



IJCRR

Section: Healthcare

Sci. Journal Impact

Factor: 6.1 (2018)

ICV: 90.90 (2018)



Copyright@IJCRR

Risk Prediction for Diabetes Mellitus - A Population Based Approach

Rajesh Kumar¹, Prakash Khunte², Uday Shankar Chandrawanshi², Priyadarshini Rangari³

¹Assistant Professor, Department of Community Medicine, Prasad Institute of Medical Sciences, Lucknow, India; ²Assistant Professor, Department of General Medicine, Bharat Ratna Late Shri Atal Bihari Vajpayee Memorial Government Medical College, Rajnandgaon, Chhattisgarh, India; ³Associate Professor, Department of Dentistry, Sri Shankaracharya Institute of Medical Sciences, Bilai, Chhattisgarh, India.

ABSTRACT

Introduction: Etiological models use identical estimation procedures as most predictive modelling (i.e., regression) to quantify the relative risk related to a selected exposure on an outcome. Though regression is usually used for both purposes, the way within which the model is built will differ thanks to the goals of the model. The goal of a prediction model differs in several important ways.

Methods: Using a cohort design that links baseline risk factors to a validated population-based diabetes registry, a model (Diabetes Population Risk Tool or DPoRT) to predict risk factors for diabetes using commonly-collected national survey data was developed and validated. The event cohort was the National Population Health Survey (NPHS) linked to the validated Diabetes Database, a provincial component of the National Diabetes closed-circuit television (NDSS). Variables were restricted to factors routinely measured within the population. The probability of developing diabetes was modelled using sex-specific survival functions for those > 20 years, without diabetes and not pregnant at baseline (N = 19,000).

Results: The age-standardized 5-year incidence rates in the development cohorts were 6.52 % for males and 5.42 % for females. The 3-year age-standardized incidence rates in the development cohort were 3.42 % for males and 2.41% for females. The age-standardized 5-year incidence rates in the development cohorts were 6.42 % for males and 4.20 % for females. The age-standardized 3-year incidence rates for validation cohort was 3.45 % for males and 3.22 % for females.

Conclusion: Determinants of weight and weight change are essential when developing strategies to prevent or reduce the future diabetes burden. In monitoring trends over time researchers are often faced with the dilemma of separating trends between individuals and trends within individuals. Multilevel growth models allow us to model both these aspects which strengthen the ability to model trends that vary between and within individuals.

Key Words: Diabetes Burden, Framingham Heart Score, Multilevel Growth Models, National Population Health Survey, Prediction Models

INTRODUCTION

In many scientific disciplines, the study that predicts or forecast what is going to happen within the future has contributed to our understanding of the planet. The worth of scientific studies that provide models to tell strategies that may modify and possibly mitigate future events is of importance to society. Examples include estimating the impact of climate or environmental changes on the earth's ecosystems or the impact of policy changes on the economy.¹⁻³ These prediction models are accepted as valuable tools by scientists and have provided critical information for the event of strategies to

change predicted trends^{4,5} within the field of epidemiology, prediction models are underrepresented and also the concept of risk prediction is overshadowed by the estimation of relative risk measures to clarify etiological perspectives of disease. Etiological models use identical estimation procedures as most predictive modelling (i.e., regression) to quantify the relative risk related to a selected exposure on an outcome. Though regression is usually used for both purposes, the way within which the model is built will differ thanks to the goals of the model. The goal of a prediction model differs in several important ways. First, the result which must be optimized

Corresponding Author:

Dr. Prakash Khunte, Department of General Medicine, Bharat Ratna Late Shri Atal Bihari Vajpayee Memorial Government Medical College, Rajnandgaon, Chhattisgarh, India; Phone: 7879101234; Email: panhealth121013@gmail.com

ISSN: 2231-2196 (Print)

ISSN: 0975-5241 (Online)

Received: 09.09.2020

Revised: 07.10.2020

Accepted: 28.10.2020

Published: 12.11.2020

is an absolute measure of risk, often expressed as percentage or probability (versus a risk or hazard ratio). Second, the goal of a prediction model is to maximise the flexibility to discriminate between in danger groups and to properly classify true risk, referred to as discrimination and calibration. Typically these indices don't seem to be evaluated in etiological models. Thirdly, prediction models must be generalizable in other populations to which the model is applied. Typically etiological models fit the information accustomed to generate the relative risk estimate as tightly as possible and as a result, might not be reproducible using data in other settings or might not be applicable to explain risk in another population. These goals change the factors for model assessment and concomitantly the methodological framework is employed.

