International Journal of Diabetes in Developing Countries (April–June 2025) 45(2):424–432
https://doi.org/10.1007/s13410-024-01374-w

ORIGINAL ARTICLE

Healthy lifestyle index development and its association with type 2 diabetes mellitus status among teachers

Yit Han Ng[1] · Foong Ming Moy[1,2] · Noran Naqiah Hairi[1,2] · Awang Bulgiba[3]

Received: 26 March 2024 / Accepted: 27 June 2024 / Published online: 12 July 2024

Abstract

Background Type 2 diabetes mellitus (T2DM) is a significant health concern and imposes a substantial burden on society. Although leading a healthy lifestyle is an effective means of reducing the risk of T2DM, the complex interplay of diverse lifestyle factors necessitates a comprehensive assessment.

Objective This study aims to develop and investigate the association between a healthy lifestyle index (HLI) and T2DM among teachers.

Methods A total of 11412 teachers from schools in Peninsular Malaysia were recruited. The HLI score was constructed using both unweighted and regression-weighted methods. A comparison was conducted and the index with the lowest AIC was considered the best. The final model with an index and other covariates was then optimised, followed by model validations.

Results All teachers ($n = 11412$) who completed the baseline survey were included in the analysis. Majority were females, Malays, married, and had at least a bachelor's degree. Among the developed indexes comprised of waist circumference, physical activity, sleep duration, and mental health status (e.g. stress, anxiety, and depression). The B HLI index is the best proxy for teachers' lifestyles associated with T2DM. The optimised regression showed that an increase of an HLI score reduced the odds of T2DM by 59%. The model had an adjusted R^2 of 26.8% and an area under the curve of 0.837. The model was externally validated with the validation dataset by achieving an adequate C-statistic value of 0.798.

Conclusion A higher HLI score was associated with a reduced odd of T2DM. This finding highlights the importance of integrating multidimensional lifestyle modification rather than singular ones when developing health promotion strategies.

Keywords Lifestyles · Composite index · T2DM · Teachers · CLUSTer

Introduction

Understanding the link of type 2 diabetes mellitus (T2DM) and cardiovascular diseases (CVD) is crucial in current times, as CVD has emerged as a major cause of mortality worldwide, surpassing infectious diseases [1]. The significance of T2DM in precipitating the onset of many cardiovascular conditions, including coronary artery disease, heart failure, atrial fibrillation, and sudden death, highlights the urgent need for effective preventive strategies [2]. Adopting healthy lifestyles such as healthy eating, weight management, active living, and emotional resilience has been recognised as an approach for T2DM [1, 3]. However, lifestyles differ between communities and vary across occupational groups, necessitating tailored interventions to address the specific needs of these individual populations.

Considering this, the development of a healthy lifestyle index (HLI) emerges as a promising solution. The HLI concept involves providing a proxy for an individual's overall lifestyle through a combination of relevant lifestyle variables. By adopting this approach, individuals' lifestyle can be comprehensively understood, rather than relying on individual or single exposure risk factors.

Researchers need to explore the interactions between all lifestyle behaviours and the onset of T2DM. However, the presence of such complexity presents methodological challenges. Implementing a modelling approach with a high level of interaction

✉ Yit Han Ng
hanzhai2574@gmail.com

[1] Department of Social and Preventive Medicine, Faculty of Medicine, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

[2] Centre for Epidemiology and Evidence-Based Practice, Kuala Lumpur, Malaysia

[3] Academy of Sciences Malaysia, Kuala Lumpur, Malaysia

among the lifestyle variables, including third-order interactions and above, is particularly difficult [4]. To overcome this, researchers typically rely on advanced and high-end statistical techniques to model the data. An effective balance between granularity and practicability is important for this to be useful.

