Received: 20 April 2023 | Revised: 3 September 2023 | Accepted: 4 September 2023

DOI: 10.1111/dom.15291

WILEY

ORIGINAL ARTICLE

Differential effect of interventions in patients with prediabetes stratified by a machine learning-based diabetes progression prediction model

Xiantong Zou MD[1] | Yingying Luo MD[1] | Qi Huang MD[1] | Zhanxing Zhu PhD[2,3,4] | Yufeng Li MD[5] | Xiuying Zhang MD[1] | Xianghai Zhou MD[1] | Linong Ji MD[1]

[1]Peking University People's Hospital, Beijing, China

[2]School of Mathematical Sciences, Peking University, Beijing, China

[3]Center for Data Science, Peking University, Beijing, China

[4]Beijing Institute of Big Data Research, Beijing, China

[5]Department of Endocrinology, Beijing Friendship Hospital Pinggu Campus, Capital Medical University, Beijing, China

Correspondence
Linong Ji, Peking University People's Hospital, Beijing, 100044, China.
Email: jiln@bjmu.edu.cn

Funding information
Beijing Municipal Science and Technology Commission, Grant/Award Number: Z201100005520013; Beijing Nova Cross program, Grant/Award Number: Z211100002121169; Beijing Nova Program of Science and Technology, Grant/Award Number: Z191100001119026; Clinical Medicine Plus X - Young Scholars Project, Peking University, the Fundamental Research Funds for the Central Universities, Grant/Award Number: PKU2022LCXQ004; National Natural Science Foundation of China, Grant/Award Number: 81970708; Peking University People's Hospital Research And Development Funds, Grant/Award Number: RDH2021-10

Abstract

Aim: To investigate whether stratifying participants with prediabetes according to their diabetes progression risks (PR) could affect their responses to interventions.

Methods: We developed a machine learning-based model to predict the 1-year diabetes PR (ML-PR) with the least predictors. The model was developed and internally validated in participants with prediabetes in the Pinggu Study (a prospective population-based survey in suburban Beijing; $n = 622$). Patients from the Beijing Prediabetes Reversion Program cohort (a multicentre randomized control trial to evaluate the efficacy of lifestyle and/or pioglitazone on prediabetes reversion; $n = 1936$) were stratified to low-, medium- and high-risk groups using ML-PR. Different effect of four interventions within subgroups on prediabetes reversal and diabetes progression was assessed.

Results: Using least predictors including fasting plasma glucose, 2-h postprandial glucose after 75 g glucose administration, glycated haemoglobin, high-density lipoprotein cholesterol and triglycerides, and the ML algorithm XGBoost, ML-PR successfully predicted the 1-year progression of participants with prediabetes in the Pinggu study [internal area under the curve of the receiver operating characteristic curve 0.80 (0.72–0.89)] and Beijing Prediabetes Reversion Program [external area under the curve of the receiver operating characteristic curve 0.80 (0.74–0.86)]. In the high-risk group pioglitazone plus intensive lifestyle therapy significantly reduced diabetes progression by about 50% at year l and the end of the trial in the high-risk group compared with conventional lifestyle therapy with placebo. In the medium- or low-risk group, intensified lifestyle therapy, pioglitazone or their combination did not show any benefit on diabetes progression and prediabetes reversion.

Conclusions: This study suggests personalized treatment for prediabetes according to their PR is necessary. ML-PR model with simple clinical variables may facilitate personal treatment strategies in participants with prediabetes.

Xiantong Zou, Yingying Luo and Qi Huang authors are equal contribution.

KEYWORDS
lifestyle, machine learning, pioglitazone, prediabetes, progression

1 | INTRODUCTION

More than one-third of the Chinese population has prediabetes,[1] a condition with mildly elevated fasting glucose and postprandial blood glucose levels yet not meeting the diagnostic criteria of diabetes. Prediabetes carries high risks of developing type 2 diabetes, cardiovascular diseases and premature death.[2,3] According to the Daqing Study, prediabetes is associated with an increased risk of mortality in China, and much of this excess risk is attributable to the development of type 2 diabetes.[4] Accordingly, controlling the development of diabetes in prediabetes may facilitate the prevention of premature death in China.

Several interventions were shown to be effective in diabetes prevention in participants with prediabetes, such as intensive lifestyle therapy and medications, including metformin,[5] pioglitazone[6] and acarbose.[7] In fact, only ~5%–10% of participants without intervention progressed to diabetes every year,[8,9] yet the progression to type 2 diabetes in participants with prediabetes was highly heterogeneous.[10] Currently, clinical guidelines only suggest metformin treatment for prediabetes in a few patients [e.g. those with body mass index (BMI) ≥35 kg/m^2, those aged <60 years or women with gestational diabetes].[11] There is a clinical demand for detailed algorithms to evaluate intricately the diabetes progression risks (PRs) and to recommend further corresponding therapies in participants with prediabetes.

Recently, the STOP DIABETES study used glucose levels, insulin resistance, β-cell function and other cardiovascular risk factors to stratify participants with prediabetes into different PRs and treated with different regimens.[12] The implementation of personalized interventions requires precise methodologies for diabetes prediction. In recent decades, the adoption of machine learning (ML) in diabetes prediction has seen a significant upswing because of its capacity to handle complex datasets, incorporate multimodal variables, deliver enhanced prediction accuracy and foster personalized health care.[13–16] Current advances in artificial intelligence, particularly ML algorithms, in diabetes prediction were summarized and reviewed.[17] In brief, a myriad of variables were chosen for diabetes prediction, encompassing anthropometric measurements,[18] laboratory test results[19] and even radiographic images.[20] Widely employed algorithms, including logistic regression, random forest (RF), decision tree, gradient boosting machine and deep neural network, have been recognized for their superior predictive accuracy[17,21] (details reviewed in the Supporting Information section). The model performance may depend on the coherence between the developing and validation datasets of the model.[14] Here, we developed an ML-based model to evaluate the risk of diabetes progression in a population-based cohort and used this model to stratify patients in a clinical trial to investigate whether participants with different PRs required different treatment regimens.

2 | MATERIALS AND METHODS

2.1 | Pinggu study

The Pinggu Study was a population-based prospective study conducted in the Pinggu district of Beijing, China, to investigate the natural history of metabolic disorders. Between March 2012 and May 2013, participants were recruited using a stratified, random, two-stage cluster sampling process according to sex and age, and complete baseline data were collected from 3350 participants. From September 2013 to July 2014, 2846 individuals completed the 1-year follow-up investigation. The detailed study design and participant characteristics were presented previously.[22] An oral glucose tolerance test (OGTT) was performed in all participants without previously diagnosed diabetes at baseline and during the 1-year follow-up. In total, 630 participants were diagnosed with prediabetes at baseline, and 622 participants were included in the final analysis after excluding three individuals without follow-up OGTT data and five participants with key missing measurements. All patients with prediabetes in the Pinggu cohort were not on antidiabetic drugs; those who were on antidiabetic drugs were excluded. All eligible patients would be on conventional lifestyle therapy once they were diagnosed according to current Chinese guidelines,[23] yet the detailed lifestyle of these patients was undocumented. The collection of clinical parameters was described previously.[22,24]

The baseline survey was approved by the Ethics and Human Subject Committee of the Peking University People's Hospital, and the follow-up survey (Pinggu Metabolic Disease Study) was approved by the ethics committee of the Peking University Health Science Center. All participants provided written informed consent.

