

RESEARCH ARTICLE

Probit Models to Investigate Prevalence of Total Diagnosed and Undiagnosed Diabetes among Aged 45 Years or Older Adults in China

Minghui Yin, Balekouzou Augustin, Chang Shu, Tingting Qin, Ping Yin*

Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

* pingyintj@163.com



OPEN ACCESS

Citation: Yin M, Augustin B, Shu C, Qin T, Yin P (2016) Probit Models to Investigate Prevalence of Total Diagnosed and Undiagnosed Diabetes among Aged 45 Years or Older Adults in China. PLoS ONE 11(10): e0164481. doi:10.1371/journal.pone.0164481

Editor: Susanne Kaser, Medical University Innsbruck, AUSTRIA

Received: April 2, 2016

Accepted: September 25, 2016

Published: October 10, 2016

Copyright: © 2016 Yin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: China Health and Retirement Longitudinal Study (CHARLS), harmonized with the U.S. Health and Retirement Study (HRS) family of surveys, is publicly available and de-identified. The data file is available at <https://figshare.com/s/1a140270fe2f813652eb>.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

The aims of this study are to identify the most important predictors of total diagnosed and undiagnosed diabetes and estimate the mean change in the predicted probability among aged 45+ adults in China. We used baseline data collected from 2011 wave of the China Health and Retirement Longitudinal Study (CHARLS) (n = 9,513). First, we estimated the prevalence of diagnosed, measured, total diagnosed, and undiagnosed diabetes. Second, we used probit models to determine whether individual attributes, socioeconomic characteristics and behavioral health factors, including smoking, alcohol consumption, obesity, central obesity, are associated with total diagnosed and undiagnosed diabetes. We also consider other factors, including contact with medical system, hypertension and urban/rural settings. Third, we estimated average marginal effects of variables in probit models. Among Chinese people aged 45+, the prevalence of diagnosed, measured, total diagnosed and undiagnosed diabetes were 5.8% (95%CI, 5.3%-6.3%), 14.7% (95%CI, 14.0%-15.4%), 17.0% (95%CI, 16.3%-17.7%), 11.3% (95%CI, 10.6%-12.0%), respectively. The probability of total diagnosed diabetes is 3.3% (95% CI, 1.2%-5.3%) and 10.2% (95% CI, 7.0%-13.5%) higher for overweight and obesity than normal BMI, 5.0% (95% CI, 3.0%-7.1%) higher for central obesity than normal waist circumference, 5.4% (95% CI, 3.7%-7.0%) higher for hypertensive than normotensive and 1.8% (95% CI, 0.8%-2.7%) higher in urban areas than in rural areas, respectively. The probability of undiagnosed diabetes is 2.7% (95% CI, 1.2%-4.2%) and 7.2% (95% CI, 4.7%-9.6%) higher for overweight and obesity than normal BMI, 2.6% (95% CI, 0.9%-4.4%) higher for central obesity than normal waist circumference and 2.6% (95% CI, 1.2%-4.0%) higher for hypertensive than normotensive, respectively, and -1.5% (95% CI, -2.5% to -0.5%) lower for individuals who were in contact with the medical system. Greater focus on prevention of diabetes is necessary for obesity, central obesity, hypertensive and in urban areas for middle-aged and older in China.

Introduction

Cardiovascular disease has become the leading cause of death in China, a development that has followed rapid economic growth, an increase in life expectancy, and changes in lifestyle.[1] Diabetes is one of the leading causes of mortality and morbidity worldwide, which increases the risk for cardiovascular and kidney diseases[2], and the prevalence of diabetes is high and increasing in China.[3–6]

In China, approximately 114 million, or 11.6% of adults, had diabetes in 2010.[6] However, less than one-third (30.1%) of these patients were aware of their disease condition[6], and this may affect the estimate of diabetes prevalence. The management of diabetes in China has been ineffective for many years.[7] Although these studies have documented a marked increase in the prevalence of diabetes in China and many researchers have studied the risk factors related to diabetes, there has been little investigation to study risk factors leading to the effects of higher probability of diabetes.

In many survey studies with a component or a focus on health, prevalence of diabetes is according to "self-reports": respondents are asked to report whether they have diabetes at present and/or have ever been diagnosed with diabetes. However, the actual prevalence, based on self-reports, might be seriously underestimated. This is because the signs and symptoms of type 1 diabetes are usually obvious and develop very quickly, often over a few weeks, but type 2 diabetes tends to develop more slowly, usually over a period of months or even years, and prevalence of undetected or undiagnosed diabetes are generally high.[8]

An effective way to assess the extent to which diabetes is undiagnosed is to collect both objective measures and self-reported for the same respondents. This is, however, rarely done in survey studies. China Health and Retirement Longitudinal Study (CHARLS) collected both of the information.

Our first aim is to identify the most important factors leading to higher probability of total diagnosed and undiagnosed diabetes using baseline data from a national survey among aged 45 + adults in China. Our second aim is to estimate the mean change in the predicted probability.

Methods

Data

We used data collected from the CHARLS of Chinese people aged 45 years or older.[9, 10] CHARLS, harmonized with the U.S. Health and Retirement Study (HRS) family of surveys, is publicly available and de-identified.[10] The data were collected in a survey in which four-stage, stratified, cluster sampling was used to select eligible individuals. A total of 17,708 individual participants (10,069 aged 45 years or older and 7,639 spouses of eligible individuals) completed a computer-assisted personal interview (CAPI) in their home.[9–11] A structured questionnaire with several main sections was used to collect data from each respondent. Some demographical variables were gathered by a face-to-face interview. As explained in detail by Zhao et al.[9]

For our analysis, we used the first wave (2011) of CHARLS which was conducted between June 2011 and March 2012 among aged 45 years or older adults in China. Information on various aspects of the respondents' lives were collected, including individual characteristics, behavioral health, health status, the economic dimension, the social domain, blood pressure measurements and urban/rural settings. Also, measurement of blood glucose and/or HbA1c was collected. Of 17,708 respondents, 6,173 were not measured glucose and glycosylated hemoglobin, 101 without self-reported diabetes answers, 40 without measured blood pressure, 22 missing gender and education, 1,620 without BMI and waist circumference measurements, 239

without smoking and drinking answers. The final sample size was 9,513 for our study. The coded data is available in [S1 Table](#).

Measures of diabetes prevalence

Diagnosed. CHARLS collected information on individual self-reports of specific conditions with the general question: "Have you been diagnosed with diabetes or high blood sugar by a doctor?". We classified respondents as having "diagnosed diabetes" if they answered "yes" to the questions.

