

RESEARCH

Open Access



# Establishment of dynamic nomogram and risk score models for T2DM: a retrospective cohort study in Beijing

Chao Tong<sup>1</sup>, Yumei Han<sup>2</sup>, Shan Zhang<sup>1</sup>, Qiang Li<sup>2</sup>, Jingbo Zhang<sup>2</sup>, Xiuhua Guo<sup>1</sup>, Lixin Tao<sup>1</sup>, Deqiang Zheng<sup>1</sup> and Xinghua Yang<sup>1\*</sup>

## Abstract

**Background:** Health interventions can delay or prevent the occurrence and development of diabetes. Dynamic nomogram and risk score (RS) models were developed to predict the probability of developing type 2 diabetes mellitus (T2DM) and identify high-risk groups.

**Methods:** Participants ( $n = 44,852$ ) from the Beijing Physical Examination Center were followed up for 11 years (2006–2017); the mean follow-up time was  $4.06 \pm 2.09$  years. Multivariable Cox regression was conducted in the training cohort to identify risk factors associated with T2DM and develop dynamic nomogram and RS models using weighted estimators corresponding to each covariate derived from the fitted Cox regression coefficients and variance estimates, and then undergone internal validation and sensitivity analysis. The concordance index (C-index) was used to assess the accuracy and reliability of the model.

**Results:** Of the 44,852 individuals at baseline, 2,912 were diagnosed with T2DM during the follow-up period, and the incidence density rate per 1,000 person-years was 16.00. Multivariate analysis indicated that male sex ( $P < 0.001$ ), older age ( $P < 0.001$ ), high body mass index (BMI,  $P < 0.05$ ), high fasting plasma glucose (FPG,  $P < 0.001$ ), hypertension ( $P = 0.015$ ), dyslipidaemia ( $P < 0.001$ ), and low serum creatinine (sCr,  $P < 0.05$ ) at presentation were risk factors for T2DM. The dynamic nomogram achieved a high C-index of 0.909 in the training set and 0.905 in the validation set. A tenfold cross-validation estimated the area under the curve of the nomogram at 0.909 (95% confidence interval 0.897–0.920). Moreover, the dynamic nomogram and RS model exhibited acceptable discrimination and clinical usefulness in subgroup and sensitivity analyses.

**Conclusions:** The T2DM dynamic nomogram and RS models offer clinicians and others who conduct physical examinations, respectively, simple-to-use tools to assess the risk of developing T2DM in the urban Chinese current or retired employees.

**Keywords:** Diabetes mellitus, Dynamic nomogram, Risk score, Real world study

## Background

Type 2 diabetes mellitus (T2DM) is a complex metabolic disease prevalent worldwide [1–3]. The World Health Organization estimated that, globally, 422 million adults aged over 18 years were living with diabetes in 2014 [4]. Currently, China has the largest number of patients with diabetes worldwide, and the prevalence of diabetes in

Chao Tong and Yumei Han should be considered joint first author.

\*Correspondence: xinghuayang@ccmu.edu.cn

<sup>1</sup> School of Public Health, Capital Medical University, NO.10 Xitoutiao, Youanmen, Beijing 100069, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

China ranks among the top in the world [5, 6]. Indeed, several previous studies succeeded have proven that life-style interventions can be effective in the prevention of T2DM [7, 8]. However, Heidemann et al. found that some individuals perceived their diabetes risk to be low, even though the actual risk was high [9].

In recent years, numerous diabetes risk score (RS) models have been developed to predict the risk of T2DM incidence [10, 11]. However, various diabetes risk models and scores were rarely used because they were developed without a specific user or explicit use in mind [12]. Thus, several investigators have developed risk scores specifically for rural or urban population, respectively [13–15]. In addition, several diabetes RS models have been constructed for specific age groups [16, 17]. However, no study has developed an RS model specially for the current or retired employees. Moreover, based on normal nomogram, dynamic nomogram is an online scoring system that provides clinicians with a simple-to-use tool to tailor clinical decisions [18]. To our knowledge, the dynamic nomogram has not yet been developed for predicting the probability of T2DM incidence.

This study aimed to develop a dynamic nomogram and an RS model based on physical examination data to

16 to 68 years, with an average age of  $38.48 \pm 11.87$  years. The project was approved by the ethics committee of the Beijing Physical Examination Center (Beijing, China; ethics approval ID: 201,802 & 202,008). All methods were carried out in accordance with relevant guidelines and regulations, and informed consent was obtained from 11,145 participants. As this was a retrospective study, the ethics committee of the Beijing Physical Examination Center approved the remaining participants with waived written informed consent.

### Measurements and surveys

Data from January 2006 to December 2017 were used in this study. The project data included anthropometric measurements (height, weight, and blood pressure [BP]), blood biochemical indicators (FPG, serum creatinine [sCr], blood uric acid [SUA], estimated glomerular filtration rate [eGFR], total cholesterol, triglycerides [TGs], low-density lipoprotein cholesterol [LDL-C], and high-density lipoprotein cholesterol [HDL-C]), and questionnaire survey (sex, age, history of DM and hypertension, and family history of diabetes). eGFR was calculated using the abbreviated Modification of Diet in Renal Disease equation:

$$eGFR = 175 * \text{creatinine}^{-1.154} * \text{age}^{-0.203} * 1.212(\text{if black}) * 0.742(\text{if female})$$

predict the probability of developing T2DM and identify high-risk populations, and to provide a basis for health management and risk communication in primary care practice.