2.2 | Beijing Prediabetes Reversion Program

The Beijing Prediabetes Reversion Program (BPRP) is a prospective, multicentre, randomized, 2 × 2 factorial designed study clinical trial to investigate whether intensive lifestyle intervention and/or pioglitazone may revert patients with prediabetes to normal glucose regulation (NGR; which is also called normoglycaemia) from 2007 to 2014. In total, 1945 participants with prediabetes were randomized into four treatment groups: conventional lifestyle (CON) + placebo (PLA), CON + pioglitazone hydrochloride 30 mg daily (PIO), intensive lifestyle therapy (INT) + PLA, and INT + PIO 30 mg daily. The follow-up time for this study was 3 years, and OGTT was conducted in each patient every year to define the glucose tolerance status. The follow-up was terminated if patients reverted to NGR, progressed to diabetes, or dropped out. The major finding of BPRP is that all the other three groups did not show any significant benefit in prediabetes

regression compared with CON + PLA, although INT + PIO mildly reduced the risk for diabetes progression. Our study included 1936 participants after excluding nine patients without key variables. The complete inclusion and exclusion criteria, randomization, follow-up strategy and methods for clinical measurements were published previously.[25,26]

Approval of the protocol and consent forms by the local institutional review board was obtained at Peking University Health Science Center. Trial registration: http://www.chictr.org.cn: ChiCTR-PRC-06000005.

2.3 | Definition of the conditions

Prediabetes, diabetes and normal glucose status were all defined according to WHO 1999[27] criteria using fasting plasma glucose (FPG) and 2-h postprandial glucose after 75 g glucose administration (PG2h) during the OGTT. Prediabetes: FPG ≥6.1 and <7.0 mmol/L or PG2h ≥7.8 and <11.0 mmol/L; diabetes: FPG ≥7.0 or PG2h ≥11.1 mmol/L; NGR: FPG <6.1 and PG2h <7.8 mmol/L. We did not use glycated haemoglobin (HbA1c) as a diagnostic criterion and did not change the FPG cut-off value to 5.6 mmol/L according to the American Diabetes Association[11] because (a) HbA1c is neither sensitive nor specific for detecting prediabetes,[5] and (b) current clinical guidelines in China primarily use the WHO 1999 criteria to diagnose prediabetes.

Missing data were imputed by using multiple imputation techniques. DBP, LDL-C and estimated glomerular filtration rate were excluded because of high collinearity with other covariates. Seventeen variables were included to develop ML-based models to predict whether the patient progressed to diabetes at the 1-year follow-up. Algorithms including logistic regression, RF, support vector machine, simple decision tree and XGBoost were adopted to develop the models. As the dataset is imbalanced with the predicted outcome, we used an upsampling method to replicate cases with diabetes progression and increase their representation in advance. Parameters were tuned by five-fold cross-validation in the derivation cohort. To achieve the best clinical predictive efficacy using the fewest parameters, we reconstructed the models using the five most important variables and calibrated the models to achieve the best performance in the internal validation dataset, as evaluated by the highest area under the curve of the receiver operating characteristic curve (ROC AUC). We also tested the ROC AUC of progression prediction using conventional risk assessment tools, including the Framingham score and Cambridge Diabetes Risk Score, as described previously.[29,30]

The equation of the Cambridge Score is:

$$1/1+e^{-\left(\begin{array}{c}-6.322-0.879\times \text{Female}+1.222\times I(\text{Prescribed antihypertensive medication})+2.191\times I(\text{Prescribed steriods})+\\ 0.699\times I(25\le \text{BMI}<27.5)+1.970\times I(27.5\le \text{BMI}<30)+2.518\times I(\text{BMI}\ge 30)+0.728\times I(\text{Parent or sibilings with diabetes})\\ +0.753\times I(\text{Parent and sibiling with diabetes})-0.218\times I(\text{Ex smoker})+0.855\\ \times I(\text{current smoker})\end{array}\right)}$$

2.4 | Development and validation of a machine learning-based 1-year diabetes progression model

In total, 622 patients with prediabetes were randomly divided 2:1 into the derivation cohort ($N = 413$) and the internal validation cohort ($N = 209$). Variables related to diabetes, including age, sex, BMI, smoking status, family history of diabetes, systolic blood pressure (SBP), diastolic blood pressure (DBP), alanine transaminase, aspartate transaminase, serum creatinine, triglycerides (TGs), total cholesterol, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), estimated glomerular filtration rate, albumin-to-creatinine ratio, FPG, PG2h, HbA1c and fasting insulin, were selected as predictors to develop the model. Laboratory tests were conducted at the central lab of each trial as previously described.[26,28] The interassay coefficient of variation of HbA1c was 1.31% at a mean value of 4.74%,[28] and the interassay coefficient of variation for biochemical tests, including glucose, HDL-C and TG, ranged from 1.5% to 2.2%. We assume that the small error would not affect the model performance.

The equation of the Framingham score is

$$\begin{aligned}\text{Risk} = {} & 10\times I(6.1\le \text{FPG}<7.0)+2\times I(25\le \text{BMI}<30)+5\times I(\text{BMI}\ge 30)\\ & +2\times I(\text{SBP}\ge 130)+2\times I(\text{SBP}<130\&\text{DBP}\ge 85)+3\\ & \times I(\text{Male}\&\text{HDL}<1.03)+3\times I(\text{Female}\&\text{HDL}<1.29).\end{aligned}$$

Other metrics to describe the model, including accuracy, sensitivity, specificity, positive predictive value, negative predictive value and F1 score, were assessed to select the best model, and the precision-recall curve was presented. For clinical convenience, we also tried to construct the models using five variables that can be obtained by a simple fasting blood sample. The model with the highest ROC AUC in the internal validation cohort was selected as the ML-based model to predict the 1-year diabetes PR (ML-PR) model. The output of the ML-PR model was a score presenting the PR of a patient.

Patients in the CON + PLA arm in the BPRP cohort ($N = 478$) were selected as the external validation cohort of the ML-PR mode to avoid the interference of interventions. Model performance was assessed by ROC AUC in the prediction of 1-year diabetes progression. We divided the patients into tertiles according to the score predicted by the ML-PR model (Figure 1).

2.5 | Outcomes

The primary outcome of BPRP is the proportion of participants who reverted to a normal glucose state during follow-up. In our specific post hoc study, the primary outcomes are the difference among four treatment arms in the effect of prediabetes reversal and diabetes progression within each risk group at the end of the first year and the end of the trial.

