


Performance and costs of multiple screening strategies for type 2 diabetes: two population-based studies in Shanghai, China

Yanyun Li,¹ Huiru Jiang,^{2,3} Minna Cheng,¹ Weiyuan Yao,^{2,3} Hua Zhang,^{2,3} Yan Shi,¹ Wanghong Xu ^{2,3}

To cite: Li Y, Jiang H, Cheng M, *et al*. Performance and costs of multiple screening strategies for type 2 diabetes: two population-based studies in Shanghai, China. *BMJ Open Diab Res Care* 2020;**8**:e001569. doi:10.1136/bmjdr-2020-001569

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2020-001569>).

YL and HJ contributed equally.

Received 15 May 2020
Revised 27 June 2020
Accepted 6 July 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Shanghai Municipal Center for Disease Control and Prevention, Shanghai, China

²School of Public Health, Fudan University, Shanghai, China

³Key Lab of Health Technology Assessment (National Health Commission), Fudan University, Shanghai, China

Correspondence to

Dr Wanghong Xu;
wanghong.xu@fudan.edu.cn
and Ms Yan Shi;
shiyanyan@scdc.sh.cn

ABSTRACT

Introduction To compare the performance and the costs of various assumed screening strategies for type 2 diabetes mellitus (T2DM) among Chinese adults, and identify an optimal one for the population.

Research design and methods Two multistage-sampling surveys were conducted in Shanghai, China, in 2009 and 2017. All participants were interviewed, had anthropometry, measured fasting plasma glucose (FPG), hemoglobin A1c (A1c) and/or postprandial glucose. The 1999 WHO diagnostic criteria was used to identify undiagnosed T2DM. A previously developed Chinese risk assessment system and a specific risk assessment system developed in this study were applied to calculate diabetes risk score (DRS) 1 and 2. Optimal screening strategies were selected based on the sensitivity, Youden index and the costs using the 2009 survey data as the training set and the 2017 survey data as the validation set. A twofold cross-validation was also performed.

Results Of numerous assumed strategies, FPG ≥ 5.6 mmol/L alone performed well (Youden index of 71.8%) and cost least (US\$18.4 for each case detected), followed by the strategy of DRS2 ≥ 8 combining with FPG ≥ 5.6 mmol/L (Youden index of 71.7% and US\$20.2 per case detected) and the strategy of DRS1 ≥ 17 combining with FPG ≥ 5.6 mmol/L (Youden index of 72.0% and US\$21.6 per case detected). However, FPG alone resulted in more subjects requiring oral glucose tolerance test (OGTT) than did combining with DRS. The strategy of FPG ≥ 5.6 mmol/L combining with A1c $\geq 4.7\%$ achieved a Youden index of 72.1%, but had a cost as high as US\$48.8 for each case identified. Twofold cross-validation also supported the use of FPG alone, but with an optimal cut-off of 6.1 mmol/L.

Conclusions Our results support the use of FPG alone in T2DM screening in Chinese adults. DRS may be used combining with FPG in populations with available electronic health records to reduce the number of OGTT and save costs of screening.

INTRODUCTION

Type 2 diabetes mellitus (T2DM) is a major global health problem affecting over 451 million individuals worldwide.¹ Over the past decades, a continuous increase in prevalence of T2DM has been observed in both

Significance of this study

What is already known about this subject?

- Almost half Chinese patients with undiagnosed diabetes have a normal fasting plasma glucose (FPG), but an elevated postprandial glucose level, making oral glucose tolerance test (OGTT) an essential test in type 2 diabetes mellitus (T2DM) screening in China.
- Our previous study suggests that combining use of FPG and hemoglobin A1c (A1c) had potential to replace OGTT in T2DM screening in Chinese adults; however, the strategy is not the best choice due to the high cost of A1c assay.
- Risk assessment has been used as a complementary to FPG or A1c assay in T2DM screening to reduce the number of subjects requiring blood tests and decrease costs of screening, showing its potential in T2DM screening in Chinese adults.
- So far, however, very few studies have focused on both performance and costs of screening strategies for T2DM in Chinese population.

What are the new findings?

- A specific risk score system developed in this study includes age, sex, body mass index, waist circumference, systolic blood pressure and family history of T2DM, which are similar to those in the Chinese risk score system developed based on a nationwide survey in China.
- Assumed screening strategies are established using a stepwise approach for the first time, with the risk score as the first stage (by each score), followed by FPG (by 0.1 mmol/L) or A1c level (by 0.1%); or with FPG (by 0.1 mmol/L) as the first stage, followed by A1c level (by 0.1%).
- The strategies of risk score followed by a blood test are not superior to FPG alone in screening T2DM among Chinese adults when taking both performance and costs into consideration, but may reduce the number of subjects requiring OGTT.

high-income² and low-income and middle-income countries.³ In Chinese adults, the prevalence of T2DM was 9.7% in 2007–2008 according to the 1999 WHO diagnostic

Significance of this study

How might these results change the focus of research or clinical practice?

- ▶ FPG assay, the currently used screening strategy for T2DM in China, may be a good choice for Chinese adults under the WHO diagnostic criteria for T2DM.
- ▶ Combining use of risk score and FPG level may be recommended in populations with available electronic health record to reduce workload of OGTT and save costs of screening.

criteria,⁴ and reached 11.6% in 2010 based on the 2010 American Diabetes Association diagnostic criteria.⁵ Due to lack of specific symptoms at early stage of T2DM, 60.7% of patients with diabetes remained undiagnosed in China,⁴ and a considerable proportion of patients suffered from at least one complication at first diagnosis.^{6–8} T2DM screening in general population aiming at early diagnosis is of great significance in clinic and public health in China.

