Comparative study on risk prediction model of type 2 diabetes based on machine learning theory: a cross-sectional study

Shu Wang, Rong Chen, Shuang Wang, Danli Kong, Rudai Cao, Chunwen Lin, Ling Luo, Jialu Huang, Qiaoli Zhang, Haibing Yu, Yuan Lin Ding

ABSTRACT

Objectives To compare the prediction effects of six models based on machine learning theories, which can provide a methodological reference for predicting the risk of type 2 diabetes mellitus (T2DM).

Setting and participants This study was based on the monitoring data of chronic disease risk factors in Dongguan residents from 2016 to 2018. The multistage cluster random sampling method was adopted at each monitoring site, and 4157 people were finally selected. In the initial population, we excluded individuals with more than 20% missing data and eventually included 4106 subjects.

Design K nearest neighbour algorithm and synthetic minority oversampling technique were used to process the data. Single factor analysis was used for preliminary selection of variables. The 10-fold cross-validation was used to optimise the parameters of some models. The accuracy, precision, recall and area under receiver operating characteristic curve (AUC) were used to evaluate the prediction effect of models, and Delong test was used to analyse the differences of AUC values of each model.

Results After balancing data, the sample size increased to 8013, of which 4023 are patients with T2DM and 3990 in control group. The comparison results of the six models showed that back propagation neural network model has the best prediction effect with 93.7% accuracy, 94.6% accuracy, 92.8% recall and the AUC value of 0.977, followed by logistic model, support vector machine model, CART decision tree model and C4.5 decision tree model. Deep neural network has the worst prediction performance, with 84.5% accuracy, 86.1% precision, 82.9% recall and the AUC value of 0.845.

Conclusions In this study, six types of risk prediction models for T2DM were constructed, and the predictive effects of these models were compared based on various indicators. The results showed that back propagation neural network based on the selected data set had the best prediction effect.

INTRODUCTION

Diabetes mellitus (DM) is a metabolic disease characterised by a disorder of blood glucose metabolism, which is one of the major public health problems in the 21st century. In 2021, the number of patients with diabetes in the world has been 536.6 million, and this number is expected to increase to 783.2 million by 2045, which will bring a great burden to the healthcare system. Type 2 DM (T2DM) is the most common form of diabetes. Early lifestyle changes or pharmacological interventions have been shown to be effective in delaying or preventing T2DM and its complications. Therefore, it is very important to accurately predict the risk of T2DM in advance. However, T2DM has a slow onset and a long clinical incubation period. Resulting in a possible delay of more than 10 years from the onset to the diagnosis of T2DM. Therefore, timely screening and management of diabetes high-risk groups is of great significance to reduce the incidence of diabetes.

In recent years, the application of artificial intelligence (AI) in the medical field has developed rapidly. The use of machine learning methods and mathematical models to predict the risk of disease has been widely used in clinical medicine. Machine learning...
models have been successfully applied to the study of diabetes, chronic obstructive pulmonary disease, hepatitis B and other diseases. Machine learning represents a powerful set of methods for characterising, adjusting, learning, predicting and analysing data. These methods use large amounts of data input and output to recognise patterns and learn efficiently to train machines to make autonomous recommendations or decisions. After sufficient repetition and modification, the machine can receive the input and predict the output. The output results were compared with known results to judge their accuracy, and then iteratively adjusted to improve their ability to predict disease. Machine learning can be divided into three types: supervised learning, unsupervised learning and reinforcement learning. Some of the most common supervised learning methods include logistic regression, K-nearest neighbour (KNN), naive Bayes (NB), decision tree, artificial neural network (ANN) and support vector machine (SVM). At present, the supervised learning method is widely used in T2DM-related research.

Currently, in the research on the risk prediction model of T2DM, the commonly used machine learning methods include logistic regression, CART decision tree, C4.5 decision tree, SVM, back propagation (BP) neural network and deep neural network (DNN). No systematic comparisons of the predictive effectiveness at present have been reported for the six models mentioned above. Based on the above background, this study used the monitoring data of chronic disease risk factors of Dongguan residents from 2016 to 2018 to construct six type 2 diabetes risk prediction models, including logistic regression, CART, C4.5, BP neural network, SVM and DNN, and compared the prediction results of each model, so as to provide methodological reference for the risk prediction of T2DM.

