

BMJ Open Comparison of self-reported and biomedical data on hypertension and diabetes: findings from the China Health and Retirement Longitudinal Study (CHARLS)

Meng Ning,¹ Qiang Zhang,¹ Min Yang^{1,2}

To cite: Ning M, Zhang Q, Yang M. Comparison of self-reported and biomedical data on hypertension and diabetes: findings from the China Health and Retirement Longitudinal Study (CHARLS). *BMJ Open* 2016;**6**:e009836. doi:10.1136/bmjopen-2015-009836

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-009836>).

Received 26 August 2015
Revised 16 November 2015
Accepted 18 November 2015



CrossMark

¹West China School of Public Health, Sichuan University, Chengdu, China

²School of Medicine, University of Nottingham, Nottingham, UK

Correspondence to

Dr Min Yang;
yangmin2013@scu.edu.cn

ABSTRACT

Objectives: We examined the level of agreement between biomedical and self-reported measurements of hypertension and diabetes in a Chinese national community sample, and explored associations of the agreement and possible contextual effects among provinces and geographic regions in China.

Design: Secondary analysis of a cohort sample.

Setting and participants: Community samples were drawn from the national baseline survey of the China Health and Retirement Longitudinal Study (CHARLS, 2011–2012) through multistage probability sampling, which included households with members 45 years of age or above with a total sample size of 17 708 individuals.

Outcome measures: Sensitivity, specificity and κ were used as measurements of agreements or validity; variance of validity measures among provinces and communities was estimated using random-effects models.

Results: Self-reports for hypertension and diabetes showed high specificity (96.3% and 98.3%, respectively) but low sensitivity (56.3% and 61.5%, respectively). Agreement between self-reported data and biomedical measurements was moderate for both hypertension (κ 0.57) and diabetes (κ 0.65), with respondents who were older, of higher socioeconomic status, better educated and who had hospital admissions in the past 12 months showing stronger agreements than their counterparts. Large and significant variations in the sensitivity among provinces for hypertension, and among communities for both hypertension and diabetes, could neither be attributed to the effects of respondents' characteristics nor to the contextual effects of city–village differences.

Conclusions: As a considerable number of people in the overall sample were unaware of their conditions, self-reports will lead to an underestimation of the prevalence of hypertension and diabetes. However, in more developed communities or provinces, self-reported data can be a reliable estimate of the prevalence of the two conditions. Further investigations of contextual effects at provincial and community levels could highlight public health strategies to improve awareness of the two conditions.

Strengths and limitations of this study

- The results of the present study, with a response rate of over 80%, are based on the latest national cohort (China Health and Retirement Longitudinal Study, CHARLS) undertaken with international collaborators in accordance with internationally acceptable standards.
- This is the first study in China to use advanced multilevel modelling to examine potential context effects, such as provincial and community effects, to increase the reliability and generalisability of the findings.
- The reliability of the data could be affected by the fact that the three measurements of blood pressures were taken within a single physical examination rather than by way of multiple measurements over time, and that only one glycated haemoglobin test for diabetes was used in the CHARLS baseline survey.
- Also, there were missing blood sample data from about one-third of the eligible participants. However, the statistical power in the estimates of validity measures and associations are less likely affected in the large sample analyses.

INTRODUCTION

Hypertension and diabetes are two major chronic conditions that contribute considerably to the burden of disease.¹ Reliable prevalence estimates based on biomedical measurements of both diseases are essential for research on the determinants of health and for the planning and implementation of public health policy aimed at prevention. Given the high costs and low efficiency of clinical diagnosis in survey studies, much of the epidemiological investigation of hypertension and diabetes has been and, particularly at the national level, will likely continue to be, based on self-reported data.²

To use self-reported data with confidence, a number of studies have attempted to assess the accuracy of self-reported data using test characteristics such as sensitivity, specificity and κ coefficients. However, their approaches and results vary.^{3–14} Criterion standards for the presence of diseases have been inconsistent, such as biomedical measurements from physical examinations or medical records, and data have been collected in different ways: through mailed self-report questionnaires, telephone-based interviews and face-to-face interviews, using a variety of questionnaire instruments. The characteristics of the cohorts used in these studies also varied widely, and many of the research results on this topic seem to have been based on small, unrepresentative samples.

Almost all studies on the sensitivity and specificity of self-reporting for non-communicable diseases, such as hypertension and diabetes, have been carried out in Western Europe and North America, but health service surveys have become critical sources of data in developing countries, too. Recent studies in China^{15–16} have provided sensitivity estimates equivalent to ‘awareness of disease’ for diabetes and hypertension. However, at the time of this research, few studies had looked at the factors associated with increased health awareness among diabetes or evaluated the validity of self-reported hypertension, and no studies had analysed national-level data to look specifically for between-region variations in people’s levels of awareness regarding these two conditions.

Although previous reports found relatively low overall awareness rates regarding hypertension in developing countries,¹⁷ high awareness rates were observed in some studies conducted, particularly in developed regions of China.¹⁸ Since the management of non-communicable diseases has become part of China’s policy agenda and part of the large-scale reform of China’s national health system in recent years, it is more important than ever to understand the extent to which self-reported cases reflect actual, diagnosable cases of disease in China and to understand how differing levels of disease awareness might be accounted for by variations in geographic region or by other factors.