Measured. The key advantage in using data collected from CHARLS is that blood samples were measured in the survey. Nearly two-thirds blood samples of individuals were collected by medically trained staff from the China Center for Disease Control and Prevention. Participants were asked to fast overnight. After collection, plasma for glucose assay was separated from blood samples and stored at -20°C , [10] and whole blood for HbA1C assay was stored immediately and during shipment at 4°C . [10] All the blood samples were transported within 2 weeks to the China Center for Disease Control and Prevention, where samples were placed at -80°C in a deep freezer before assay. [10] Blood assays were performed at the Youanmen Center for Clinical Laboratory of Capital Medical University during February 2013 and June 2013. [10] The laboratory used quality control samples daily during the testing of the CHARLS study samples, and all test results were within the target range (within two SDs of mean quality control concentrations). Glucose was measured using an enzymatic colorimetric test, and HbA1c was analyzed using boronate affinity chromatography. [10]

We used a binary variable for measured diabetes according to measurement of blood glucose and/or HbA1c. We divided respondents as diabetics if their fasting plasma glucose ≥ 126 mg/dL and/or HbA1c $\geq 6.5\%$. The cut-off points for diagnosis of diabetes were based on current recommendations from the American Diabetes Association. [12]

Total diagnosed. The total diagnosis coded respondents as diabetics if either they self-reported to be diabetic and/or the measured value above the diagnostic threshold, which is fasting plasma glucose ≥ 126 mg/dL and/or HbA1c $\geq 6.5\%$.

Undiagnosed. We divided respondents as having "undiagnosed diabetes" if they did not report having been told by a doctor that they have diabetes but were diabetics according to the more comprehensive total diagnosis. The prevalence of undiagnosed diabetes is the fraction of total prevalence that is not diagnosed.

Control variables. Our study included age, gender and marital status as individual attributes. We used two variables of socioeconomic status: level of education and household income. We classified level of education as illiterate, primary education, secondary education and at least college level based on Chinese education system. As for the income, we employed information from a single comprehensive question about income according to all household members. We corrected household income via dividing it by the square root of the number of persons in the household. [13–15] Then, we assigned individuals into corresponding income tertiles: first tertile (low income), second tertile (middle income) and third tertile (high income). [16]

CHARLS collected information on several health-related behaviors. We identified three categories for smoking: current smoker, past smoker and never smoked according to the respondents' tobacco use answer; and respondents were grouped into three categories of drinking—never, less than once a month, more than once a month according to their alcohol consumption answer.

The survey also collected information on body mass index (BMI), which is the ratio of weight in kilograms to height in meters squared. We use BMI to identify whether respondents

are underweight/normal ($BMI < 23$), overweight ($BMI \geq 23$ and < 27.5) and obese ($BMI \geq 27.5$) based on the suggested categories for Asian populations.[17]

We defined the respondents as central obesity according to the male's waist circumference ≥ 90 cm or the female's waist circumference ≥ 80 cm based on the International Diabetes Federation suggested categories for Chinese populations.[18]

We also included a variable meant to capture the extent of "contact with the medical system". We created a dichotomous variable equal to one if the respondent had visited a public hospital, private hospital, public health center, clinic, or health worker's or doctor's practice, or been visited by a health worker or doctor for outpatient care or having received inpatient care at least once.

We coded respondents as hypertensive if either they self-reported to be hypertensive and/or had blood pressure value above the diagnostic threshold, which is $SBP \geq 140$ mmHg and/or $DBP \geq 90$ mmHg.

We also considered each respondent's urban/rural settings according to he or she was recorded as a long-term urban resident or long-term rural resident.

Although other risk factors might be important, we used some key variables based on other studies in our study.

Statistical methods

We used unweighted probit models of determinants of total diagnosed and undiagnosed diabetes prevalence. We used R version 3.2.3 (R Core Team 2015, Vienna, Austria) to conduct the analyses[19] and set a prior level of significance at 0.05. We transformed the parameter estimates of probit models to estimates of average marginal effects AMEs.[20–23] AMEs are an effective means by which the effects of variables in nonlinear models can be made more intuitively meaningful.[24] Briefly, the AMEs of a categorical variable is the mean change in the predicted probability that the outcome is equal to one as the categorical variable changes from 0 to 1, holding all other covariates at their observed values.[25]

Results

Baseline characteristics

Table 1 listed the prevalence of diagnosed, measured, total diagnosed and undiagnosed diabetes and 95% confidence intervals. The prevalence of diagnosed, measured, total diagnosed and undiagnosed diabetes were 5.8% (95%CI, 5.3%-6.3%), 14.7% (95%CI, 14.0%-15.4%), 17.0% (95%CI, 16.3%-17.7%), 11.3% (95%CI, 10.6%-12.0%), respectively. The total prevalence of diabetes is high in China. Table 1 also indicated levels in key diabetes risk factors.

Predictors of diabetes prevalence

Table 2 listed estimates of probit models for total diagnosed and undiagnosed diabetes prevalence. The model has two functions. The first is to identify the key factors leading to higher probability of total diagnosed and undiagnosed diabetes, and the second is to estimate AMEs of variables in probit models (Table 2, Figs 1 and 2).

In China, for 45+ adults, age, overweight or obesity, central obesity, "contact with the medical system", hypertension and urban area were positively associated with the probability of total diagnosed diabetes. The probability of being diabetic based on the more comprehensive total diagnosis was higher for individuals who were older, overweight or obese, central obesity, contact with the medical system, hypertensive and living in urban area.

Table 1. Diabetes prevalence rates and diabetes risk factors in China (2011–2012): 45+ years old.