**RESEARCH DESIGN AND METHODS**

**Research object**

This study was based on the monitoring data of chronic disease risk factors in Dongguan residents from 2016 to 2018. The survey subjects were residents aged 18 and above who had lived at the monitoring site for more than 6 months in the 12 months before the survey. The multistage cluster random sampling method was adopted at each monitoring site, and 4157 people were finally selected. In the initial population, we excluded individuals with more than 20% missing data and eventually included 4106 subjects, including 149 patients with T2DM as the case group and the remaining 3957 as the control group. The case group included patients previously diagnosed with T2DM and patients newly diagnosed with T2DM during the investigation period. The diagnostic criteria for T2DM used the 1999 WHO diagnostic criteria: fasting plasma glucose level (FPG) ≥ 7.0 mmol/L or 2-hour post-prandial blood glucose (2hPG) ≥ 11.1 mmol/L.

**Data collection**

The monitoring data included three parts. The first part is the questionnaire surveys, which includes gender, age, education level, smoking, drinking, weekly frequency of consumption of various foods. The second part is body measurements, which includes body mass index, waist-line, systolic blood pressure and diastolic blood pressure (DBP). The last part is laboratory testing, which includes FPG, total cholesterol (TC), triglyceride (TG), high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C).

**Data cleaning**

The data distribution was checked using R Studio and the remaining missing values were filled in using the KNN method.

**Process unbalanced data**

In the original data, the proportion of patients with type 2 diabetes and normal population samples reached 1:27, and the data were in an unbalanced state. The synthetic minority oversampling technology (SMOTE) was used to process the unbalanced data by R studio software, and 70% of the processed data were randomly selected as the training set and 30% as the test set.

**Variable selection**

Normality test was conducted on the balanced data, and single factor analysis was conducted by Mann-Whitney U rank sum test and χ² test to screen out the target variables.

**Ten-fold cross-validation**

The 10-fold cross-validation is when a dataset is randomly divided into 10 equally sized subsets. One subset is used as the test set, and the other nine are used as the training set. In this study, the 10-fold cross-validation method was applied to 70% of the training sets, and the included variables were statistically significant variables in univariate analysis. By continuously adjusting the model parameters and comparing the corresponding 10-fold cross-validation results, the optimal parameters were selected to construct the risk prediction model of type 2 diabetes.

**Model construction**

Seventy per cent of the processed data were randomly selected as the training set and the remaining 30% as the test set. The logistic regression, CART, C4.5, BP neural network, SVM and DNN were used in the training set to construct the risk prediction models of type 2 diabetes, respectively, and the prediction effects of each model were evaluated and compared by accuracy, precision, recall and the area under receiver operating characteristic (ROC) curve (AUC) value. These evaluation indicators are mainly based on a confusion matrix, as shown in online supplemental table 1. The higher the accuracy, the higher the correctness of the model for true positive and true negative samples. The higher the recall is, the better the model can identify positive samples. In the ROC curve, the horizontal axis...
is the false positive rate, the vertical axis is the true positive rate, and the AUC value of the AUC the ROC curve refers to the random selection of a positive sample and a negative sample. If the AUC value is large, more positive samples will be accurately predicted with a greater probability, and the probability of negative samples being predicted as positive samples will be smaller. The accuracy of AUC between 0.5 and 0.7 is low, the accuracy above 0.9 is high, and the accuracy below 0.5 is not consistent with the real situation.\textsuperscript{11}

Patient and public involvement
No patient involved.

RESULTS
Statistical description of baseline data
Subjects with many missing values and obvious data errors were deleted, and the final sample size was 4106, including 149 patients with type 2 diabetes and 3957 people in the normal population. The difference in sample size between the two groups was huge, and the SMOTE method was adopted to process the unbalanced data. The parameters between 0.7 and 0.9 is general, the accuracy above 0.9 is the best when the number of hidden layers is 2, the prediction effect is the best. For DNN, the number of neurons in hidden layers was 8–12, and the number of neurons in hidden layers was 25–35. The number of neurons in each hidden layer was set to be equal in this study. The results showed that the prediction effect was best when the number of hidden layers was 9 and the number of neurons in each hidden layer was 33.