In medicine, prediction models, within the type of risk algorithms, are used as tools for patient decision-making. A risk algorithm could be a tool wont to estimate absolutely the risk of an outcome for a person as a function of their baseline characteristics. Typically, the risk is expressed because of the probability of dying or developing a disease in a very given fundamental measure.⁶ One in every of the foremost utilized risk tools is that the Framingham Heart Score.⁷ This tool is employed to calculate the probability that a patient will develop coronary heart condition in 5 or 10 years and has been widely integrated into disorder prevention and management throughout the planet.⁸⁻¹² Risk algorithms are widely recommended by medical societies for appropriate identification of patients that may have the benefit of specific interventions. This is often exemplified in clinical guidelines for pharmacologic interventions like cholesterol-lowering medications.¹³ Several potential benefits are also realized by extending the appliance of those tools to the population level. Just like the individual level, at the population setting predictive risk tools have the potential of providing insight into the longer-term burden of a disease in a complete region or nation and therefore the influence of specific risk factors. These tools can support health care higher cognitive process, including the effective and efficient allocation and distribution of health care resources and plan for effective disease prevention interventions. To date, prediction tools specifically designed to be used at the population level are neither created nor used for planning.

MATERIALS AND METHODS

Using a cohort design that links baseline risk factors to a validated population-based diabetes registry, a model (Diabetes Population Risk Tool or DPoRT) to predict risk factors for diabetes using commonly-collected national survey data was developed and validated. The event cohort was the National Population Health Survey (NPHS) linked to the validated Diabetes Database, a provincial component of the National Diabetes closed-circuit television (NDSS). Variables were

restricted to factors routinely measured within the population. The probability of developing diabetes was modelled using sex-specific survival functions for those > 20 years, without diabetes and not pregnant at baseline (N = 19,000). [Ethical clearance no. BRLSABV7/07/2019]

The model was validated in two external validation cohorts, both linked to administrative data for NDSS-defined physician-diagnosed diabetes. Predictive accuracy was assessed by comparing observed physician-diagnosed diabetes rates with predicted risk estimates from DPoRT. Discrimination of the model was assessed employing a C statistic and calibration was assessed with the chi-square statistic.

RESULTS

In the development cohort, 700 males and 665 females developed physician-diagnosed diabetes within the 5 year follow-up period. The age-standardized 5-year incidence rates in the development cohorts were 6.52% for males and 5.42% for females. The 3-year age-standardized incidence rates in the development cohort were 3.42% for males and 2.41% for females. The age-standardized 5-year incidence rates in the development cohorts were 6.42% for males and 4.20 % for females. The age-standardized 3-year incidence rates for validation cohort was 3.45 % for males and 3.22 % for females. All baseline population characteristics in the derivation cohort and two validation cohorts are shown in table 1. Both the validation cohorts differed from the derivation cohort. There were similar in age distribution, however, both had a higher proportion of obese individuals. Compared to the derivation cohort had a higher baseline prevalence of hypertension and heart disease but a lower prevalence of smoking while the other cohort had higher levels of hypertension and heart disease compared to the derivation cohort in women only.

