Development of a binomial logistic regression model for non-invasive prediction of Type 2 Diabetes.

Abstract ID:382

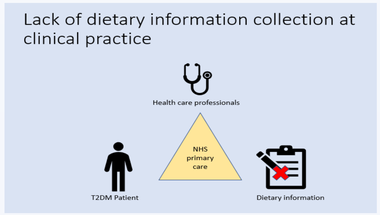
Authors: Joshi Kshama; Vartak Manjiri; Antonio Peña-Fernández; Parvez I. Haris

Institutions: School of Allied Health Sciences, De Montfort University, Leicester, UK.

Introduction

Evidence suggested that early detection of Type 2 Diabetes (T2DM) can be useful in delaying or even preventing T2DM.

Diet is one of the prime modifiable risk factors of T2DM. Yet collection of dietary information, and identification of dietary risk factors for T2DM as pre-screening tool is lacking at clinical settings.

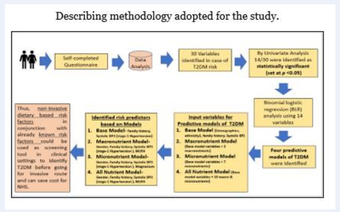


Objective

Developing predictive models to identify T2DM using binomial logistic regression.

Methods

- The study developed a culturally sensitive questionnaire based on an amended version of EPIC- Norfolk FFQ.
- Data analysis was carried out within 382 participants (184 with and 208 without T2DM) both male and female with aged 18 to 80 living within Leicester City, UK.
- Study questionnaire collected information on demographics, anthropometric measurements, hypertension, health and lifestyle information, along with dietary intake.
- The daily dietary output was obtained in form of daily consumption of 22 nutrients (specific to T2DM) and fourteen food groups at each participant level by accessing 380 foods.
- Participant's nutrient values were then compared with UK Government Dietary Nutrition Recommended values (as reference standard) set by Public Health of England to get nutrient status in terms of 'high, low, or as per recommended level'.
- The Chi-Square test was used to find statistical significance (set at p=0.05) and binary logistic regression was selected to develop predictive models at multivariate levels using IBM SPSS (version 26).
- 14 out of 30 variables were statistically significant at univariate level (set p<0.05),
- Performance matrix was used to measured model accuracy such as AUROC & AUPRC.



Results

Characteristics of study population

- The study conducted in most of ethnic minority group consisting 86.73% South Asian (Indians, Pakistani & Bangladeshi) and 4.08% African (including with & without T2DM participants) as compared to 9.18%White British.
- South Asian Indians in specific showed high prevalence of T2DM at young age (below 40 years) as compared to White British & African counterpart.
- T2DM was more pronounced among male (54.19%) than female (42.19%) participants.
- Family history of T2DM was prominent within 68.48% T2DM participants and 51.44% without T2DM.
- Hypertension (stage 1 & 2 systolic BP) was present among 47.83% of T2DM and 29.33% of without T2DM participants.

Development of four predictive models at multivariate level

Label	Base Model	Macronutrient Model	Micronutrient Model	All Nutrient Model
Variables included	Gender, Ethnicity1, Systolic BP, Family history	Gender, Family history, Ethnicity1, Systolic BP, Carbohydrate, Fat-total, MUFA (Monounsaturated fatty acids)	Family history, Ethnicity1, Systolic BP, Potassium, Magnesium, Copper, Iodine, Selenium, Zinc, Folate	Gender, Family history, Ethnicity1, Systolic BP, Carbohydrate, Fat-total, MUFA (Monounsaturated fatty acids), Potassium, Magnesium, Copper, Iodine, Selenium, Zinc, Folate
Specificity	68.3%	72.1%	70.7%	72.1%
Sensitivity	54.3%,	57.6%	58.2%	57.6%
NPV	62.83%.	65.79%	65.63%	65.79%
PPV	60.24%,	64.63%	63.69%	64.63%
AuROC	0.663 (with 95% CI 0.610 to 0.717)	0.703 (with 95% CI 0.652 to 0.755)	0.704 (with 95% CI 0.652 to 0.755)	0.703 (with 95% CI 0.652 to 0.755)
AuPRC	0.61	0.65	0.65	0.65

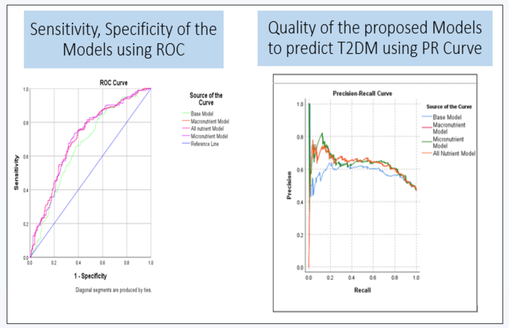
Note: Ethnicity1- White British versus South Asian, PPV- Positive predictive value, NPV- Negative predictive value, AuROC- Area under receiver operating characteristic curve, AuPRC- Area under precision recall curve.