Screening results of variables in univariate analysis
The normal test found that the distribution of each characteristic attribute in the two groups of samples was mostly skewed distribution, so the Mann-Whitney U rank sum test and $\chi^2$ test in SPSS V.25.0 were used to analyse the quantitative and qualitative data, respectively. The results of univariate analysis were shown in online supplemental table 3. The results showed that except the educational level, the distribution of the other 23 characteristic variables between the case group and the control group was statistically different.

Parameter tuning results
The prediction effect of SVM, BP neural network and DNN is closely related to the setting of relevant model parameters. The corresponding 10-fold cross-validation results were compared by continuously adjusting the model parameters. For SVM, linear kernel function, radial basis function, polynomial kernel function and Sigmoid kernel function are used to perform 10-fold cross-validation. The results showed that the linear kernel function is the best predictor. For BP neural network, the maximum number of iterations is set to 3000, and 10-fold cross-validation was performed on the number of hidden layer neurons within the range of 5–20, respectively. The results showed that when the number of hidden layer neurons is 2, the prediction effect is the best. For DNN, the range of hidden layers was 8–12, and the number of neurons in hidden layers was 25–35. The number of neurons in each hidden layer was set to be equal in this study. The results showed that the prediction effect was best when the number of hidden layers was 9 and the number of neurons in each hidden layer was 33.

Model construction results
Logistic regression model
Fit all data except education level, build logistic regression model, and then used stepwise regression method to screen variables based on Akaiake information criterion. A total of 16 variables were finally screened, which were age, drinking, consumption frequency of cereals, potatoes, beans, fruits, eggs, dairy, poultry and fish, DBP, FPG, TC, TG, HDL-C and LDL-C (online supplemental table 4). Variables screened by stepwise regression were applied to the training set to build a logistic model, as shown in table 1. In this model, the factors had a greater influence on T2DM included potato consumption frequency, fish consumption frequency, TC, FPG, HDL-C. In addition, the frequency of cereal consumption and TC were negatively correlated with the incidence of T2DM, while the other variables were positively correlated with the incidence of T2DM.

The logistic model equation is:

$$\logit(P) = -17.486 + 0.027 \text{age} + 0.173 \text{drinking} - 0.236 \text{cereals} + 0.442 \text{potatoes} + 0.176 \text{Beans} + 0.199 \text{fruits} + 0.294 \text{eggs} + \ldots$$

Table 1 Logistic regression analysis results

| Variable       | Estimate | SE    | Z-value | Pr(>|z|) |
|----------------|----------|-------|---------|----------|
| (Intercept)    | -17.468  | 0.768 | -22.57  | <0.001   |
| Age            | 0.027    | 0.004 | 6.880   | <0.001   |
| DBP            | 0.026    | 0.005 | 5.041   | <0.001   |
| FPG            | 2.112    | 0.075 | 28.271  | <0.001   |
| TC             | -0.724   | 0.111 | -6.525  | <0.001   |
| TG             | 0.249    | 0.035 | 7.143   | <0.001   |
| HDL-C          | 0.573    | 0.159 | 3.610   | <0.001   |
| LDL-C          | 0.303    | 0.129 | 2.355   | <0.05    |
| Drinking       | 0.173    | 0.097 | 1.776   | 0.076    |
| Cereals        | -0.236   | 0.078 | -3.033  | <0.01    |
| Potatoes       | 0.442    | 0.068 | 6.481   | <0.001   |
| Beans          | 0.176    | 0.068 | 2.579   | <0.01    |
| Fruits         | 0.199    | 0.059 | 3.384   | <0.001   |
| Eggs           | 0.294    | 0.063 | 4.656   | <0.001   |
| Milk           | 0.154    | 0.054 | 2.832   | <0.01    |
| Poultry        | 0.373    | 0.064 | 5.814   | <0.001   |
| Fish           | 0.491    | 0.067 | 7.416   | <0.001   |

DBP, diastolic blood pressure; FPG, fasting plasma glucose; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride.
0.154milk+0.373poultry+0.491fish+0.026DBP+2.112FP-