Usually, T2DM screening is based on biochemical glycaemic assays such as the fasting plasma glucose (FPG) and glycosylated hemoglobin A1c (A1c) and the 2-hour postprandial oral glucose tolerance test (OGTT). With high sensitivities and specificities, these invasive measurements have been used either alone or combined to identify T2DM in high-income countries and areas.^{9–10} As almost half undiagnosed patients with diabetes had a normal FPG but an elevated postprandial glucose level in China,^{11–12} OGTT, the uncomfortable, inconvenient and time-consuming test, is still widely used for screening and diagnosis purpose in the country. In our previous report, we found that combining use of FPG and A1c performed well and had potential to replace OGTT in Chinese population.¹³ Due to the high cost of A1c assay, however, the strategy may not be the best choice for a country with a huge population but limited healthcare resources.

Relative to blood tests, risk assessment using a score system is a cheap, convenient and non-invasive approach to identify patients with T2DM. Several risk score systems, such as Finnish,¹⁴ Danish,¹⁵ Canadian,¹⁶ Oman¹⁷ and Thailand¹⁸ systems, have been developed and widely used for diabetes screening. The diverse risk score systems across populations and their low sensitivities and specificities have greatly limited their utilities. However, risk assessment can be used as a complementary to FPG or A1c assay to reduce the number of subjects requiring blood tests and remarkably decreased costs of screening.^{19–21} In China, a diabetes risk score (DRS) system was established based on a nationwide diabetes survey. The performance of the DRS system was externally validated in two populations in Qingdao, achieving an area under receiver operating curve (AUC) >0.70 in predicting incidence of T2DM in both exploratory and validation settings.²² So far, only one study has evaluated the performance and the costs of screening tests for undiagnosed diabetes in

Chinese adults, but only focused on the fasting capillary glucose test alone and the Chinese DRS system alone.²³

In this study, taking advantage of two population-based diabetes surveys conducted in Shanghai, China, we established a new risk assessment system, and evaluated the performance and costs of multiple assumed screening strategies based on two DRS systems, FPG level and A1c level in identifying patients with T2DM. We aimed to seek a valid and economical screening strategy for T2DM in Chinese community settings.

METHODS

Study design and populations

Two different population-based diabetes surveys were conducted among Chinese adults in Shanghai, China in 2009 and 2017. Both surveys were based on a multistage sampling process. The survey in 2009 was described in our previous report.¹³ Briefly, 4 districts and 2 counties were randomly selected from all 12 districts and 7 counties. Then one to two subdistricts or towns were randomly selected from each selected district or county. And then, one to two communities or villages were randomly selected from each selected subdistrict or town. Finally, 1000–2000 eligible subjects (permanent residents of Shanghai, 35–74 years of age and having resided in the city for at least 5 years) were randomly selected from each selected community or village and invited to participate in the survey (online supplementary figure 1A). A similar sampling process was conducted in the 2017 survey (online supplementary figure 1B). A total of 2500 participants of the 2009 survey were included in the 2017 survey.

Pregnant women, individuals with type 1 diabetes and those physically or mentally disabled were excluded from participation. In the survey of 2009, of 11 844 eligible adults, 7964 participated in the survey, resulting in a response rate of 67.2%. In the 2017 survey, a total of 23 993 subjects aged 35 years or above were recruited and 22 246 subjects (8898 men and 13 348 women) participated in the survey.

Data collection

In-person interview was conducted to collect information on demographic characteristics, lifestyle factors and previous diagnosis of diabetes in both surveys. At the interview, body weight, standing height, waist circumference (WC) and blood pressure were measured for each participant according to a standard protocol, as described previously.¹² Two measurements were taken and the mean value was used in the analyses. Body mass index (BMI) was calculated as weight in kilograms divided by the square of height in meters (kg/m^2). Hypertension was defined as systolic/diastolic blood pressure (SBP/DBP) $\geq 140/90$ mm Hg or using antihypertensive drugs.

Laboratory assays

After at least 10 hours overnight fasting, a 1.0–1.5 mL venous blood sample was collected for each subject in a

vacuum tube containing sodium fluoride to measure FPG level. For those with an FPG <7.0 mmol/L, a standard 75 g glucose load was given and a second blood sample was collected to measure 2-hour postprandial blood glucose (2hPG).

Biochemical assay was conducted in one lab according to a standardized protocol. Plasma glucose was tested using glucose oxidase-peroxidase method. A1c level was assayed using high-performance liquid chromatography, which was recommended by the National Glycohemoglobin Standardization Programme.²⁴

Costs of screening

Individuals with T2DM were diagnosed according to the 1999 WHO. Costs of screening were estimated based on the costs in risk assessment, FPG, A1c and OGTT assay, which were determined as US\$0.2, US\$0.9, US\$9.3 and US\$1.4 (¥1.5, ¥6, ¥65 and ¥10) per assay, respectively, according to the charge standard for biochemistry assay in Shanghai, China (<http://wsjkw.sh.gov.cn/yysfbz/index.html>, access date: August 22, 2019). Costs per case detected were calculated as the total costs divided by the number of patients with diabetes detected.