The analysis presented here, using data from the China Health and Retirement Longitudinal Study (CHARLS), estimates the validity of self-reported hypertension and diabetes in a nationally representative sample of China’s ageing population. We seek not only to determine the respondents’ characteristics associated with increased validity of self-reported hypertension and diabetes, but also to explore the variation in the validity of such data between different geographic regions in China by using multilevel models.¹⁹

METHODS

Data

Data for this study are drawn from the national baseline survey of CHARLS, a 10-year, up-to-date follow-up study

run by the National School for Development (China Center for Economic Research). This nationally representative longitudinal survey was conducted between June 2011 and March 2012, and includes 17 708 individuals. Samples of households with members 45 years of age or above were chosen through multistage probability sampling, but non-community-based individuals or those in collective dwellings, such as military bases, schools, dormitories or nursing homes, were excluded. A total of 435 villages and urban communities from 28 provinces were selected so that the survey could include a mix of urban and rural settings, as well as a wide variety of levels of economic development. The respondent rate of CHARLS is 80.5%, with an 8.8% refusal, 8.2% inability to contact sample residents and 2% other reasons. The included information is from three sources: (1) a questionnaire administered through face-to-face interviews in the sampled household; (2) a physical examination carried out by trained interviewers in the household and (3) laboratory blood tests after the interview at designated local healthcare institution. The interview used a computer-assisted personal interview approach by trained interviewers, to minimise interviewer bias. Detailed information about the CHARLS sampling procedure and data quality management has been published previously.^{20–21}

Self-reported data on hypertension and diabetes were obtained by the question, ‘Have you been diagnosed with hypertension/diabetes by a doctor?’. Blood pressure was measured three times (approximately 45 s apart) on a single occasion, using an electronic monitor (Omron model HEM-7112), as part of the physical examination. The average of these blood pressure readings was used to determine each respondent’s blood pressure level. Hypertension was defined as a systolic blood pressure ≥ 140 mm Hg and/or a diastolic blood pressure ≥ 90 mm Hg and/or current use of antihypertensive medication, following the WHO guideline.²² Blood samples provided estimates of glycated haemoglobin (HbA1c), which were obtained using a boronate affinity high-performance liquid chromatography assay. The diagnostic criterion for diabetes in our study was defined as HbA1c values $\geq 6.5\%$.²³ Although HbA1c may not be the most widely used screening test, it has been suggested as an alternative means of screening for diabetes and has been used this way in many surveys.²³ Biomedical measurements corrected for current medication use are referred to in the remainder of this paper simply as ‘biomedical measurements’.

Statistical methods

To assess the difference in prevalence estimates based on the data collection method used, the prevalence of hypertension and diabetes was calculated according to self-reported information, as well as according to the results of biomedical measurements obtained from the CHARLS. To assess the accuracy of self-reported data, sensitivity, specificity and κ coefficients were each

calculated with SAS software V.9.2. Results of the biomedical measurements were treated as reference or the 'gold standard' for a diagnosis of hypertension or diabetes. Sensitivity was defined as the percentage of respondents who self-reported having hypertension/diabetes among those with diagnosed hypertension/diabetes. This value is thus equivalent to 'hypertension/diabetes awareness'. Specificity was defined as the percentage of individuals who self-reported not having hypertension/diabetes among those with 'normal' or 'healthy' biomedical measurements. As the κ coefficient takes into account the agreement occurring by chance, and since it is a more robust measure than simple percent agreement, we calculated it to determine the degree of agreement between self-reported questionnaire data and biomedical measurements.⁷ The κ coefficients can be divided into four grades, namely ≤ 0.40 (poor-to-fair agreement), $0.41\sim 0.60$ (moderate agreement), $0.61\sim 0.80$ (substantial agreement) and $0.81\sim 1.0$ (excellent agreement).

A 95% CI was calculated for sensitivity and specificity estimates across different subgroups. To determine which participant characteristics independently predicted the likelihood of correctly self-reporting the condition, multivariate logistic regression analysis was considered. Because the multistage sampling design of the CHARLS produces samples with respondents nested within 435 communities (neighbourhoods in urban and villages in rural areas) and communities nested within 28 provinces, the outcomes of interest are inherently geographically clustered. Hence we used a multilevel logistic model in which respondents were at level 1, communities at level 2 and provinces at level 3. Such a model can examine associated factors with the outcome in its fixed part of parameters, and decompose total variance into different levels in its random part of parameters to measure the clustering effects if they exist. The model can be simplified to a two-level model if the variation, in this case between provinces, is not significant. Second order predictive/penalised quasi-likelihood estimation,¹⁹ as implemented within MLwiN software V.2.3, was used to examine associated factors for sensitivity and specificity whenever appropriate. Participants with missing data on a particular variable were listwise excluded from that analysis, although the fraction of missing values on relevant variables is very small in CHARLS ($<0.2\%$). The analyses were carried out separately for participants with and without a diagnosis in order to examine sensitivity and specificity, respectively.

Independent variables were derived from questionnaires and were included in the fixed part of models as both associated factors with and adjustment for the likelihood of sensitivity and specificity. Variables at respondent level included age in years, gender, education (organised into three levels: no formal education, primary education only, secondary education and higher) and hospital admission in the past 12 months. Annual household spending was used as the measure of

economic status and categorised into three levels: low (household spending of $<¥1000/\text{year}$), moderate ($¥1000\text{--}¥5000$) and high ($¥5000$ or more). According to National Bureau of Statistics of China, community type, a community-level variable, was categorised into city, village and others, which consisted of combined urban-rural areas, town centre areas, combined town-township areas, township centre areas and special districts.

RESULTS

Study population

Among the 17 708 respondents in the CHARLS cohort, the percentages of people younger than 50, over 50, and 60 years of age or above, were 24.4%, 35.9% and 39.6%, respectively, with a median age of 58 years (25th and 75th centile: 51–65 years). The cohort was 52.1% female and 47.9% male. Among rural respondents, 33.8% lacked formal education, and 42.2% had never gone beyond elementary school, whereas among urban residents, only 9.9% lacked formal education and only 25.4% had not continued their education beyond elementary school. Of the people in the CHARLS cohort, 9% had at least one hospital admission during the past 12 months, 13 610 (76.9%) had anthropometric data collected as part of the study and 11 546 (65.2%) supplied blood samples. Characteristics of the cohort sample are summarised in [table 1](#).