C4.5 decision tree
The 23 significant characteristic variables from univariate analysis in the training set were put into the C4.5 decision tree model. As shown in online supplemental figure 2, the decision tree model output by C4.5 algorithm includes six root nodes and nine leaf nodes. According to the model, T2DM was diagnosed when FPG was 5.61 mmol/L; when FPG ≤5.61 mmol/L; (1) potatoes=0 was diagnosed as non-T2DM; (2) potatoes=1, Age ≤54 was diagnosed as non-T2DM; (3) potatoes=3, Age >54, TC ≤5.11 mmol/L was diagnosed as T2DM; (4) potatoes=1, Age >54, TC >5.11 mmol/L was diagnosed as non-T2DM; (5) potatoes=2, DBP ≤81 mm Hg was diagnosed as non-T2DM; (6) potatoes=2, DBP >81 mm Hg was diagnosed as T2DM; (7) potatoes=3, Age ≤34 was diagnosed as non-T2DM; (8) potatoes=3, Age >34 was diagnosed as T2DM. The C4.5 model was applied to the test set for verification, and the correlation confusion matrix table and ROC curve were obtained. As shown in online supplemental table 5, the accuracy of this model was 88.6%, the precision was 92.7%, the recall was 84.9% and the AUC value was 0.888.

DNN model
The 23 characteristic variables that were significant for univariate analysis in the training set were substituted into the DNN model. The number of hidden layers was 9, with 33 neurons in each layer and the correlation confusion matrix table and ROC curve were obtained. As shown in online supplemental table 5, the accuracy of this model was 84.5%, the precision was 86.1%, the recall was 82.9%, and the AUC value was 0.845.

Comparison of model performance
DeLong test in R Studio was used to compare the AUC values of each model, as shown in figure 2 and online supplemental table 6. Based on the dataset and incorporating the robustness of the model and the prediction effect of type 2 diabetes, BP neural network model is the best, the accuracy was as high as 93.7%, the precision was 94.6%, the recall was 92.8%, and the AUC value was 0.977, followed by logistic regression model, the SVM model, CART decision tree model, C4.5 decision tree model, depth of neural network model. The prediction effect of SVM and CART was similar, and the difference was not statistically significant.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Performance indicators of each model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.894</td>
</tr>
<tr>
<td>SVM</td>
<td>0.912</td>
</tr>
<tr>
<td>BP</td>
<td>0.937</td>
</tr>
<tr>
<td>CART</td>
<td>0.887</td>
</tr>
<tr>
<td>C4.5</td>
<td>0.886</td>
</tr>
<tr>
<td>DNN</td>
<td>0.845</td>
</tr>
</tbody>
</table>

AUC, area under the curve; BP, back propagation; DNN, deep neural network.
DISCUSSION

According to the 2021 survey results of the International Diabetes Association, the number of patients with diabetes in China has reached 140.9 million, making China the country with the largest number of patients with diabetes in the world. Global diabetes-related health expenditures were estimated at US$966 billion in 2021, and are projected to reach US$1054 billion by 2045. More and more researchers have realised that the early diagnosis or prediction of DM through model is of great significance for the prevention of T2DM, the improvement of life quality of patients with T2DM and the prevention of related complications.

In recent years, machine learning techniques have been widely used to predict the risk of T2DM. Logistic regression, as a classical classification model, has the advantages of low calculation cost, easy implementation and understanding, and is suitable for solving the
binary classification problem of relatively simple datasets. For researchers in the field of clinical medicine, most of them are still inclined to use regression-based methods for disease correlation analysis in order to obtain an intuitive and clear result under limited conditions. Logistic regression is essentially a linear classifier. If the relationship between data is linearly separable, this traditional method will be more suitable. However, it has poor effect on data processing of multiple categories of related features, which may oversimplify the complex relationship between factors with non-linear interaction, leading to potential loss of important relevant information. However, diagnosis models based on machine learning, such as BP neural network, SVM, decision tree and DNN, are often more suitable for classification tasks with multidimensions and large sample sizes. They have better predictive performance in solving non-linear problems and better fitting ability for complex datasets, which shows the importance of selecting appropriate models according to the characteristics of datasets.