	China (%)
Diabetes prevalence	
Diagnosed	5.8 (5.3, 6.3)
Measured	14.7 (14.0, 15.4)
Total diagnosed	17.0 (16.3, 17.7)
Undiagnosed	11.3 (10.6, 12.0)
Individual characteristics	
Age, mean	59.2
Male	45.7 (44.7, 46.7)
Married	87.7 (87.1, 88.3)
Behavioural health	
Smoking	
Non-smoker	61.4 (60.4, 62.4)
Past smoker	8.7 (8.2, 9.2)
Current smoker	29.9 (28.9, 30.9)
Alcohol consumption ^a	
Abstainer	67.6 (66.7, 68.5)
Current drinker (>0 but ≤1)	7.8 (7.3, 8.3)
Current drinker (>1)	24.5 (23.7, 25.3)
BMI category	
Under/normal weight	48.3 (47.3, 49.3)
Overweight	37.7 (36.7, 38.7)
Obesity	14.0 (13.4, 14.6)
Waist Circumference	
Normal waist circumference	48.7 (47.7, 49.7)
Central obesity	51.3 (50.3, 52.3)
Socioeconomic gradient	
Education	
Illiterate	28.9 (28.0, 29.8)
Primary	41.5 (40.5, 42.5)
Secondary	28.3 (27.4, 29.2)
College and above	1.3 (1.1, 1.5)
Adjusted household income ^b	
First tercile	33.4 (32.4, 34.4)
Second tercile	35.3 (34.3, 36.3)
Third tercile	31.4 (30.4, 32.4)
Medical system	
Not contact with medical system	72.0 (71.1, 72.9)
Contact with medical system	28.0 (27.1, 28.9)
Blood pressure ^c	
Normotension	59.2 (58.2, 60.2)
Hypertension	40.8 (39.9, 41.7)
Urban/rural settings	
Rural area	65.0 (64.1, 65.9)
Urban area	35.0 (34.0, 36.0)

(Continued)

Table 1. (Continued)

	China (%)
N	9,513

BMI body mass index

^a Drinking episodes per month.

^b First tertile (low income), second tertile (middle income) and third tertile (high income).

^c Hypertensive, if either they self-reported to be hypertensive and/or had a blood pressure value above the diagnostic threshold, which is SBP \geq 140 mmHg and/or DBP \geq 90 mmHg.

doi:10.1371/journal.pone.0164481.t001

Age, overweight or obesity, central obesity and hypertension were positively associated with the probability of undiagnosed diabetes. "Contact with the medical system" was negatively associated with the probability of undiagnosed diabetes. The probability of undiagnosed diabetes was higher for who were older, overweight or obese, central obesity and hypertensive, and lower for individuals who were contact with the medical system.

A number of differences in the sign or significance of diabetes risk factors also emerged. Age had significant influence on diabetes. Notably, the probability of total diagnosed diabetes is 3.3% (95% CI, 1.2%-5.3%) and 10.2% (95% CI, 7.0%-13.5%) higher for overweight and obesity than normal BMI, 5.0% (95% CI, 3.0%-7.1%) higher for central obesity than normal waist circumference, 5.4% (95% CI, 3.7%-7.0%) higher for hypertensive than normotensive and 1.8% (95% CI, 0.8%- 2.7%) higher for in urban areas than in rural areas, respectively. The probability of undiagnosed diabetes is 2.7% (95% CI, 1.2%-4.2%) and 7.2% (95% CI, 4.7%-9.6%) higher for overweight and obesity than normal BMI, 2.6% (95% CI, 0.9%-4.4%) higher for central obesity than normal waist circumference and 2.6% (95% CI, 1.2%-4.0%) higher for hypertensive than normotensive, respectively, and -1.5% (95% CI, -2.5% to -0.5%) lower for individuals who were contact with the medical system.

To investigate further on the four key factors of the results discovered by the probit models: BMI, waist circumference, hypertension and urban/rural settings. We compared the prevalence of total diagnosed diabetes for each group (Fig 3). The prevalence of total diagnosed diabetes for underweight/normal, overweight and obesity were 12.8%, 18.6% and 27.7%, respectively. The prevalence of total diagnosed diabetes for normal waist circumference and central obesity were 12.5% and 21.3%, respectively. The prevalence of total diagnosed diabetes for normotensive and hypertensive were 13.4% and 22.4%, respectively. The prevalence of total diagnosed diabetes for rural area and urban area were 15.9% and 19%, respectively.

Discussion

Probit models is an importance as a method to determine key factors on public health and used to estimate AMEs of variables. In this study, using CHARLS data on ageing, we took advantage of AMEs in probit models to investigate the prevalence of total diagnosed and undiagnosed diabetes among Chinese people aged 45 years or older. We identified the most important predictors of diabetes leading to higher probability of diabetes through probit models. It is worth noting that diabetes is a major risk factor for cardiovascular disease, and the prevalence of diabetes is high and is increasing in China. China has also strengthened monitoring and adopted diabetes prevention strategies.

Our results show that, based on the more comprehensive total diagnosis that combines diagnosed and measured diabetes, the prevalence of diabetes is high in China. Six results are outstanding to the predictors of diabetes prevalence for marginal effects in probit models.

Table 2. Average marginal effects (95% confidence intervals) in probit models for diabetes prevalence (total diagnosed and undiagnosed) in China: 45+ years old.

	Total diagnosed		Undiagnosed	
	A.M.E. (95% CI) (%)	p	A.M.E. (95% CI) (%)	p
Individual characteristics				
Age	0.3 (0.2, 0.3)	0.000***	0.2 (0.1, 0.2)	0.000***
Male	1.6 (-0.8, 3.9)	0.192	1.2 (-0.8, 3.2)	0.236
Married	2.1 (-0.2, 4.4)	0.070	1.0 (-0.9, 2.9)	0.317
Behavioral health				
Smoking				
Non-smoker	Ref.		Ref.	
Past smoker	1.6 (-1.5, 4.7)	0.313	1.1 (-1.6, 3.7)	0.426
Current smoker	0.4 (-1.9, 2.6)	0.744	0.9 (-1.1, 2.8)	0.382
Drinking				
Abstainer	Ref.		Ref.	
Current drinker (>0 but ≤1)	-0.6 (-3.5, 2.2)	0.670	0.0 (-2.5, 2.4)	0.982
Current drinker (>1)	-0.7 (-2.7, 1.3)	0.518	1.0 (-0.8, 2.7)	0.272
BMI category				
Under/normal weight	Ref.		Ref.	
Overweight	3.3 (1.2, 5.3)	0.002**	2.7 (1.2, 4.2)	0.000***
Obesity	10.2 (7.0, 13.5)	0.000***	7.2 (4.7, 9.6)	0.000**
Waist Circumference				
Normal waist circumference	Ref.		Ref.	
Central obesity	5.0 (3.0, 7.1)	0.000***	2.6 (0.9, 4.4)	0.003**
Health status				
Medical system				
Not contact with medical system	Ref.		Ref.	
Contact with medical system	2.7 (1.0, 4.4)	0.002**	-1.5 (-2.5, -0.5)	0.005**
Blood pressure				
Normotension	Ref.		Ref.	
Hypertension	5.4 (3.7, 7.0)	0.000***	2.6 (1.2, 4.0)	0.000***
Socioeconomic gradient				
Education				
Illiterate	Ref.		Ref.	
Primary	0.8 (-1.2, 2.7)	0.434	-0.2 (-1.9, 1.4)	0.773
Secondary	1.2 (-1.2, 3.6)	0.315	0.5 (-1.5, 2.5)	0.638
College and above	-1.9 (-8.4, 4.6)	0.568	-2.8 (-7.9, 2.4)	0.294
Adjusted household income				
First tercile	Ref.		Ref.	
Second tercile	-0.9 (-2.7, 0.9)	0.336	-0.6 (-2.1, 0.9)	0.423
Third tercile	-0.3 (-2.2, 1.7)	0.784	-1.4 (-3.0, 0.2)	0.087
Urban/rural settings				
Rural area	Ref.		Ref.	
Urban area	1.8 (0.8, 2.7)	0.000***	-0.7 (-2.1, 0.7)	0.309