2.6 | Statistical analysis

Data are presented as the mean ± SD for normally distributed continuous variables and as the median (interquartile range) for variables with skewness. Baseline characteristics among the three risk groups were compared using one-way ANOVA or the Kruskal-Wallis test for data with skewness or unequal variance if applicable. Sankey diagrams were used to visualize the prognosis of patients (who dropped out, progressed to diabetes, reverted to NRG, and remained prediabetic), with different risk stratifications at 1 year and the end of the trial. As other outcomes would preclude the occurrence of the primary event of interest, we applied an estimated competing-risks regression model and estimated the subdistribution hazard ratio (SHR) and 95% confidence interval (CI) of participants who reverted to normoglycaemia or progressed to diabetes at 1 year and the end of the trial in each risk group. Statistical analysis was conducted using R version 4.0.2 (R Foundation for Statistical Computing). Multiple imputations were constructed using the *mice* package.

3 | RESULTS

3.1 | Performance of the machine learning-based model to predict the 1-year diabetes progression risk model

Our study encompassed 622 patients with prediabetes from the Pinggu cohort and 1936 participants from the BPRP. The Pinggu set

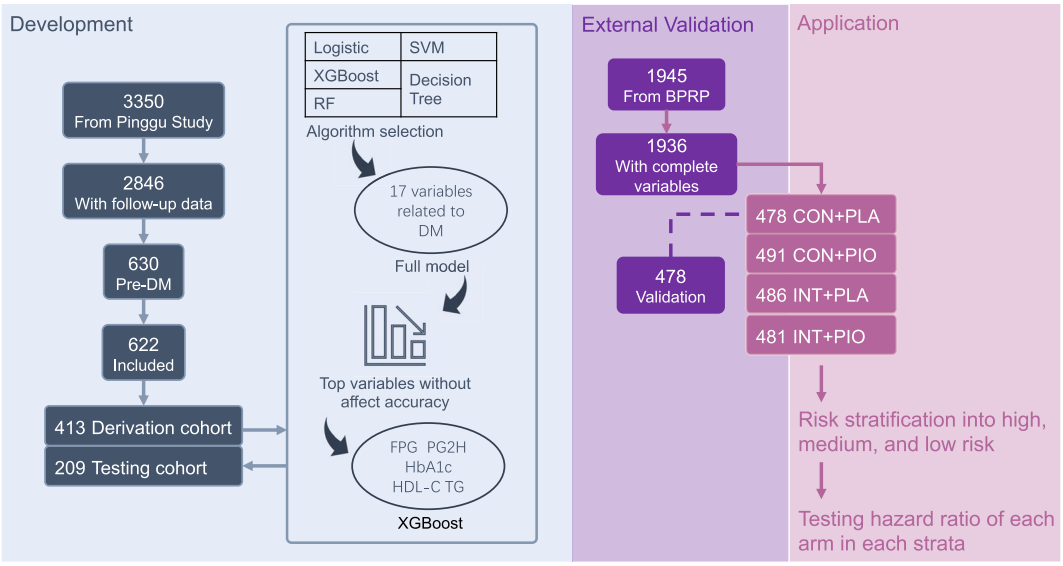

FIGURE 1 Study diagram. In the Pinggu cohort study, eight patients with prediabetes were excluded because of the absence of crucial variables. Consequently, a total of 622 patients were included in the final analysis. This cohort was divided in a 2:1 ratio, with 413 patients forming the derivation cohort for model development and the remaining 209 patients constituting the testing cohort for internal validation. Several algorithms, namely, logistic regression, XGBoost, RF, SVM and decision tree, were evaluated for their efficacy in predicting 1-year diabetes progression. The original models incorporated 17 diabetes-related variables, and they were subsequently refined using only the top five variables without compromising its accuracy. Selection of variables and algorithms was based on model accuracy, area under the curve of the receiver operating characteristic curve and F1 score evaluated in the internal validation cohort. Ultimately, XGBoost with variables FPG, PG2h, HbA1c, HDL-C and TG was chosen and termed as ML-PR. For the BPRP trial, nine patients were excluded, and those in the placebo combined with conventional lifestyle therapy arm served as the external validation dataset ($n = 478$). Within the BPRP, participants were categorized into low-, medium- and high-risk groups based on ML-PR predictions. The variation in progression and reversion risks across different treatments was examined for each risk stratum. BPRP, Beijing Prediabetes Reversion Program; CON, conventional lifestyle therapy; DM, diabetes mellitus; FPG, fasting plasma glucose; HbA1c, haemoglobin A1c; HDL-C, high-density lipoprotein cholesterol; INT, intensified lifestyle therapy; ML-PR, machine learning-based model for diabetes progression risk; PG2h, 2 h-post prandial plasma glucose after 75 g glucose administration; PIO, pioglitazone; PLA, placebo; RF, random forest; SVM, support vector machine; TG, triglycerides; XGBoost, machine learning algorithm.