Table 1: Baseline characteristics of development and validation cohorts with DPoRT multivariate-adjusted hazard ratios and 95% confidence intervals for 5-year physician diagnosed diabetes for males

Risk Factor	Subset	CI
Intercept		10.5861
Hypertension	No	-
	Yes	-0.2413
Non-white Ethnicity	No	-
	Yes	-0.6214
Heart Disease	No	-
	Yes	-0.5344
Current Smoker	No	-
	Yes	-0.1642

Table 1: (Continued)

Risk Factor	Subset	CI
Education	< Post-secondary	-
	Secondary	0.2222
Age-BMI category	BMI<23*Age<45	-
	23≤BMI<25*Age<45	-1.1364
	25≤BMI<30*Age<45	-1.4280
	30≤BMI<35*Age<45	-2.4324
	BMI≥35*Age<45	-3.3704
	BMI<23*Age≥45	-1.8463
	23≤BMI<25*Age≥45	-2.2225
	25≤BMI<30*Age≥45	-2.5466
	30≤BMI<35*Age≥45	-3.2163
	BMI≥35*Age≥45	-3.4555
Scale		0.6026

Table 2: Baseline characteristics of development and validation cohorts with DPoRT multivariate-adjusted hazard ratios and 95% confidence intervals for 5-year physician-diagnosed diabetes for females

Risk Factor	Subset	CI
Intercept		10.3272
Hypertension	Yes	-
	No	-0.2642
Non-white Ethnicity	Yes	-
	No	-0.3207
Immigrant Status	< Post-secondary	-
	Secondary	0.2130
Education	< Post-secondary	-
	Secondary	0.2130
Age-BMI category	BMI<23*Age<45	-
	23≤BMI<25*Age<45	-0.4230
	25≤BMI<30*Age<45	-0.4361
	30≤BMI<35*Age<45	-1.2102
	BMI≥35*Age<45	-2.0440
	BMI = missing*Age<45	-1.1224
	BMI<23*45≤Age<65	0.0610
	23≤BMI<25*45≤Age<65	-0.6021
	25≤BMI<30*45≤Age<65	-1.2132
	30≤BMI<35*45≤Age<65	-2.1140
	BMI≥35*45≤Age<65	-2.1665
	BMI = missing*45≤Age<65	-1.6240
	BMI<23*Age≥65	-1.0413
23≤BMI<25*Age≥65	-1.1215	
25≤BMI<30*Age≥65	-1.5765	

Table 2: (Continued)

Risk Factor	Subset	CI
	30≤BMI<35*Age≥65	-1.6142
	BMI≥35*Age≥65	-1.6552
	BMI = missing*Age≥65	-1.6552
Scale		0.6302

DISCUSSION

This study demonstrated that diabetes risk can be accurately predicted at the population level using self-reported age, sex, Body Mass Index and other measures available in population health surveys. In addition to displaying good discrimination, DPoRT-predicted rates closely agreed with observed rates for both males and females in both external validation cohorts, and this agreement was generally maintained across deciles and quintiles of risk. To my knowledge, DPoRT is the first validated risk tool that is integrated into commonly-collected population health survey data. DPoRT offers advantages over existing methods used to estimate future diabetes risk in populations. Previous studies that estimate future diabetes burden have either extrapolated overall trends in diabetes prevalence or indirectly incorporated information on the influence of risk factors with various assumptions.¹³ Other studies focus on overall diabetes burden, a useful approach, but one which does not enable users to directly assess the impact of risk factors, such as BMI, on future diabetes. Furthermore, these studies did not assess how future diabetes can be prevented by targeting risk factors since they do not directly quantify the influence of risk factors on baseline risk or diabetes incidence.

Complex modelling and simulation studies differ from the approach used in this study in that they use additional information on how populations and risk factors change over time.^{14,15} Other simulation studies add more detailed clinical information such as fasting blood sugar level or information on diabetes family history, data not available at the population level. Strength of these simulation models is that they can combine different data sources and study findings¹⁶ However, these models are complex and often represent clinical or theoretical populations, making their estimates difficult to validate in external populations that are meaningful for population health planning. DPoRT could be incorporated into simulation models that consider future changes in population composition and risk factors. The nature of diabetes risk allowed us to discriminate and explain risk using a limited number of variables – most importantly BMI. Discrimination of DPoRT is as high as or higher than many clinical risk prediction tools used in clinical practice. The algorithm was further calibrated using population means, which may atten-

uate differences between populations since risk estimates are relative to baseline risk in the population.