AME average marginal effect, CI confidence interval, REF reference category the AME of a categorical variable is the mean change in the predicted probability that the outcome is equal to one as the categorical variable changes from 0 to 1, holding all other covariates at their observed values. To illustrate, the probability of being measured is 4.6 percentage points higher for Central obesity than Normal waist circumference (95% CI, 2.8–6.4), holding all other covariates at their observed values.

* p < 0.05,
 ** p < 0.01,
 *** p < 0.001

doi:10.1371/journal.pone.0164481.t002

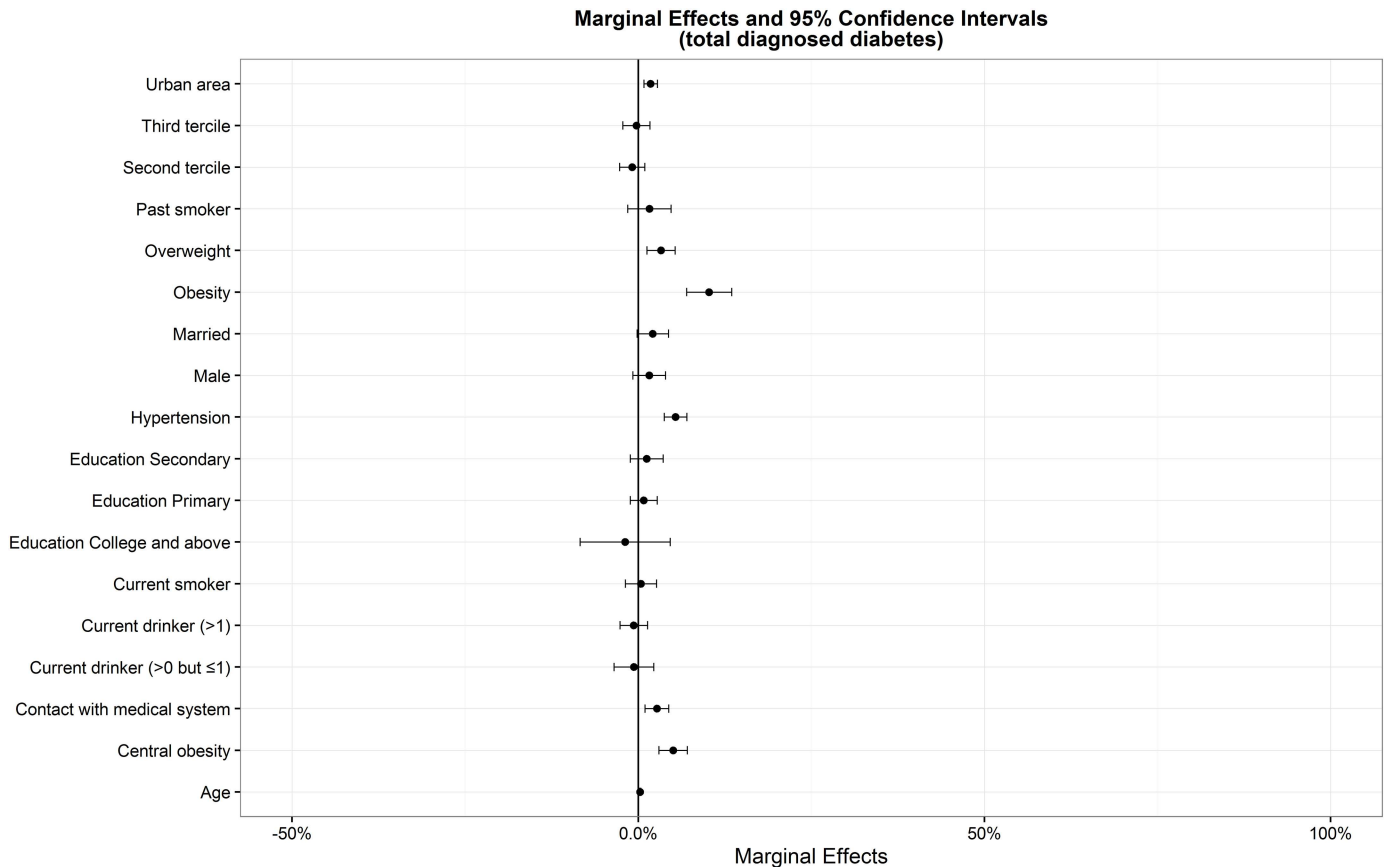


Fig 1. Average marginal effects and 95% confidence intervals from probit models for the prevalence of total diagnosed diabetes.

doi:10.1371/journal.pone.0164481.g001

First, age is positively associated with the probability of total diagnosed and undiagnosed diabetes, respectively. This finding is in line with many studies [26, 27], as they confirm increasing age was an important risk factor for diabetes.

Second, obesity and central obesity are positively associated with total diagnosed and undiagnosed diabetes, respectively. Being overweight or obese is the main modifiable risk factor for type 2 diabetes, and a large waist circumference is associated with increased likelihood of developing type 2 diabetes. [28] Both a high proportion of body fat and a predominance of central obesity are associated with insulin resistance. [29, 30] The precise mechanisms linking central obesity to insulin resistance remain, however, unclear.

Third, "contact with the medical system" is positively associated with the probability of total diabetes and negatively associated with the probability of undiagnosed diabetes. It may be that the more visits are, the more opportunity physicians detecting undiagnosed diabetes will be, or the causality may be reversed, and those who have already been diagnosed make more visits to monitor treatment effectiveness or refill prescriptions.