Many HLIs include factors such as diet, physical activity, weight management, smoking, and alcohol consumption [5–9]. However, developing an HLI for specific populations, such as teachers, may require additional components such as sedentary behaviour, sleep quality, and mental health issues [10]. Therefore, adding these components is necessary to build an index that can predict health issues, such as T2DM or other lifestyle-related illnesses among teachers. In addition, the methodology used to construct an HLI can impact the index's accuracy in estimating disease risk. Building an HLI using a regression-weighted technique on all lifestyle indicators can outperform the unweighted method [11]. Although the weighted index demonstrates greater discriminant power, several studies have found no difference between the weighted and unweighted indexes when predicting health status [12]. To develop an ideal index, a combination of multiple lifestyle components, along with an appropriate index construction method, is necessary.

The development of an HLI based on rigorous methodology has the potential to account for interactions between multiple lifestyle factors and T2DM. It can also direct tailored interventions in specific occupational groups such as teachers. This study aimed to develop and investigate the association between an HLI and T2DM status among teachers.

Methods and Materials

Study design

This was a cross-sectional study that utilised the baseline data from the CLUSTer cohort study. Briefly, CLUSTer is a Malaysian teacher cohort, in which participants were teachers from public primary and secondary schools in Peninsular Malaysia. The study was initiated in 2013, with the primary objective of determining the clustering of lifestyle risk factors and work-related factors among teachers in Peninsular Malaysia, and follow-up for non-communicable diseases such as diabetes, hypertension, cardiovascular diseases, and cancers [13].

Study participants and study variables

A total of 11412 primary and secondary school teachers were enrolled in this study. These teachers were free from psychiatric illness and completed the baseline questionnaire, which contained demographic information, medical history, and lifestyle-related behaviours. Additionally, waist circumference, blood pressure, and blood samples were collected to screen for fasting blood glucose and lipid levels. All variables related to T2DM were summarised into three groups: demographic characteristics, lifestyle-related indicators, and laboratory-measured parameters (Supplementary Table 1).

Statistical analysis

The dataset was randomly split into training ($n = 8559$) and validation ($n = 2853$) subsets in a 70:30 ratio. Under descriptive analysis, the mean with standard deviation or median with inter-quartile range was used to present all scaled variables, subjected to the normality of the data. Frequency with percentage was used for all categorical variables.

The training dataset was used for the development of the lifestyle index. Lifestyle variables such as physical activity (categorical: low/moderate/high), waist circumference, central obesity status (based on the Asian abdominal obesity cut-off values: ≥ 90 cm for men; ≥ 80 cm for women), sleep hours, and anxiety/depression/stress scores were selected to construct the index, according to a study that investigated lifestyle factors associated with T2DM among teachers in the CLUSTer study [14].

An unweighted index was constructed through dichotomising the variables (e.g. central obesity: yes vs no), followed by combining them into a single score. Concurrently, five established regression-weighted methods were used to compute the coefficient-based indexes [11, 15] and compared with the unweighted index. Akaike information criterion (AIC) was used as a parameter to evaluate the indexes' performance. Notably, the index that scored the lowest AIC, along with the corrected Akaike information criterion (AICc), and with the highest weighted AIC was considered the ideal index to estimate T2DM status among teachers [16]. Additionally, a sensitivity analysis was performed to compare the best index with lifestyle indicators including fruit and vegetable consumption, sitting duration (sedentary/non-sedentary), smoking status, and alcohol consumption.

The HLI was further modelled via multiple logistic regression, which incorporated other covariates with p-values < 0.25. The model was optimised through exploration of potential interactions or non-linear terms between the HLI and other covariates in the multivariable regression analysis. The final model was validated internally using the bootstrap method with 300 replicates [17]. A minimum deviation from the bootstrap on the tested parameters suggested that the index was validated internally.