Application of proposed model

- The current study identified four predictive models of T2DM: Base, Macro, Micro & All Nutrient model.
- All three dietary models (using nutrients) can be adopted as a pre-screening tool for early detection of the T2DM.
- Negative predictive power is almost similar to all three dietary models.
- Micronutrient model was proposed as the best suited model for population based screening to identify T2DM incidences,
  - ✓ Due to high sensitivity (58.2%)
  - ✓ Increased predictive power with AUROC 0.704 (95%CI 0.652 to 0.755) and AUPRC 0.65 as compared to Base model AUROC 0.663 (95%CI 0.610 to 0.717) and AUPRC 0.61.
  - ✓ Less complexity in terms of data storage, handling and analysis.
  - ✓ Thus, optimally utilising NHS resources.
- Micronutrient model suggested that males had 1.647 (1/0.607) times higher odds to exhibit T2DM than females. Individual with family history of T2DM showed 2.320 high chances for T2DM risk as compared with without family history. Hypertension (stage-1) was associated with 1.025 times an increased likelihood of exhibiting T2DM. Increasing intake of magnesium was associated with a reduction in T2DM risk.

Sensitivity, Specificity of the Models using ROC

Quality of the proposed Models to predict T2DM using PR Curve



Discussion

Three strengths were identified by the study when compared with currently available diabetes risk score such as American Diabetes Association Risk score (ADART), the Finnish Diabetes Risk score (FINDRISC), the German Diabetes Risk Score (GDRS), The Leicester Risk Assessment score (LRA),

1. Study conducted among most of ethnic minority group such as South Asian and African accounting over 90% of the studied population as compared to other studies.
2. Dietary information captured was comprehensive by accessing and assessing 22 nutrients (specific for T2DM) and 14 food groups through 380 foods.
3. Wider age range was covered (from 18-80 years).

Diagnostic accuracy for predicting T2DM

Label	ADART	FINDRISC	GDRS	LRA Score	The cureent study Micronutrient based model
Model Type	LR	LR	LR	LR	BLR- Forward
Sample size	1021	1069	3625	6186	382
Sample population	Taiwanese	Peruvian population	German Population	Multiethnic population, UK (73% white European, 22% South Asian)	South Asian (86.73%), White British (9.18%), African (4.08%)
Use of Questionniare	Yes	Yes	Yes	Yes	Yes
Variables considered	Age, BMI, family history of diabetes, ethnicity, Physical activity, IFG or IGT, hypertension, HDL cholesterol and/or triglyceride level, history of vascular disease, lifestyle behaviors	Age, BMI, WC, physical activity, family history of T2DM, daily consumption of fruits, berries or vegetables,history of anti-hypertensive drug treatment, and history of high blood glucose	Age, WC, height, history of hypertension, physical activity, smoking, family history of diabetes, intake of red meat, wholegrain and muesli, and coffee consumption.	Age, Sex, ethnicity (White European vs. other),first degree family history, WC, BMI, family history of cardiovascular disease, antihypertensive medication or high blood pressure, statin use, smoking status, alcohol consumption and previous diagnosis of any health	Age, Gender, ethnicity, family history, BMI, WC, hypertension (systolic BP),awareness towards T2DM reversal, Carb, englyst fibre, protein, fat-total,SFA, MUFA, PUFA, Na, K, Ca, Mg, Cu, I, Fe, Se, Zn, P, Vit-A (Retinol), Vit-B12, Vit-C, Vitamin-D, folate
Addition of dietary variables	No	Yes	Yes	Yes	Yes
Age	40–88 years	35 to 64 years	18–79 years	40–75 years	18-80 years
Model used for	Predicting the three-year incidence of pre-diabetes and diabetes	Predicting the incidence of undiagnosed T2DM	Five-year risk prediction of diagnosed type 2 diabetes	High risk of impaired glucose regulation and T2DM	Identifying T2DM risk
AUROC	Male 0.600 (0.54–0.66) & Female 0.720 (0.65–0.77)	0.690 (0.64– 0.74)	0.860(0.82 to 0.89)	0.720 (0.69 to 0.74)	0.704 (0.652 to 0.755)
Reference	(Li et al., 2011)	(Bernabe-Ortiz et al., 2018)	(Paprott et al., 2016)	(Gray et al., 2010)	The current study