Prevalence, sensitivity, specificity and κ coefficient

The prevalence of hypertension was 38.5% according to biomedical test and 24.0% according to the self-reported data, which indicated that self-reporting led to an underestimation of hypertension by 38%. Likewise, the prevalence of diabetes was 6.9% according to the biomedical data and 5.8% according to self-report, which indicated an underestimated prevalence of diabetes by 16% from self-reported data. The overall sensitivity of self-reported hypertension and diabetes was 56.3% and 61.5%, respectively. The difference in sensitivity between subgroups for each of the interested factors was significant to varying degrees ([table 2](#)). The overall specificities of self-reported hypertension and diabetes were high, 96.3% and 98.3%, with no significant difference between subgroups of interested factors ([table 2](#)). The overall κ coefficient of self-reported hypertension was 0.57 and of diabetes, 0.65. The difference in κ coefficient between subgroups of interested factors was significant to varying degrees ([table 2](#)). Individuals in the lowest economic group had the lowest sensitivity of self-reporting for both conditions, 45.7% for hypertension and 39.2% for diabetes. Individuals <50 years of age had the next lowest sensitivity of self-reported hypertension (49.4%) and the third lowest for diabetes (54.9%). In contrast, individuals living in a city or having had a hospital admission in the past 12 months or coming from good economic conditions had higher sensitivity of self-reported hypertension than their counterparts. A similar

Table 1 Characteristics of the Cohort, China Health and Retirement Longitudinal Study, 2011–2012

	Total % (N=17 708)	Physical examination participants % (N=13 610)	Blood samples test participants % (N=11 546)
Age (years)			
<50	24.42	23.05	22.84
50–59	35.90	36.16	36.58
≥60	39.68	40.79	40.59
Sex			
Male	47.88	46.89	46.49
Female	52.12	53.10	53.50
Education			
No formal education	27.21	28.34	28.75
Primary education	39.38	40.77	40.51
Secondary education and above	33.41	30.89	30.74
Hospital admission			
Yes	9.06	8.77	9.06
No	90.94	91.23	90.94
Economic status			
<1000	13.55	14.40	13.54
¥1000–4999	35.83	37.77	37.02
≥5000	42.51	41.62	42.99
Not clear	8.11	6.22	6.45
Community type			
City	15.41	11.87	11.60
Others	28.60	28.78	28.55
Village	55.99	59.35	59.86

pattern was presented for self-reported diabetes, too. It was noticed that female individuals had higher sensitivity of self-reported hypertension but lower sensitivity of self-reported diabetes than males.

Respondent characteristics associated with sensitivity

The partial effects of respondents' characteristics on the sensitivity of self-reported hypertension based on analysis of a multilevel multivariate logistic model are presented in [table 3](#). The results further confirmed that individuals aged ≥50 years, living in city areas, having had a hospital admission in the past 12 months and being in good economic status, are strongly and independently associated with having more accurately self-reported hypertension than their counterparts. For self-reported diabetes, age and community type were not associated factors. Female participants with higher levels of education were also more likely to accurately self-report hypertension than were their respective counterparts.

The parameter estimates of the random part in the model without adjusting for independent variables ([table 3](#)) suggested that variations in the sensitivity of self-reported hypertension between communities and provinces were statistically significant, and remained so even after controlling for a range of individual factors, as well as the contextual effects of city–village differences. To examine provincial effects, we calculated the residuals from the model without independent variables for all 28 provinces in the sample and produced a 'caterpillar plot' ([figure 1](#)) that showed provincial effects ranked in order, as well as the 95% CIs for each value.

The sensitivity of self-reported hypertension in Guizhou, Guangdong, Guangxi, Yunnan, Hubei and Sichuan, were significantly lower than the overall average (residuals fell below the zero line), and they showed accurate self-reporting of hypertension at rates of 25.6%, 39.4%, 39.5%, 41.3%, 44.5% and 46.8%, respectively. In contrast, the sensitivity of self-reported hypertension in Beijing, Zhejiang and Anhui, were significantly above average (residuals fell above the zero line) with accurate hypertension self-reporting rates of 85.7%, 68.0% and 69.5%, respectively. Although accurate hypertension self-reporting rates were higher in Shanghai, Tianjin and Heilongjiang, at 85.0%, 74.5% and 72.2%, respectively, the wide CIs for these values overlapped with the zero line because small sample sizes in these provinces led to large SEs for their estimated residuals. A map was created showing the accuracy of self-reported hypertension (equivalent to 'awareness of hypertension') in provinces of China ([figure 2](#)), which indicates that provinces in northern regions have more accurate hypertension self-reports than those in the southwest region of the country. To explore possible sources of variation among provinces, we further examined the difference among three officially defined economic zones in China: the Eastern zone, which consists of 11 provinces; the Central zone, which consists of 8 provinces; and the Western zone, which consists of 12 provinces. The Eastern zone is economically most advanced, and the Western zone is the least economically developed, with distinct policy and geographic circumstances. Although not reported in tables, an adjusted model analysis

Table 2 Sensitivity, specificity and κ coefficients of self-reported hypertension and diabetes compared with biomedical data, China Health and Retirement Longitudinal Study, 2011–2012