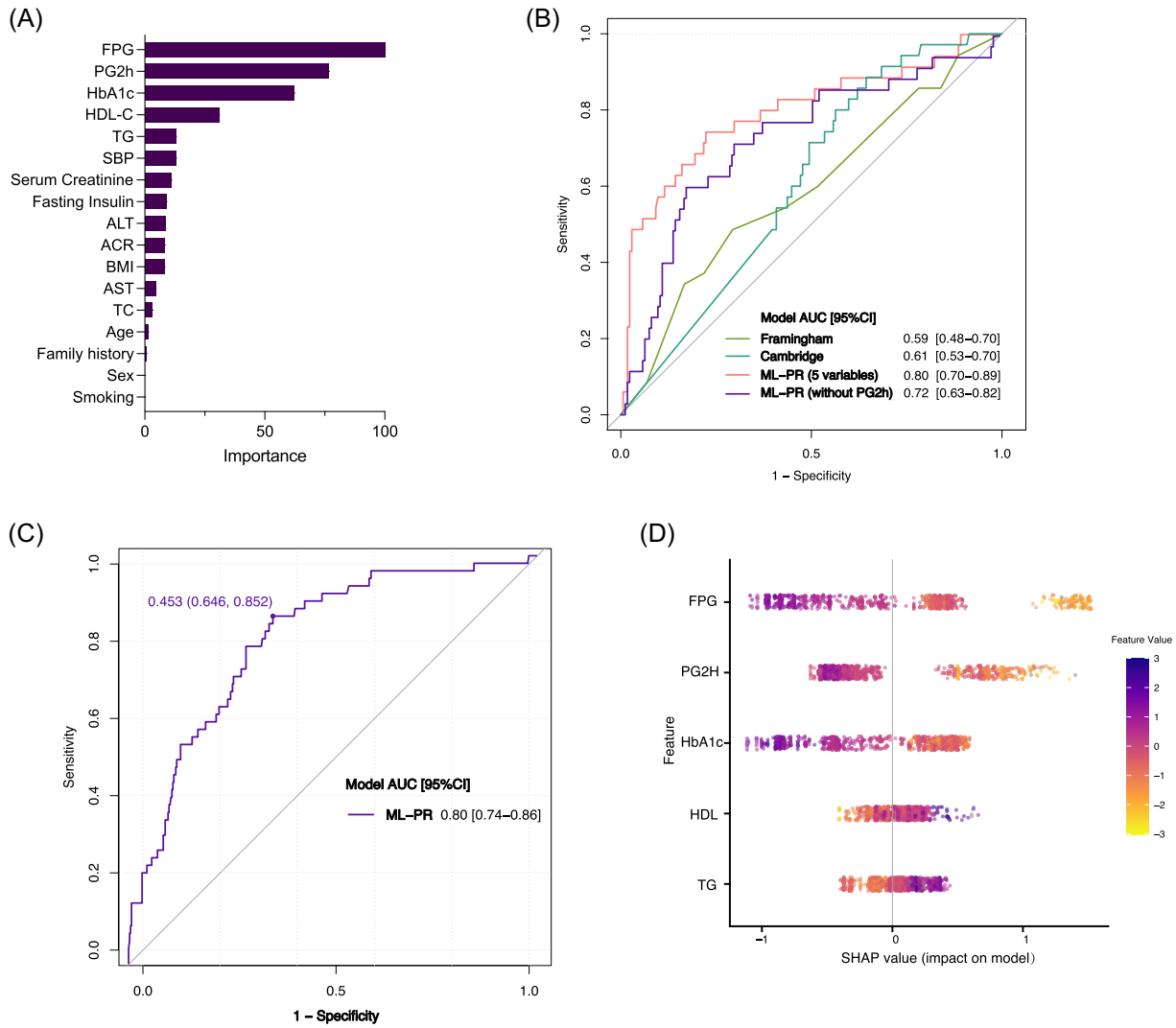

FIGURE 2 Model performance and feature importance of ML-PR. An ML-PR was developed using XGBoost. (A) The importance of 17 original variables outputted by XGBoost in the derivation cohort of the Pinggu study. (B) The receiver operating characteristic curve of previous Framingham score (light green) and Cambridge score (dark green), ML-PR with top five most important variables (pink) and top five most important variables without PG2h (dark purple) in the internal validation cohort ($n = 209$). (C) The receiver operating characteristic curve of ML-PR model with top five most important variables (including FPG, PG2h, HbA1c, HDL-C and TGs) in the external validation cohort ($N = 478$, the conventional lifestyle therapy plus placebo arm of the Beijing Prediabetes Reversion Program). The Youden index, calculated as the maximum sensitivity + specificity − 1, of ML-PR is 0.453, with a sensitivity of 0.646 and specificity of 0.852. (D) The SHAP feature importance for variables in ML-PR model with top five most important variables. In this SHAP summary plot, the *y*-axis ranks all predictors by importance, and the *x*-axis represents impact of features on model. Each point represents a participant in cohorts. ACR, albumin-to-creatinine ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; BMI, body mass index; CI, confidence interval; FPG, fasting plasma glucose; HbA1c, glycated haemoglobin; HDL-C, high-density lipoprotein; ML-PR, machine learning-based diabetes progression risk; PG2h, 2-h postprandial glucose after 75 g glucose administration; SBP, systolic blood pressure; SHAP, Shapley additive explanations; TC, total cholesterol; TG, triglyceride; XGBoost, machine learning algorithm.

was randomly split in a 2:1 ratio, forming a derivation dataset ($N = 413$) and an internal validation dataset ($N = 209$). The CON + PLA arm from the BPRP was utilized as the external validation dataset ($N = 478$) (see Figure 1). Baseline characteristics such as age, BMI and HbA1c were consistent between both groups, but the Pinggu cohort had a higher male percentage. Within the Pinggu cohort, 105 (16.7%) developed diabetes, and 199 (31.7%) reverted to NRG. Of the 1936 patients in the BPRP, 195 (10%) progressed to diabetes, and 729 (37.5%) did so by the end of the first year (detailed in Table S1).

The distribution of the 17 variables within the Pinggu cohort is shown in Figure S1. Employing these variables as predictors, both XGBoost and RF algorithms showed high ROC AUC (95% CI) for predicting the 1-year diabetes progression in the derivation cohort of the Pinggu study: XGBoost yielded 0.92 (0.90–0.93) using 17 variables and 0.90 (0.89–0.91) using the top five variables, while RF achieved

TABLE 1 Baseline characteristics of patients by ML-PR strata in the Beijing Prediabetes Reversion Program study

Characteristic	Low-risk	Medium-risk	High-risk	*p*
N	646	645	645	
Range of ML-PR score	0.07–0.30	0.30–0.60	0.60–0.91	<.001
CON + PLA	165 (25.5)	156 (24.8)	157 (24.3)	.94
CON + PIO	160 (24.8)	162 (25.1)	169 (26.2)	
INT + PLA	164 (25.4)	169 (26.2)	153 (23.7)	
INT + PIO	157 (24.3)	158 (24.5)	166 (25.7)	
Calorie intake per day, kcal	1650 (630)	1622 (579)	1593 (527)	.21
Age, years	50.6 (10.1)	52.6 (9.5)	53.7 (9.2)	<.001
Male, %	244 (37.7)	283 (43.9)	295 (45.7)	.009
BMI, kg/m^2	26.05 (3.1)	26.08 (3.0)	26.69 (2.9)	<.001
Body weight, kg	69.7 (11.3)	70.5 (10.9)	72.0 (10.8)	<.001
CRP, mg/L	2.01 (3.3)	2.21 (6.3)	2.27 (3. 8)	.750
HbA1c, %	5.57 (0.3)	5.75 (0.3)	6.14 (0.49)	<.001
HbA1c, mmol/mol	37.2 (3.1)	39.1 (3.7)	43.4 (4.3)	<.001
FPG, mmol/L	5.53 (0.6)	6.00 (0.5)	6.21 (0.5)	<.001
PG2h, mmol/L	8.24 (1.2)	8.67 (1.4)	9.61 (0.9)	<.001
FINS[a], μU/ml	9.99 (6.0)	10.81 (7.4)	11.03 (6.5)	.011
HOMA-B[a]	108.49 (86.7)	88.19 (84.8)	85.58 (58.2)	.001
HOMA-IR[a]	2.47 (1.50)	2.88 (2.0)	3.04 (1.80)	.001
SBP, mmHg	121.28 (13.3)	122.55 (14.4)	121.67 (13.3)	.656
DBP, mmHg	77.00 (8.6)	77.24 (8.9)	76.93 (9.1)	.646
Triglycerides[a], mmol/L	1.67 (1.1)	1.79 (1.11)	1.83 (1.1)	.001
LDL-C, mmol/L	3.09 (0.8)	3.22 (0.8)	3.25 (0.8)	.003
HDL-C, mmol/L	1.28 (0.3)	1.22 (0.3)	1.19 (0.3)	<.001
Current smoker, %	127 (19.7)	146 (22.5)	162 (25.2)	.057

