

Concordance of self-reporting of diabetes compared with medical records: A comparative study using polyclinic data in Singapore

Khai Wei Tan¹ *MMed (FM)*, Jeremy Kaiwei Lew² *MA (Biomedical Informatics)*, Poay Sian Sabrina Lee² *PhD*,
Sin Kee Ong² *BSocSci Hons (Psychology)*, Hui Li Koh² *BSc (Psychology) (Hons)*, Doris Yee Ling Young³ *MD*, Eng Sing Lee^{2,4} *PhD*

ABSTRACT

Introduction: Studies of concordance between patients' self-report of diseases and a criterion standard (e.g. chart review) are usually conducted in epidemiological studies to evaluate the agreement of self-reported data for use in public health research. To our knowledge, there are no published studies on concordance for highly prevalent chronic diseases such as diabetes and pre-diabetes. The aims of this study were to evaluate the concordance between patients' self-report and their medical records of diabetes and pre-diabetes diagnoses, and to identify factors associated with diabetes concordance.

Method: A cross-sectional, interviewer-administered survey was conducted on patients with chronic diseases after obtaining written consent to assess their medical notes. Interviewers were blinded to the participants' profiles. Concordance was evaluated using Cohen's kappa (κ). A multivariable logistic regression model was used to identify factors associated with diabetes concordance.

Results: There was substantial agreement between self-reported and medical records of diabetes diagnoses ($\kappa=0.76$) and fair agreement for pre-diabetes diagnoses ($\kappa=0.36$). The logistic regression model suggested that non-Chinese patients had higher odds of diabetes concordance than Chinese patients (odds ratio [OR]=4.10, 95% confidence interval [CI] 1.19–14.13, $P=0.03$). Patients with 3 or more chronic diseases (i.e. multimorbidity) had lower odds of diabetes concordance than patients without multimorbidity (OR=0.21, 95% CI 0.09–0.48, $P<0.001$).

Conclusion: Diabetes concordance was substantial, supporting the use of self-report of diabetes by patients with chronic diseases in the primary care setting for future research. Pre-diabetes concordance was fair and may have important clinical implications. Further studies to explore and improve health literacy and patient-physician communication are needed.

Ann Acad Med Singap 2022;52:62-70

Keywords: Concordance, diabetes, multimorbidity, primary care, self-reported data

INTRODUCTION

Approximately 422 million people worldwide have diabetes and 1.6 million deaths are attributed to diabetes each year,¹ contributing to high economic costs worldwide. Diabetes education and awareness of the disease contribute significantly to minimising complications and reducing morbidity and mortality.² In addition, there is also a strong impetus to enhance the accessibility to programmes on the prevention of diabetes, pre-diabetes and associated risk factors.³

In Singapore, the Ministry of Health reported that more than 400,000 Singaporeans have type 2 diabetes and the condition is costing the healthcare system more than SGD1 billion a year.⁴ According to the Singapore National Health Survey (SNHS) 2010 on residents, 1 in 3 patients with diabetes was not aware of having the disease.⁵ There are approximately 430,000 (14%) Singaporeans aged 18–69 years diagnosed with pre-diabetes⁶ and there have been no published studies on the awareness level of pre-diabetes in the country.

¹ Toa Payoh Polyclinic, National Healthcare Group Polyclinics, Singapore

² Clinical Research Unit, National Healthcare Group Polyclinics, Singapore

³ Division of Family Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

⁴ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

Correspondence: Dr Khai Wei Tan, Toa Payoh Polyclinic, 2003 Lor 8 Toa Payoh, Singapore 319260.

Email: khaiwei_tan@nhgp.com.sg

CLINICAL IMPACT

What is New

- Concordance between self-reported and medical records of disease diagnoses was substantial for diabetes but not pre-diabetes.
- Diabetes concordance is less likely among Chinese patients, and patients with 3 or more chronic conditions (i.e. multimorbidity).

Clinical Implications

- Patients' self-reports of diabetes diagnoses are a reliable source of data for primary care research.
- Poorer pre-diabetes concordance could be indicative of factors such as the lack of health literacy and patient-physician communication.

Self-reported health is commonly used as a data source for epidemiological studies and national health surveys, as it generally involves lower costs than clinical assessments.^{7,8} However, the agreement of such self-reported data has been mixed, and concordance with clinical records can vary based on the patients and disease factors.⁹⁻¹² Additionally, concordance could be indicative of the patients' health literacy and quality of patient-physician communication.¹³ In Western countries, the concordance for diabetes ranged from 0.75 (substantial agreement) to 0.92 (almost perfect agreement).^{9,14-20} In Asia, the trend was similar, although studies were limited and there were no local studies on the concordance of diabetes.