## Methods

### Study population

We used retrospective cohort data from the Beijing Health Management Cohort (BHMC), which covers current or retired employees in Beijing, China [19]. Individuals were included in the dynamic cohort if they were older than 16 years old, had records at least three medical examinations, without T2DM at the time of entry into the cohort. The members with (a) missing data on important variables such as age, sex, FPG levels, and clinical laboratory indicators ( $n=134$ ) and (b) diabetes ( $n=4536$ ), cardiovascular and cerebrovascular diseases ( $n=49$ ), liver and kidney diseases (37), and family history of diabetes ( $n=134$ ) were excluded. According to the Chinese Guidelines for the Prevention and Treatment of T2DM (2017 edition), participants were diagnosed with diabetes and discontinued from follow up when the fasting plasma glucose (FPG) level was  $\geq 7.0$  mmol/L [20]. Finally, a total of 44,852 individuals were enrolled in the study population, which including 24,817 men and 20,035 women. The age of the population ranged from

In accordance with the guidelines for prevention and control of overweight and obesity in Chinese adults, body mass index (BMI) was categorized into four groups: underweight ( $< 18.5$  kg/m<sup>2</sup>), normal (18.5 to  $< 24.0$  kg/m<sup>2</sup>), overweight (24.0 to  $< 28.0$  kg/m<sup>2</sup>), and obesity ( $\geq 28.0$  kg/m<sup>2</sup>) [21]. Hypertension was diagnosed as an average systolic BP (SBP)  $\geq 140$  mmHg, or diastolic BP (DBP)  $\geq 90$  mmHg, based on the US Seventh Joint National Committee (JNC7) [22]. Dyslipidaemia was defined as total cholesterol (TC)  $\geq 6.2$  mmol/L or TG  $\geq 2.3$  mmol/L or LDL-C  $\geq 4.1$ , or HDL-C  $< 1.0$  mmol/L, according to the 2016 Chinese Guideline for the Management of Dyslipidaemia in Adults [23]; Hyperuricaemia was defined as SUA  $\geq 360$   $\mu\text{mol/L}$  in women or  $\geq 420$   $\mu\text{mol/L}$  in men under normal purine diet [24]. eGFR was divided into two categories: eGFR  $\geq 90$  mL/min/1.73 m<sup>2</sup> and eGFR  $< 90$  mL/min/1.73 m<sup>2</sup> [25]. sCr was transformed into categorical variables according to quartiles.

### Statistical analysis

Categorical variables were expressed as numbers (percentages) and compared using the chi-square test. For nomogram construction and internal validation, we randomly divided the study population into training ( $n=31,391$ ) and validation ( $n=13,461$ ) cohorts in a ratio

of 7:3. The univariate and multivariate Cox proportional hazards models were conducted in the training cohort to develop a nomogram using weighted estimators corresponding to each covariate derived from the fitted Cox regression coefficients and variance estimates [26]. An RS was calculated by summing the risk points corresponding to each of the weighted covariates used to perform the dynamic nomogram. Individuals were classified into three groups according to the risk of developing diabetes. The dynamic nomogram was evaluated using ten-fold cross-validation. Kaplan–Meier survival curves were plotted to estimate the probability of remaining free of T2DM during the follow-up for each group of subjects in the training and validation sets. Besides, exploratory subgroup analysis was used to examine the effect of employment status on the RS model. According to the general retirement age in China, retirement was defined as age  $\geq 50$  years in women or  $\geq 60$  years in men [27]. Sensitivity analysis was conducted in a cohort that excluded the first two years of follow-up to validate the accuracy of the model.

For all analyses, a two-tailed  $p$ -value of  $<0.05$  was defined as statistically significant. SPSS software (version 26.0; IBM, Chicago, United States) was used for general descriptive analysis and  $\chi^2$ -tests. Univariate and multivariate Cox proportional hazards regression models, dynamic nomogram, tenfold cross-validation, and Kaplan–Meier survival curves were performed using R statistical software (version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria).

## Results

### Baseline characteristics of participants

The baseline characteristics of the study cohort are presented in Table 1. A total of 44,852 participants without diabetes at baseline were assigned into two groups based on whether they had diabetes during the follow-up period. For over a median follow-up of 3.58 years (interquartile range, 2.31–5.10 years), 2,912 (6.49%) participants developed new-onset T2DM (16.00 events per 1,000 person-years at follow-up). Individuals with new-onset diabetes during follow-up tended to be older, men and retired, have higher levels of BMI, FPG, sCr and eGFR, and were more likely to have hypertension, dyslipidaemia, and hyperuricaemia.

### Independent high-risk factors associated with incidence

Further multivariate Cox regression analysis showed that male sex (HR=1.76,  $P<0.001$ ), age (30–39 years) (HR=1.54,  $P<0.001$ ), age (40–49 years) (HR=1.99,  $P<0.001$ ), age (50–59 years) (HR=2.45,  $P<0.001$ ), age ( $\geq 60$  years) (HR=3.41,  $P<0.001$ ), underweight (HR=0.46,  $P=0.029$ ), overweight (HR=1.45,  $P<0.001$ ),

obesity (HR=2.22,  $P<0.001$ ), FPG (5.6–6.1 mmol/L) (HR=9.90,  $P<0.001$ ), FPG (6.2–7.0 mmol/L) (HR=21.69,  $P<0.001$ ), hypertension (HR=1.15,  $P=0.015$ ), dyslipidaemia (HR=1.28,  $P<0.001$ ), sCr (57.2–68.1  $\mu\text{mol/L}$ ) (HR=0.92,  $P=0.321$ ), sCr (68.2–78.9  $\mu\text{mol/L}$ ) (HR=0.68,  $P<0.001$ ), and sCr  $>78.9$   $\mu\text{mol/L}$  (HR=0.63,  $P<0.001$ ) were independent influential factors associated with the development of T2DM (Table 2).

### Predictive dynamic nomogram and RS on the incidence of diabetes

Figure 1 presents a nomogram established for the 3-, 5- and 10-year incidence of diabetes based on the above multivariable Cox proportional hazard model in the training cohort. The nomogram demonstrated that FPG levels contributed the most to the incidence, followed by sex, age, BMI, dyslipidaemia, hypertension, and sCr, which had good accuracy in estimating the risk of developing diabetes, with a concordance index (C-index) of 0.909 in the training set and 0.905 in the validation set. Ten-fold cross-validation estimated the AUC of the dynamic nomogram at 0.909 (95% CI 0.897–0.920). Additionally, a diabetes prediction dynamic nomogram application (<https://nomogramxhy.shinyapps.io/DynNomapp/>) was produced, which can be conveniently available to clinicians in primary care. Using the application, one individual's risk probability of remaining T2DM-free in the following years can be obtained immediately when imputing the information of the seven identified variables. The R code and data for the application were attached in Supplement Data.