Note: Data were presented as mean ± SD for normally distributed continuous variables and n (%) for categorical variables and significance among groups was tested using ANOVA and χ^2 test, respectively.
Abbreviations: BMI, body mass index; CON, conventional lifestyle therapy; CRP, C-reaction protein; DBP, diastolic blood pressure; FINS, fasting insulin; FPG, fasting plasma glucose; HbA1c, haemoglobin A1c; HDL-C, high-density lipoprotein cholesterol; HOMA-B, Homeostasis Model Assessment of β-cell function; HOMA-IR, Homeostasis Model Assessment of insulin resistance; INT, intensified lifestyle therapy; LDL-C, low-density lipoprotein cholesterol; ML-PR, machine learning-based model to predict the 1-year diabetes progression risk; PG2h, 2h-postprandial plasma glucose after 75 g glucose administration; PIO, pioglitazone; PLA, placebo; SBP, systolic blood pressure; WC, waist circumference.
[a]Skewed continuous variables, which were expressed as medium (IQR) and tested using Kruskal-Wallis test.

0.98 (0.98–0.99) with 17 variables and 0.99 (0.99–1.00) with the top five. In internal validation validations, the ROC AUC of both methods was comparable, yet XGBoost displayed superior accuracy and F1 score compared with RF (Table S2). Consequently, XGBoost, leveraging the five pivotal variables, FPG, PG2h, HbA1c, HDL-C and TG, was designated ML-PR. Blood glucose levels emerged as the most influential variable within ML-PR (Figure 2A). Notably, the ML-PR (utilizing five variables) outperformed traditional risk metrics, such as the Cambridge and Framingham scores, in the internal validation cohort ($n = 209$) (Figure 2B). To explore further the clinical utility, we attempted to omit PG2h from the model, integrating FPG, HbA1c, HDL-C, TG and serum creatinine. However, this alteration caused the ROC AUC of the XGBoost model to drop to 0.72 (0.63–0.82) (Figure 2B). The external ROC AUC for ML-PR stood at 0.80 (0.74–0.86) (Figure 2C), with the SHAP feature importance of the ML-PR model depicted in Figure 2D. The parameter calibration for all ML models can be found in Tables S3 and S4, and the calibration curve is presented in Figure S2.

3.2 | Stratifying the Beijing Prediabetes Reversion Program into three progression risk groups

We applied the ML-PR model, designed to predict the 1-year risk of diabetes progression, to the BPRP cohort. These individuals were subsequently stratified into tertiles based on their PR scores as computed

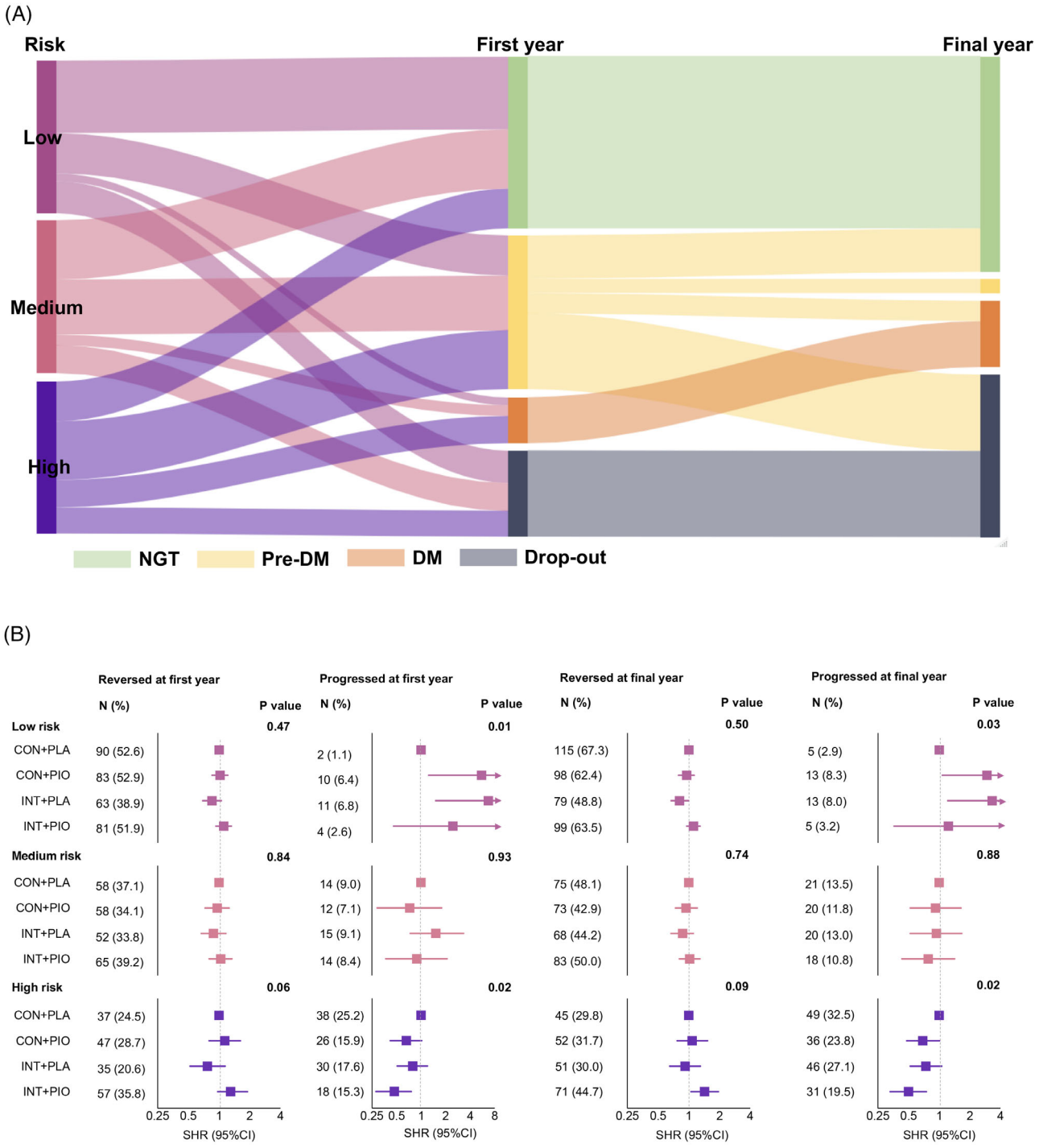

FIGURE 3 Progression and reversion of pre-DM under intervention in the Beijing Prediabetes Reversion Program study. (A) Sankey diagram visualizing the prognosis of patients (dropped out, progressed to DM, reverted to normoglycaemia and remained pre-DM). (B) SHR and 95% confidence interval in reference to CON-PLA for reversion to normoglycaemia and progression to diabetes in three risk groups at year 1 and at the end of the trial. *p* value was compared among four treatment groups. CON, conventional lifestyle therapy; DM, diabetes mellitus; PLA, placebo; INT, intensified lifestyle therapy; NGT, normal glucose tolerance; PIO, pioglitazone; SHR, sub-distribution hazard ratio.

by the ML-PR model. The baseline attributes of each risk category are outlined in Table 1. As PRs escalated, we observed increases in age, HbA1c, FPG, PG2h, homeostatic model assessment of insulin resistance and LDL-C, and a decline in the homeostatic model assessment of β-cell function. WC, SBP, DBP and caloric intake remained consistent across the different subgroups. Both pharmacological treatments and lifestyle interventions were equally represented across all three risk categories. Notably, the high-risk group consistently exhibited elevated FPG and PG2h throughout the follow-up duration (as depicted in Figure S3).