The index was validated externally (validation dataset that did not involve in the model construction) via a calibration plot. A plot that is distributed near a diagonal line was considered acceptable. Lastly, the HLI index was further categorised into tertiles, with additional analyses aimed at

Table 1 CLUSTer cohort teachers' characteristics stratified by training and validation datasets

Characteristics	Training data, $N=8559$[1]	Validation data, $N=2853$[1]	Total, $N=11412$[1]
Age	40.61 ± 8.58	40.73 ± 8.59	40.64 ± 8.59
Sex			
Male	1555 (18%)	550 (19%)	2105 (18%)
Female	7004 (82%)	2303 (81%)	9307 (82%)
Ethnicity			
Malay	6571 (77%)	2189 (77%)	8760 (77%)
Chinese	1390 (16%)	475 (17%)	1865 (16%)
Indian	543 (6.3%)	161 (5.6%)	704 (6.2%)
Other races	55 (0.6%)	28 (1.0%)	83 (0.7%)
Education			
Diploma and below	518 (6.1%)	157 (5.5%)	675 (5.9%)
Degree	6525 (76%)	2202 (77%)	8727 (76%)
Master and above	1516 (18%)	494 (17%)	2010 (18%)
Marital status			
Single	966 (11%)	315 (11%)	1281 (11%)
Married	7400 (86%)	2471 (87%)	9871 (86%)
Divorced/widowed	193 (2.3%)	67 (2.3%)	260 (2.3%)
Family history of T2DM (yes)	4244 (50%)	1425 (50%)	5669 (50%)
Fasting blood glucose (mmol/L)	5.09 ± 1.26	5.11 ± 1.27	5.10 ± 1.26
Systolic blood pressure (mmHg)	125.05 ± 16.19	125.41 ± 16.44	125.14 ± 16.25
Diastolic blood pressure (mmHg)	76.98 ± 10.39	77.11 ± 10.06	77.01 ± 10.31
Total cholesterol (mmol/L)	5.34 ± 0.90	5.33 ± 0.90	5.34 ± 0.90
Triglycerides (mmol/L)	1.16 ± 0.54	1.17 ± 0.56	1.16 ± 0.55
High-density lipoprotein (mmol/L)	1.44 ± 0.32	1.45 ± 0.32	1.44 ± 0.32
Low-density lipoprotein (mmol/L)	3.32 ± 0.80	3.31 ± 0.79	3.31 ± 0.80
Fruit and vegetable servings			
Adequate	302 (3.5%)	93 (3.3%)	395 (3.5%)
Smoking status (yes)	234 (2.7%)	58 (2.0%)	292 (2.6%)
Alcohol consumption (yes)	267 (3.1%)	85 (3.0%)	352 (3.1%)
Physical activity			
Low	4550 (53%)	1510 (53%)	6060 (53%)
Moderate	2275 (27%)	706 (25%)	2981 (26%)
High	1734 (20%)	637 (22%)	2371 (21%)
Sedentary behaviour			
Sedentary	131 (1.5%)	42 (1.5%)	173 (1.5%)
Waist circumference (cm)			
Men	89.32 ± 10.59	88.88 ± 10.17	89.20 ± 10.48
Women	79.75 ± 10.5	79.78 ± 10.40	79.76 ± 10.48
Central obesity status (by sex)			
Central obese	4033 (47%)	1335 (47%)	5368 (47%)
Sleep duration (hours)	6.31 ± 0.92	6.29 ± 0.89	6.31 ± 0.91
Sleep hours			
< 7 h	5996 (70%)	2007 (70%)	8003 (70%)
7 h and above	2563 (30%)	846 (30%)	3409 (30%)
Depression score	6.40 ± 6.55	6.53 ± 6.48	6.44 ± 6.53
Depression			
Symptomatic (mild symptoms and above)	2338 (27%)	819 (29%)	3157 (28%)
Anxiety score	8.47 ± 6.73	8.54 ± 6.68	8.49 ± 6.71
Anxiety			
Symptomatic (mild symptoms and above)	4421 (52%)	1481 (52%)	5902 (52.0)

Table 1 (continued)

Characteristics	Training data, $N=8559$[1]	Validation data, $N=2853$[1]	Total, $N=11412$[1]
Stress score	9.90 ± 7.08	10.01 ± 7.15	9.93 ± 7.09
Stress			
Symptomatic (mild symptoms and above)	2622 (31%)	897 (31%)	3519 (31%)
T2DM status (yes)	743 (8.7%)	270 (9.5%)	1013 (8.9%)
Impaired fasting glucose (yes)	477 (6.1%)	152 (5.9%)	629 (6.0%)
Hypertension (yes)	682 (8.0%)	228 (8.0%)	910 (8.0%)
Dyslipidaemia (yes)	1043 (12%)	303 (11%)	1346 (12%)

[1]Mean ± SD; *n* (%)

SD standard deviation

revealing potential trends in estimation between teachers who practiced healthier lifestyles compared to their counterparts on other health issues, such as impaired fasting glucose (IFG), hypertension, and dyslipidaemia.