Variables	Hypertension			Diabetes		
	Sensitivity % (95% CI)	Specificity % (95% CI)	κ (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	κ , (95% CI)
Overall	56.3 (55.0 to 57.7)	96.3 (95.9 to 96.7)	0.57 (0.55 to 0.58)	61.5 (58.1 to 64.8)	98.3 (98.1 to 98.6)	0.65 (0.62 to 0.68)
Age (years)						
<50	49.4 (45.8 to 53.0)	96.9 (96.2 to 97.6)	0.54 (0.50 to 0.58)	54.9 (46.1 to 63.8)	98.7 (98.2 to 99.1)	0.59 (0.51 to 0.66)
50–59	57.6 (55.3 to 60.0)	96.6 (96.0 to 97.2)	0.59 (0.57 to 0.62)	62.6 (57.1 to 68.0)	98.4 (98.0 to 98.8)	0.66 (0.61 to 0.71)
≥60	57.4 (55.6 to 59.2)	95.4 (94.6 to 96.2)	0.52 (0.50 to 0.55)	62.7 (57.8 to 67.6)	98.1 (97.7 to 98.5)	0.65 (0.61 to 0.7.0)
Sex						
Male	53.6 (51.5 to 55.6)	96.4 (95.8 to 97.0)	0.54 (0.52 to 0.57)	63.4 (58.3 to 68.6)	98.7 (98.4 to 99.0)	0.67 (0.63 to 0.72)
Female	58.7 (56.9 to 60.5)	96.2 (95.6 to 96.7)	0.58 (0.57 to 0.60)	60.0 (55.5 to 64.5)	98.0 (97.7 to 98.4)	0.63 (0.59 to 0.66)
Education						
No formal education	53.1 (50.7 to 55.5)	95.6 (94.7 to 96.4)	0.51 (0.49 to 0.54)	53.6 (47.0 to 60.1)	98.2 (97.7 to 98.7)	0.57 (0.52 to 0.63)
Primary education	57.2 (55.1 to 59.3)	96.5 (95.8 to 97.1)	0.58 (0.55 to 0.60)	66.9 (61.6 to 72.2)	98.4 (98.0 to 98.8)	96.36 (0.64 to 0.73)
Secondary education and above	59.0 (56.4 to 61.5)	96.6 (96.0 to 97.3)	0.61 (0.58 to 0.63)	61.9 (56.1 to 67.7)	98.4 (98.0 to 98.8)	0.66 (0.61 to 0.71)
Hospital admission						
Yes	70.1 (66.2 to 73.9)	94.6 (92.8 to 96.3)	0.66 (0.62 to 0.70)	76.2 (68.8 to 83.6)	97.6 (96.6 to 98.6)	0.76 (0.70 to 0.82)
No	54.7 (53.3 to 56.2)	96.4 (96.0 to 96.8)	0.55 (0.54 to 0.57)	58.7 (55.0 to 62.4)	98.4 (98.2 to 98.7)	0.62 (0.59 to 0.66)
Economic status						
<1000	45.7 (42.2 to 49.1)	95.8 (94.7 to 97.0)	0.45 (0.41 to 0.49)	39.2 (28.1 to 50.3)	98.8 (98.2 to 99.4)	0.46 (0.35 to 0.57)
¥1000–4999	55.2 (53.1 to 57.4)	96.4 (95.7 to 97.0)	0.55 (0.53 to 0.58)	59.4 (53.6 to 65.1)	98.3 (97.9 to 98.7)	0.62 (0.57 to 0.67)
≥5000	62.6 (60.5 to 64.7)	96.1 (95.5 to 96.8)	0.63 (0.61 to 0.65)	66.9 (62.4 to 71.5)	98.3 (97.9 to 98.7)	0.7 (0.66 to 0.73)
Not clear	51.9 (47.1 to 56.8)	97.7 (96.5 to 98.9)	0.54 (0.48 to 0.59)	61.5 (46.3 to 76.8)	98.2 (97.2 to 99.2)	0.61 (0.48 to 0.74)
Community type						
City	67.6 (64.3 to 70.9)	96.8 (95.6 to 98.0)	0.65 (0.62 to 0.69)	69.4 (62.2 to 76.5)	97.1 (96.2 to 98.1)	0.69 (0.63 to 0.76)
Others	56.0 (53.5 to 58.5)	96.7 (95.9 to 97.4)	0.56 (0.54 to 0.59)	61.2 (55.4 to 67.0)	98.4 (98.0 to 98.9)	0.66 (0.61 to 0.71)
Village	53.6 (51.8 to 55.4)	96.0 (95.5 to 96.0)	0.54 (0.52 to 0.56)	58.2 (53.1 to 63.3)	98.5 (98.2 to 98.8)	0.61 (0.57 to 0.66)

Table 3 Respondent characteristics associated with the sensitivity and specificity of self-reports in multilevel logistic regression modelling, presented by OR and its 95% CI, China Health and Retirement Longitudinal Study, 2011–2012

	Hypertension OR (95% CI)		Diabetes OR (95% CI)	
	With condition (sensitivity)	Without condition (specificity)	With condition (sensitivity)	Without condition (specificity)
<i>Model without covariates</i>				
Level 3: province	0.218 (0.072)**	0.176 (0.083)*	0.192 (0.121)	0.063 (0.070)
Level 2: community	0.193 (0.039)**	0.283 (0.112)*	0.641 (0.208)**	0.618 (0.207)**
<i>Model with covariates</i>				
Fixed part				
Age (years)				
<50	1	1	1	1
50–59	1.54 (1.28 to 1.86)**	0.88 (0.64 to 1.20)	1.56 (0.95 to 2.56)	0.80 (0.51 to 1.26)
≥60	1.62 (1.35 to 1.95)**	0.66 (0.48 to 0.91)*	1.59 (0.96 to 2.63)	0.66 (0.42 to 1.04)
Sex				
Male	1	1	1	1
Female	1.42 (1.25 to 1.61)**	0.70 (0.48 to 1.01)	0.94 (0.64 to 1.38)	0.68 (0.49 to 0.95)*
Education				
No formal education	1	1	1	1
Primary education	1.3 (1.12 to 1.51)**	1.05 (0.79 to 1.39)	1.79 (1.16 to 2.77)**	1.07 (0.72 to 1.58)
Secondary education and above	1.3 (1.08 to 1.56)**	1.01 (0.69 to 1.48)	1.34 (0.81 to 2.22)	1.10 (0.68 to 1.78)
Hospital admission				
Yes	1.73 (1.41 to 2.12)**	0.89 (0.70 to 1.13)	2.10 (1.26 to 3.49)**	0.72 (0.45 to 1.15)
No	1	1	1	1
Economic status				
<1000	1	1	1	1
¥1000–4999	1.42 (1.19 to 1.69)**	1.15 (0.82 to 1.62)	2.22 (1.21 to 4.06)**	0.68 (0.40 to 1.15)
≥5000	1.72 (1.43 to 2.07)**	1.06 (0.77 to 1.46)	2.94 (1.61 to 5.38)**	0.73 (0.43 to 1.24)
Not clear	1.07 (0.82 to 1.40)	2.01 (1.02 to 3.95)*	2.39 (0.95 to 6.02)	0.68 (0.33 to 1.40)
Community type				
City	1.57 (1.25 to 1.97)**	1.18 (0.75 to 1.86)	1.70 (0.98 to 2.96)	0.49 (0.30 to 0.80)**
Others	1.08 (0.91 to 1.28)*	1.22 (0.91 to 1.64)	1.08 (0.69 to 1.69)	0.91 (0.60 to 1.38)
Village	1	1	1	1
Random part†				
Level 3: province	0.145 (0.052)**	0.175 (0.082)*	–	–
Level 2: community	0.163 (0.037)**	0.245 (0.108)*	0.779 (0.217)**	0.647 (0.207)**

Significant differences are indicated as * $p < 0.05$, ** $p < 0.01$.