There were also limited studies showing how patient factors affected concordance. Hansen et al. reported that multimorbidity (higher disease count) was associated with better concordance for diabetes.²¹ Muggah et al. reported an inverse relationship between morbidity burden and concordance.²⁰ Chun et al. reported that those aged 60–69 years paradoxically had better concordance for diabetes compared to those aged 50–59 years.²² Clinically, poor concordance led to poor doctor-patient relationships, which affected the quality of care.¹³

For pre-diabetes, there were many studies comparing objective laboratory data and their concordance with pre-diabetes. To our knowledge, there are no studies that looked at the concordance of self-reported data.

This study aimed to determine the concordance between patients' self-reports of diabetes and their medical records of diabetes diagnoses, among patients

visiting a primary care centre for chronic conditions management. It also explored the associations of various patient factors with concordance for diabetes. The secondary objective was to determine the concordance of pre-diabetes between patients' self-reports of pre-diabetes and their medical records of pre-diabetes diagnoses, among patients without diabetes diagnoses.

METHOD

Study setting

The public primary healthcare services in Singapore are provided through an island-wide network of outpatient polyclinics, which offer subsidised care. For this study, the recruitment of participants was based at Toa Payoh Polyclinic, which is located in central Singapore. Toa Payoh Polyclinic has an estimated base pool of more than 40,000 multiethnic patients who come for regular follow-up of their chronic disease management, of whom approximately 12% have diabetes.

Study design and data collection

Self-reported data

We conducted a cross-sectional, interviewer-administered survey on patients who come for regular follow-up of their chronic disease management at Toa Payoh Polyclinic. Patient recruitment commenced between December 2019 and January 2020, then halted due to COVID-19 safety measures, but subsequently resumed and completed from March to April 2021. Study approval was obtained from the National Healthcare Group (NHG) Domain Specific Review Board (Reference no: 2019/00719) with written informed consent from all patients who participated in the study.

Based on the inclusion criteria, the Information Management and Analytics (IMA) department of NHG Polyclinics assisted in generating a list of patients who were eligible for participation on a weekly basis. The recruiting team then approached these patients when they were at the polyclinic waiting for their consultation. They were screened for eligibility before being invited to participate in the study. The inclusion criteria were patients who were aged 21 years and above, and had made 2 or more visits to Toa Payoh Polyclinic for chronic condition management in the past one year. The exclusion criteria were patients without mental capacity or for whom communication and decisions were made via a proxy.

The questions from the survey were adapted from the SNHS 2010.⁵ The following baseline characteristics

were collected: age, sex, ethnicity, marital status, education level, housing type, employment status and multimorbidity status (having 3 or more chronic conditions) (Table S1 in online Supplementary Materials). Patient responses and participant identification would be entered into the NHG Research Electronic Data capture (REDCap) system, without any patient identifiers.

As a feedback loop, IMA took into account the recruited number of patients with and without diabetes (in medical records), and attempted to keep the proportion of recruited patients with diabetes to 50%, to minimise the effect of prevalence. The recruiting team members were aware of this recruitment proportion but were blinded as to which participants had diabetes or no diabetes recorded in their medical records (Fig. 1).

Data from medical records

Data on patients' disease diagnoses were extracted by IMA from existing medical records and reconciled with the patients' survey responses collected in REDCap. The combined data spreadsheet was then de-identified before returning to the study team for analysis.

Statistical analysis

Concordance for diabetes was quantified using Cohen's kappa.^{23,24} The kappa (κ) coefficient is frequently used to determine the strength of agreement between 2 raters (self-reported data and medical records). A κ value of <0.40 is considered as indicating poor to a fair agreement, 0.41 – 0.60 moderate agreement, 0.61 – 0.80 substantial agreement, and 0.81 – 1.00 almost perfect agreement.²⁵ As kappa is affected by prevalence and bias,^{23,26,27} we also present the prevalence-adjusted bias-adjusted kappa (PABAK).

In determining our study sample size, we assumed a kappa of 0.70 as the expected agreement, based on comparisons with similar studies in Asia.^{11,22,28} Under our null hypothesis, kappa was set at 0.60 (high moderate agreement) as we believe that a poorer concordance ($\kappa < 0.60$) is clinically unacceptable for an impactful disease such as diabetes. We further targeted to recruit equal proportions of patients with and without diabetes to mitigate any predisposed prevalence effect that lowers kappa.^{23,26,27} In a test for agreement between 2 raters using the kappa statistic, a sample size of 472 subjects achieves 80% power at the 0.05 significance level to detect a true kappa value of 0.70 , in a test where kappa was set at 0.60 under the null hypothesis with an assumed 50% proportion of positive ratings^{23,29} (given that we target to recruit an equal number of patients with and without diabetes). Accounting for

a 30% non-response rate, we aimed to approach 472 patients.