All statistical analyses were conducted using R software version 4.2.3 on Mac. The list of R packages used and their functions includes the following: gtsummary—data tabulation [18]; Tidyverse—data wrangling [19]; AICcmodavg—AIC calculation and summary [16]; rms—statistical regression modelling [17]; and ggeffects—multivariable regression effect plot [20].

Results

Teacher's sociodemographic characteristics, health profiles, and lifestyles are presented in Table 1. The mean age of the teachers was 40.64 years, and the majority of them were females and Malays. More than two-thirds graduated with degrees and were married with approximately half having a family history of T2DM. The majority of teachers had normal blood pressure, fasting blood glucose levels, and lipid profiles. However, about 8.9% had T2DM (both clinically diagnosed and undiagnosed T2DM). Meanwhile, approximately 6.0% of the teachers were prediabetic. The proportions of hypertension and dyslipidaemia among the teachers were 8% and 12%, respectively.

In terms of lifestyle habits, only 3.5% of the teachers consumed the recommended daily intake of two servings of fruits and three servings of vegetables. In addition, only a minority of teachers were smokers (2.6%) and alcohol drinkers (3.1%). More than half of the teachers were classified as having low physical activity by the IPAQ-short questionnaire, whereas 173 (1.5%) sat for more than 10 h daily. The average waist circumference measured was 89.2 cm for males and 79.76 for females, with 47% of teachers being categorised as abdominally obese. Regarding sleep duration, teachers only had an average sleep duration of 6.4 h, with 70% of them had less than optimal 7 h of sleeping time daily. Lastly, approximately 52% of teachers had acute anxiety, followed by stress (31%) and depression (28%).

Since there was low adherence of lifestyle behaviours such as fruit and vegetable consumption, smoking, alcohol consumption, and sitting durations among teachers, only indicators including waist circumference, physical activity, sleep durations, and acute mental health were involved in the index development. Of all these lifestyle variables, waist circumference had the largest negative absolute beta value to the B HLI index associated with the T2DM status among teachers, followed by sleep hours and physical activity. No noticeable pattern was observed in the acute mental health variables (Table 2). All regression-weighted indexes (HLI $=\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6$, where $\beta=$regression coefficient) had lower AICc values than the unweighted HLI (Table 3). Of these regression-weighted indexes (B HLI, BI HLI, B10 HLI, BSch HLI, and B Su HLI), the B HLI scored the lowest AICc, lowest delta AICc, and highest AIC weights compared to other indexes. Therefore, the B HLI index was recalibrated and subsequently used for downstream regression modelling. A sensitivity analysis between the models with B HLI index and an index with additional lifestyle components showed not much difference in AIC values ($AICc_{BHLI}=4000.01$; $AICc_{HLIAdds}=3996.72$), although the index with additional lifestyle components scored a lower AIC value.

Table 4 displays a comparison between the saturated and optimised models. Model 1 included all covariates (saturated model), whereas Model 2 and Model 3 were optimised models comprising of reduced model and reduced model with interaction terms, respectively. Although Model 1 had more variables than Model 2 and Model 3, these models did not show much difference in *R* squared compared to Model 1 (26.6–26.8%), as supported by the likelihood ratio tests ($p>0.05$). Model 3 was selected as it had the lowest AIC value compared with the other models. In Model 3, for every one score increased in the HLI, there was a reduction of the odds of developing T2DM by 59% (95% confidence interval 50, 67), adjusted for other covariates. An internal validation of Model 3 demonstrated no substantial deviation from the

Table 2 Assigned scores for lifestyle indicators that relevant to teachers (training data, $n = 8559$)