Our RS was calculated based upon the weighted of diabetes risk factors as follows:  $(1.90 \times I [\text{male sex}]) + (1.39 \times I [30\text{--}39 \text{ years age}]) + (2.18 \times I [40\text{--}49 \text{ years age}]) + (2.86 \times I [50\text{--}59 \text{ years age}]) + (3.92 \times I [\geq 60 \text{ years age}]) + (2.55 \times I [\text{normal}]) + (3.76 \times I [\text{overweight}]) + (5.16 \times I [\text{obesity}]) + (7.45 \times I [5.6 \leq \text{FPG} < 6.1 \text{ mmol/L}]) + (10.00 \times I [6.1 \leq \text{FPG} < 7.0 \text{ mmol/L}]) + (0.45 \times I [\text{hypertension}]) + (0.81 \times I [\text{dyslipidaemia}]) + (0.37 \times I [68.2\text{--}78.9 \mu\text{mol/L sCr}]) + (1.33 \times I [57.2\text{--}68.1 \mu\text{mol/L sCr}]) + (1.67 \times I [<57.2 \mu\text{mol/L sCr}])$ , where  $I []$  denotes the indicator function that is equal to 1 if the condition in the parentheses is satisfied, and 0 otherwise.

Furthermore, RS was classified into three levels of risk according to the risk of developing diabetes: those with 0–12 points who had less than 10% incidence of diabetes been considered low risk (Class A), 12–18 points with 10%–50% incidence of diabetes was mid-risk (Class B), and 18–26 points with over 50% incidence was high risk (Class C). The Kaplan–Meier survival curve for both cohorts according to RS was shown in Fig. 2. Each set of RS class appeared well separated, indicating reasonable discrimination in the cohort ( $P<0.001$ ). Using this predictive model,

**Table 1** Baseline characteristics of participants

	Overall (n = 44,852)	Diabetes-free (n = 41,940)	New diabetes (n = 2,912)	P-value
Gender (n, %)				< 0.001
Male	24,817 (55.3)	22,648 (54.0)	2,169 (74.5)	
Female	20,035 (44.7)	19,292 (46.0)	743 (25.5)	
Age, years (n, %)				< 0.001
< 30	15,067 (33.6)	14,866 (35.4)	201 (6.9)	
30–39	11,440 (25.5)	10,949 (26.1)	491 (16.9)	
40–49	9,648 (21.5)	8,731 (20.8)	917 (31.5)	
50–59	6,631 (14.8)	5,737 (13.7)	894 (30.7)	
≥ 60	2,066 (4.6)	1,657 (4.0)	409(14.0)	
Employment status				< 0.001
Employed	39,812 (88.8))	37,580 (89.6)	2,232 (76.6)	
Retired	5,040 (11.2)	4,360 (10.4)	680 (23.4)	
BMI status (n, %)				< 0.001
Underweight	2,093 (4.7)	2,084 (5.0)	9 (0.3)	
Normal	21,628 (48.2)	21,063 (50.2)	565 (19.4)	
Overweight	15,258 (34.0)	13,931 (33.2)	1,327 (45.6)	
Obesity	5,873 (13.1)	4,862 (11.6)	1,011 (34.7)	
FPG, mmol/L (n, %)				< 0.001
< 5.6	37,789 (84.3)	37,240 (88.8)	549 (18.9)	
5.6–6.1	3,381 (7.5)	2,632 (6.3)	749 (25.7)	
6.1–7.0	3,682 (8.2)	2,068 (4.9)	1,614 (55.4)	
Hypertension (n, %)				< 0.001
No	41,600 (92.7)	39,271 (93.6)	2,329 (80.0)	
Yes	3,252 (7.3)	2,669 (6.4)	583 (20.0)	
Dyslipidaemia (n, %)				< 0.001
No	33,355 (74.4)	31,903 (76.1)	1,452 (49.9)	
Yes	11,497 (25.6)	10,037 (23.9)	1,460 (50.1)	
Hyperuricemia (n, %)				< 0.001
No	38,995 (86.9)	36,762 (87.7)	2,233 (76.7)	
Yes	5,857 (13.1)	5,178 (12.3)	679 (23.3)	
sCr, $\mu\text{mol/L}$ (n, %)				< 0.001
< 57.2	11,178 (24.9)	10,723(25.6)	455(15.6)	
57.2–68.1	11,274 (25.1)	10,571(25.2)	703(24.1)	
68.2–78.9	11,246 (25.1)	10,397(24.8)	849(29.2)	
> 78.9	11,154 (24.9)	10,249(24.4)	905(31.1)	
eGFR, mL/min/1.73m <sup>2</sup> (n, %)				< 0.001
≥ 90	33,670 (75.1)	31,786 (75.8)	1,884 (64.7)	
< 90	11,182 (24.9)	10,154 (24.2)	1,028 (35.3)	

Categorical variables are presented as numbers (percentages) and compared using the chi-square test

Abbreviations: BMI, body mass index; FPG, fasting plasma glucose; sCr, serum creatinine; eGFR, estimated glomerular filtration rate

people can directly calculate their RS and categorize themselves into different risk levels and obtain a simple understanding of the risk of developing T2DM and how they may develop in the following years by reviewing the nomogram and Kaplan–Meier survival curve.