A multivariable logistic regression model was used to determine patient factors associated with diabetes concordance while controlling for other variables (Table S2 of online Supplementary Materials). Concordance is dichotomised into "yes" or "no", with "yes" being that the patients' self-reported diabetes matches their medical records and "no" otherwise. The multicollinearity test was performed to check for high correlations among independent variables. The variance inflation factors of all independent variables are less than 5 (i.e. not highly correlated).³⁰ Statistical analyses were conducted using R software version 3.6.3.³¹ A P value of <0.05 was considered statistically significant.

RESULTS

A total of 751 patients were approached. Of these, 26 were excluded as they were unable to give consent due to a lack of mental capacity or whose decisions were communicated via a proxy caregiver, and 247 patients declined participation. The final number of patients recruited was 478, giving a response rate of 65.9%.

Among the 478 patients, there were 239 patients with medical records of diabetes and 239 with no medical records of diabetes. From the self-reported data, 187 (78.2%) and 234 (97.9%) were concordant for having diabetes and not having diabetes, respectively. These numbers are presented in Fig. 2.

Participant characteristics

Table 1 shows the baseline characteristics of the patients who were recruited. There were 80.6% of patients aged 60 years and above (mean age 67.9 ± 11.0 years). There were more male patients (52.5%). The majority of the patients were Chinese (86.6%), married (70.1%), had secondary school or lower education (68.6%), and lived in public housing³² (87.5%). Less than half of them were actively employed (42.3%) and more than half had multimorbidity (59.4%).

Concordance between self-reported and medical records for diabetes and pre-diabetes

There was substantial agreement between patients' self-reports of diabetes and medical records of diabetes diagnoses ($\kappa = 0.76$, 95% CI 0.67 – 0.85 , $P < 0.001$; PABAK = 0.76). The results are shown in Table 2. The value of kappa (0.76) for diabetes concordance was very close to PABAK (0.76). This