Statistical analyses

A validated Chinese DRS system with a summary risk score ranging from 0 to 51²² was used as a risk assessment tool in this study (DRS1). A specific risk assessment system was also developed in this population with a specific risk score (DRS2). First, univariate regression was used to identify potential risk factors for T2DM based on the 2009 survey data. All variables with p value <0.10 were included in a multivariable logistic regression model. After excluding variables with p value >0.05, we established a final model and estimated β -coefficient for each variable. The accuracy and fitness of the model was evaluated using the Brier scores. DRS2 was calculated by multiplying the β -coefficients by 10 and rounding to the nearest integer,²² and was validated using the 2017 survey data. The optimal cut-off points of the DRS1 and DRS2 were identified based on the Youden index, which was at the maximum sum of the sensitivity and specificity-1.²⁵

Non-linear relationships of the DRS1 or DRS2 with the levels of FPG, 2hPG or A1c were evaluated by restricted cubic splines using the 5th, 25th, 75th and 95th percentiles as fixed knots.

Stepwise approach was used to establish combining screening strategies, with DRS1 or DRS2 (by each score) as the first stage to identify individuals at high risk, followed by FPG (by 0.1 mmol/L) or A1c assay (by 0.1%); or with FPG level (by 0.1 mmol/L) as the first stage, followed by A1c level (by 0.1%). Only positive subjects in the last step and with FPG <7.0 mmol/L (if available) would take an OGTT as a diagnostic test (figure 1). The performance of assumed screening strategies using DRS or blood tests alone or combined were tested in participants of the 2009 survey and validated in participants of the 2017 survey. Of hundreds of assumed stepwise strategies, we only focused

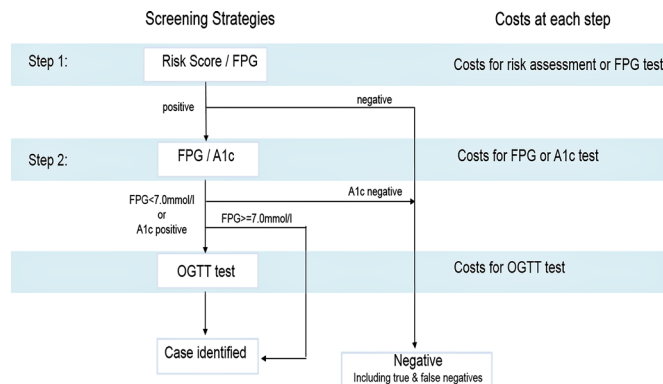


Figure 1 Diabetes screening strategies and costs at each step. A1c, hemoglobin A1c; FPG, fasting plasma glucose; OGTT, oral glucose tolerance test

on those with sensitivity $\geq 85\%$ and Youden index $\geq 70\%$. We also performed a twofold cross-validation by using randomly selected half participants of the two surveys as the training set and the remaining half as the validation set. A bootstrap resampling method was conducted 100 times to obtain CIs for all related estimates.

Diagnostic accuracy of screening strategies was assessed using AUC.²⁶ An AUC >0.9 indicates a high diagnostic value, $0.7 < \text{AUC} \leq 0.9$ indicates a moderate diagnostic value and $0.5 < \text{AUC} \leq 0.7$ indicates a low diagnostic value.

We used SAS V.9.2 for all statistical analyses, and considered $p < 0.05$ being statistically significant for a two-sided test.

RESULTS

Characteristics of the participants

After excluding subjects with incomplete questionnaires, at age of 75 years or above, having a prior history of T2DM or with missing values of A1c or BMI, 6649 subjects (3050 men and 3599 women) of the 2009 survey and 16 103 subjects (6253 men and 9850 women) of the 2017 survey were finally included in this analysis (online supplementary figure 1A,B).

Table 1 presents characteristics of study participants in the 2009 and the 2017 surveys by sex. The subjects in the 2017 survey were older than those in the 2009 survey, with an average age of 60.6 years in men and 59.9 years in women vs 54.3 years in men and 54.2 years in women. Compared with the participants of the 2009 survey, the subjects of the 2017 survey obtained more education, had higher levels of BMI, WC, SBP, DBP, and were more likely to drink alcohol and have a family history of T2DM (all $p < 0.001$). The women in the 2017 survey were less likely to smoke while the men were more likely to smoke than those in the 2009 survey ($p < 0.0001$).

DRS in subjects with diabetes and without diabetes

Online supplementary table 1 shows the risk score for each significant variable in the specific risk assessment system developed based on the 2009 survey data. The summary risk score, namely DRS2 in this study, ranged