Fourth, hypertension is positively associated with total diagnosed and undiagnosed diabetes, respectively. A meta-analysis shows that a high blood pressure could increase the risk of developing type 2 diabetes by around 50%. [31] And there is evidence to suggest that lowering blood pressure significantly reduces diabetes related deaths, strokes, heart failure and microvascular complications. [32, 33] It may be that diabetes and hypertension share common pathways such as sympathetic nervous system, renin-angiotensin-aldosterone system, oxidative stress,

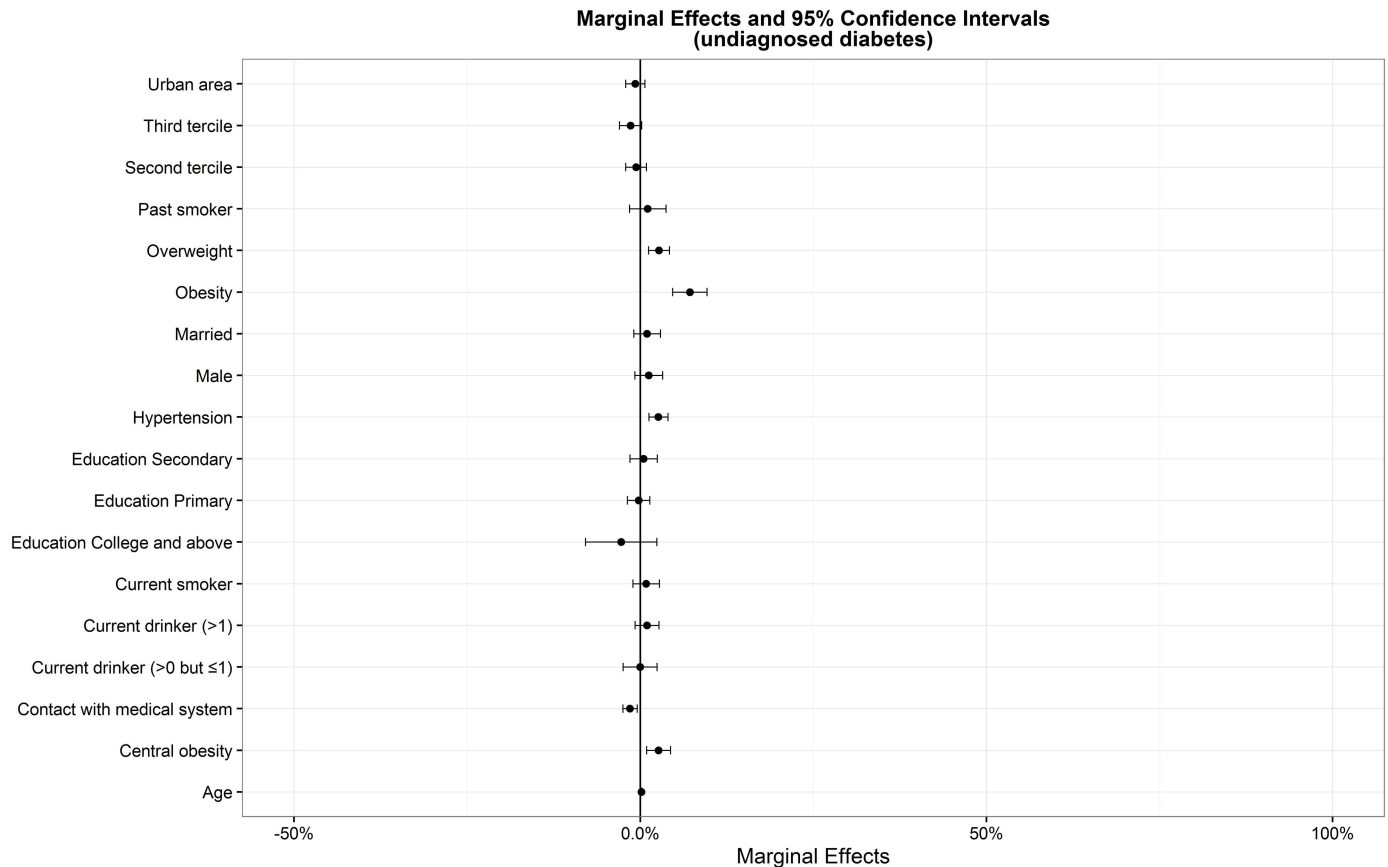


Fig 2. Average marginal effects and 95% confidence intervals from probit models for the prevalence of undiagnosed diabetes.

doi:10.1371/journal.pone.0164481.g002

adipokine, insulin resistance, and peroxisome proliferator-activated receptors. These pathways interact and influence each other and may even cause a vicious cycle. Hypertension and diabetes are both end results of the metabolic syndrome.[34]

Fifth, the probability of total diabetes is higher in urban areas than in rural ones. This finding is similar to Zhou's study[35], which using geographic visualization at the provincial level indicated widespread variation in diabetes prevalence and detection across China. Urban people are more likely to suffer from diabetes than rural ones, the prevalence of diabetes is 13.1% in urban areas, while 8.7% in rural areas.[35] Peng et al.[36] investigated the relationship between changing lifestyles and non-communicable disease, and they found urbanization has led to changes in patterns of human activity, diet, and social structures in China, with profound implications for non-communicable, including diabetes.[36] He and colleagues' study[37] in a geographically and socially isolated ethnic minority group in southwest China provided early evidence of the effect of urbanization on chronic disease.[37] Since then, researchers have suggested that Chinese urban environments promote lifestyles that place people at risk of diabetes, obesity and central obesity.

Finally, education did not play a significant role in predicting the prevalence of total diagnosed and undiagnosed diabetes. A possible explanation is that the higher level of education completed is not correlated with obtaining medical knowledge regarding the prevention and treatment of diabetes. This suggests that educational programs to increase the awareness of diabetes and associated chronic diseases should target adults across all education levels.[38]