†Estimate and SE in logit scale.

indicated that incidences of self-reported hypertension in the Eastern and Central zones of China were more accurate than those in the Western zone (OR=1.53, 95% CI 1.07 to 2.09; OR=1.57, 95% CI 1.09 to 2.25). In fact, provinces with highly accurate self-reported hypertension were mostly located in the Eastern zone (Beijing, Zhejiang, Shanghai and Tianjin) or in the Central zone (Anhui, Heilongjiang). In contrast, provinces with the lowest self-reported hypertension were located in the Western zone (Guizhou, Guangxi, Yunnan and Sichuan).

For self-reported diabetes, variance estimates in the random part of the model without covariates (table 3) suggested that variations in the accuracy of diabetes self-reporting between communities were statistically significant, while those between provinces were not. As was the case in the full model adjusted for respondent characteristics, the estimate of between-community variance and its SE was 0.598 (0.209) with $p < 0.01$, while the between-

province variance and SE was 0.197 (0.124) with $p > 0.05$. The intracommunity correlation was calculated as 19.5% ($((0.197+0.598)/(0.197+0.598+3.29))$), which was higher than the calculation of the hypertension as 8.6%. Because there appeared to be no significant variation between provinces, no further examination by economic zone was carried out, and we reduced our covariate adjustment model down to the two-level model presented in table 3. The intracommunity correlation was then recalculated as 0.191 ($(0.779/(0.779+3.29))$), indicating that 19.1% of the residual variation in the accuracy of people's self-reported diabetes could be attributed to unobserved community characteristics.

Respondent characteristics associated with specificity

The overall specificity of self-reported cases of both, hypertension and diabetes, was high. The results in table 3 suggest that no respondent characteristic was significantly associated with more accurate reporting

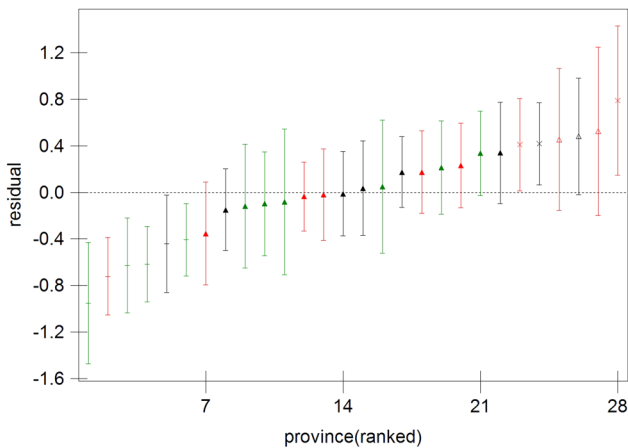


Figure 1 Plot of province residuals with 95% CIs for sensitivity of self-reported hypertension, China Health and Retirement Longitudinal Study, 2011–2012 (sequencing from left to right: Guizhou, Guangdong, Guangxi, Yunnan, Hubei and Sichuan tagged by ‘-’; Anhui, Zhejiang and Beijing tagged by ‘x’; Heilongjiang, Tianjin and Shanghai tagged by ‘△’). Three separate colours—red, black and green—indicating provinces located in the Eastern zone, Central zone and Western zone, respectively).

except for age group, where elderly people were more likely to erroneously report the absence of hypertension than those younger than 50 years of age. As [figure 3](#) shows, only two provinces indicated significant variation in the specificity of self-reported data (ie, they did not overlap the horizontal line at zero): the specificity of self-reported data in Anhui (92.5%) was significantly below average, while the specificity of self-reported data from Guangxi (99.7%) was significantly above average.

Figure 2 Map showing the estimated awareness rates of hypertension in each province of China, China Health and Retirement Longitudinal Study, (CHARLS) 2011–2012. The colour from light to dark indicates four categories: category 1 (light blue): 25.6–50.0%; category 2: 50.1–58%; category 3: 58.1–66.0% and category 4 (dark blue): 66.1–85.7%. Three colours are used to mark the edge of provinces—red, black and green—indicating provinces located in the Eastern zone, Central zone and Western zone, respectively. Tibet, Ningxia and Hainan were not included in CHARLS and are left as blank on the map.