Table 1 Demographic and lifestyle characteristics of study participants

Characteristics	Men			Women		
	The 2009 survey (n=3050)	The 2017 survey (n=6253)	P value	The 2009 survey (n=3599)	The 2017 survey (n=9850)	P value
Age (years, mean±SD)	54.3±9.5	60.6±8.9	<0.0001	54.2±9.0	59.9±8.5	<0.0001
BMI (kg/m ² , mean±SD)	24.3±3.1	25.1±3.2	<0.0001	24.2±3.5	24.6±3.3	<0.0001
WC (cm, mean±SD)	85.0±8.8	88.2±8.8	<0.0001	80.5±9.3	83.1±8.8	<0.0001
SBP (mm Hg, mean±SD)	126.6±16.9	139.4±17.7	<0.0001	124.0±17.1	136.8±18.7	<0.0001
DBP (mm Hg, mean±SD)	80.5±9.7	84.7±10.2	<0.0001	78.1±9.3	81.9±10.2	<0.0001
Education (n, %)			<0.0001			<0.0001
Primary school or below	548 (18.0)	1292 (20.7)		923 (25.6)	2746 (27.9)	
Middle school	2236 (73.3)	4390 (70.2)		2532 (70.4)	6513 (66.1)	
College or above	266 (8.8)	571 (9.13)		144 (4.0)	591 (6.0)	
Family history of T2DM (n, %)			<0.0001			<0.0001
Yes	418 (13.7)	943 (15.1)		611 (17.0)	1709 (17.4)	
No	2632 (86.3)	5310 (84.9)		2988 (83.1)	8141 (82.7)	
Smoking (n, %)			<0.0001			<0.001
Yes	1929 (63.3)	4248 (67.9)		64 (1.8)	116 (1.2)	
No	1121 (36.7)	2005 (32.1)		3535 (98.2)	9734 (98.8)	
Alcohol drinking (n, %)			<0.0001			<0.0001
Yes	1368 (44.8)	3222 (51.5)		121 (3.4)	353 (3.6)	
No	1682 (55.2)	3031 (48.5)		3478 (96.6)	9497 (96.4)	

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure; T2DM, type 2 diabetes mellitus; WC, waist circumference.

Table 2 Comparison of risk scores and related factors between diabetes and non-diabetes

Characteristics	The 2009 survey			The 2017 survey		
	Diabetes	Non-diabetes	P value	Diabetes	Non-diabetes	P value
No. of subjects	454	6195		1868	14 235	
Age (year, mean±SD)	57.3±8.7	54.0±9.2	<0.0001	62.4±7.9	59.9±8.7	<0.0001
Sex (n, %)			0.012			<0.0001
Men	234 (51.5)	2817 (45.5)		845 (45.2)	5408 (38.0)	
Women	220 (48.5)	3378 (54.5)		1023 (54.8)	8827 (62.0)	
Family history of T2DM (n, %)			<0.0001			<0.0001
Yes	109 (24.0)	920 (14.9)		439 (23.5)	2213 (15.6)	
No	345 (76.0)	5275 (85.2)		1429 (76.5)	12022 (84.5)	
BMI (kg/m ² , mean±SD)	26.1±3.8	24.1±3.2	<0.0001	26.0±3.4	24.7±3.2	<0.0001
WC (cm, mean±SD)	88.2±9.8	82.1±9.2	<0.0001	88.7±9.2	84.6±9.1	<0.0001
SBP (mm Hg, mean±SD)	134.7±17.3	124.5±16.8	<0.0001	145.2±18.3	136.8±18.2	<0.0001
DBP (mm Hg, mean±SD)	83.9±10.8	78.9±9.4	<0.0001	85.3±10.4	82.7±10.2	<0.0001
DRS1	32.1±6.3	26.5±7.2	<0.0001	38.5±6.0	35.3±6.9	<0.0001
DRS2	23.4±7.8	16.3±8.5	<0.0001	26.7±7.3	21.9±8.0	<0.0001

DRS1 was the risk score calculated based on the Chinese risk score system; DRS2 was the risk score calculated based on the system developed in this study.

BMI, body mass index; DBP, diastolic blood pressure; DRS, diabetes risk score; SBP, systolic blood pressure; T2DM, type 2 diabetes mellitus; WC, waist circumference.

from 0 to 49 and had an optimal cut-off point of 22 in subjects of the 2009 survey and 24 in participants of the 2017 survey.

Table 2 shows the comparison of risk score and related risk factors between subjects with diabetes and without diabetes. In the 2009 survey, a total of 454 patients with diabetes were identified, with prevalence of T2DM being 6.8%. In the 2017 survey, the prevalence reached 11.6%. Generally, patients with diabetes were older, tended to have a family history of diabetes and had higher average levels of BMI, WC, SBP, DBP and DRS comparing with non-diabetes in both surveys (all p values <0.05). The mean DRS1 was 5~6 points higher in patients with diabetes than those without in the 2009 survey, but only about 3 points higher in the 2017 survey. By contrast, the mean DRS2 was 6~7 points higher in patients with diabetes than those without in the 2009 survey, it was about 4~5 points higher in the 2017 survey.

Diagnostic value of the DRS

The non-linear does-response relationship was observed for DRS with FPG, 2hPG and A1c levels, with all p values for non-linear association <0.0001 . As shown in online supplementary figure 2A,B, the levels of FPG, 2hPG and A1c were observed to increase with increasing DRS1 in both surveys in a non-linear pattern, with p values for non-linear association <0.0001 . A similar non-linear does-response relationship was observed for FPG, 2hPG and A1c levels with DRS2 (online supplementary figure 2C,D).

FPG level ranked first in AUC in diagnosis of T2DM, followed by A1c, DRS2 and DRS1 in both surveys (figure 2A,B). The optimal cut-off point was 34 for DRS1, 22 for DRS2, 5.7 mmol/L for FPG and 6.1% for A1c. Generally, the two risk score systems had a moderate diagnostic value in this population, with an AUC of 0.728 and 0.671 for DRS2 in the 2009 and the 2017 surveys, respectively, slightly higher than 0.678 and 0.633 for DRS1 in the two surveys.