Given current data in most countries, DPoRT is a more balanced approach to estimating diabetes risk than methods used in previous research. Several important clinical values are excluded from DPoRT, such as hip to waist ratio, waist circumference, fasting blood glucose, and family history.¹⁷⁻¹⁹ Although these variables may be clinically important for assessing diabetes risk, adding these, or other detailed anthropometric measures is not feasible because they are not routinely collected in most populations. These omitted variables are unlikely to have a major impact on the performance characteristics of the model due to the clustering of risk factors, particularly when dealing with abnormalities of the metabolic system.²⁰⁻²¹ Variables not included in DPoRT, such as the family history of diabetes or poor diet, are also associated with the clustering of metabolic risk factors that are included in the algorithm such as hypertension and BMI. Obesity is the most important factor in predicting diabetes risk. BMI is the most commonly used marker of obesity; however, measures of central obesity may capture the entire risk domain more comprehensively and be more meaningful across all age groups.²² A recent meta-analysis has shown that there is no evidence of a difference in estimates associated with incident diabetes between BMI, waist circumference and waist/hip ratios.²³ Furthermore, algorithms to identify individuals for weight loss in populations did not differ if using BMI or waist circumference.²⁴ To ensure DPoRT can be applied in different populations, we gave preference to variables that were: based on established evidence, remained stable over time, were unlikely to be subject to serious measurement error (such as alcohol and dietary habits), and were easily captured using survey data in different populations. For example, physical activity has been shown to have a protective effect on diabetes incidence²⁵ but was removed from the final algorithm due to the inability to capture this in a reliable and reproducible manner across studies, and because of its marginal improvement in the discrimination of diabetes risk in our creation cohort. Despite placing considerable constraints on variable selection as a means of ensuring maximum feasibility, DPoRT maintained good discrimination. The effect of self-reported BMI may depend on the population where the algorithm is being applied since these patterns have been shown to vary across gender and socioeconomic status.²⁶⁻²⁸ The ability of DPoRT's predictive estimates to agree with observed diabetes risk in different populations will be reduced if systematic errors associated with responses vary across populations or time.

If diabetes testing/screening increases over time predicted estimates could be lower than the observed estimates (under the assumption that this would lead to increased case detection). DPoRT is accurate in different populations for different periods; however, DPoRT could be recalibrated to predict

total diabetes cases using revised information on screening/testing or using estimates of the number of undiagnosed cases relative to diagnosed cases in the population. Finally, the potential for inaccuracy increases the longer into the future the predictions are made or when unforeseen changes occur; therefore, it is recommended that predictions from DPoRT are updated frequently by using the most recent data, limiting predictive calculations to 5 years or less, and validating the risk tool in the population where it is being applied.

CONCLUSION

In building the risk tool for diabetes it was demonstrated that that BMI (a relative measure of weight for height) overwhelmingly influences the predictions for developing diabetes in the future. For that reason, clarifying determinants of weight and weight change is essential when developing strategies to prevent or reduce the future diabetes burden. In monitoring trends over time researchers are often faced with the dilemma of separating trends between individuals and trends within individuals. Multilevel growth models allow us to model both these aspects which strengthen the ability to model trends that vary between and within individuals.

Acknowledgement: None

Conflicts of Interest: Nil

Source of funding: Self

REFERENCES

1. Colditz G, Willett WC, Stampfer MJ, Manson JE, Hennekens CH, Arky RA, et al. Weight as a risk factor for clinical diabetes in women. *Am J Epidemiol* 1990;132:501-513.
2. Colditz G, Willett WC, Rotnitzky A, Manson JE. Weight gain as a risk factor for clinical diabetes Mellitus women. *Ann Intern Med* 1995;122:481-486.
3. Perry IJ, Wannamethee SG, Walker MK, Thomson AG, Whincup PH, Shaper AG. Prospective study of risk factors for development of non-insulin dependent diabetes in middle aged British men. *BMJ* 1995;310:560-564.
4. Vanderpump MPJ, Tunbridge WM, French JM, Appleton D, Bates D, Rodgers H, et al. The incidence of diabetes mellitus in an English community: a 20-year follow-up of the Wickham Survey. *Diabet Med* 1996;13:741-747.
5. Wilson PWF, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr. Prediction of incident diabetes mellitus in middle-aged adults. *Arch Intern Med* 2007;167:1068-1074.
6. Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, Koplan JP. The continuing epidemics of obesity and Diabetes in the United States. *JAMA* 2001;286:1195-1200.
7. Orpana HM, Tremblay MS, Fines P. Trends in weight change among Canadian adults. *Health Reports* 2007;18:9-16.
8. Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *J Edu Behav Stat* 1998;23:323-355.
9. Statistics Canada. National population health survey household component cycle 6 (2005-2006) longitudinal documentation. 2006. Ottawa.