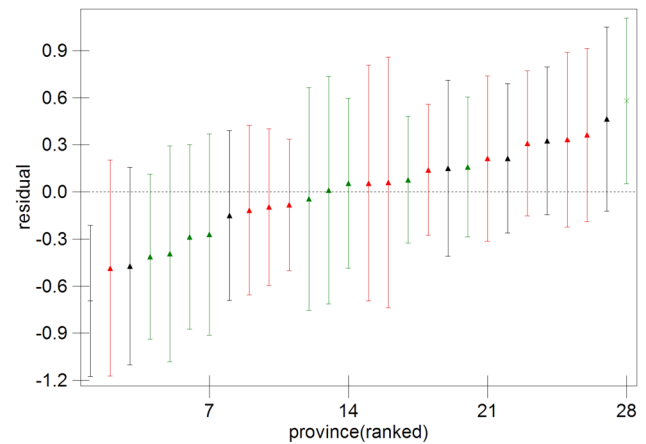
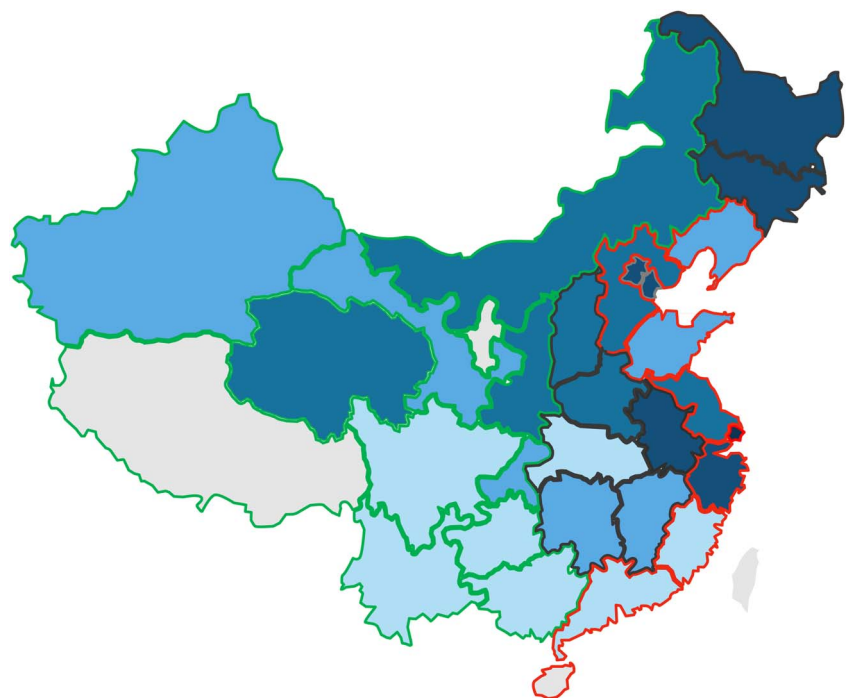


Figure 3 Plot of province residuals with 95% CIs for specificity of self-reported hypertension, China Health and Retirement Longitudinal Study, 2011–2012 (Anhui tagged by ‘-’; Guangxi tagged by ‘x’). Three separate colours—red, black and green—indicating provinces located in the Eastern zone, Central zone and Western zone, respectively).

The specificity of self-reported diabetes was slightly higher than that of hypertension. Of all participants without diabetes, 98.3% accurately reported an absence of the condition. Multivariate analyses showed that men and village residents had slightly more accurate reporting than their counterparts ([table 3](#)).

DISCUSSION

Although self-reports tend to underestimate the prevalence of hypertension and diabetes, our study showed moderate agreement between participants’ questionnaire responses and their biomedical measurements for

both conditions among Chinese people. And yet, while most respondents with normal measurements classified themselves correctly, nearly half of all respondents with abnormal measurements did not report their conditions. Individuals' age, community type, hospital care in the last year and economic status are factors associated with accurate self-reports of both conditions.

Sensitivity and specificity

When comparing our findings to those of other studies, it is important to bear in mind that the cohort characteristics, study methodologies and results vary widely among them.⁷ For example, previous studies^{3 5 6 8 10 24} showed that using biomedical measurements as reference would reveal a range in the sensitivity of self-reported hypertension from 34.5% to 83.9%, a range in specificity from 90.0% to 98.0% and a trend where higher sensitivity of self-reported hypertension correlated with the lower specificity of such self-reports.

In comparison with a comparable study in the USA,⁶ which used data from the 2006 and 2008 Health and Retirement Study, and showed that hypertension self-reports had a sensitivity of 83.9% and a specificity of 92%, our study presented a much lower sensitivity of 56.3% and a higher specificity of 96.3%. After raising the threshold values for systolic and diastolic blood pressure from the current definition of 140/90 mm Hg to the 160/95 mm Hg value, the sensitivity of the self-reported hypertension in our study increased markedly from 56.3% to 76.4%, and its specificity decreased slightly from 96.3% to 94.9%. These changes suggest that a considerable proportion of people with blood pressure between 140/90 and 160/95 mm Hg were not aware of their condition, and the validity of self-reported hypertension cases may depend, at least in part, on the proportion of people with mild hypertension. It is worth mentioning that the diagnostic criteria for hypertension changed in 1999 from 160/95 to 140/90 mm Hg. Nonetheless, some medical staff may continue to use the old, higher cut-off value (160/95 mm Hg) in their diagnoses and thereby wrongly classify current hypertensive patients as not having the disease.⁵

A study in Taiwan³ also provides an interesting case of comparison for our study. It included a substantial elderly population of illiterate individuals and those lacking formal education. The accuracy of self-reported hypertension in our study was greater than that in Taiwan, with a sensitivity, specificity and κ coefficient of 56.3%, 96.3% and 0.57, respectively, in comparison to Taiwan values of 49.4%, 95.3% and 0.41, respectively. In contrast, the accuracy of self-reported diabetes in our study was lower than that in Taiwan, with the three measures at 61.5%, 98.3% and 0.65, in comparison to 85.2%, 98.3% and 0.86 in Taiwan, respectively. It is important to note that the diagnostic criterion for diabetes in our study was defined as HbA1c values $\geq 6.5\%$. This is lower than the criterion used in the Taiwan study, where it was defined as $\geq 7.0\%$. Using an HbA1c

threshold of 6.5% is quite common in studies involving diabetes diagnosis.^{12 14} When raising the threshold values of HbA1c from the definition used in our study of 6.5% to the 7.0%, the sensitivity and κ coefficient of diabetes self-reports in the CHARLS data set would increase markedly to 71.5% and 0.70%, and the specificity would decrease slightly from 98.3% to 98.0%, respectively. Even with these adjustments, the accuracy of self-reported diabetes in our study would still be lower than that in the Taiwan study. The reason for this is not clear, but one possible explanation is that the different socioeconomic conditions of sampled areas in these two studies may have impacted their results. As observed in our analysis, better economic conditions correlated with more accurately self-reported information. Our findings also supported other previous research indicating that the accuracy of self-reported diabetes tends to be higher than that of hypertension.^{3 5 13} In other studies that used biomedical measurements as reference,^{3 5 12-14} the accuracy of self-reported diabetes varied in sensitivity from 67.1% to 83.9%, and varied in specificity from 96.0% to 99.6%. They all showed an overall correlation between higher sensitivity and lower specificity. Compared to these previous studies, the accuracy of self-reported diabetes in our study is relatively low.