3.3 | Reversion and progression of patients in three subgroups

In total, 48%, 37.7% and 26.0% of participants reverted to normoglycaemia, and 4.6%, 7.6% and 17.8% of participants progressed to diabetes at the end of year 1 in the low-risk, medium-risk and high-risk groups, respectively. Approximately 20% of participants dropped out at the 1-year follow-up, and the number increased to 35.4% at the end of the trial. Finally, 60.5%, 48.1% and 32.2% of participants reverted to NGR, and 6.3%, 10.7% and 26.4% of participants progressed to diabetes in the low-, medium- and high-risk subgroups at the end of the third year (Figure 3A and Table S5).

Within the low-risk group, no statistically significant differences were observed among the four distinct treatments regarding the risk of regression from prediabetes. However, higher progression rates were evident in those administered INT + PLA and CON + PIO ($p = .01$), as shown in Figure 3B. Among the medium-risk group, therapeutic interventions did not elicit any discernible differences in diabetes progression rates. In contrast, in the high-risk category, the four treatments exhibited significant variances in their efficacy to deter diabetes onset ($p = .019$) and suggested a potential difference in facilitating NGR restoration ($p = .06$). In comparison with those treated with CON + PLA, the INT + PIO regimen lessened the diabetes PR by 50% [SHR (95% CI) of 0.47 (0.28–0.76)] and augmented the probability of reverting from prediabetes by 30% [SHR 1.30 (0.97–1.80)]. Analogous results persisted until the end of the trial, with INT + PIO manifesting a pronounced influence in stymieing diabetes progression [SHR 0.49 (0.32, 0.74)] and a nearly significant impact on bolstering the regression [SHR 1.45 (1.05, 1.95)] when compared against CON + PLA. Upon accounting for body weight reductions during the initial year, the risk associated with progression diminished by 50% under INT + PIO, although data for 20% of the cases were inaccessible because of missing weight loss data (as detailed in Table S8).

4 | DISCUSSION

In this study, we stratified patients with prediabetes in a clinical trial using an ML-based model to predict the 1-year diabetes PR. A differential effect of interventions was observed in prediabetes with high PR in preventing the progression towards diabetes.

In this study, we constructed an ML model to predict diabetes progression in patients with prediabetes. Historically, the Cambridge and Framingham scores were pivotal benchmarks for predicting the likelihood of diabetes progression over a span of 5–10 years in European and American populations, respectively.[29,30] However, when these traditional scoring systems were applied to the Pinggu studies, their predictive value was diminished, possibly because of the distinct lifestyles, dietary habits and societal norms that differentiate the Chinese from Western populations. As discussed in the supplementary review, ML has gained traction in the realm of diabetes prediction. Notably, the variables prioritized by our model were intricately linked to blood glucose levels: the top three were FPG, PG2h and HbA1c. We are not surprised that FPG and PG2h ranked higher than HbA1c because, in our study, prediabetes and diabetes were defined by OGTT rather than HbA1c. In addition, lipid metrics, namely, HDL-C and TG, were identified as the fourth and fifth most crucial variables predicting diabetes progression. This aligns with recent findings positing a strong correlation between elevated TG/HDL-C ratios and diabetes.[31] While other determinants, such as creatinine, age, BMI and familial medical history, also enriched the model's robustness, gender played a negligible role. This is intriguing considering a marginally higher diabetes prevalence in men.[23] In fact, the prevalence of diabetes in men and women was largely affected by age. The age bracket predominantly represented in our study (50–59 years) evinces comparable diabetes rates among genders.[32] Furthermore, in a multivariate analysis, the impact of sex on diabetes progression may be overshadowed by more dominant predictors, such as blood glucose levels and lipid profiles.

We selected XGBoost with five variables as our final model, ML-PR. Both XGBoost and RF showed excellent internal ROC AUCs, but XGBoost outshone in accuracy and F1 score. The average ROC AUC for diabetes prediction using ML is 0.86, showing our model's strength yet room for perfection.[33] By adding more diabetes-associated variables, possibly incorporating ECG or chest X-ray data and optimizing with deep learning, performance might be enhanced. Aiming for broader clinical utility, we streamlined our predictors to the five-variable ML-PR model. We considered omitting PG2h to circumvent OGTT, improving clinical convenience. However, doing so risks prediction accuracy, given the critical correlation of PG2h with diabetes progression. As such, our final ML-PR model retained five variables, including PG2h. This model predicts 1-year diabetes progression at baseline, the end of year 1 and year 2 with an ROC AUC of 0.80–0.93. It also predicts 2- and 3-year progression with an ROC AUC of 0.77–0.79 (see Tables S6 and S7). In clinics with OGTT facilities, the ML-PR model can provide both baseline and dynamic prediabetes risks. The model is shared on the website (http://diabetesmodels.com:9001/phone?type=1).

Our study emphasized the importance of personalized treatments for prediabetes based on individual PRs in clinical contexts. In the initial analysis of BPRP,[26] all interventions showed a negative effect on the reversal to normoglycaemia, although INT + PIO showed a mild effect in reducing the risk of diabetes progression. Our post hoc analysis suggested that for those with low or medium PRs, there may not be a need for rigorous lifestyle interventions or pioglitazone treatments. For the low-risk group, using PIO or INT alone might even raise diabetes risks, although these outcomes might be influenced by the group's inherently low progression rates. When accounting for weight reduction in the first year, the impact of INT + PIO remained consistent in the high-risk category (refer to Table S8). Similar to our study, the STOP DIABETES trial suggested that intense medication can successfully improve the restoration rate in participants with higher progressive risk.[12] Nevertheless, the question remains: is it worthwhile to treat those with mild or moderate diabetes risk, particularly with medications? Critics argue against broadening the diagnostic criteria for prediabetes because of potential excessive

commitments to the pharmaceutical sector.[34] Supporting a more targeted approach, our study indicates that by categorizing patients based on their risk levels and focusing treatments on high-risk individuals, we could potentially optimize health care investments and spare approximately two-thirds of patients from unwarranted treatments.