Lifestyle indicators	Unweighted HLI	B HLI	BI HLI	B10 HLI	B Sch HLI	B Su HLI
Physical activity (vs low)						
Moderate	1	−0.05	0	−1	0	−1
High	2	−0.14	0	−1	0	−4
Waist circumference (vs non-central obese)	1	-	-	-	-	-
WC = 60–69.9 cm vs 110–109.9 cm	-	−3.14	−3	−34	−11	−97
WC = 70–79.9 cm vs 110–109.9 cm	-	−2.25	−2	−23	−7	−64
WC = 80–89.9 cm vs 110–109.9 cm	-	−1.45	−1	−14	−5	−41
WC = 90–99.9 cm vs 110–109.9 cm	-	−0.67	−1	−7	−2	−19
WC = 100–109.9 cm vs 110–109.9 cm	-	−0.30	0	−3	−1	−9
Sleep duration (vs ≥ 7 h)	1	-	-	-	-	-
Sleep hours = 5 h vs 4 h	-	−0.35	0	−4	−1	−10
Sleep hours = 6 h vs 4 h	-	−0.40	0	−4	−1	−11
Sleep hours = 7 h vs 4 h	-	−0.56	−1	−6	−2	−16
Sleep hours = 8 h vs 4 h	-	−0.70	−1	−7	−2	−20
Sleep hours = 9 h vs 4 h	-	−0.66	−1	−7	−2	−19
Stress (vs not symptomatic)	1	-	-	-	-	-
Stress = normal vs very severe	-	1.65	2	17	5	47
Stress = mild vs very severe	-	1.11	1	11	4	31
Stress = moderate vs very severe	-	1.38	1	14	5	39
Stress = severe vs very severe	-	1.00	1	10	3	28
Anxiety (vs not symptomatic)	1	-	-	-	-	-
Anxiety = normal vs very severe	-	−0.50	0	−5	−2	−14
Anxiety = mild vs very severe	-	−0.63	−1	−6	−2	−18
Anxiety = moderate vs very severe	-	−0.44	0	−4	−1	−13
Anxiety = severe vs very severe	-	−0.32	0	−3	−1	−9
Depression (vs not symptomatic)	1	-	-	-	-	-
Depression = normal vs very severe	-	0.10	0	1	0	3
Depression = mild vs very severe	-	0.11	0	1	0	3
Depression = moderate vs very severe	-	−0.04	0	0	0	−1
Depression = severe vs very severe	-	0.47	0	5	2	13

Note:

1) Negative regression-weighted score suggestive the lifestyle indicators were protective against T2DM

2) The best total score for all regression coefficient weighted (B HLI, BI HLI, B10 HLI, B Sch HLI, and B Su HLI) were reversed and calibrated from 0-maximum to compare with the HLI using ordinary arithmetic addition

WC waist circumference,

HLI healthy lifestyle index,

B HLI regression coefficient weighted index,

BI HLI regression coefficient weighted index (rounded to the nearest integer),

B10 HLI regression coefficient weighted (multiply by 10 and rounded) index,

B Sch HLI Schneeweiss's scoring system,

B Su HLI Sullivan's scoring system

original estimated index parameters with tested values via a bootstrap method; thus, the model was internally valid. In addition, the external validation also revealed the model could estimate the T2DM status among teachers adequately, by scoring an adequate C-statistic value of 0.798 (Fig. 1). Figure 1 summarises both internal and external validations of Model 3.