### Subgroup analysis

We conducted subgroup analysis to examine the effect of employment status on the RS model. A total of 39,812 participants were included into the employed cohort, with a median follow-up of 3.62 years (interquartile

**Table 2** Univariate and multivariate Cox proportional hazards regression analysis of new-onset diabetes in the training cohort

Variable names	Univariate Cox regression model		Multivariable Cox regression model	
	HR	95%CI	HR	95%CI
Gender				
Female	1		1	
Male	2.63	2.38–2.91	1.76	1.49–2.07
Age (years old)				
< 30	1		1	
30–39	2.98	2.45–3.62	1.54	1.27–1.88
40–49	5.96	4.97–7.15	1.99	1.64–2.40
50–59	9.44	7.88–11.32	2.45	2.02–2.98
≥ 60	14.95	12.23–18.26	3.41	2.75–4.24
BMI status				
Normal	1		1	
Underweight	0.20	0.10–0.41	0.46	0.23–0.92
Overweight	3.31	2.94–3.72	1.45	1.28–1.64
Obesity	7.19	6.36–8.13	2.22	1.95–2.54
FPG (mmol/L)				
< 5.6	1		1	
5.6–6.1	15.37	13.46–17.51	9.90	8.65–11.33
6.2–7.0	39.50	35.18–44.35	21.69	19.13–24.60
Hypertension				
No	1		1	
Yes	3.47	3.12–3.86	1.15	1.03–1.28
Dyslipidaemia				
No	1		1	
Yes	3.02	2.77–3.29	1.28	1.17–1.40
Hyperuricemia				
No	1		1	
Yes	2.17	1.96–2.40	1.04	0.94–1.16
sCr (μmol/L)				
< 57.2	1		1	
57.2–68.1	1.57	1.36–1.81	0.92	0.77–1.09
68.2–78.9	1.86	1.63–2.14	0.68	0.56–0.83
> 78.9	2.05	1.79–2.34	0.63	0.48–0.82
eGFR (mL/min/1.73m <sup>2</sup> )				
≥ 90	1		1	
< 90	1.52	1.39–1.66	0.96	0.82–1.12

HR was calculated with female vs. male sex, age < 30 years vs. age 30–39 years or age 40–49 years or age 50–59 years or age ≥ 60 years, normal vs. overweight or obesity, FPG < 5.6 mmol/L vs. FPG 5.6–6.1 mmol/L or 6.2–7.0 mmol/L, with vs. without hypertension, with vs. without dyslipidaemia, with vs. without hyperuricaemia, sCr < 57.2 μmol/L vs. sCr 57.2–68.1 μmol/L or sCr 68.2–78.9 μmol/L, eGFR ≥ 90 mL/min/1.73 m<sup>2</sup> vs. eGFR < 90 mL/min/1.73 m<sup>2</sup>

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; FPG, fasting plasma glucose; sCr, serum creatinine; eGFR, estimated glomerular filtration rate

range, 2.28–5.10 years), 2,232 (5.6%) employed developed new-onset T2DM. The nomogram had good accuracy in estimating the risk of developing diabetes in the

employed cohort, with a C-index of 0.910. In addition, 5,040 individuals were included into the retired cohort, with a median follow-up of 3.22 years (interquartile range, 2.61–5.08 years), 680 (13.5%) retired developed T2DM. The nomogram had good accuracy in the retired cohort, with a C-index of 0.867. The Kaplan–Meier survival curves for the employed and retired cohorts according to RS were shown in Figure S2, and each set of RS classes appeared well discriminated in the two cohorts ( $P < 0.001$ ).

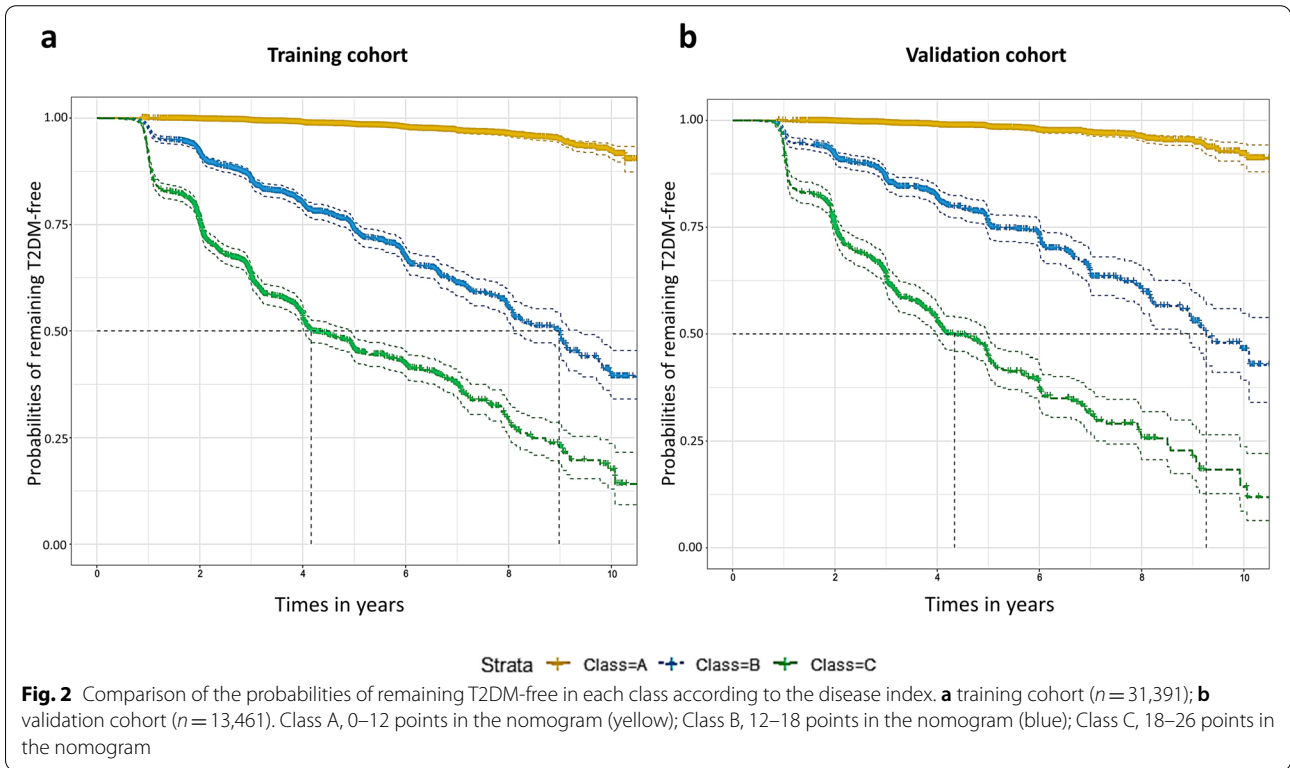
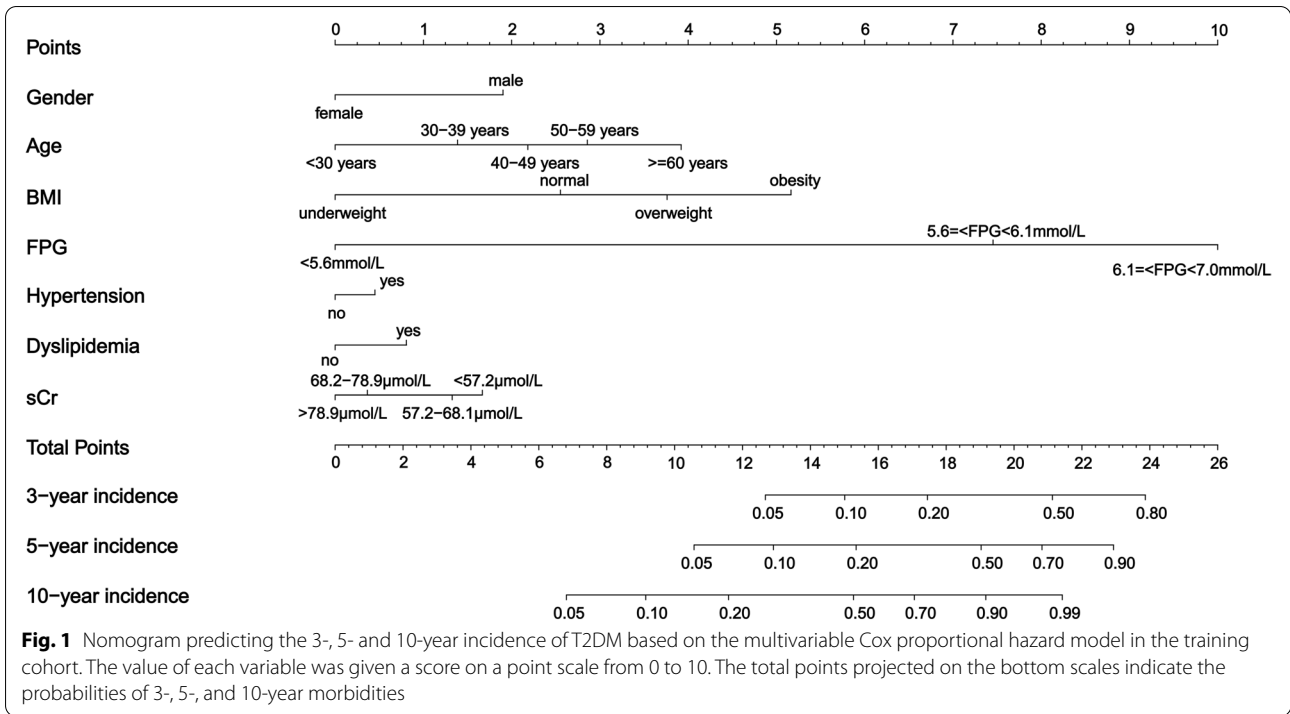
### Sensitivity analysis

