diabetes treatments observed in younger adult females may also reflect the changing management of gestational diabetes [23].

Third, the management of diabetes is not restricted to medications, and may consist of diet and lifestyle modifications only [24]. Hence the frequency data on treatments for diabetes will not capture all persons with a clinical diagnosis of diabetes within a population. This may account for the observed underestimates of DT diabetes prevalence in the more elderly cohorts where management may be more conservative. The accuracy of diabetes case definition may be improved if the dispensing of blood glucose monitoring strips are used to complement the pharmaceutical signal. Another point to consider is that some people may obtain diabetes treatments from non-PBS sources, and consequently will not be incorporated into these analyses of disease frequencies [25]. However, it is likely that this will only be a small proportion of the total number of persons with diabetes.

Finally, the comparison data used in this study is the NHS survey. This is likely not an ideal reference source as

the population coverage of this survey is acknowledged to be restricted [6]. In addition, the prevalence estimates for diabetes have published margins of errors that vary across age and gender cohorts. Younger age groups in particular have higher error values of more than 50% the diabetes prevalence estimates and may reflect limitations in the extent of NHS coverage in these age groups [7]. This critique of NHS data may be an advantage in favour of PBS data which readily counts persons with DT diabetes from these age cohorts.

Estimating diabetes incidence has previously proved to be logistically difficult and expensive to perform [24, 27]. The incidence of diabetes is in the region of 0.1% per annum and as a consequence, a large population is required to provide sufficient power to identify and assess change [28]. Survey methodology requires the same population to be assessed over time, whereas this study, because of the extent of the PBS coverage, utilised the Australian population as an open cohort and the denominator value for incidence estimates [26]. The data has the advantage of being freely available, and can be analysed using straightforward data interrogation techniques, allowing for repeated estimates on the same Australian cohort. This has been considered a prerequisite to maintain the awareness of societal trends for diabetes, which is especially important in a country such as Australia where demographics are changing [1–3].

The results presented on the frequencies of DT diabetes treatments provide a population perspective of diabetes. Persons recorded with DT diabetes in the PBS data are compliant, exchanging prescriptions and actively consuming health care. This viewpoint presents a context as to what is an appropriate epidemiological definition of a diabetes diagnosis. New diagnostic criteria that define the subgroup of diabetes, prediabetes, gestational diabetes, impaired fasting glycaemia and impaired glucose tolerance for individuals are in evolution and have been modified over the last twenty years [28]. These syndromes of diabetes all have different methods for definition such that a straightforward epidemiological technique that could adequately identify all sub groups would be logistically difficult to develop. Furthermore, it is recognised that diabetes is a changing disease with increasing heterogeneity such that even the definition of the polar forms of type 1 and type 2 diabetes are becoming increasingly blurred [28]. These evolutions in concepts as to what constitutes a diagnosis of diabetes or one of its subgroups makes the simplicity of a proxy diagnosis of diabetes based on the treatment received an attractive one.

The comparison of this simplified definition of diabetes using treatments prescribed with corresponding results of the NHS surveys in 2004, 2008, 2011 and 2014 of self-reported diagnoses show excellent correlations which has improved over time. The enhanced result in 2014 may be the consequence of the PBS bill in 2012 allowing greater coverage of PBS data from that year [9, 26]. In our opinion the PBS data and the methodology utilised in this study may be employed as an instrument to assess the population perspectives of diabetes. Estimates of DT diabetes and calibrated diabetes disease frequency from administrative pharmaceutical data may also be useful for trend analyses. These data may complement NHS community survey estimates of diabetes incidence and prevalence that can be used for triangulated

estimates and provide an improved overall evaluation of diabetes in Australia [11].

Acknowledgements

We are grateful to Professor Jonathan Shaw of the Baker Institute, Australia for critique and input into the manuscript and La Trobe research ethics committee for authorisation of the study.

Ethics Statement

Ethics approval for the study was obtained from the La Trobe University ethics committee (S17-198).

Statements of conflicts of interest

The authors have no conflicts of interest to declare.