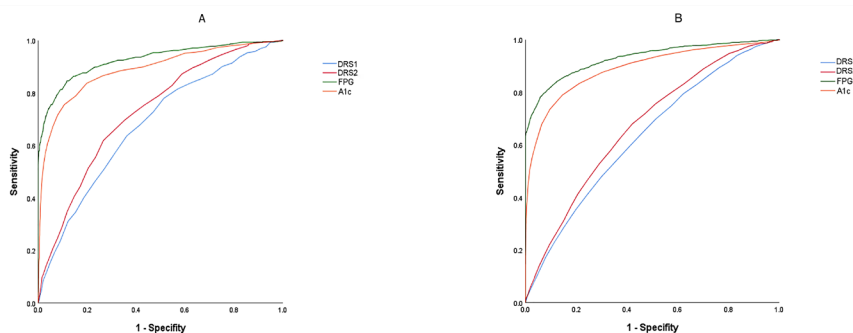


Figure 2 AUC of DRS1, DRS2, FPG and A1c level in the 2009 survey (A) and the 2017 survey (B). DRS1 was the risk score calculated based on the Chinese diabetes risk score system; DRS2 was the risk score calculated based on the specific system developed in this study. A1c, glycosylated hemoglobin A1c; AUC, area under receiver operating curve; FPG, fasting plasma glucose; DRS, diabetes risk score.

Performance and costs of assumed screening strategies

Listed in online supplementary table 2 are the performance of hundreds of assumed screening strategies in identifying T2DM in the 2009 survey. When used alone, DRS1, DRS2, FPG and A1c achieved a Youden index of 27.4%, 35.3%, 72.5% and 64.9%, respectively. In combining screening strategies, those with A1c level cost much more than those without.

We only focused on screening strategies with sensitivity $\geq 85\%$ and Youden index $\geq 70\%$. For screening strategies using same index in each step, the one with the lowest cost was selected. As shown in table 3, FPG ≥ 5.6 mmol/L alone performed well with Youden index of 71.8% and cost least (US\$18.4 for each case detected), followed by the strategy of DRS2 ≥ 8 followed by FPG ≥ 5.6 mmol/L (Youden index of 71.7% and US\$20.2 per case detected) and the strategy of DRS1 ≥ 17 followed by FPG ≥ 5.6 mmol/L (Youden index of 72.0% and US\$21.6 per case detected). The strategy of FPG ≥ 5.6 mmol/L followed by A1c $\geq 4.7\%$ performed well but had a higher cost than the strategies without A1c. The strategy of DRS2 ≥ 8 followed by FPG ≥ 5.6 mmol/L had the fewest subjects taking OGTT compared with other screening strategies.

Twofold cross-validation was further performed using randomly selected half subjects in the 2009 and the 2017 survey as training sets and the remaining half subjects as validation sets. Of screening strategies with sensitivity $\geq 80\%$ and Youden index $\geq 70\%$, we identified four screening strategies with the lowest cost, as shown in table 4. Similarly, FPG alone performed well and cost least, but with an optimal cut-off point of 6.1 mmol/L.

DISCUSSION

In this study of Chinese adults randomly selected from community settings of Shanghai in 2009 and in 2017, we established a DRS system to detect undiagnosed T2DM and compared the performance and costs of various assumed screening strategies. We did not find that strategies of DRS followed by a blood test were superior to FPG alone in screening diabetes among Chinese adults when

Table 3 Validity and costs of screening strategies in the 2009 survey (as the training set) and in the 2017 survey (as a validation set)

Screening strategy		No. of positive subjects in step 1	No. of positive subjects in step 2	No. of subjects taking OGTT	No. of cases detected	No. of cases missed	Sensitivity (%)	Specificity (%)	Youden index (%)	Positive predictive value (%)	Total costs (US\$)	Costs per case detected (US\$)
Step 1	Step 2											
Training set (n=6649)												
FPG ≥ 5.6 mmol/L	-	1287	-	1048	392	62	86.3	85.5	71.8	30.5	7196	18.4
DRS1 ≥ 17	FPG ≥ 5.6 mmol/L	6501	1281	1042	392	62	86.3	85.6	72.0	30.6	8486	21.6
DRS2 ≥ 8	FPG ≥ 5.6 mmol/L	5788	1225	990	387	67	85.2	86.5	71.7	31.6	7800	20.2
FPG ≥ 5.6 mmol/L	A1c $\geq 4.7\%$	1287	1273	1034	392	62	86.3	85.8	72.1	30.8	19 127	48.8
Validation set (n=16 103)												
FPG ≥ 5.6 mmol/L	-	6912	-	5729	1747	121	93.5	63.7	57.2	25.3	21 987	12.6
DRS1 ≥ 17	FPG ≥ 5.6 mmol/L	16 009	6897	5714	1708	160	91.4	70.2	61.6	28.7	25 336	14.8
DRS2 ≥ 8	FPG ≥ 5.6 mmol/L	15 595	6824	5650	1734	134	92.8	64.2	57.1	25.4	24 889	14.4
FPG ≥ 5.6 mmol/L	A1c $\geq 4.7\%$	6912	6888	5708	1744	124	93.4	63.9	57.2	25.3	86 140	49.4

DRS1 was risk score calculated based on the Chinese risk score system; DRS2 was risk score calculated based on the system developed in this study. A1c, glycosylated hemoglobin A1c; DRS, diabetes risk score; FPG, fasting plasma glucose; OGTT, oral glucose tolerance test.