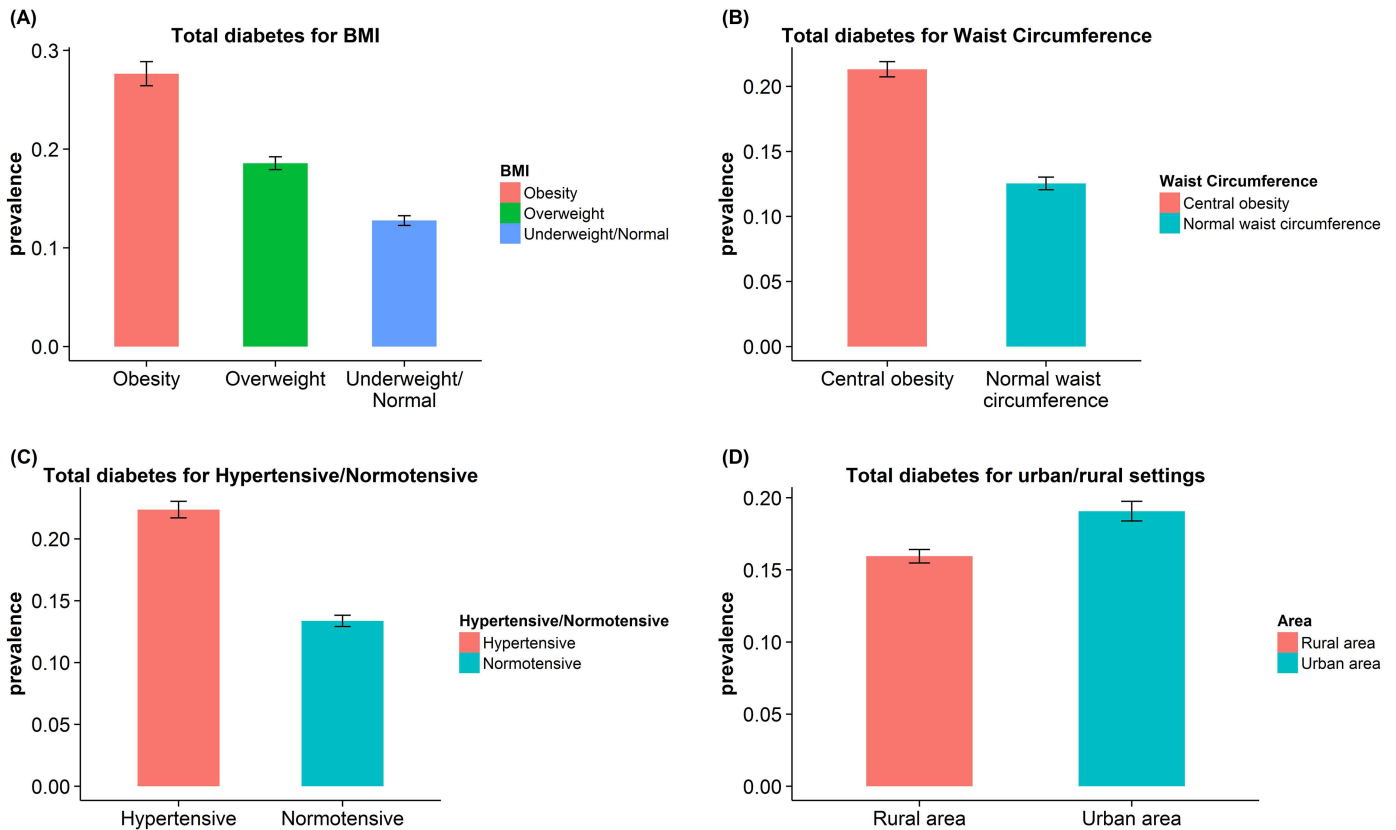


Fig 3. Comparisons of the prevalence of total diabetes by BMI, waist circumference, blood pressure and urban/rural settings.

doi:10.1371/journal.pone.0164481.g003

Turning to boronate affinity chromatography we used, it is possible that the values of HbA1c obtained from the samples which were frozen upon arrival in Beijing are lower than true values. Amount of studies reported that frozen samples can be used, but it is not the standard clinical laboratory practice. A number of studies have shown that the results of HbA1c assays which use the affinity high performance liquid chromatography (HPLC) method are stable on whole blood specimens that have been stored at -70°C for more than 2 decades.[39] However, it has been shown that the results can be affected by the ambient conditions and the length of time whole blood samples are stored during transport from the field to the lab.[40] Storage at freezing temperature has also been related to assay value; the average frozen HbA1c was lower than fresh samples after a year, and the relative error ratio between HbA1c from frozen samples and from fresh samples was significantly higher at high levels of fresh HbA1c.[41]

Our findings should be considered in the context of several limitations.

First, the data analyses are the first wave for CHARLS, we investigated associations and were unable to confidently identify the causes of prevalence of diabetes.

Second, the measured individual attributes and risk factors are incomplete. CHARLS have not collected information on respondents' dietary patterns, which may offer additional explanatory power.

Third, only a random subsample of households was asked about the amount of time you spend on different types of physical activities. Because of this, we could not adequately analyze the effects of physical activity contribute to the outcomes.

Despite these limitations, it is noteworthy that this study is unique in using the survey data by probit models to identify the key factors leading to higher probability of total diagnosed and undiagnosed diabetes, which provided useful information on health outcome for Chinese people aged 45 years or older.

In conclusion, our results indicate that the probability of total diagnosed diabetes among Chinese people aged 45 years or older adults are significantly higher for obesity than normal BMI, higher for central obesity than normal waist circumference, higher for hypertensive than normotensive and higher in urban areas than rural ones, respectively. Our results suggest that greater focus on prevention of diabetes are necessary for obesity, central obesity, hypertensive and in urban areas among middle-aged and older in China.

Supporting Information

S1 Table. Data coding and variable description.
(XLSX)

Acknowledgments

We would like to acknowledge the China Health and Retirement Longitudinal Study (CHARLS) team for providing data. Collection of the data was supported by the National Institute on Aging, Natural Science Foundation of China and the World Bank. The contents of this publication are the sole responsibility of the authors.

Author Contributions

Conceptualization: MHY PY.

Data curation: CS TTQ.

Formal analysis: MHY PY.

Methodology: MHY PY.

Project administration: CS TTQ.

Software: MHY BA.

Writing – original draft: MHY PY.

Writing – review & editing: MHY PY.