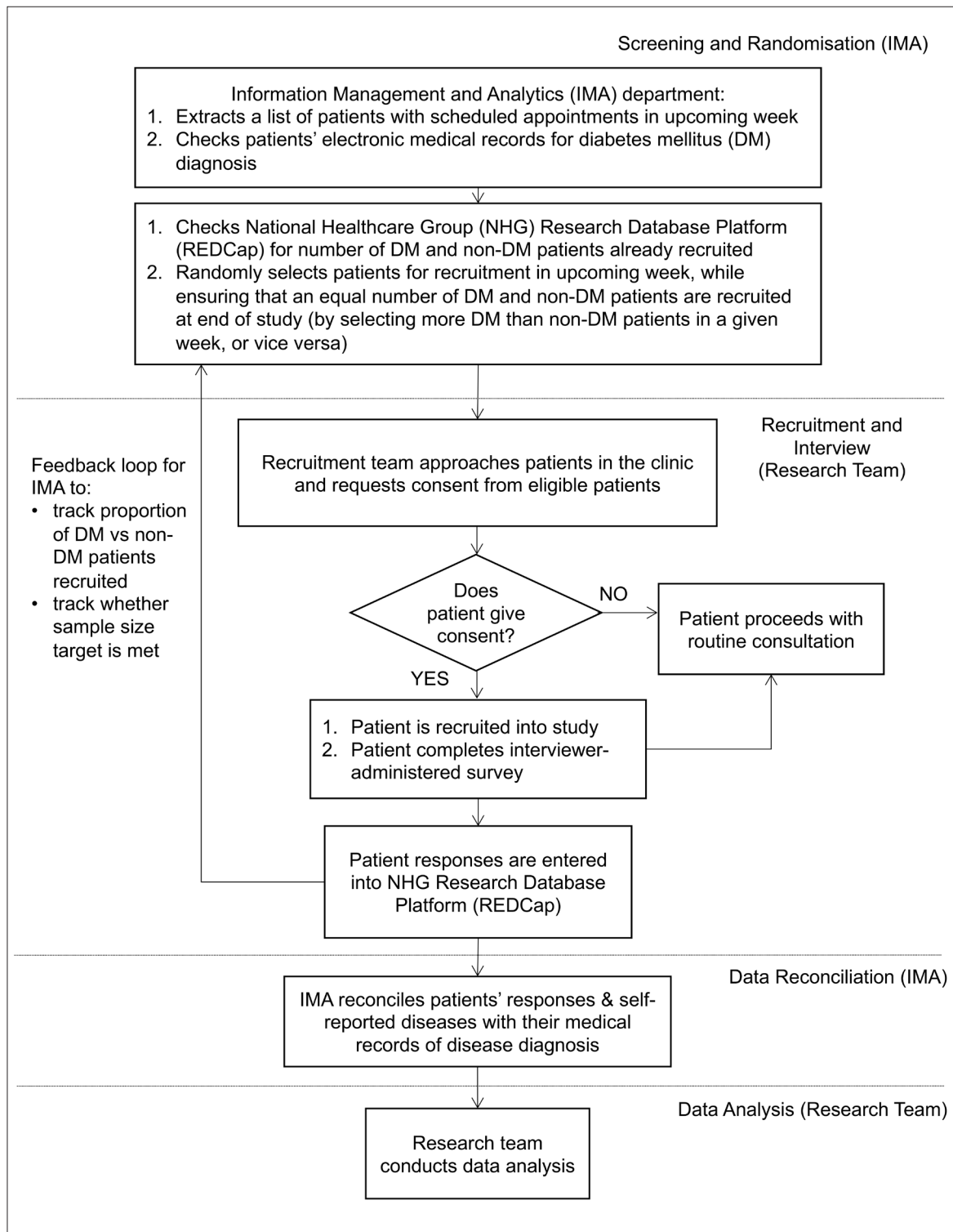


Fig. 1. Flow diagram of patient recruitment from screening to data analysis.

was expected as our proportions of patients with and without diabetes was 1:1, which minimised the effect of prevalence (PI=0.10) on the value of kappa.^{26,33}

Among patients without diabetes, there was a fair agreement between patients' self-reports of pre-

diabetes and medical records of pre-diabetes diagnoses ($\kappa=0.36$, 95% CI 0.24–0.48, $P<0.001$; PABAK=0.60). However, after adjustment for prevalence and bias using PABAK, the concordance was found to be moderate. The results are shown in Table 3.

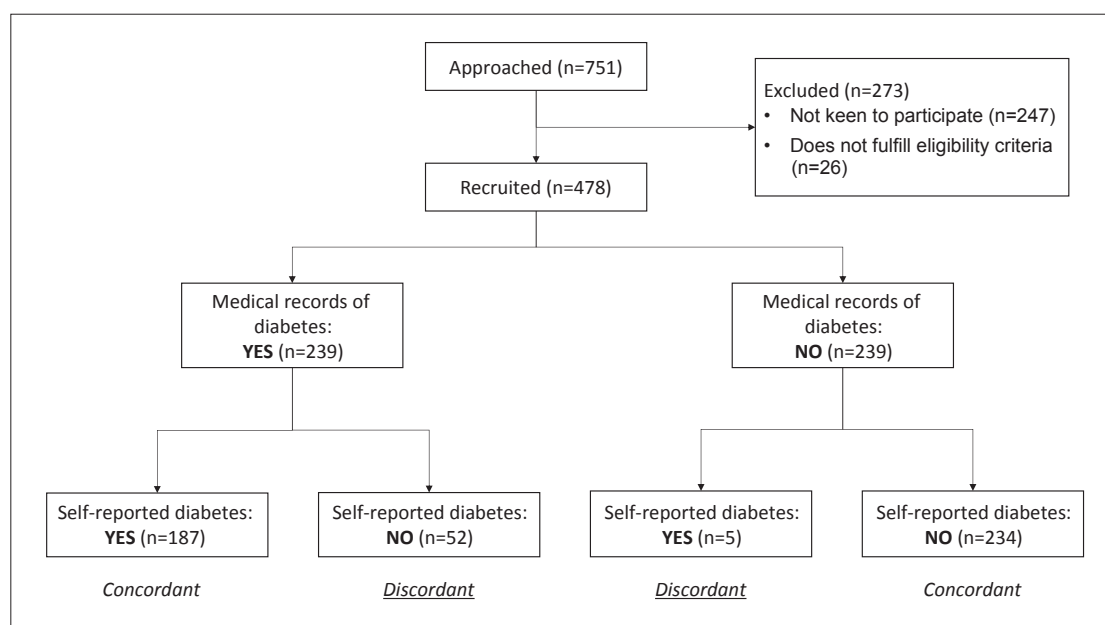


Fig. 2. Breakdown of recruited patients according to concordant and discordant status.

Logistic regression of patient factors associated with diabetes concordance

Table 4 presents the results of the logistic regression model that included all the variables listed in Table 1. All independent variables showed no statistically significant association with diabetes concordance, except for ethnicity and multimorbidity.

Chinese patients were associated with poorer diabetes concordance. The odds of diabetes concordance for Chinese patients was 0.24 times that of non-Chinese patients ($P=0.03$). Having multimorbidity was associated with poorer diabetes concordance ($P<0.001$). The odds of diabetes concordance for patients with multimorbidity (≥ 3 chronic diseases) was 0.21 times that of patients without multimorbidity (< 3 chronic diseases).

DISCUSSION

Our study sought to quantify the concordance among patients with diabetes and to determine the patient-related factors associated with poorer concordance. Our results showed substantial agreement ($\kappa=0.76$, PABAK=0.76) between self-reported diabetes and medical records of diabetes diagnoses. This kappa value for diabetes concordance ($\kappa=0.76$) in our study was similar to other studies in Asia, namely South Korea ($\kappa=0.82$)²² and Taiwan ($\kappa=0.76$).²⁸ This suggests that self-reported data for diabetes can potentially be a cost-efficient method of studying diabetes prevalence and trends in the population.