We conducted sensitivity analysis by excluding the first two years of follow-up. A total of 37,694 participants were included into the medium- and long-term cohort. For over a median follow-up of 3.93 years (interquartile range, 2.95–5.69 years), 1,778 (4.6%) participants developed new-onset T2DM (10.39 events per 1,000 person-years at follow-up). The nomogram had good accuracy in estimating the risk of developing diabetes in the medium- and long-term cohort, with a C-index of 0.888. The Kaplan–Meier survival curve for the cohort according to RS was shown in Figure S1, and each set of RS classes appeared well discriminated in the cohort ( $P < 0.001$ ).

### Discussion

The primary aim of this study was to develop a dynamic nomogram and an RS model based on physical examination data to predict the probability of T2DM incidence. Multiple Cox regression demonstrated that male sex, older age, high BMI, high FPG, hypertension, dyslipidaemia, and low sCr at presentation were risk factors for T2DM. In addition, dynamic nomogram and RS models were constructed using weighted estimators corresponding to each covariate derived from the fitted Cox regression coefficients and variance estimates. The dynamic nomogram achieved a high C-index of 0.909 in the training set and 0.905 in the validation set. A tenfold cross-validation estimated the AUC of the nomogram at 0.909 (95% CI: 0.897–0.920).

According to the International Diabetes Federation Diabetes Atlas 9th edition (2019), diabetes is one of the fastest growing health challenges in the twenty-first century, and the number of adults with diabetes has more than tripled in the past 20 years [28]. Currently, China has the largest number of diabetes cases worldwide, and this number continues to grow, putting constant strains on healthcare resources [6]. Yang et al. interviewed 1482 adults with diabetes and 1553 adults with glucose tolerance in the normal range to estimate the expenditures for medical care caused by diabetes in China [29]. The study showed that expenditures for healthcare were 3.38



times higher among people with diabetes than among those with normal glucose tolerance. Fortunately, many studies have found that lifestyle interventions, such as

weight-loss, improved diet, and physical activity, substantially decreased the incidence of diabetes among high-risk individuals [30, 31]. Therefore, it is important



to identify individuals at risk of developing T2DM and provide targeted health guidance.

In the current study, for over a median follow-up of 3.58 years (interquartile range, 2.31–5.10 years), 2,912 (6.49%) participants developed new-onset T2DM, which is based on the Beijing Health Management Cohort. This cohort includes current or retired employees who are representative of the urban population in China, with a good follow-up stability. According to previous studies, the prevalence of diabetes is higher in Beijing compared to the overall Chinese population [32–34], which may be due to the fact that Beijing has the highest prevalence of overweight and obesity among children and adults in China [35].