Note: AUROC- Area under receiver operating curve, LR- Logistic Regression, BLR- Binary Logistic regression, ADART- American Diabetes Association Risk score , FINDRISC- the Finnish Diabetes Risk score, GDRS- the German Diabetes Risk Score, LRA- The Leicester Risk Assessment score, , BMI- Body mass index, WC- Waist circumference, IFG- Impaired fasting glucose, IGT - impaired glucose tolerance, HDL- High-density lipoprotein, Carb- carbohydrate, SFA- Saturated fatty acids, MUFA- Monounsaturated fatty acids, PUFA- Poly unsaturated fatty acids, Na- Sodium, K- Potassium, Ca- Calcium, Mg- Magnesium, Cu- Copper, I- Iodine, Fe Iron, Se- Selenium, Zn- Zinc, P- Phosphorus,

Conclusion

1. The study questionnaire can be adopted as a "briefing tool" to set personalised dietary goal among T2DM patients, and high risk individuals for not only in management but prevention of T2DM at clinical settings.
2. The study showed that by adding dietary based factors to already known risk predictors, predictive power of the models increased to discriminate T2DM.
3. Thus, the study questionnaire can be adopted as a non-invasive screening tool to identify T2DM at clinical settings, before applying costly invasive route.

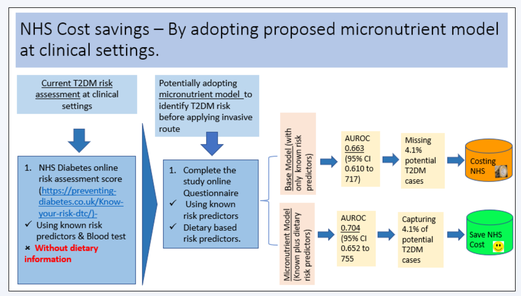
Strengths-

1. Culturally sensitive online dietary questionnaire was developed using EPIC- FFQ as base template to assess daily habitual dietary intake over one year.
2. All three dietary models (using nutrients) can be adopted as a pre-screening tool for early detection of the T2DM.
3. Study models were developed with majority of ethnic minority individuals (86.73% South Asian and African 4.08%) and White British (9.18%), unlike other current diabetes risk score.
4. Able to assess twenty-two nutrients specific to T2DM at individual and ethnicity level through 14 food groups and 380 foods specific to studied cohort.

Limitations-

1. White Other and Asian other ethnicities were underrepresented in this study.
2. Multi centre studies are needed with larger sample size with biometric measurements and time factor.
3. Models need to be tested at clinical settings.
4. Use of questionnaire can cause under or over representation, memory loss.

Overview of proposed approach



References

1. Mulligan, A. A. et al. (2014) 'A new tool for converting food frequency questionnaire data into nutrient and food group values: FETA research methods and availability', BMJ Open, 4(3). doi: 10.1136/bmjopen-2013-004503.
2. Goff, L. M. (2019) 'Ethnicity and Type 2 diabetes in the UK', Diabetic Medicine, 36(8), pp. 927–938. doi: 10.1111/dme.13895.
3. Public Health England (2016) 'Government recommendations for energy and nutrients for males and females aged 1 – 18 years and 19+ years. Government Dietary Recommendations', Government Dietary Recommendations, pp. 1–12. Available at: www.gov.uk/phe%0Awww.gov.uk/phe%0Awww.gov.uk/phe.
4. Bernabe-Ortiz, A. et al. (2018) 'Diagnostic accuracy of the Finnish Diabetes Risk Score (FINDRISC) for undiagnosed T2DM in Peruvian population', Primary Care Diabetes. Primary Care Diabetes Europe, 12(6), pp. 517–525. doi: 10.1016/j.pcd.2018.07.015.
5. Gray, L. J. et al. (2010) 'The Leicester Risk Assessment score for detecting undiagnosed Type 2 diabetes and impaired glucose regulation for use in a multiethnic UK setting', Diabetic Medicine, 27(8), pp. 887–895. doi: 10.1111/j.1464-5491.2010.03037.x.
6. Li, C. I. et al. (2011) 'Prospective validation of american diabetes association risk tool for predicting pre-diabetes and diabetes in Taiwan-Taichung community health study', PLoS ONE, 6(10), pp. 1–7. doi: 10.1371/journal.pone.0025906.
7. Paprott, R. et al. (2016) 'Validation of the German Diabetes Risk Score among the general adult population: Findings from the German Health Interview and Examination Surveys', BMJ Open Diabetes Research and Care, 4(1), pp. 1–9. doi: 10.1136/bmjdrc-2016-000280.

Acknowledgement

This academic research is self-funded. We liked to thank the volunteers who took part in this study and completing questionnaires.

Sincere thanks to Leicester city Mayor Sir. Peter Soulsby for voluntarily participating and appreciating DMU's efforts within Leicester's diverse ethnic communities.

Contact

Kshama Joshi: kshamarjoshi@gmail.com
ORCID ID: https://orcid.org/ 0000-0003-1301-5263
LinkedIn: www.linkedin.com/in/kshamajoshi2022