Other studies have postulated that: good concordance is observed when the disease is well defined;¹⁰ has clear diagnostic criteria; is non-episodic;^{10,20,28} requires frequent monitoring³⁴ or repeated engagement with the healthcare system;¹⁰ affects daily function;³⁴ and when the disease is not confused with another (e.g. stroke and transient ischaemic attack).²⁰ Clinically, the aim is to achieve as close to the perfect agreement ($\kappa=1.00$) as possible, though perfect agreement is seldom achieved especially in healthcare research.³⁵ We find that the above factors are consistent with our experience with diabetes management in the primary care context, and this could explain the substantial diabetes concordance observed in our study.

In the US, the estimated awareness of pre-diabetes was lower than 14% across all population subgroups.³⁶ In China, a study on the Suzhou community found that only 38.5% knew they had pre-diabetes.³⁷ In our study, 22 out of 57 patients (38.6%) knew that they had pre-diabetes (Table 3). This proportion was very similar to the findings in China.

The concordance for patients with pre-diabetes ($\kappa=0.36$, PABAK=0.60) was lower than for diabetes. Here, kappa and PABAK values differed more, compared to the concordance results for the diabetes aforementioned. This was expected and the key contributory reason was that the prevalence of pre-diabetes in this sample was significantly lower than 50%. This does not mean that the actual kappa for pre-diabetes is 0.60. PABAK should be interpreted

Table 1. Characteristics of study participants

Characteristic	N=478 No.	%
Age (as at recruitment)		
<60 years	93	19.5
60–69 years	170	35.6
70–79 years	149	31.2
≥80 years	66	13.8
Sex		
Male	251	52.5
Female	227	47.5
Ethnicity		
Chinese	414	86.6
Non-Chinese	64	13.4
Marital status		
Married	335	70.1
Single/divorced/widowed	143	29.9
Education level		
No formal education	39	8.2
Primary	113	23.6
Secondary	176	36.8
Post-secondary	150	31.4
Housing type		
HDB 1-room/2-room flat	54	11.3
HDB 3-room flat	136	28.5
HDB 4-room flat	122	25.5
HDB 5-room/HUDC flat	106	22.2
Condominium/landed property	60	12.6
Employment status		
Employed	202	42.3
Homemaker/retired/unemployed	276	57.7
Multimorbidity status		
Non-multimorbid (<3 chronic diseases)	194	40.6
Multimorbid (≥3 chronic diseases)	284	59.4

HDB flat: Housing and Development Board flat (housing by public housing authority in Singapore); HUDC flat: Housing and Urban Development Company flat (public housing, now privatised, for middle-income households in Singapore)

alongside kappa, and it helps to contextualise the effect of prevalence on the sample. What our data highlight, however, is that the concordance of pre-diabetes is poorer than diabetes.

A contributory factor to poor pre-diabetes concordance could be the incomplete capture of pre-diabetes in the medical records as pre-diabetes is not as strictly coded for, unlike diabetes. These patients may also have a lack of understanding about pre-diabetes, which requires the doctor's effort to form an equal partnership with the patient to explain the condition and the necessary treatment plans to prevent progression to diabetes.³⁸ Moreover, patients with pre-diabetes may not be as actively engaged as patients with diabetes within the healthcare system. However, individuals with pre-diabetes are at high risk for developing type 2 diabetes, which accounts for 90–95% of all cases of diabetes per year.³⁹ Each year, 11% of individuals with pre-diabetes who do not lose weight and do not engage in moderate physical activity will progress to type 2 diabetes during an average follow-up of 3 years.⁴⁰ Managing this group of patients well can potentially reduce the incidence rate of diabetes significantly.

Our logistic regression results showed that being of Chinese ethnicity was statistically significantly associated with poorer diabetes concordance (odds ratio [OR]=0.24, i.e. lower odds of concordance compared to non-Chinese). This finding is supported by a cross-sectional survey of 2,895 participants in the Singapore general population, where the Chinese ethnicity had significantly higher odds of inadequate health literacy on diabetes compared to non-Chinese.⁴¹ A plausible explanation is that the patients who visit Toa Payoh Polyclinic may also be seeking traditional Chinese medicine (TCM) treatment elsewhere, and the way diabetes is categorised under the TCM framework could be different from the Western medicine practice in primary care. For example, patients could consider themselves to have high blood sugar instead of diabetes under the TCM framework. A similar instance was reported by Goldman et al.¹¹ in the case of hypertension; however, our conjecture can only be confirmed in future studies.