The κ coefficient of paradoxes was known, that is, high agreement but low κ , or lack of predictability,²⁵ which could only be traced to certain conditions,²⁶ such as a substantial imbalance in the fourfold table's marginal totals, or used in rare diseases. We felt that our data did not have either condition, and hence resulted in proper use of the measure.

Associates with sensitivity and specificity

As in previous studies,^{3 5 8} our research showed that factors associated with more accurate self-reported data were older age, female sex, recent utilisation of medical services, higher levels of education and better economic conditions.

An early study of the accuracy of self-reported hypertension in China was carried out in 2004.²⁴ This study reported the sensitivity of hypertension self-reports among people ≥ 50 and ≥ 60 years of age was 33.8% and 39.9%, respectively. Our study was conducted 7 years later, in 2011, and produced corresponding age-based, self-reported hypertension sensitivity estimates of 57.6% and 57.4%. Thus, we can say that the awareness of hypertension among elderly Chinese people appears to have improved significantly in recent years with a relatively stable sensitivity observed among those ≥ 50 years of age. A possible explanation for this could be that elderly people tend to have more frequent contacts with healthcare systems, in general. As a result, they might be better informed about their chronic conditions than younger individuals, who may only rarely use health services, are. Our results also confirmed the findings of previous studies indicating that individuals who had used healthcare recently are more likely to self-report their

conditions correctly.^{3 5 9} These findings may contribute to the implementation of China's national primary healthcare programme at the community level, where individual health records are taken as mandatory, and citizens are entitled to receive free health check-ups, including blood pressure monitoring.¹⁶ However the sensitivity of hypertension self-reports in younger age groups appeared to be lower and to increase with age. This finding is consistent with many published studies.^{5 8} Our findings suggest that young and middle-aged Chinese people—the working age population—were less aware of their hypertensive conditions. This destabilises the validity of self-reported hypertension prevalence rates and suggests that hypertension estimates may actually fluctuate widely in response to public awareness programmes up until high levels of awareness are attained in the population at large.⁹

Not surprisingly, our study showed that self-reports for both diseases tend to be more valid among those living in urban areas or cities than in rural areas or villages. This could be due to the fact that China's cities have better economic resources and healthcare accessibility than its countryside areas. Socioeconomic factors may also influence the quality and quantity of healthcare received by those living in rural areas, particularly if they have insufficient medical services available and lack of health knowledge regarding self-care due to their generally lower education levels. One previous study showed that the rate of accurate hypertension self-reporting between the ages of 15 and 69 years in Shanghai was 85.5%, which closely matches our finding that this figure was 85%, and it suggests that reliably valid hypertension self-reports can be attained in large and developed Chinese cities.¹⁸ At the same time, raising awareness, diagnosing and treating early, creating programmes for disease screening and increasing education around hypertension are all still needed in rural areas.

Variation among provinces and among communities

The study found a considerably large variation in awareness of hypertension among provinces, ranging from 25.6% in Guizhou to 85.7% in Beijing, and a much higher awareness of hypertension in the Eastern zone of China than in the Central and Western zones. Even after adjusting for the contextual effects of urban–rural differences and individual characteristics, variations in the awareness rates for both diseases between communities were still significant. Such variations may be attributed to unobserved contextual factors in communities of participants, such as level of health education, performance of healthcare institutions, accessibility of healthcare resources and the degree of economic development within the community, etc. The much higher variance among communities in the awareness of diabetes than that of hypertension could be attributable to the contextual effects of community characteristics that were far more apparent in the accuracy of diabetes self-reports, which may be due to the fact that the accessibility and

affordability of diabetes diagnosis is much more strongly influenced by community environment. This finding also implies that, in developed communities or provinces, especially in Eastern China, self-reported information might be reliably used for hypertension surveillance.

STUDY LIMITATIONS

Our study had several potential limitations. First, it relied on three biomedical measurements taken during a single physical examination instead of using multiple measurements over longer time intervals. This may have led to biased estimates of the prevalence of hypertension. Second, the HbA1c test is an imperfect standard for ascertaining the presence of diabetes. Although HbA1c is increasingly used to screen and diagnose diabetes in clinical settings, further follow-up tests are typically used to confirm these diagnoses.¹⁴ Third, of the people in the CHARLS cohort, 23.1% of the interviewees had missing anthropometric measurements, and 35% did not provide blood samples. Although the overall response rate of 80.5% was reasonably good, the response rate of 76.9% and 65% for the two outcome measures, respectively, could potentially have caused a selection bias in the data. Women are more likely to get physical examinations and to have blood samples taken, whereas younger men are less likely to do so. It is possible that younger men were more likely to be working than were women, and thus were less likely to complete these aspects of the CHARLS. It is important to also note, however, that many of our study's potential limitations were mitigated by the fact that CHARLS interviewees are fairly evenly distributed in terms of many of their background characteristics and the quality of our study was guaranteed by the stringent quality control implemented in every stage of CHARLS.²⁰ Finally, elderly residents of nursing homes were excluded in the CHARLS baseline data. This is unlikely to cause a major problem, however, because the fraction of nursing home residents is very small in China,²⁰ with an estimate of 1.5% among elderly population.²⁷

CONCLUSION

The results of this study suggest that the agreement of self-reported data with biomedical measurements is moderate for both diabetes and hypertension among the population over 45 years of age and older in China. Although the awareness of those conditions among individuals is positively associated with older people, living in cities, better education and higher economic status—the large variation in the awareness of the two conditions between communities and provinces cannot be explained by those characteristics. Further research should try to explore the contextual effects at both community and provincial levels, which may generate evidence for developing effective intervention strategies to further improve the population's awareness regarding these conditions in China.