Table 4 Performance and costs of screening strategies in the training set (randomly selected half subjects in the 2009 and the 2017 survey) and the validation set (the remaining half subjects in the 2009 and the 2017 survey)

Screening strategy		No. of positive subjects (95% CI) in step 1	No. of positive subjects (95% CI) in step 2	No. of subjects (95% CI) taking OGTT	No. of cases detected (95% CI)	No. of cases missed (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Youden index (%) (95% CI)	Positive predictive value (%) (95% CI)	Total costs (95% CI) (US\$)	Costs per case detected (95% CI) (US\$)
Step 1	Step 2											
Training set (n=11 376)												
FPG ≥6.1 mmol/L		1952 (1947 to 1957)	-	1242 (1238 to 1247)	950 (946 to 955)	208 (206 to 210)	82.0 (81.9 to 82.2)	90.2 (90.1 to 90.2)	72.3 (72.1 to 72.3)	48.7 (48.5 to 48.9)	11 525 (11 519 to 11 532)	12.1 (12.1 to 12.2)
DRS1 ≥24 mmol/L	FPG ≥6.1 mmol/L	10 403 (10399 to 10406)	1902 (1897 to 1907)	1209 (1205 to 1213)	929 (925 to 933)	229 (227 to 231)	80.2 (80.0 to 80.3)	90.5 (90.4 to 90.5)	70.7 (70.5 to 70.8)	48.8 (48.7 to 49.0)	13 081 (13 074 to 13 088)	14.1 (14.0 to 14.2)
DRS2 ≥8 mmol/L	FPG ≥6.1 mmol/L	10 029 (10 024 to 10034)	1912 (1907 to 1917)	1215 (1211 to 1219)	937 (933 to 940)	225 (223 to 227)	80.6 (80.4 to 80.8)	90.4 (90.4 to 90.5)	71.1 (70.9 to 71.1)	49.0 (48.8 to 49.1)	12 770 (12 762 to 12 779)	13.6 (13.6 to 13.7)
FPG ≥6.1 mmol/L	A1c ≥4.3%	1952 (1947 to 1957)	1950 (1945 to 1955)	1240 (1236 to 1245)	950 (946 to 955)	208 (206 to 210)	82.0 (81.9 to 82.2)	90.2 (90.2 to 90.3)	72.3 (72.1 to 72.3)	48.7 (48.6 to 48.9)	29 647 (29 593 to 29 700)	31.2 (31.1 to 31.3)
Validation set (n=11 376)												
FPG ≥6.1 mmol/L		1965 (1960 to 1970)	-	1253 (1249 to 1257)	956 (951 to 960)	208 (206 to 210)	82.1 (82.0 to 82.3)	90.1 (90.1 to 90.2)	72.2 (72.1 to 72.3)	48.6 (48.5 to 48.8)	11 541 (11 535 to 11 547)	12.1 (12.0 to 12.1)
DRS1 ≥24 mmol/L	FPG ≥6.1 mmol/L	10 406 (10403 to 10410)	1914 (1909 to 1919)	1218 (1214 to 1222)	934 (930 to 938)	230 (228 to 232)	80.3 (80.1 to 80.4)	90.4 (90.3 to 90.4)	70.7 (70.5 to 70.7)	48.8 (48.6 to 49.0)	13 098 (13 091 to 13 105)	14.0 (14.0 to 14.1)
DRS2 ≥8 mmol/L	FPG ≥6.1 mmol/L	10 039 (10 035 to 10044)	1926 (1921 to 1931)	1227 (1223 to 1231)	939 (935 to 944)	224 (222 to 226)	80.8 (80.6 to 80.9)	90.3 (90.3 to 90.4)	71.1 (71.0 to 71.2)	48.8 (48.6 to 48.9)	12 796 (12 788 to 12 804)	13.6 (13.6 to 13.7)
FPG ≥6.1 mmol/L	A1c ≥4.3%	1965 (1960 to 1970)	1964 (1959 to 1969)	1252 (1247 to 1256)	956 (951 to 960)	208 (206 to 210)	82.1 (82.0 to 82.3)	90.1 (90.1 to 90.2)	72.2 (72.1 to 72.3)	48.7 (48.5 to 48.8)	29 787 (29 734 to 29 840)	31.2 (31.1 to 31.3)

DRS1 was risk score calculated based on the Chinese risk score system; DRS2 was risk score calculated based on the system developed in this study. A1c, glycosylated hemoglobin A1c; DRS, diabetes risk score; FPG, fasting plasma glucose; OGTT, oral glucose tolerance test.

taking both validity and costs into consideration. To the best of our knowledge, this is the first large-scale study to evaluate the validity and costs of diabetes screening strategies in Chinese population.