Our logistic regression results also showed that having multimorbidity (defined as having 3 or more chronic diseases) was significantly associated statistically with poorer diabetes concordance (OR=0.21, i.e. lower odds of concordance compared to patients without multimorbidity). This association was similarly reported among the Canadian population⁴² and Minnesota residents in the US,¹⁰ although their definition of multimorbidity (i.e. list of diseases considered) differs from our study. Okura et al.¹⁰ suggested that the association of having

Table 2. Concordance analysis for diabetes for study participants

Diabetes	Status from clinical diagnoses recorded data							κ	<i>P</i> value of κ	BI	PI	PABAK
	N (%)	No		Yes		Total						
		n	%	n	%	n	%	0.76	<0.001	0.10	0.10	0.76
Self-reported status	No	234	48.95	52	10.88	286	59.83					
	Yes	5	1.05	187	39.12	192	40.17					
	Total	239	50.00	239	50.00	478	100					

Table 3. Concordance analysis for pre-diabetes for study participants without diabetes

Pre-diabetes	Status from clinical diagnoses recorded data							κ	P value of κ	BI	PI	PABAK
	N (%)	No		Yes		Total						
		n	%	n	%	n	%	0.36	<0.001	0.09	0.62	0.60
Self-reported status	No	169	70.71	35	14.64	204	85.36					
	Yes	13	5.44	22	9.21	35	14.64					
	Total	182	76.15	57	23.85	239	100					

κ : kappa; BI: Bias Index; PI: Prevalence Index; PABAK: prevalence-adjusted bias-adjusted kappa

PABAK = $\kappa + (1 - \kappa) PI^2 + (\kappa - 1) BI^2$

Refer to online Supplementary Material S3 for more details

multimorbidity with poorer concordance could be due to the increased awareness of diseases as a result of more frequent engagements with the healthcare system. This results in over-reporting and poorer concordance. However, our data showed otherwise with patients with multimorbidity under-reporting (i.e. patients who indicated that they have not been told by a doctor to have diabetes even though the medical records showed otherwise). Further studies are needed to explore our patients' understanding and acceptance of diabetes, especially for patients with multimorbidity and diabetes.

Limitations and strengths

Our main limitation was that in comparing self-reported diabetes with medical records of diabetes, we regarded medical records as the source of truth. While this is a widely accepted standard, we acknowledge that there are instances when clinicians code diagnoses incorrectly in the system, or even misdiagnose diseases, leading to inaccuracies in the medical records.^{43,44} On the other hand, the accuracy of self-reported data collected from surveys is limited by recall bias, social desirability effect,^{28,45,46} the way the questions are phrased and asked,^{8,47} and the comprehension ability of the participant,⁴⁷ including factors that may impair judgement such as mild cognitive impairment.

The differences in the profile of patients between our centre (elderly with multimorbidity) and the general population also limit the generalisability of our results to the rest of the population.

Lastly, we only studied non-modifiable patient factors. Chun et al. explored the impact of health-related behaviours such as smoking, drinking and exercise on concordance, although they did not find statistical significance.²²

Despite the limitations, this is one of the first studies in Singapore to explore diabetes concordance among patients with chronic conditions in the primary care setting. In contrast, such concordance had been studied in Western countries^{9,14-20} and some countries in Asia.^{22,28} We took into account the limitations of kappa as a statistical tool, such as prevalence and bias, and tried to keep equal proportions of participants with and without diabetes as much as possible.^{26,33} This study was also an initial attempt to understand the pre-diabetes concordance among our primary care patients.

Modifiable patient factors such as health behaviours of smoking, drinking or exercise, which could affect concordance indirectly, could be explored. The population base for the study could also be extended to cover the wider community, or recruitment could focus on a nationally representative sample for greater

Table 4. Multivariable logistic regression of independent variables with diabetes concordance

Independent variables	Adjusted odds ratio ^a	95% confidence interval ^b	P value ^c
Age (as at recruitment)			
60–69 years	REF		
<60 years	1.45	0.51–4.12	0.48
70–79 years	1.37	0.66–2.81	0.40
≥80 years	1.52	0.60–3.88	0.38
Sex			
Male	REF		
Female	0.82	0.43–1.54	0.53
Ethnicity			
Chinese	REF		
Non-Chinese	4.10	1.19–14.13	0.03
Marital status			
Married	REF		
Single/divorced/widowed	1.04	0.54–2.01	0.91
Education level			
Secondary	REF		
No formal education	0.59	0.22–1.57	0.29
Primary	1.18	0.54–2.57	0.69
Post-secondary	1.13	0.49–2.62	0.78
Housing type			
HDB 3-room flat	REF		
HDB 1-room/2-room flat	0.79	0.33–1.92	0.61
HDB 4-room flat	2.08	0.91–4.76	0.08
HDB 5-room/HUDC flat	1.83	0.76–4.38	0.18
Condominium/landed property	3.34	0.89–12.50	0.07
Employment status			
Employed	REF		
Homemaker/retired/unemployed	0.88	0.44–1.78	0.73
Multimorbidity status			
Non-multimorbid (<3 chronic diseases)	REF		
Multimorbid (≥3 chronic diseases)	0.21	0.09–0.48	<0.001

HDB flat: Housing and Development Board flat (housing by public housing authority in Singapore); HUDC flat: Housing and Urban Development Company flat (public housing, now privatised, for middle-income households in Singapore)

^a The odds ratios are all adjusted in the same multivariable logistic regression model

^b Confidence interval

^c $P < 0.05$ is considered statistically significant

generalisability. Qualitative research could also be conducted among patients with diabetes discordance, to explore perceptions and aspects of diabetes they have difficulty understanding.