As a considerable number of people were unaware of their conditions, and the degree of inaccuracy in self-reporting varied between subgroups, self-reported data are likely leading to an underestimation of the prevalence of the two diseases in China. They might also bias the observed associations between self-reported diseases and the characteristics of individuals. However, in developed communities or provinces, self-reported data can be a reliable source for the prevalence of the two conditions.

Acknowledgements The authors would like to thank the CHARLS team, who collected data and assisted with data access for the study, Lillian Pureher of Washington University, who edited and improved the English in the manuscript, and the journal reviewers whose valuable comments helped to shape the manuscript to meet the high standards of the journal.

Contributors MN and MY developed the study design. MN was responsible for data management, statistical analysis and drafting of this paper. MY was responsible for the theoretical and methodological framework and provided critical knowledge in drafting of the paper. QZ provided oversight and advice for the design and interpretation of the statistical analyses.

Funding Part of MY's time was supported by China Medical Board under a grant (CMB 12-106). The original CHARLS was supported by the Behavioral and Social Research division of the National Institute on Aging of the National Institute of Health (grants 1-R21-AG031372-01, 1-R01-AG037031-01 and 3-R01AG037031-03S1); the National Science Foundation of China (grants 70773002, 70910107022 and 71130002), the World Bank (contracts 7145915 and 7159234) and Peking University.

Competing interests None declared.

Patient consent Obtained.

Ethics approval The original CHARLS was approved by the Ethical Review Committee of Peking University, and all participants signed informed consent at the time of participation.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

1. Yang GH, Kong LZ, Zhao WH, *et al.* Emergence of chronic non-communicable diseases in China. *Lancet* 2008;372:1697–705.
2. Wu F, Guo Y, Kowal P, *et al.* Prevalence of major chronic conditions among older Chinese adults: the Study on Global AGEing and Adult Health (SAGE) wave 1. *PLoS ONE* 2013;8:e74176.
3. 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.
4. Alonso A, Beunza JJ, Delgado-Rodriguez M, *et al.* Validation of self-reported diagnosis of hypertension in a cohort of university graduates in Spain. *BMC Public Health* 2005;5:94.
5. Molenaar EA, Van Ameijden EJ, Grobbee DE, *et al.* Comparison of routine care self-reported and biometrical data on hypertension and diabetes: results of the Utrecht Health Project. *Eur J Public Health* 2007;17:199–205.
6. White K, Avendano M, Capistrant BD, *et al.* Self-reported and measured hypertension among older US- and foreign-born adults. *J Immigrant Minor Health* 2012;14:721–6.
7. 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.
8. Taylor A, Dal Grande E, Gill T, *et al.* Comparing self-reported and measured high blood pressure and high cholesterol status using data from a large representative cohort study. *Aust N Z J Public Health* 2010;34:394–400.
9. Vargas CM, Burt VL, Gillum RF, *et al.* Validity of self-reported hypertension in the National Health and Nutrition Examination Survey III, 1988–1991. *Prev Med* 1997;26:678–85.
10. de Menezes TN, Oliveira EC, de Sousa Fischer MA. Validity and concordance between self-reported and clinical diagnosis of hypertension among elderly residents in northeastern Brazil. *Am J Hypertens* 2014;27:215–21.
11. Hansen H, Schafer I, Schon 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;15:39.
12. Goto A, Morita A, Goto M, *et al.* Validity of diabetes self-reports in the Saku diabetes study. *J Epidemiol* 2013;23:295–300.
13. Huerta JM, Tormo MJ, Egea-Caparrós JM, *et al.* Accuracy of self-reported diabetes, hypertension and hyperlipidemia in the adult Spanish population. DINO study findings. *Rev Esp Cardiol* 2009;62:143–52.
14. White K, Mondesir FL, Bates LM, *et al.* Diabetes risk, diagnosis, and control: do psychosocial factors predict hemoglobin A1c defined outcomes or accuracy of self-reports? *Ethn Dis* 2014;24:19–27.
15. Li M, Su L, Liang B, *et al.* Trends in prevalence, awareness, treatment, and control of diabetes mellitus in Mainland China from 1979 to 2012. *Int J Endocrinol* 2013;2013:753150.
16. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. *Bull World Health Organ* 2014;92:29–41.
17. Basu S, Millett C. Social epidemiology of hypertension in middle-income countries: determinants of prevalence, diagnosis, treatment, and control in the WHO SAGE study. *Hypertension* 2013;62:18–26.
18. Li XJ, Xu JY, Yao HH, *et al.* Epidemiological status of hypertension in residents aged between 15 and 69 in Shanghai. *Chin J Prev Contr Chron Dis* 2010;18:233–5.
19. Rodriguez G, Goldman N. Improved estimation procedures for multilevel models with binary response: a case-study. *J Royal Stat Soc A* 2001;164:339–55.
20. Zhao Y, Hu Y, Smith JP, *et al.* Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). *Int J Epidemiol* 2014;43:61–8.
21. Zhao Y, Strauss J, Yang G, *et al.* China health and retirement longitudinal study—2011–2012 national baseline users' guide. Beijing: Peking University, 2013. http://charls.ccer.edu.cn/uploads/document/2011-charls-wave1/application/Chinese_users_guide_20130407_.pdf
22. Global status report on noncommunicable diseases 2010. Geneva: World Health Organization, 2011.
23. Bennett CM, Guo M, Dharmage SC. HbA(1c) as a screening tool for detection of type 2 diabetes: a systematic review. *Diabetic Med* 2007;24:333–43.
24. Wang Y, Jiang Y, Zhang M, *et al.* Validity of self-reported hypertension among Chinese adults. *Zhonghua Liu Xing Bing Xue Za Zhi* 2009;30:414–16.
25. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;43:543–9.
26. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: resolution of the two paradoxes. *J Clin Epidemiol* 1996;49:431–4.
27. Chu LW, Chi I. Nursing homes in China. *J Am Med Dir Assoc* 2008;9:237–43.