Risk assessment has been widely used to identify individuals at high risk of T2DM.^{27 28} A proper risk score system may reduce screening costs without loss of much validity.¹⁴ In this study, we established a specific diabetes risk score system including age, sex, BMI, WC, SBP and family history of T2DM, which was consistently with the Chinese diabetes risk score system developed based on a nationwide survey in China.²² We found that both DRS were positively associated with FPG, 2hPG and A1c levels in a non-linear fashion. However, the Chinese DRS system achieved an AUC <0.700 in our population, much less than those in original exploratory Qingdao population (0.748) and the two validation populations (0.725 and 0.717).²² The optimal cut-off point of DRS1, which was based on the Chinese diabetes risk score system, was 34 in the 2009 survey, much higher than 25 in Qingdao population, and had sensitivity of only 63.7% and 64.8% in the 2009 and the 2007 surveys, and specificity of only 63.8% and 53.7%, respectively. The specific DRS system developed in this study performed better than did the Chinese DRS system, but with an AUC of only 0.728 in the 2009 survey and 0.671 in the 2017 survey. Our results did not support the application of DRS alone in diabetes screening. As a non-invasive tool, it may be used as the first step in stepwise screening strategies combining with blood tests to balance the validity and the costs.

A number of previous studies have used simulation models to evaluate the cost-effectiveness of screening strategies for T2DM.^{14–16} However, none of these studies was conducted in China, the country holding approximately half of the world's diabetes population.^{4,5} In this study, we found that FPG ≥ 5.6 mmol/L alone or ≥ 6.1 mmol/L alone performed well and cost least in detecting T2DM, providing supportive evidence for the screening strategy currently used in China.²⁹ When using the Chinese DRS system, the strategy of DRS1 ≥ 17 followed by FPG ≥ 5.6 mmol/L had comparable performance and costs, and had 148 (2.23%) patients free from invasive blood tests and 6 (0.09%) free from OGTT relative to the strategy of FPG ≥ 5.6 mmol/L alone. Whereas, the strategy of DRS2 ≥ 8 followed by FPG ≥ 5.6 mmol/L had 861 (13.0%) patients free from blood tests and 58 (0.87%) patients free from further OGTT. Considering that many cities in China have established electronic health record systems (eHR),^{30–33} risk assessment based on eHR system may save enormous costs on manpower and materials to identify high-risk population. In this case, combining use of DRS and FPG assay can be used in Chinese adults. Otherwise, FPG alone may be good enough for diabetes screening in the population.

The strengths of the study include large sample size, random sampling process and various screening strategies established at the level of 1.0 for risk score, 0.1 mmol/L for FPG and 0.1% for A1c. However, the response rate was relatively low (67.2%) in the 2009 survey, and lack of

information on non-participants' characteristics limited our ability to evaluate the potential selection bias. Moreover, the accuracy of DRS, and FPG and A1c measurements was the key for stepwise screening. In the real world, the risk assessment and blood tests would be conducted in different labs. The possible interlab bias in measurements may influence the effect of related screening strategies. Finally, the screening costs were estimated mainly based on the costs of laboratory assay. It may lead to underestimation of screening costs, because the costs for manpower, phone calls and transportation were not taken into account in our analyses.

CONCLUSIONS

In conclusion, in view of performance and costs, the strategy of FPG assay alone is comparable to the strategies of non-invasive risk assessment combining with a blood test in Chinese adults. Our results support the use of FPG alone, the currently used strategy in China and suggest that combining use of risk score and FPG level may be a good choice for Chinese adults with available eHR system. Further epidemiological and clinical studies are warranted to validate our results.

Acknowledgements The authors would like to thank study participants of the two cross-sectional surveys and the healthcare workers in all communities involved. The authors would also like to thank the research assistants for their great contributions to data collection and data entry.

Contributors YL and HJ drafted the manuscript. WX and YS coordinated the study and contributed to study design, statistical analysis, YL and MC contributed to data acquisition, HZ contributed to data cleaning and statistical analysis. All authors contributed to the interpretation of data and revision of the manuscript. WX and YS are the guarantors of this work and had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding This study was funded by the Three-year Action Plan on Public Health, Phase IV, Shanghai, China (15GWZK0801).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The Institutional Review Board (#IORG0000630) of the Shanghai Municipal Center of Disease Control and Prevention approved both surveys (#2016–25 for the 2017 survey). Informed written consent was obtained from each participant before interview and bio-specimen collection.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. De-identified data collected for this study and a data dictionary are available from the corresponding author on reasonable request.

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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Wanghong Xu <http://orcid.org/0000-0003-2045-2218>

REFERENCES

- 1 Cho NH, Shaw JE, Karuranga S, *et al*. IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res Clin Pract* 2018;138:271–81.