References

1. Zimmet P, Alberti KG, Magliano DJ, Bennett PH. Diabetes mellitus statistics on prevalence and mortality: facts and fallacies. *Nature Reviews Endocrinology*. 2016 Oct;12(10):616. <https://doi.org/10.1038/nrendo.2016.105>
2. Magliano DJ, Peeters A, Vos T, Sicree R, Shaw J, Sindall C, Haby M, Begg SJ, Zimmet PZ. Projecting the burden of diabetes in Australia—what is the size of the matter? *Australian and New Zealand journal of public health*. 2009 Dec;33(6):540–3. <https://doi.org/10.1111/j.1753-6405.2009.00450.x>
3. Australian Institute of Health and Welfare. Diabetes prevalence in Australia: An assessment of national data sources. Shaw, J. and Tanamas, S., 2012. Diabetes: the silent pandemic and its impact on Australia. *Melbourne: Baker IDI Heart and Diabetes Institute*. Retrieved from <https://static.diabetesaustralia.com.au/s/fileassets/diabetes-australia/e7282521-472b-4313-b18e-be84c3d5d907.pdf>
4. Zimmet PZ, Magliano DJ, Herman WH, Shaw JE. Diabetes: a 21st century challenge. *The lancet Diabetes & endocrinology*. 2014 Jan 1;2(1):56–64. [https://doi.org/10.1016/s2213-8587\(13\)70112-8](https://doi.org/10.1016/s2213-8587(13)70112-8)
5. National Health Survey First Results Australia 2014–15 4364.0.55.001 Australian Bureau of Statistics. National health survey: First results, 2014–15. ABS Cat. No. 4364.0. 55.001). 2015. <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012014-15?OpenDocument> last accessed 07/09/2020
6. Catanzariti L, Faulks K, Moon L, Waters AM, Flack J, Craig ME. Australia's national trends in the incidence of Type 1 diabetes in 0–14-year-olds, 2000–2006. *Diabetic Medicine*. 2009 Jun;26(6):596–601. <https://doi.org/10.1111/j.1464-5491.2009.02737.x>
7. 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>
8. 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>
9. 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>
10. 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>
11. Purkiss SF, Keegel T, Vally H, Wollersheim D. A comparison of Australian chronic disease prevalence estimates using administrative pharmaceutical dispensing data with international and community survey data. *International Journal of Population Data Science*. 2020 Dec 11;5(1).
12. Benrimoj SI, Frommer MS. Community pharmacy in Australia. *Australian Health Review*. 2004;28(2):238–46. <https://doi.org/10.1071/AH040238>
13. Moles RJ, Stehlik P. Pharmacy practice in Australia. *The Canadian Journal of Hospital Pharmacy*. 2015 Sep;68(5):418. <http://doi.org/10.4212/cjhp.v68i5.1492>
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. Moll E, van der Veen F, van Wely M. The role of metformin in polycystic ovary syndrome: a systematic review. *Human reproduction update*. 2007 Sep 1;13(6):527–37. <https://doi.org/10.1093/humupd/dmm026>
16. Seifarth C, Schehler B, Schneider HJ. Effectiveness of metformin on weight loss in non-diabetic individuals with obesity. *Experimental and Clinical Endocrinology & Diabetes*. 2013 Jan;121(01):27–31. <https://doi.org/10.1055/s-0032-1327734>
17. Australian Bureau of Statistics 2015 Census data cubes. Retrieved from <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3101.0Jun%202015?OpenDocument>

18. Furey E. Percentage Error Calculator. [online] Calculator soup. 2020. Available at: <https://www.calculatorsoup.com>. Last accessed 07/09/2020
19. 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>
20. Bullock S, Pearce J, Byng K. Incidence of Insulin-treated Diabetes in Australia: 2014. Australian Institute of Health and Welfare; 2016. Retrieved from <https://www.aihw.gov.au/getmedia/c17ddab6-1977-41e2-8019-48c8cc4c5a7f/19899.pdf.aspx?inline=true>
21. Sarrazin MS, Rosenthal GE. Finding pure and simple truths with administrative data. Jama. 2012 Apr 4;307(13):1433–5. <https://doi.org/10.1001/jama.2012.404>
22. Johnson EK, Nelson CP. Utility and pitfalls in the use of administrative databases for outcomes assessment. The Journal of urology. 2013 Jul;190(1):17. <https://doi.org/10.1016/j.juro.2013.04.048>
23. Hod M, Kapur A, Sacks DA, Hadar E, Agarwal M, Di Renzo GC, Roura LC, McIntyre HD, Morris JL, Divakar H. The International Federation of Gynecology and Obstetrics (FIGO) Initiative on gestational diabetes mellitus: A pragmatic guide for diagnosis, management, and care. International Journal of Gynecology & Obstetrics. 2015 Oct 1;131:S173–211. [https://doi.org/10.1016/s0020-7292\(15\)30033-3](https://doi.org/10.1016/s0020-7292(15)30033-3)
24. Magliano DJ, Barr EL, Zimmet PZ, Cameron AJ, Dunstan DW, Colagiuri S, Jolley D, Owen N, Phillips P, Tapp RJ, Welborn TA. Glucose indices, health behaviors, and incidence of diabetes in Australia: the Australian Diabetes, Obesity and Lifestyle Study. Diabetes care. 2008 Feb 1;31(2):267–72. <https://doi.org/10.2337/dc07-0912>
25. Webbie K, O'Brien K. Use of Medicines by Australians with Diabetes. Australian Institute of Health and Welfare; 2006. Retrieved from <https://www.aihw.gov.au/getmedia/c17ddab6-1977-41e2-8019-48c8cc4c5a7f/19899.pdf.aspx?inline=true>
26. Duckett SJ. Drug policy down under: Australia's pharmaceutical benefits scheme. Health care financing review. 2004;25(3):55. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4194861/>
27. Gilbert-Ouimet M, Ma H, Glazier R, Brisson C, Mustard C, Smith PM. Adverse effect of long work hours on incident diabetes in 7065 Ontario workers followed for 12 years. BMJ Open Diabetes Research and Care. 2018 Jun 1;6(1):e000496. <https://doi.org/10.1136/bmjdr-2017-000496>
28. Tuomi T, Santoro N, Caprio S, Cai M, Weng J, Groop L. The many faces of diabetes: a disease with increasing heterogeneity. The Lancet. 2014 Mar 22;383(9922): 1084–94. [https://doi.org/10.1016/s0140-6736\(13\)62219-9](https://doi.org/10.1016/s0140-6736(13)62219-9)

Abbreviations

ABS:	Australian Bureau of Statistics
ATC:	Anatomic Therapeutic Chemical Codes
CI:	Confidence Interval
DT:	Drug Treated
M:	Mean
MPE:	Mean Percentage Error
NHS:	National Health Survey
PBS:	Pharmaceutical Benefits Scheme
PE:	Percentage Error
SD:	Standard deviation
SPSS:	Statistical Package for the Social Sciences
WHO:	World Health Organization



Supplementary appendices

Supplementary Appendix 1: Adjusted Drug Treated diabetes person incidence estimates to reflect NHS calibrated Australian diabetes incidence for the period 2006–2014