CONCLUSION

The findings from this study suggested substantial concordance between self-reported and medical records for diabetes in our study population. This lends support that self-reported diabetes is a valid source of data in public health research. Further research is required to understand the association of poorer diabetes concordance among patients who have multimorbidity and are of Chinese ethnicity. Fair concordance was found between self-reported and medical records for pre-diabetes. This has important medical implications as a significant proportion of patients with pre-diabetes progress to diabetes. Further studies to explore and improve the health literacy and patient-physician communication may be important in managing patients with pre-diabetes and diabetes.

Funding

This research was supported by the Singapore Ministry of Health's National Medical Research Council under the Centre Grant Programme (reference number: NMRC/CG/C019/2017).

REFERENCES

1. World Health Organization. Diabetes Fact Sheet, updated 16 September 2022. <https://www.who.int/news-room/fact-sheets/detail/diabetes>. Accessed 5 October 2021.
2. Nazar CM, Bojerenu MM, Safdar M, et al. Effectiveness of diabetes education and awareness of diabetes mellitus in combating diabetes in the United Kingdom; a literature review. *J Nephropharmacol* 2015;5:110-5.
3. O'Connell JM, Manson SM. Understanding the Economic Costs of Diabetes and Prediabetes and What We May Learn About Reducing the Health and Economic Burden of These Conditions. *Diabetes Care* 2019;42:1609-11.
4. Ministry of Health, Singapore. Speech by Minister for Health, Mr Gan Kim Yong, At the MOH Committee of Supply Debate 2016, 13 April 2016. <https://www.moh.gov.sg/news-highlights/details/speech-by-minister-for-health-mr-gan-kim-yong-at-the-moh-committee-of-supply-debate-2016>. Accessed 5 October 2021.
5. Ministry of Health, Singapore. National Health Survey 2010, November 2011. <https://www.moh.gov.sg/resources-statistics/reports/national-health-survey-2010>. Accessed 5 May 2021.
6. Ministry of Health, Singapore. Ministry of Health Studying Measures to Better Support Persons With Pre-diabetes, 14 November 2017. <https://www.moh.gov.sg/news-highlights/details/ministry-of-health-studying-measures-to-better-support-persons-with-pre-diabetes>. Accessed 5 October 2021.
7. Sherbourne CD, Meredith LS. Quality of Self-Report Data: A Comparison of Older and Younger Chronically Ill Patients. *J Gerontol* 1992;47:204-11.
8. Gordis L. Assuring The Quality of Questionnaire Data In Epidemiologic Research. *Am J Epidemiol* 1979;109:21-4.