Hong\(^1\) built a diabetes prediction model based on BP neural network, SVM and integrated learning, and the results showed that the prediction effect of BP neural network was better than that of SVM. Gao\(^13\) used BP neural network and logistic regression to construct the prediction model of T2DM complications, and the study found that the prediction effect of BP neural network was higher than that of logistic regression. Liu et al\(^24\) constructed logistic model, BP neural network model and decision tree model to analyse the risk factors of T2DM, and the results showed that the prediction effects of the three models from high to low were BP neural network, logistic regression model and decision tree model, respectively. Dwivedi\(^15\) used six algorithms of classification trees, SVM, ANN, NB, logistic and KNN to predict diabetes, and the results showed that the prediction effect of logistic regression model was better than that of SVM. The above conclusions are consistent with the results of this study. However, the BP neural network model also has certain limitations, as it cannot determine the direction of variables and cannot determine whether variables are protective or risk factors. Although the predictive performance of the logistic model is weaker than that of the BP neural network, it has strong interpretability for the results and can reflect the relationship between various factors and T2DM.

As a simple and easy to use non-parametric classifier, decision tree model has high computational speed and robustness. The SVM model overcomes the problem of dimensionality disaster and non-linear divisibility by using kernel function method, and has strong generalisation ability, which can deal with high dimensional problems. The results of this study showed that the prediction effect of SVM and CART was similar, and the difference was not statistically significant. However, Faruque et al\(^16\) and Kandhasamy and Balamurali\(^17\) found that the prediction effect of C4.5 decision tree model was significantly higher than that of SVM model. In practical application, compared with SVM model, CART decision tree model and C4.5 decision tree model can present the variables included in the model more intuitively. Cheruku et al\(^18\) built a prediction model for T2DM based on PIDD dataset by using a variety of machine learning methods. The results showed that the prediction effect of C4.5 decision tree model was better than that of CART decision tree model, and the accuracy was 74.2% and 70.7%, respectively. Althunayan et al\(^19\) compared the performance of nine algorithms such as Naive Bayes, C4.5, CART and random forest in the prediction of T2DM, and the results showed that the accuracy of random forest was the highest, and the accuracy of C4.5 algorithm was much higher than that of CART algorithm. Meng et al\(^20\) used ANN, logistic regression and C4.5 data mining technology to predict diabetes, and finally concluded that C4.5 machine learning technology is more effective and accurate than other methods. The results of this study showed that the prediction effect of CART decision tree model is slightly better than that of C4.5 decision tree model. The CART decision tree model is a binary tree, and compared with C4.5 decision tree model, its operation is faster and the model formed is more concise. Therefore, CART is more suitable for the processing of large sample data.

DNN is a kind of multilayer unsupervised neural network, which can represent complex non-linear problems more carefully and efficiently. Islam Ayon and Milon Islam\(^21\) used DNN method to build a prediction model for T2DM based on PIDD dataset, and the accuracy of the model reached 98.35%. Mohapatra et al\(^22\) applied DNN to the prediction of T2DM, and the accuracy of the algorithm was as high as 97.11%. The results of this study showed that compared with the other five T2DM prediction models, DNN had the worst prediction effect, with the accuracy of only 84.5% and the AUC value of 0.845. The results of this study are different from those of previous related studies, which may be caused by the

---

**Figure 2** Comprehensive ROC curve of six models. BP, back propagation; DNN, deep neural network; ROC, receiver operating characteristic; SVM, support vector machine.
differences in sample size, data quality, included characteristic variables, definition of related variables and construction techniques of different datasets. In a word, the quality of the model is affected by data characteristics, training time, parameter complexity and other factors, so the most appropriate model should be selected according to specific problems.

In addition, screening and early diagnosis and treatment of high-risk T2DM groups are effective measures to reduce their mortality. Based on the results of various model construction, it is recommended to implement targeted T2DM screening for people aged 34 and above, as well as comprehensive smoking and alcohol cessation. For people aged 54 and above, comprehensive screening should be carried out. Early intervention measures should be implemented for the identified high-risk T2DM population, and high attention should be paid to risk factors that can be changed by behaviour, such as controlling sugar and oil, reducing the intake of red and processed meat, reducing the frequency of potato consumption, supplementing dietary fibre, etc, and strengthening the detection of fasting and postprandial blood glucose levels, so as to reduce the incidence of T2DM.