Our study showed that intensive lifestyle therapy combined with pioglitazone had a significant effect in preventing diabetes progression in the high-risk group, yet neither intensive lifestyle intervention nor pioglitazone treatment alone displayed efficacy in the high-risk group. In fact, intensive lifestyle intervention and drug treatment are both effective ways to prevent diabetes progression. The Daqing study[35] and DPP study[36] showcased lifestyle interventions as powerful preventive treatments. The limited efficacy of INT, as suggested by our research, might be rooted in the higher base conversion rates witnessed in the control group. In the preliminary BPRP analysis, the INT + PLA segment displayed a 1-year regression rate of 30.8%, notably lower than the rates seen in the other three groups (ranging from 38% to 42%).[26] It is worth noting that in the modern age, physicians are not the sole source of healthy lifestyle guidance for the Chinese populace. Such advice is readily accessible via various media outlets. The ACT NOW study showed that pioglitazone reduced the risk of conversion of prediabetes to type 2 diabetes mellitus and bolstered the chances of reverting to NGR.[6] Beyond ethnic distinctions, it is important to recognize that while the ACT NOW study focused on participants with impaired glucose tolerance, the BPRP targeted those with impaired fasting glucose as well. To derive definitive conclusions, more comprehensive studies are needed to ascertain the individual and combined efficacies of INT and pioglitazone across the varying risk strata within the prediabetic population.

In current practice, the endorsement of intensive lifestyle therapy remains consistent across numerous guidelines. The American Diabetes Association recommends considering metformin to prevent diabetes in patients with prediabetes and with BMI ≥35 kg/m^2, aged <60 years, and women with previous gestational diabetes mellitus.[11] The Chinese Association of Diabetes recommends considering metformin or acarbose when lifestyle therapy is ineffective.[37] For the integration of ML-PR guided therapies in clinical settings and the revamp of current guidelines, it is essential to conduct a randomized clinical trial that contrasts ML-PR guided strategies with standard treatments as per existing guidelines. It is crucial to highlight that while implementing drug-based interventions, potential side effects cannot be ignored. For instance, pioglitazone might cause oedema or weight gain, and metformin has been linked with gastrointestinal issues and biochemical vitamin B12 deficiency.[11] Therefore, it is imperative to balance potential benefits against risks in clinical practice.

Our study leverages ML models for predicting diabetes, offering insight into clinical decision-making. However, some limitations persist. First, the ML-PR model was generated and tested in a singular, moderately scaled study. For enhanced validation, a larger, multicentre study would be beneficial. Secondly, the mechanics of ML-PR are difficult to interpret. While the results are clear, the processes leading to these conclusions remain veiled. In addition, the model lacks the finesse to track the dynamic risks in patients with prediabetes in different follow-up periods. Thirdly, in the high-risk group, the enduring impacts of pioglitazone and intensive lifestyle therapy postcessation are unclear. Extended follow-ups could provide clarity on post-treatment implications. Fourthly, our model, tailored for Asian participants, requires further validation for its applicability across other ethnicities, such as Caucasians and Africans. Finally, the elevated dropout rate of the BPRP could introduce biases, potentially affecting the study's comprehensive representation. To capitalize fully on our model's potential, these constraints warrant further exploration in subsequent research.

5 | CONCLUSION

An ML-based model using simple clinical parameters associated with diabetes progression can successfully predict the PR of diabetes. Lifestyle and drug interventions were only effective in participants with high PR to delay diabetes progression in participants with prediabetes. Highlighting the importance of patient stratification before treatment, our findings pave the way for a more targeted, precise therapeutic approach in individuals with prediabetes.

AUTHOR CONTRIBUTIONS

XZou designed the study and wrote the manuscript. YLuo and QH conducted the study and analysis; ZZ, YLi, XZhang and XZhou were involved in the data collection, conception, and the interpretation of the results. LJ designed the study and revised the manuscript. All authors reviewed and approved the final version of the manuscript.

ACKNOWLEDGMENTS

Beijing Nova Cross program (Z211100002121169) (to Xiantong Zou) and Beijing Nova Program of Science and Technology (Z191100001119026), Peking University People's Hospital Research And Development Funds (RDH2021-10 to Xiantong Zou); the National Natural Science Foundation of China (81 970 708 to Linong Ji); Clinical Medicine Plus X—Young Scholars Project, Peking University, the Fundamental Research Funds for the Central Universities (PKU2022LCXQ004 to Xiantong Zou) and Beijing Municipal Science and Technology Commission (Z201100005520013 to Linong Ji).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/dom.15291.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are not publicly available because of human resource regulations of China, but are available from the corresponding author

upon reasonable request. Data Access and Responsibility: Linong Ji is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