References

1. He J, Gu D, Wu X, Reynolds K, Duan X, Yao C, et al. Major causes of death among men and women in China. *New England Journal of Medicine*. 2005; 353(11):1124–34. doi: [10.1056/NEJMsa050467](https://doi.org/10.1056/NEJMsa050467) PMID: [16162883](https://pubmed.ncbi.nlm.nih.gov/16162883/)
2. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *The Lancet*. 2011; 378(9785):31–40. doi: [10.1016/S0140-6736\(11\)60679-X](https://doi.org/10.1016/S0140-6736(11)60679-X) PMID: [21705069](https://pubmed.ncbi.nlm.nih.gov/21705069/)
3. Pan X-R, Yang W-Y, Li G-W, Liu J. Prevalence of diabetes and its risk factors in China, 1994. *Diabetes care*. 1997; 20(11):1664–9. doi: [10.2337/diacare.20.11.1664](https://doi.org/10.2337/diacare.20.11.1664) PMID: [9353605](https://pubmed.ncbi.nlm.nih.gov/9353605/)
4. Group NDR. Diabetes mellitus survey of 300,000 in fourteen provinces and cities of China. *Chin Med J*. 1981; 20:678–81.
5. Gu D, Reynolds K, Duan X, Xin X, Chen J, Wu X, et al. Prevalence of diabetes and impaired fasting glucose in the Chinese adult population: International Collaborative Study of Cardiovascular Disease

- in Asia (InterASIA). *Diabetologia*. 2003; 46(9):1190–8. doi: [10.1007/s00125-003-1167-8](https://doi.org/10.1007/s00125-003-1167-8) PMID: [12879248](https://pubmed.ncbi.nlm.nih.gov/12879248/)
6. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, et al. Prevalence and control of diabetes in Chinese adults. *Jama*. 2013; 310(9):948–59. doi: [10.1001/jama.2013.168118](https://doi.org/10.1001/jama.2013.168118) PMID: [24002281](https://pubmed.ncbi.nlm.nih.gov/24002281/)
 7. Shen H. Effectiveness of a peer-led self-management program for older people with type 2 diabetes in China: Queensland University of Technology; 2008.
 8. Beagley J, Guariguata L, Weil C, Motala AA. Global estimates of undiagnosed diabetes in adults. *Diabetes research and clinical practice*. 2014; 103(2):150–60. doi: [10.1016/j.diabres.2013.11.001](https://doi.org/10.1016/j.diabres.2013.11.001) PMID: [24300018](https://pubmed.ncbi.nlm.nih.gov/24300018/)
 9. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: The China health and retirement longitudinal study (CHARLS). *International journal of epidemiology*. 2012;dys203. doi: [10.1093/ije/dys203](https://doi.org/10.1093/ije/dys203) PMID: [23243115](https://pubmed.ncbi.nlm.nih.gov/23243115/)
 10. Zhao Y, Strauss J, Yang G, Giles J, Hu P, Hu Y, et al. China health and retirement longitudinal study—2011–2012 national baseline users' guide. Beijing: National School of Development, Peking University. 2013.
 11. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. *Bulletin of the World Health Organization*. 2014; 92(1):29–41. doi: [10.2471/BLT.13.124495](https://doi.org/10.2471/BLT.13.124495) PMID: [24391298](https://pubmed.ncbi.nlm.nih.gov/24391298/)
 12. Association AD. Diagnosis and classification of diabetes mellitus. *Diabetes care*. 2010; 33(Supplement 1):S62–S9 doi: [10.2337/dc14-S081](https://doi.org/10.2337/dc14-S081) PMID: [24357215](https://pubmed.ncbi.nlm.nih.gov/24357215/).
 13. Buhmann B, Rainwater L, Schmaus G, Smeeding TM. Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates across Ten Countries Using the Luxembourg Income Study (LIS) Database. *Review of income and wealth*. 1988; 34(2):115–42. doi: [10.1111/j.1475-4991.1988.tb00564.x](https://doi.org/10.1111/j.1475-4991.1988.tb00564.x)
 14. Avendano M, Glymour MM. Stroke Disparities in Older Americans Is Wealth a More Powerful Indicator of Risk Than Income and Education? *Stroke; a journal of cerebral circulation*. 2008; 39(5):1533–40. doi: [10.1161/STROKEAHA.107.490383](https://doi.org/10.1161/STROKEAHA.107.490383) PMID: [18436891](https://pubmed.ncbi.nlm.nih.gov/18436891/)
 15. Avendano M, Glymour MM, Banks J, Mackenbach JP. Health disadvantage in US adults aged 50 to 74 years: a comparison of the health of rich and poor Americans with that of Europeans. *American Journal of Public Health*. 2009; 99(3):540–8. doi: [10.2105/AJPH.2008.139469](https://doi.org/10.2105/AJPH.2008.139469) PMID: [19150903](https://pubmed.ncbi.nlm.nih.gov/19150903/)
 16. Mosca I, Kenny RA. Exploring differences in prevalence of diagnosed, measured and undiagnosed hypertension: the case of Ireland and the United States of America. *International journal of public health*. 2014; 59(5):759–67. doi: [10.1007/s00038-014-0573-7](https://doi.org/10.1007/s00038-014-0573-7) PMID: [24942996](https://pubmed.ncbi.nlm.nih.gov/24942996/)
 17. WHO EC. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet (London, England)*. 2004; 363(9403):157 doi: [10.1016/S0140-6736\(03\)15268-3](https://doi.org/10.1016/S0140-6736(03)15268-3) PMID: [14726171](https://pubmed.ncbi.nlm.nih.gov/14726171/).
 18. Alberti KGM, Zimmet P, Shaw J, Group IETFC. The metabolic syndrome—a new worldwide definition. *The Lancet*. 2005; 366(9491):1059–62. doi: [10.1016/S0140-6736\(05\)67402-8](https://doi.org/10.1016/S0140-6736(05)67402-8) PMID: [16182882](https://pubmed.ncbi.nlm.nih.gov/16182882/)
 19. R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria. Available at: <http://www.R-project.org/>. (Accessed: 30th March 2016).
 20. Christofides LN, Stengos T, Swidinsky R. On the calculation of marginal effects in the bivariate probit model. *Economics Letters*. 1997; 54(3):203–8. doi: [10.1016/S0165-1765\(97\)00025-6](https://doi.org/10.1016/S0165-1765(97)00025-6)
 21. Murphy A. Score tests of normality in bivariate probit models. *Economics Letters*. 2007; 95(3):374–9. doi: [10.1016/j.econlet.2006.11.010](https://doi.org/10.1016/j.econlet.2006.11.010)
 22. Anderson S, Newell RG. Simplified marginal effects in discrete choice models. *Economics Letters*. 2003; 81(3):321–6. doi: [10.1016/S0165-1765\(03\)00212-X](https://doi.org/10.1016/S0165-1765(03)00212-X)
 23. Hasebe T. Marginal effects of a bivariate binary choice model. *Economics Letters*. 2013; 121(2):298–301. doi: [10.1016/j.econlet.2013.08.028](https://doi.org/10.1016/j.econlet.2013.08.028)
 24. Fox J, Hong J. Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software*. 2009; 32(1):1–24.
 25. Delfino RJ, Tjoa T, Gillen DL, Staimer N, Polidori A, Arhami M, et al. Traffic-related air pollution and blood pressure in elderly subjects with coronary artery disease. *Epidemiology (Cambridge, Mass)*. 2010; 21(3). doi: [10.1097/EDE.0b013e3181d5e19b](https://doi.org/10.1097/EDE.0b013e3181d5e19b) PMID: [20335815](https://pubmed.ncbi.nlm.nih.gov/20335815/)
 26. Harris MI, Flegal KM, Cowie CC, Eberhardt MS, Goldstein DE, Little RR, et al. Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in US adults: the Third National Health and Nutrition Examination Survey, 1988–1994. *Diabetes care*. 1998; 21(4):518–24. PMID: [9571335](https://pubmed.ncbi.nlm.nih.gov/9571335/)
 27. King H, Aubert RE, Herman WH. Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. *Diabetes care*. 1998; 21(9):1414–31. PMID: [9727886](https://pubmed.ncbi.nlm.nih.gov/9727886/)