A nationwide population-based survey in Germany showed that 78.9% of high-risk individuals perceived themselves as having diabetes risk as almost absent or slight, demonstrating that providing effective risk communication is essential [9]. In the past 20 years, experts have been striving to develop weighted models that can be presented as scores to reflect the complexity of prediction models while being perceived as sufficiently simple, credible, affordable, and widely implementable in practice [10, 11]. However, Schmid et al. found that the prevalence of participants at risk for T2DM varied considerably depending on the scoring system used [36]. Most of the RS models were described as “simple” or “easily implemented”, and less were specific about target users and environments [12]. To adequately prevent T2DM, a risk scoring system should be constructed for each population. Several investigators established special RS models of T2DM for the rural Chinese population, but the AUC values for all models were less than 0.800 [14, 15, 37]. Pires de Sousa et al. developed a simple risk score for the Brazilian urban population, and the model achieved an AUC of 0.720 through external validation [13]. Furthermore, a study following a cohort of 6342 middle-aged adults developed a risk score derived from simple biochemical examination, with an AUC of the model at 0.77. To our knowledge, few diabetes risk models have been constructed based on primary healthcare. Our dynamic nomogram and RS model are simple-to-use tools for assessing the risk of developing T2DM in the current or retired employees with a high accuracy.

Meanwhile, Schulze et al. reported a diabetes RS based on age, waist circumference, height, history of hypertension, physical activity, smoking, and consumption of red meat, whole-grain bread, coffee, and alcohol to screen the individuals at high risk of developing T2DM [38]. The area under the receiver-operator characteristic (ROC) curve was 0.84 in the European Prospective Investigation into Cancer and Nutrition (EPIC) -Potsdam and 0.82 in the EPIC-Heidelberg studies. However, a systematic

review suggested that the use of non-invasive diabetes risk assessment tools is limited. The barriers of these assessment tools are common in screening and clinical tests, and include interference with physician–patient interaction, lack of training, and lack of time [39]. He et al. compared the predictive accuracy of a polyexposure score (PXS), a polygenic risk score (PGS), and a clinical risk score (CRS). The PXS factors were selected from exposure variables including physiological state, environmental exposure, and self-reported behavior, the PGS was calculated from previously derived single nucleotide polymorphism (SNP) weights of > 6 million SNPs, and the CRS factors included sex, age, family history, BMI, SBP, FPG, HDL-C, and TG. The C-statistics for the PGS, PXS, and CRS models were 0.709, 0.762, and 0.839, respectively [40]. As discussed above, clinical factors play a key role in the diabetes RS model compared to self-reported behavior, environmental exposures, and genes, while self-reported behavior are limited by recall bias and reporting bias, and genetic testing is not a widely applied procedure; therefore, Our RS model was developed based on readily available clinical indicators.

Li et al. conducted a nomogram based on an urban community located in China with 687 participants and an internally validated AUC of 0.812 [41]. This study is the first to develop diabetes risk score for the urban population in China with a good accuracy, based on a large cohort. Older age, sex, BMI, FPG, hypertension and dyslipidaemia are significant predictors of T2DM incidence [42, 43]. Bao et al. conducted a multivariable Cox proportional hazards regression analysis in a general population sample from China [44]. They found that sCr concentration is inversely related to incident T2DM in both men and women, which could be an effective indicator for discovering people at high risk of diabetes. This finding was validated in the present study. Here, we derived that a T2DM dynamic nomogram and an RS based on participants' sex, age, BMI, FPG, hypertension, dyslipidaemia, and sCr. To our knowledge, this study is the first to construct a dynamic nomogram to predict the risk of developing T2DM. The dynamic nomogram demonstrated good accuracy in estimating the risk of developing diabetes, with a C-index of 0.909 in the training set and 0.905 in the validation set.

The strength of this study was that the participants were from the BHMC, current or retired employees in fixed work environments in Beijing, which effectively improved participant compliance and the completeness of data collection, and selection bias was controlled by the cohort. Although the cohort has not collected information on the employment status of the participants, according to the general retirement age in China, retirement was defined as age  $\geq 50$  years in women or  $\geq 60$  years in men. we conducted a subgroup

analysis for different employment status. In addition, sensitivity analysis was conducted in a cohort that excluded the first two years of follow-up to validate the accuracy of the model. Meanwhile, based on normal nomogram, dynamic nomogram was developed, which provides clinicians with a simple-to-use tool to tailor clinical decisions. Our study has several limitations. First, it should be noted that reporting bias in diseases exclusion stage may have limited their ability to obtain accurate risk estimates, a prospective study is needed to confirm the accuracy and reliability of the T2DM RS model in the real world. Second, other specific indicators, such as HbA1c and Homeostatic Model Assessment of Insulin Resistance (HOMA-IR), and can be added to further improve sensitivity and specificity. Thirdly, the models have not been externally validated, therefore, the model can currently be used in primary care and external validation is required for its clinical use. Finally, the cohort should be followed up, verifying the accuracy of the models.

## Conclusions

In summary, diabetes is one of the fastest growing health challenges in the twenty-first century, and health interventions can delay or prevent the occurrence and development of diabetes. Therefore, identifying high-risk groups is essential for the early initiation of health education and therapeutic interventions. The T2DM dynamic nomogram and RS model consisting of seven clinical characteristics that are routinely collected in primary healthcare offer clinicians and others who conduct physical examination, respectively, simple-to-use tools for assessing the risk of developing T2DM in health check-up population. We recommend that T2DM dynamic nomogram and RS model be used for the urban Chinese current or retired employees.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-022-14782-6>.

**Additional file 1.**

**Additional file 2.**

## Acknowledgements

The authors thank the data management team of the Beijing Physical Examination Center and the medical staff who collected the data.

## Authors' contributions

All authors contributed to the manuscript preparation. C.T. analyzed the data, contributed to the interpretation of the results, and wrote the manuscript; S.Z., Q.L., and J.B.Z. organized the data and conducted statistical analyses; X.H.G., L.X.T., and D.Q.Z. reviewed and edited the manuscript before submission; Y.M.H. and X.H.Y. conceived the analysis plan and reviewed the manuscript before submission. All authors have read and approved the final manuscript.

## Funding

This study was supported by Capital's Funds for Health Improvement and Research [grant numbers 2018–2–2242]; the National Key Research and

Development Program of China [grant numbers 2016YFC0900603]; the Natural Science Foundation of Beijing Municipality [grant numbers 7202011].

## Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Ethical approval and consent to participate

The project was approved by the ethics committee of the Beijing Physical Examination Center (Beijing, China; ethics approval ID: 201802 & 202008). All methods were carried out in accordance with relevant guidelines and regulations, and informed consent was obtained from 11,145 participants. As this was a retrospective study, the ethics committee of the Beijing Physical Examination Center approved the remaining participants with waived written informed consent.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>School of Public Health, Capital Medical University, NO.10 Xitoutiao, Youanmen, Beijing 100069, China. <sup>2</sup>Beijing Physical Examination Center, No. 59, Beiwei Road, Xicheng District, Beijing, China.

Received: 25 November 2021 Accepted: 30 November 2022

Published online: 09 December 2022

## References

- Marx N, Davies MJ, Grant PJ, Mathieu C, Petrie JR, Cosentino F, Buse JB. Guideline recommendations and the positioning of newer drugs in type 2 diabetes care. *Lancet Diabetes Endocrinol.* 2021;9:46–52.
- Park HY, Han K, Kim Y, Kim YH, Sur YJ. The risk of hip fractures in individuals over 50 years old with prediabetes and type 2 diabetes - A longitudinal nationwide population-based study. *Bone.* 2020;142:115691.
- Zafra-Tanaka JH, Miranda JJ, Gilman RH, Checkley W, Smeeth L, Bernabe-Ortiz A. Obesity markers for the prediction of incident type 2 diabetes mellitus in resource-poor settings: The CRONICAS Cohort Study. *Diabetes Res Clin Pract.* 2020;170:108494.
- Global report on diabetes. In: World Health Organization; 2016. <https://apps.who.int/iris/handle/10665/204871>. Accessed Nov 2020.
- Liu M, Liu SW, Wang LJ, Bai YM, Zeng XY, Guo HB, Liu YN, Jiang YY, Dong WL, He GX, et al. Burden of diabetes, hyperglycaemia in China from 2016: Findings from the 1990 to 2016, global burden of disease study. *Diabetes Metab.* 2019;45(3):286–93.
- Wu Z, Jin T, Weng J. A thorough analysis of diabetes research in China from 1995 to 2015: current scenario and future scope. *Sci China Life Sci.* 2019;62(1):46–62.
- Aujla N, Yates T, Dallosso H, Kai J. Users' experiences of a pragmatic diabetes prevention intervention implemented in primary care: qualitative study. *BMJ Open.* 2019;9(8):e028491.
- Gilis-Januszewska A, Lindström J, Tuomilehto J, Piwońska-Solska B, Topór-Mądry R, Szybiński Z, Peltonen M, Schwarz PE, Windak A, Hubalewska-Dydejczyk A. Sustained diabetes risk reduction after real life and primary health care setting implementation of the diabetes in Europe prevention using lifestyle, physical activity and nutritional intervention (DE-PLAN) project. *BMC Public Health.* 2017;17(1):198.
- Heidemann C, Paprott R, Stühmann LM, Baumert J, Mühlenbruch K, Hansen S, Schiborn C, Zahn D, Gellert P, Scheidt-Nave C. Perceived diabetes risk and related determinants in individuals with high actual diabetes risk: results from a nationwide population-based survey. *BMJ Open Diabetes Res Care.* 2019;7(1):e000680.
- Doi Y, Ninomiya T, Hata J, Hirakawa Y, Mukai N, Iwase M, Kiyohara Y. Two risk score models for predicting incident Type 2 diabetes in Japan. *Diabet Med: a journal of the British Diabetic Association.* 2012;29(1):107–14.