10. Statistics Canada. 1996-7 National Population Health Survey: Derived Variable Specifications. 1999. Ottawa.
11. Yeo D, Mantel H, Lui TP. Bootstrap variance estimation for the National Population Health Survey, Baltimore, American Statistical Association 1999;778-783.
12. Demidenko E, Stukel TA. Efficient estimation of generalized linear mixed-effects models. *J Stat Plann Inference* 2002;104:197-219.
13. Stukel TA, Demidenko E. Two-stage method of estimation for general linear growth curve models. *Biometrics* 1997;53:720-728.
14. Singer JD, Willett JB. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford: Oxford University Press, 2003
15. Statistics of Canada. NPHS Public Use Microdata Documentation. 1999. Ottawa, Canada, StatisticsCanada.
16. Heo M, Faith MS, Mott JW, Gorman BS, Redden DT, Allison DB. Hierarchical linear models for the development of growth curves: an example with body mass index in overweight/obese adults. *Statistics in Medicine* 2003;22:1911-1942.
17. Streiner DL, Norman GR. *Measuring Change. Health Measurement Scales*. New York: Oxford University Press, 2003:194-212.
18. Haapanen N, Miilunpalo S, Pasanen M, Oja P, Vuori I. Association between leisure-time physical activity and 10-year body mass change among working-aged men and women. *Int J Obes* 1997;21:288-296.
19. Hu FB, Sigal RJ, Rich-Edwards JW, Colditz GA, Solomon CG, Willett WC, et al. Walking Compared With Vigorous Physical Activity and Risk of Type 2 Diabetes in Women. *JAMA* 1999;282:1433-1439.
20. Ahmad MS. Prevalence of risk factors for diabetes among construction workers in majmaah city, Saudi Arabia. *Int J Cur Res Rev* 2018; 10: 4:55-60.
21. Williamson DF, Madans J, Anda RF, Kleinman JC, Kahn HS, Byers T. Recreational Physical-Activity and 10-Year Weight Change in A United-States National Cohort. *Int J Obes Relat Metab Disord* 1993;17:279-286.
22. Gordon-Larsen P, Hou N, Sidney S, Sternfeld B, Lewis CE, Jacobs DR Jr, Popkin BM. Fifteen-year longitudinal strands in walking patterns and their impact on weight change. *Am J Clin Nutr* 2009;89:19-26.
23. Petersen L, Schnohr P, Sorensen TIA. A longitudinal study of the long-term relation between physical activity and obesity in adults. *Int J Obesity* 2004;28:105-112.
24. Wilsgaard T, Jacobsen BK, Arnesen E. Determining lifestyle correlates of body mass index using multilevel analyses: The Tromso study, 1979-2001. *Am J Epidemiol* 2005;162:1179-1188.
25. Fogelholm M, Kujala U, Kaprio J, Sarna S. Predictors of weight change in middle-aged and old men. *Obes Res* 2000;8:367-373.
26. Meltzer AA, Everhart JE. Unintentional Weight-Loss in the United-States. *Am J Epidemiol* 1995;142:1039-1046.
27. Molarius A, Seidell JC, Kuulasmaa K, Dobson AJ, Sans S. Smoking and relative body weight: An international perspective from the WHO MONICA project. *Journal of Epidemiology and Community Health* 1997;51:252-260.
28. Goel MS, McCarthy EP, Phillips RS, Wee CC. Obesity among US immigrant subgroups by duration of residence. *JAMA* 2004;292:2860-2867.