9. Haapanen N, Miiunpalo S, Pasanen M, et al. Agreement between Questionnaire Data and Medical Records of Chronic Diseases in Middle-aged and Elderly Finnish Men and Women. *Am J Epidemiol* 1997;145:762-9.
10. Okura Y, Urban LH, Mahoney DW, et al. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol* 2004;57:1096-103.
11. Goldman N, Lin IF, Weinstein M, et al. Evaluating the quality of self-reports of hypertension and diabetes. *J Clin Epidemiol* 2003;56:148-54.
12. Machón M, Arriola L, Larrañaga N, et al. Validity of self-reported prevalent cases of stroke and acute myocardial infarction in the Spanish cohort of the EPIC study. *J Epidemiol Community Health* 2013;67:71-5.
13. Skinner KM, Miller DR, Lincoln E, et al. Concordance Between Respondent Self-reports and Medical Records for Chronic Conditions: Experience From the Veterans Health Study. *J Ambul Care Manage* 2005;28:102-10.
14. Galenkamp H, Huisman M, Braam AW, et al. Disease prevalence based on older people's self-reports increased, but patient-general practitioner agreement remained stable, 1992–2009. *J Clin Epidemiol* 2014;67:773-80.
15. Breitling LP, Olsen H, Müller H, et al. Self- or Physician-reported Diabetes, Glycemia Markers, and Cognitive Functioning in Older Adults in Germany. *Am J Geriatr Psychiatry* 2014;22:1105-15.
16. Kriegsman DM, Penninx BW, Van Eijk JT, et al. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. *J Clin Epidemiol* 1996;49:1407-17.
17. Oksanen T, Kivimäki M, Pentti J, et al. Self-Report as an Indicator of Incident Disease. *Ann Epidemiol* 2010;20:547-54.
18. Jiang L, Zhang B, Smith ML, et al. Concordance between Self-Reports and Medicare Claims among Participants in a National Study of Chronic Disease Self-Management Program. *Front Public Health* 2015;3.
19. Tisnado DM, Adams JL, Liu H, et al. What is the Concordance Between the Medical Record and Patient Self-Report as Data Sources for Ambulatory Care? *Medical Care* 2006;44:132-40.
20. Muggah E, Graves E, Bennett C, et al. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. *BMC Public Health* 2013;13.
21. Hansen H, Schäfer I, Schön G, et al. Agreement between self-reported and general practitioner-reported chronic conditions among multimorbid patients in primary care - results of the MultiCare Cohort Study. *BMC Fam Pract* 2014;15.
22. Chun H, Kim IH, Min KD. Accuracy of Self-reported Hypertension, Diabetes, and Hypercholesterolemia: Analysis of a Representative Sample of Korean Older Adults. *Osong Public Health Res Perspect* 2016;7:108-15.
23. Sim J, Wright CC. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Phys Ther* 2005;85:257-68.
24. Viera AJ, Garrett JM. Understanding Interobserver Agreement: The Kappa Statistic. *Fam Med* 2005;37:360-3.
25. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977;33:159-74.
26. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;46:423-9.
27. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551-8.
28. Wu CS, Lai MS, Gau SS, et al. Concordance between Patient Self-Reports and Claims Data on Clinical Diagnoses, Medication Use, and Health System Utilization in Taiwan. *PLoS ONE* 2014;9:e112257.
29. Flack VF, Afifi AA, Lachenbruch PA, et al. Sample size determinations for the two rater kappa statistic. *Psychometrika* 1988;53:321-5.
30. Sheather S. *A Modern Approach to Regression with R*. Springer New York: New York, US; 2009.
31. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria; 2017.
32. Bocquet D. Living in Singapore: Housing Policies between Nation Building Processes, Social Control and the Market. *Territorio* 2015;35-43.
33. Chen G, Faris P, Hemmelgarn B, et al. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol* 2009;9.
34. Simpson CF, Boyd CM, Carlson MC, et al. Agreement Between Self-Report of Disease Diagnoses and Medical Record Validation in Disabled Older Women: Factors That Modify Agreement. *J Am Geriatr Soc* 2004;52:123-7.
35. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22:276-82.
36. Li Y, Geiss LS, Burrows NR, et al. Awareness of Prediabetes — United States, 2005–2010. *MMWR Morb Mortal Wkly Rep* 2013;62:209-12.
37. Zhuang Q, Wu L, Lu Y, et al. Awareness and intervention status of prediabetes among Chinese adults: implications from a community-based investigation. *Int J Clin Exp Med* 2015;8:4480-6.
38. Chatterjee JS. From compliance to concordance in diabetes. *J Med Ethics* 2006;32:507-10.
39. Tabák AG, Herder C, Rathmann W, et al. Prediabetes: A high-risk state for developing diabetes. *Lancet* 2012;379:2279-90.
40. Raben A, Vestentoft PS, Brand-Miller J, et al. The PREVIEW intervention study: Results from a 3-year randomized 2 x 2 factorial multinational trial investigating the role of protein, glycaemic index and physical activity for prevention of type 2 diabetes. *Diabetes Obes Metab* 2021;23:324-37.
41. Asharani PV, Lau JH, Roystonn K, et al. Health Literacy and Diabetes Knowledge: A Nationwide Survey in a Multi-Ethnic Population. *Int J Environ Res Public Health* 2021;18:9316.
42. Lix LM, Yogendran MS, Shaw SY, et al. Population-based data sources for chronic disease surveillance. *Chronic Dis Can* 2008;29:31-8.
43. Tsopra R, Wyatt JC, Beirne P, et al. Level of accuracy of diagnoses recorded in discharge summaries: A cohort study in three respiratory wards. *J Eval Clin Pract* 2019;25:36-43.
44. Peng M, Eastwood C, Boxill A, et al. Coding reliability and agreement of international classification of disease, 10th revision (ICD-10) codes in emergency department data. *Int J Popul Data Sci* 2018;3:445.
45. Teh R, Doughty R, Connolly M, et al. Agreement between self-reports and medical records of cardiovascular disease in octogenarians. *J Clin Epidemiol* 2013;66:1135-43.
46. Tsai AC, Chang TL. Quality issues of self-report of hypertension: Analysis of a population representative sample of older adults in Taiwan. *Arch Gerontol Geriatr* 2012;55:338-42.
47. Gupta V, Gu K, Chen Z, et al. Concordance of self-reported and medical chart information on cancer diagnosis and treatment. *BMC Med Res Methodol* 2011;11.