28. Taylor J, Cottrell C, Chatterton H, Hill J, Hughes R, Wohlgemuth C, et al. Identifying risk and preventing progression to type 2 diabetes in vulnerable and disadvantaged adults: a pragmatic review. *Diabetic Medicine*. 2013; 30(1):16–25. doi: [10.1111/dme.12027](https://doi.org/10.1111/dme.12027) PMID: [23002995](https://pubmed.ncbi.nlm.nih.gov/23002995/)
29. Yoon K-H, Lee J-H, Kim J-W, Cho JH, Choi Y-H, Ko S-H, et al. Epidemic obesity and type 2 diabetes in Asia. *The Lancet*. 2006; 368(9548):1681–8. doi: [10.1016/S0140-6736\(06\)69703-1](https://doi.org/10.1016/S0140-6736(06)69703-1) PMID: [17098087](https://pubmed.ncbi.nlm.nih.gov/17098087/)
30. Li S, Xiao J, Ji L, Weng J, Jia W, Lu J, et al. BMI and waist circumference are associated with impaired glucose metabolism and type 2 diabetes in normal weight Chinese adults. *Journal of diabetes and its complications*. 2014; 28(4):470–6. doi: [10.1016/j.jdiacomp.2014.03.015](https://doi.org/10.1016/j.jdiacomp.2014.03.015) PMID: [24809931](https://pubmed.ncbi.nlm.nih.gov/24809931/)
31. Emdin CA, Anderson SG, Woodward M, Rahimi K. Usual blood pressure and risk of new-onset diabetes: evidence from 4.1 million adults and a meta-analysis of prospective studies. *Journal of the American College of Cardiology*. 2015; 66(14):1552–62. doi: [10.1016/j.jacc.2015.07.059](https://doi.org/10.1016/j.jacc.2015.07.059) PMID: [26429079](https://pubmed.ncbi.nlm.nih.gov/26429079/)
32. Group UPDS. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. *BMJ: British Medical Journal*. 1998:703–13.
33. Trivedi RB, Ayotte B, Edelman D, Bosworth HB. The association of emotional well-being and marital status with treatment adherence among patients with hypertension. *Journal of behavioral medicine*. 2008; 31(6):489–97. doi: [10.1007/s10865-008-9173-4](https://doi.org/10.1007/s10865-008-9173-4) PMID: [18780175](https://pubmed.ncbi.nlm.nih.gov/18780175/)
34. Cheung BM, Li C. Diabetes and hypertension: is there a common metabolic pathway? *Current atherosclerosis reports*. 2012; 14(2):160–6. doi: [10.1007/s11883-012-0227-2](https://doi.org/10.1007/s11883-012-0227-2) PMID: [22281657](https://pubmed.ncbi.nlm.nih.gov/22281657/)
35. Zhou M, Astell-Burt T, Bi Y, Feng X, Jiang Y, Li Y, et al. Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults. *Diabetes care*. 2014; DC_141100. doi: [10.2337/dc14-1100](https://doi.org/10.2337/dc14-1100) PMID: [25352654](https://pubmed.ncbi.nlm.nih.gov/25352654/)
36. Gong P, Liang S, Carlton EJ, Jiang Q, Wu J, Wang L, et al. Urbanisation and health in China. *The Lancet*. 2012; 379(9818):843–52. doi: [10.1016/S0140-6736\(11\)61878-3](https://doi.org/10.1016/S0140-6736(11)61878-3)
37. He J, Klag MJ, Whelton PK, Chen J-Y, Mo J-P, Qian M-C, et al. Migration, blood pressure pattern, and hypertension: the Yi Migrant Study. *American journal of epidemiology*. 1991; 134(10):1085–101. PMID: [1746519](https://pubmed.ncbi.nlm.nih.gov/1746519/)
38. Hou X. Urban—rural disparity of overweight, hypertension, undiagnosed hypertension, and untreated hypertension in China. *Asia-Pacific Journal of Public Health*. 2008; 20(2):159–69. doi: [10.1177/1010539507312306](https://doi.org/10.1177/1010539507312306) PMID: [19124310](https://pubmed.ncbi.nlm.nih.gov/19124310/)
39. Steffes M, Cleary P, Goldstein D, Little R, Wiedmeyer H-M, Rohlfing C, et al. Hemoglobin A1c measurements over nearly two decades: sustaining comparable values throughout the Diabetes Control and Complications Trial and the Epidemiology of Diabetes Interventions and Complications study. *Clinical chemistry*. 2005; 51(4):753–8. doi: [10.1373/clinchem.2004.042143](https://doi.org/10.1373/clinchem.2004.042143) PMID: [15684277](https://pubmed.ncbi.nlm.nih.gov/15684277/)
40. Little RR, Rohlfing CL, Tennill AL, Connolly S, Hanson S. Effects of sample storage conditions on glycosylated hemoglobin measurement: evaluation of five different high performance liquid chromatography methods. *Diabetes technology & therapeutics*. 2007; 9(1):36–42. doi: [10.1089/dia.2006.0055](https://doi.org/10.1089/dia.2006.0055) PMID: [17316096](https://pubmed.ncbi.nlm.nih.gov/17316096/)
41. Liotta L, Di Franco A, Pazzagli M, Luconi M. Glycated hemoglobin (HbA1c) measurement in frozen whole blood depends on baseline values of fresh samples. *Analytical and bioanalytical chemistry*. 2013; 405(1):429–34. doi: [10.1007/s00216-012-6480-y](https://doi.org/10.1007/s00216-012-6480-y) PMID: [23086085](https://pubmed.ncbi.nlm.nih.gov/23086085/)