- 2 Thomas MC, Hardoon SL, Papacosta AO, *et al*. Evidence of an accelerating increase in prevalence of diagnosed type 2 diabetes in British men, 1978-2005. *Diabet Med* 2009;26:766-72.
- 3 Ojuka EO, Goyaram V. Increasing prevalence of type 2 diabetes in sub-Saharan Africa: not only a case of inadequate physical activity. *Med Sport Sci* 2014;60:27-35.
- 4 Yang W, Lu J, Weng J, *et al*. Prevalence of diabetes among men and women in China. *N Engl J Med* 2010;362:1090-101.
- 5 Xu Y, Wang L, He J, *et al*. Prevalence and control of diabetes in Chinese adults. *JAMA* 2013;310:948-59.
- 6 Liu F, Bao Y, Hu R, *et al*. Screening and prevalence of peripheral neuropathy in type 2 diabetic outpatients: a randomized multicentre survey in 12 City hospitals of China. *Diabetes Metab Res Rev* 2010;26:481-9.
- 7 Ali A, Iqbal F, Taj A, *et al*. Prevalence of microvascular complications in newly diagnosed patients with type 2 diabetes. *Pak J Med Sci* 2013;29:899-902.
- 8 Winkley K, Thomas SM, Sivaprasad S, *et al*. The clinical characteristics at diagnosis of type 2 diabetes in a multi-ethnic population: the South London diabetes cohort (SOUL-D). *Diabetologia* 2013;56:1272-81.
- 9 American Diabetes Association. Standards of medical care in diabetes--2013. *Diabetes Care* 2013;36 Suppl 1:S11-66.
- 10 Nakagami T, Takahashi K, Suto C, *et al*. Diabetes diagnostic thresholds of the glycated hemoglobin A1c and fasting plasma glucose levels considering the 5-year incidence of retinopathy. *Diabetes Res Clin Pract* 2017;124:20-9.
- 11 Lu Q, Tong N, Liu Y, *et al*. Community-based population data indicates the significant alterations of insulin resistance, chronic inflammation and urine ACR in IFG combined IGT group among prediabetic population. *Diabetes Res Clin Pract* 2009;84:319-24.
- 12 Li R, Lu W, Jiang QW, *et al*. Increasing prevalence of type 2 diabetes in Chinese adults in Shanghai. *Diabetes Care* 2012;35:1028-30.
- 13 Mo M, Zhong W, Zhao G, *et al*. Combining glycosylated hemoglobin A1c and fasting plasma glucose for diagnosis of type 2 diabetes in Chinese adults. *BMC Endocr Disord* 2013;13:44.
- 14 Wang J, Stancáková A, Kuusisto J, *et al*. Identification of undiagnosed type 2 diabetic individuals by the Finnish diabetes risk score and biochemical and genetic markers: a population-based study of 7232 Finnish men. *J Clin Endocrinol Metab* 2010;95:3858-62.
- 15 Glümer C, Carstensen B, Sandbaek A, *et al*. A Danish diabetes risk score for targeted screening: the Inter99 study. *Diabetes Care* 2004;27:727-33.
- 16 Robinson CA, Agarwal G, Nerenberg K. Validating the CANRISK prognostic model for assessing diabetes risk in Canada's multi-ethnic population. *Chronic Dis Inj Can* 2011;32:19-31.
- 17 Al-Lawati JA, Tuomilehto J. Diabetes risk score in Oman: a tool to identify prevalent type 2 diabetes among Arabs of the middle East. *Diabetes Res Clin Pract* 2007;77:438-44.
- 18 Aekplakorn W, Bunnag P, Woodward M, *et al*. A risk score for predicting incident diabetes in the Thai population. *Diabetes Care* 2006;29:1872-7.
- 19 Khunti K, Gillies CL, Taub NA, *et al*. A comparison of cost per case detected of screening strategies for type 2 diabetes and impaired glucose regulation: modelling study. *Diabetes Res Clin Pract* 2012;97:505-13.
- 20 Chamnan P, Simmons RK, Khaw KT, *et al*. Estimating the potential population impact of stepwise screening strategies for identifying and treating individuals at high risk of type 2 diabetes: a modelling study. *Diabet Med* 2012;29:893-904.
- 21 Martin E, Ruf E, Landgraf R, *et al*. FINDRISK questionnaire combined with HbA1c testing as a potential screening strategy for undiagnosed diabetes in a healthy population. *Horm Metab Res* 2011;43:782-7.
- 22 Zhou X, Qiao Q, Ji L, *et al*. Nonlaboratory-based risk assessment algorithm for undiagnosed type 2 diabetes developed on a nationwide diabetes survey. *Diabetes Care* 2013;36:3944-52.
- 23 Zhang Y, Sun J, Pang Z, *et al*. Evaluation of two screening methods for undiagnosed diabetes in China: an cost-effectiveness study. *Prim Care Diabetes* 2013;7:275-82.
- 24 Little RR. Glycated hemoglobin standardization--National Glycohemoglobin Standardization Program (NGSP) perspective. *Clin Chem Lab Med* 2003;41:1191-8.
- 25 Schisterman EF, Faraggi D, Reiser B, *et al*. Youden index and the optimal threshold for markers with mass at zero. *Stat Med* 2008;27:297-315.
- 26 Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr* 2007;96:644-7.
- 27 Noble D, Mathur R, Dent T, *et al*. Risk models and scores for type 2 diabetes: systematic review. *BMJ* 2011;343:d7163.
- 28 Brown N, Critchley J, Bogowicz P, *et al*. Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review. *Diabetes Res Clin Pract* 2012;98:369-85.
- 29 Chinese Diabetes Society. Guidelines for the prevention and control of type 2 diabetes in China (2017 edition). *Chinese Journal of Practical Internal Medicine* 2018;38:292-344. (in Chinese).
- 30 Liu H, Hou XQ, Hu G, *et al*. Development of an EHR system for sharing - a semantic perspective. *Stud Health Technol Inform* 2009;150:113-7.
- 31 Wang Y, Wang Y, Qain Y, *et al*. Association of body mass index with cause specific deaths in Chinese elderly hypertensive patients: Minhang community study. *PLoS One* 2013;8:e71223.
- 32 Lin H, Tang X, Shen P, *et al*. Using big data to improve cardiovascular care and outcomes in China: a protocol for the Chinese electronic health records research in Yinzhou (cherry) study. *BMJ Open* 2018;8:e019698.
- 33 He D, Xu W, Su H, *et al*. Electronic health Record-Based screening for major cancers: a 9-year experience in Minhang district of Shanghai, China. *Front Oncol* 2019;9:375.