In conclusion, the T2DM disease risk prediction model based on BP neural network has the best prediction in this study, which can efficiently identify patients with T2DM. As a simple prediction model, it can be used in clinical practice and has certain guiding significance and clinical application value. The logistic regression model is second only to BP neural network in predicting effect, but it is highly explanatory to the results and can intuitively show the relationship between variables in the model, which has better operability and feasibility in clinical application. Although the decision tree model can intuitively present the classification process, it belongs to the black box algorithm just like BP neural network and DNN, and the specific relationship between independent variables and dependent variables is unknown, so it is often necessary to adopt other means to evaluate the importance of each variable in the model. Therefore, each model has its own advantages and disadvantages, and the results obtained can complement each other, suggesting that when applying classifiers in clinical decision-making, we can select appropriate classifiers according to research purposes, so as to obtain the highest value in practice.

This study still has certain limitations: (1) This prediction model has only been verified internally and lacks external verification. It will be verified in larger population samples in different regions in the future. (2) Due to the condition limitation, there was only one fasting glucose reading per survey subject, and some subjects may have been misclassified as diabetic due to random error. (3) Due to the limitation of dataset, the failure to include common disease risk factors such as genetic inheritance and self-care status (physical activity, sleeping time, etc) into the model may affect the accuracy of prediction to some extent. (4) SMOTE was used to process the unbalanced data to reduce the risk of model overfitting. However, this algorithm cannot overcome the data distribution problem of unbalanced datasets and is prone to produce distribution marginalisation problem.

CONCLUSION

Machine learning technology has high accuracy, low error rate and low cost in the early prediction of various diseases. Early diagnosis of T2DM is of great significance for improving the quality of life and preventing complications of patients with T2DM. Based on the accuracy, precision, recall, AUC value and other indicators, this study compared the prediction effects of six T2DM risk prediction models, including logistic regression, CART, C4.5, BP neural network, SVM and DNN. The results showed that the BP neural network model based on the selected dataset had the best prediction effect.

Contributors All the authors made substantial contributions to the conception and design. ShuW and RoC designed the study. ShuangW, RuC, LL and JH collected the data. ShuW, RoC and CL analysed the data. ShuW and RoC drafted the manuscript. YLD, HY, QZ and DK contributed to the critical revision of the manuscript before publication. YLD was responsible for the overall content and acted as the guarantor for this article. All authors critically revised and approved the final version of the manuscript.

Funding We appreciate all authors for their contributions and physicians and the participants. The study was funded by the Discipline Construction Project of Guangdong Medical University (grant no. 4SG21276P grant no. 1003K20220004), the Dongguan City Science and Technology Correspondent Project (grant no. 20221800500342), the Dongguan Social Development Technology Project (grant no. 20221800905642), the Natural Science Foundation of Basic and Applied Basic Research Foundation of Guangdong Province (grant no. 2022A1515012407). The Guangdong science and technology research project of traditional Chinese Medicine (grant no. 20221209, grant no. 20211215), the Medical Scientific Research Foundation of Guangdong Province (grant no. A2021398), the Innovation and entrepreneurship training programme for college students of Guangdong Medical University (grant no. S202210571088, grant no. GDMU20221112, grant no. GDMU2021138), the Basic and Applied Basic Research Foundation of Guangdong Province Regional Joint Fund Project (grant no. 2020B151512002), the Zhanjiang City science and technology development special fund competitive allocation project (grant no. 2020A01031), the Characteristic Innovation Project of Guangdong Province General University (grant no. 2020KTSCX042, grant no. 2019KTSCX046, grant no. 2019KTSCX047).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and the study was approved by the Medical Ethics Committee of the Affiliated Hospital of Guangdong Medical University (reference number: YS2021085), and permission has been obtained for the access and analysis of the original survey data. All the surveys and samples were obtained with the consents of participants in advance, and the informed consent forms were legally consented. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The data that support the findings of this study are available from the corresponding author on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability

REFERENCES