ORCID

Linong Ji https://orcid.org/0000-0002-3262-2168

REFERENCES

1. Wang L, Gao P, Zhang M, et al. Prevalence and Ethnic Pattern of Diabetes and Prediabetes in China in 2013. *JAMA*. 2017;317:2515-2523.
2. Cai X, Zhang Y, Li M, et al. Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis. *BMJ*. 2020;370:m2297. doi:10.1136/bmj.m2297
3. Schlesinger S, Neuenschwander M, Barbaresko J, et al. Prediabetes and risk of mortality, diabetes-related complications and comorbidities: umbrella review of meta-analyses of prospective studies. *Diabetologia*. 2022;65(2):275-285. doi:10.1007/s00125-021-05592-3
4. Gong Q, Zhang P, Wang J, et al. Changes in mortality in people with IGT before and after the onset of diabetes during the 23-year follow-up of the Da Qing diabetes prevention study. *Diabetes Care*. 2016;39(9):1550-1555. doi:10.2337/dc16-0429
5. Barry E, Roberts S, Oke J, Vijayaraghavan S, Normansell R, Greenhalgh T. Efficacy and effectiveness of screen and treat policies in prevention of type 2 diabetes: systematic review and meta-analysis of screening tests and interventions. *BMJ*. 2017;356:i6538. doi:10.1136/bmj.i6538
6. DeFronzo RA, Tripathy D, Schwenke DC, et al. Pioglitazone for diabetes prevention in impaired glucose tolerance. *N Engl J Med*. 2011;364(12):1104-1115. doi:10.1056/NEJMoa1010949
7. Chiasson JL, Josse RG, Gomis R, Hanefeld M, Karasik A, Laakso M. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. *Lancet*. 2002;359(9323):2072-2077. doi:10.1016/s0140-6736(02)08905-5
8. Haw JS, Galaviz KI, Straus AN, et al. Long-term sustainability of diabetes prevention approaches: a systematic review and meta-analysis of randomized clinical trials. *JAMA Intern Med*. 2017;177(12):1808-1817. doi:10.1001/jamainternmed.2017.6040
9. Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: a high-risk state for diabetes development. *Lancet*. 2012;379(9833):2279-2290. doi:10.1016/s0140-6736(12)60283-9
10. Alizadeh Z, Baradaran HR, Kohansal K, Hadaegh F, Azizi F, Khalili D. Are the determinants of the progression to type 2 diabetes and regression to normoglycemia in the populations with pre-diabetes the same? *Front Endocrinol*. 2022;13:1041808. doi:10.3389/fendo.2022.1041808
11. ElSayed NA, Aleppo G, Aroda VR, et al. 3. Prevention or delay of type 2 diabetes and associated comorbidities: standards of Care in Diabetes-2023. *Diabetes Care*. 2023;46(Suppl 1):S41-s48. doi:10.2337/dc23-S003
12. Armato JP, DeFronzo RA, Abdul-Ghani M, Ruby RJ. Successful treatment of prediabetes in clinical practice using physiological assessment (STOP DIABETES). *Lancet Diabetes Endocrinol*. 2018;6(10):781-789. doi:10.1016/S2213-8587(18)30234-1
13. Anderson JP, Parikh JR, Shenfeld DK, et al. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: An application of machine learning using electronic health records. *J Diabetes Sci Technol*. 2015;10(1):6-18. doi:10.1177/1932296815620200
14. Cahn A, Shoshan A, Sagiv T, et al. Prediction of progression from pre-diabetes to diabetes: development and validation of a machine learning model. *Diabetes Metab Res Rev*. 2020;36(2):e3252. doi:10.1002/dmrr.3252
15. Han Y, Hu H, Liu Y, Wang Z, Liu D. Nomogram model and risk score to predict 5-year risk of progression from pre-diabetes to diabetes in Chinese adults: development and validation of a novel model. *Diabetes Obes Metab*. 2022;25:675-687. doi:10.1111/dom.14910
16. Singla R, Singla A, Gupta Y, Kalra S. Artificial intelligence/machine learning in diabetes care. *Indian J Endocrinol Metab*. 2019;23(4):495-497. doi:10.4103/ijem.IJEM_228_19
17. Fregoso-Aparicio L, Noguez J, Montesinos L, García-García JA. Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetol Metab Syndr*. 2021;13(1):148. doi:10.1186/s13098-021-00767-9
18. Saberi-Karimian M, Mansoori A, Bajgiran MM, et al. Data mining approaches for type 2 diabetes mellitus prediction using anthropometric measurements. *J Clin Lab Anal*. 2023;37(1):e24798. doi:10.1002/jcla.24798
19. Mansoori A, Sahranavard T, Hosseini ZS, et al. Prediction of type 2 diabetes mellitus using hematological factors based on machine learning approaches: a cohort study analysis. *Sci Rep*. 2023;13(1):663. doi:10.1038/s41598-022-27340-2
20. Pyrros A, Borstelmann SM, Mantravadi R, et al. Opportunistic detection of type 2 diabetes using deep learning from frontal chest radiographs. *Nat Commun*. 2023;14(1):4039. doi:10.1038/s41467-023-39631-x
21. Zou Q, Qu K, Luo Y, Yin D, Ju Y, Tang H. Predicting diabetes mellitus with machine learning techniques. *Front Genet*. 2018;9:515. doi:10.3389/fgene.2018.00515
22. Chen L, Li Y, Zhang F, Zhang S, Zhou X, Ji L. Elevated serum ferritin concentration is associated with incident type 2 diabetes mellitus in a Chinese population: a prospective cohort study. *Diabetes Res Clin Pract*. 2018;139:155-162. doi:10.1016/j.diabres.2018.03.001
23. Wang L, Peng W, Zhao Z, et al. Prevalence and treatment of diabetes in China, 2013–2018. *Jama*. 2021;326(24):2498-2506. doi:10.1001/jama.2021.22208
24. Zhou X, Li Y, Zhang X, et al. Independent markers of nonalcoholic fatty liver disease in a gentrifying population-based Chinese cohort. *Diabetes Metab Res Rev*. 2019;35(5):e3156. doi:10.1002/dmrr.3156
25. Luo Y, Paul SK, Zhou X, et al. Rationale, design, and baseline characteristics of Beijing prediabetes reversion program: a randomized controlled clinical trial to evaluate the efficacy of lifestyle intervention and/or pioglitazone in reversion to Normal glucose tolerance in prediabetes. *J Diabetes Res*. 2017;2017:7602408. doi:10.1155/2017/7602408
26. Luo Y, Wang H, Zhou X, et al. A randomized controlled clinical trial of lifestyle intervention and pioglitazone for normalization of glucose status in Chinese with prediabetes. *J Diabetes Res*. 2022;2022:2971382. doi:10.1155/2022/2971382
27. WHO. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications Report of a WHO Consultation Part 1: Diagnosis and Classification of Diabetes Mellitus, Geneva, World Health Organization Department of Noncommunicable Disease Surveillance. 1999 WHO/NCD/NCS/99.2.
28. Zou X, Li Y, Cai X, et al. Decreased glycemic difference between diabetes and nondiabetes in the elderly leads to the reduced diagnostic accuracy of hemoglobin A1c for diabetes screening in an aged Chinese population. *Diabetes Technol Ther*. 2016;18(4):226-232. doi:10.1089/dia.2015.0353
29. Rahman M, Simmons RK, Harding AH, Wareham NJ, Griffin SJ. A simple risk score identifies individuals at high risk of developing type 2 diabetes: a prospective cohort study. *Fam Pract*. 2008;25(3):191-196. doi:10.1093/fampra/cmn024
30. Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr. Prediction of incident diabetes mellitus in middle-aged adults: the

Framingham offspring study. *Arch Intern Med*. 2007;167(10):1068-1074. doi:10.1001/archinte.167.10.1068
31. Yang T, Liu Y, Li L, et al. Correlation between the triglyceride-to-high-density lipoprotein cholesterol ratio and other unconventional lipid parameters with the risk of prediabetes and type 2 diabetes in patients with coronary heart disease: a RCSCD-TCM study in China. *Cardiovasc Diabetol*. 2022;21(1):93. doi:10.1186/s12933-022-01531-7
32. Group tDS. Age- and sex-specific prevalence of diabetes and impaired glucose regulation in 11 Asian cohorts. *Diabetes Care*. 2003;26(6):1770-1780. doi:10.2337/diacare.26.6.1770
33. Olusanya MO, Ogunsakin RE, Ghai M, Adeleke MA. Accuracy of machine learning classification models for the prediction of type 2 diabetes mellitus: a systematic survey and meta-analysis approach. *Int J Environ Res Public Health*. 2022;19(21):14280. doi:10.3390/ijerph192114280
34. Piller C. Dubious diagnosis. *Dubious diagnosis*. American Association for the Advancement of Science; 2019.
35. Pan XR, Li GW, Hu YH, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and diabetes study. *Diabetes Care*. 1997;20(4):537-544. doi:10.2337/diacare.20.4.537
36. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. 2002;346(6):393-403. doi:10.1056/NEJMoa012512
37. Zhu D. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition). *Chin J Pract Intern Med*. 2021;41:668-695.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zou X, Luo Y, Huang Q, et al. Differential effect of interventions in patients with prediabetes stratified by a machine learning-based diabetes progression prediction model. *Diabetes Obes Metab*. 2024;26(1):97-107. doi:10.1111/dom.15291