11. Sulaiman N, Mahmoud I, Hussein A, Elbadawi S, Abusnana S, Zimmet P, Shaw J. Diabetes risk score in the United Arab Emirates: a screening tool for the early detection of type 2 diabetes mellitus. *BMJ Open Diabetes Res Care*. 2018;6(1):e000489.
12. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. *BMJ (Clinical research ed)*. 2011;343:d7163.
13. Pires de Sousa AG, Pereira AC, Markezine GF, Marques do Nascimento-Neto R, Freitas SN, de CNRL, et al. Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population. *Eur J Epidemiol*. 2009;24(2):101–9.
14. Zhang H, Wang C, Ren Y, Wang B, Yang X, Zhao Y, et al. A risk-score model for predicting risk of type 2 diabetes mellitus in a rural Chinese adult population: A cohort study with a 6-year follow-up. *Diabetes Metab Res Rev*. 2017;33(7):e2911.
15. Zhang M, Zhang H, Wang C, Ren Y, Wang B, Zhang L, Yang X, Zhao Y, Han C, Pang C, et al. Development and Validation of a Risk-Score Model for Type 2 Diabetes: A Cohort Study of a Rural Adult Chinese Population. *PLoS One*. 2016;11(4):e0152054.
16. Lim NK, Park SH, Choi SJ, Lee KS, Park HY. A risk score for predicting the incidence of type 2 diabetes in a middle-aged Korean cohort: the Korean genome and epidemiology study. *Circ J*. 2012;76(8):1904–10.
17. Zhang X, Xu Z, Ran X, Ji L. Development and validation of a risk score model for prediction of lower extremity arterial disease in Chinese with type 2 diabetes aged over 50 years. *Endocr Connect*. 2021;10(9):1212–20.
18. Yang L, Li M, Zheng Q, Ren C, Ma W, Yang Y. A dynamic nomogram for predicting the risk of asthma: Development and validation in a database study. *J Clin Lab Anal*. 2021;35(7):e23820.
19. Liu J, Zhao Z, Mu Y, Zou X, Zou D, Zhang J, et al. Gender Differences in the Association between Serum Uric Acid and Prediabetes: A Six-Year Longitudinal Cohort Study. *Int J Environ Res Public Health*. 2018;15(7):1560.
20. Xiong R, Zhao C, Zhong M, Zhang X, Liu W. Effects of Shenqi compound on intestinal microbial metabolites in patients with type 2 diabetes: A protocol for systematic review and meta analysis. *Medicine*. 2020;99(48):e23017.
21. Chen C, Lu FC. The guidelines for prevention and control of overweight and obesity in Chinese adults. *Biomed Environ Sci: BES*. 2004;17(Suppl):1–36.
22. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, Jones DW, Materson BJ, Oparil S, Wright JT Jr, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*. 2003;289(19):2560–72.
23. Zhang M, Deng Q, Wang L, Huang Z, Zhou M, Li Y, Zhao Z, Zhang Y, Wang L. Prevalence of dyslipidemia and achievement of low-density lipoprotein cholesterol targets in Chinese adults: A nationally representative survey of 163,641 adults. *Int J Cardiol*. 2018;260:196–203.
24. Liu L, Lou S, Xu K, Meng Z, Zhang Q, Song K. Relationship between lifestyle choices and hyperuricemia in Chinese men and women. *Clin Rheumatol*. 2013;32(2):233–9.
25. Zhu B, Wang J, Chen K, Yan W, Wang A, Wang W, Gao Z, Tang X, Yan L, Wan Q, et al. A high triglyceride glucose index is more closely associated with hypertension than lipid or glycemic parameters in elderly individuals: a cross-sectional survey from the Reaction Study. *Cardiovasc Diabetol*. 2020;19(1):12.
26. Hortaigh BÓ, Gransar H, Callister T, Shaw LJ, Schulman-Marcus J, Stuijffzand WJ, Valenti V, Cho I, Szymonińska J, Lin FY, et al. Development and Validation of a Simple-to-Use Nomogram for Predicting 5-, 10-, and 15-Year Survival in Asymptomatic Adults Undergoing Coronary Artery Calcium Scoring. *JACC Cardiovasc Imaging*. 2018;11(3):450–8.
27. Hird TR, Zomer E, Owen A, Chen L, Ademi Z, Magliano DJ, Liew D. The impact of diabetes on productivity in China. *Diabetologia*. 2019;62(7):1195–203.
28. International Diabetes Federation. *Diabetes Atlas 9th edition 2019*. In., vol. 2020. 2019. <https://www.diabetesatlas.org/>. Accessed Nov 2020.
29. Yang W, Zhao W, Xiao J, Li R, Zhang P, Kissimova-Skarbek K, Schneider E, Jia W, Ji L, Guo X, et al. Medical care and payment for diabetes in China: enormous threat and great opportunity. *PLoS One*. 2012;7(9):e39513.
30. Block G, Azar KM, Romanelli RJ, Block TJ, Palaniappan LP, Dolginsky M, Block CH. Improving diet, activity and wellness in adults at risk of diabetes: randomized controlled trial. *Nutr Diabetes*. 2016;6(9):e231.
31. Costa B, Barrio F, Cabré JJ, Piñol JL, Cos X, Solé C, Bolívar B, Basora J, Castell C, Solà-Morales O, et al. Delaying progression to type 2 diabetes among high-risk Spanish individuals is feasible in real-life primary healthcare settings using intensive lifestyle intervention. *Diabetologia*. 2012;55(5):1319–28.
32. He J, Fang A, Yu S, Shen X, Li K. Dietary Nonheme, Heme, and Total Iron Intake and the Risk of Diabetes in Adults: Results From the China Health and Nutrition Survey. *Diabetes Care*. 2020;43(4):776–84.
33. Liu H, Yan S, Chen G, Li B, Zhao L, Wang Y, Hu X, Jia X, Dou J, Mu Y, et al. Association of the Ratio of Triglycerides to High-Density Lipoprotein Cholesterol Levels with the Risk of Type 2 Diabetes: A Retrospective Cohort Study in Beijing. *J Diabetes Res*. 2021;2021:5524728.
34. Xie L, Zhao X, Zhang B, Zhu H. Epidemiology and risk factors for diabetes in the suburbs of Beijing: a retrospective cohort study. *BMJ Open*. 2021;11(3):e041526.
35. Pan XF, Wang L, Pan A. Epidemiology and determinants of obesity in China. *Lancet Diabetes Endocrinol*. 2021;9(6):373–92.
36. Schmid R, Vollenweider P, Waeber G, Marques-Vidal P. Estimating the risk of developing type 2 diabetes: a comparison of several risk scores: the Cohorte Lausannoise study. *Diabetes Care*. 2011;34(8):1863–8.
37. Chen X, Wu Z, Chen Y, Wang X, Zhu J, Wang N, Jiang Q, Fu C. Risk score model of type 2 diabetes prediction for rural Chinese adults: the Rural Deqing Cohort Study. *J Endocrinol Invest*. 2017;40(10):1115–23.
38. Schulze MB, Hoffmann K, Boeing H, Linseisen J, Rohrmann S, Möhlig M, Pfeiffer AF, Spranger J, Thamer C, Häring HU, et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care*. 2007;30(3):510–5.
39. Dhappayom T, Chaiyakunapruk N, Krass I. How diabetes risk assessment tools are implemented in practice: a systematic review. *Diabetes Res Clin Pract*. 2014;104(3):329–42.
40. He Y, Lakhani CM, Rasooly D, Manrai AK, Tzoulaki I, Patel CJ. Comparisons of Polyexposure, Polygenic, and Clinical Risk Scores in Risk Prediction of Type 2 Diabetes. *Diabetes Care*. 2021;44(4):935–43.
41. Li L, Wang Z, Zhang M, Ruan H, Zhou L, Wei X, Zhu Y, Wei J, He S. New risk score model for identifying individuals at risk for diabetes in southwest China. *Prev Med Rep*. 2021;24:101618.
42. Arellano-Campos O, Gómez-Velasco DV, Bello-Chavolla OY, Cruz-Bautista I, Melgarejo-Hernandez MA, Muñoz-Hernandez L, Guillén LE, Garduño-García JJ, Alvirde U, Ono-Yoshikawa Y, et al. Development and validation of a predictive model for incident type 2 diabetes in middle-aged Mexican adults: the metabolic syndrome cohort. *BMC Endocr Disord*. 2019;19(1):41.
43. Liu S, Gao Y, Shen Y, Zhang M, Li J, Sun P. Application of three statistical models for predicting the risk of diabetes. *BMC Endocr Disord*. 2019;19(1):126.
44. Bao X, Gu Y, Zhang Q, Liu L, Meng G, Wu H, Xia Y, Shi H, Wang H, Sun S, et al. Low serum creatinine predicts risk for type 2 diabetes. *Diabetes Metab Res Rev*. 2018;34(6):e3011.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