Table 5 shows the potential trends by teachers who practiced healthier lifestyles compared to their counterparts. Taking T1 (poorer lifestyle) as a reference group, there was a significant reduction in the odds of acquiring IFG from T1 to

Table 3 Comparison of newly developed indexes in association with diabetes status among teachers in CLUSTer cohort (training data, $n = 8559$)

Lifestyle index	AICc[$]	Delta AICc	AICc weight
Unweighted HLI	4108.92	108.91	0.00
B HLI[#]	4000.01	0.00	0.59
BI HLI[#]	4062.84	62.83	0.00
B10 HLI[#]	4003.90	3.90	0.08
B Sch HLI[#]	4008.12	8.11	0.01
B Su HLI[#]	4001.28	1.27	0.31

Note: The HLI score was computed using the formula:

HLI = 1.2 + [(0)(PA = low) or (−0.05)(PA = moderate) or (−0.14)(PA = high) + (−3.14)(WC = 60–69.9) or (−2.25)(WC = 70–79.9) or (−1.45)(WC = 80–89.9) or (−0.67)(WC = 90–99.9) or (−0.30)(WC = 100.0–109.9) or (0)(WC = 110–119.9) + (0.10)(depression = normal) or (0.11)(depression = mild) or (−0.04)(depression = moderate) or (0.47)(depression = severe) or (0)(depression = very severe) + (−0.50)(anxiety = normal) or (−0.63)(anxiety = mild) or (−0.44)(anxiety = moderate) or (−0.32)(anxiety = severe) or (0)(anxiety = very severe) + (1.65)(stress = normal) or (1.11)(stress = mild) or (1.38)(stress = moderate) or (1.00)(stress = severe) or (0)(stress = very severe) + (0)(sleep = 4 h) or (−0.35)(sleep = 5 h) or (−0.40)(sleep = 6 h) or (−0.56)(sleep = 7 h) or (−0.70)(sleep = 8 h) or (−0.66)(sleep = 9 h)] × −1

AICc, Akaike information criterion corrected; *delta AICc*, differences between calculated AICc; *CI*, confidence interval;

unweighted HLI, HLI scoring using non-weighted method;

B HLI, beta coefficient weighted HLI;

BI HLI, beta coefficient weighted HLI (rounded to the nearest integer);

B10 HLI, beta coefficient weighted HLI, multiplied by 10 and rounded to the nearest integer;

B Sch HLI, Schneeweiss's scoring system;

B Su HLI, Sullivan's scoring system

*$p < 0.05$

[$]Covariates included (age, sex, ethnicity, education level, marital status, family history of DM, systolic/diastolic blood pressures, LDL, HDL, total cholesterol, and triglycerides)

[#]Regression coefficient weighted index

T3, after accounting for the covariates. The estimation was consistent in the sensitivity analyses testing other health-related outcomes such as hypertension and dyslipidaemia, with significant p-values for trend < 0.001.

Discussion

An index of health-related lifestyle indicators (HLI) was developed incorporating measures of waist circumference, physical activity, sleep durations, and acute mental health (e.g. stress, anxiety, and depression). The hold-out dataset approach was employed for both index development and subsequent regression modelling to avoid overfitting. This method differs from alternative techniques like cross-validation, where the hold-out sample serves as an external dataset for evaluating the model's accuracy. However, it should be noted that using the hold-out approach with fewer samples may not be suitable for studies with larger sample sizes [21]. Since this study recruited over 10000 teachers, therefore, it does not add bias in the outcome estimation.

In all instances, regression-weighted indexes were found to outperform the unweighted index, despite the utilisation of varying regression-weighted scores. One possible reason for this could be that the conventional method of employing the ordinary addition approach in index construction may lead to methodological inaccuracies. Compared to the unweighted index, these weighted indexes allocated more weight to variables that were strongly associated with the outcome, thus being more capable of discriminating the outcome [22, 23]. Another possible explanation is that dichotomising continuous variables was not implemented while constructing the regression-weighted indexes. Dichotomising continuous variables is widely acceptable and preferable approach in health-related research due to its ease of interpretation and clinical application. However, the method may result in a loss of statistical power, underestimation of the variations in the outcome across different groups, and obscuration of any non-linear relationships between variables and their corresponding outcomes [22].

In the construction of an index in association with T2DM among teachers, waist circumference demonstrated the largest regression coefficient among the lifestyle variables selected. This finding can primarily be attributed to the fact that half of the teachers were classified as abdominally obese. Similarly, other studies have reported that teachers have a heightened susceptibility to being overweight or abdominally obese [24–26]. These results suggest that a considerable number of teachers are at risk of developing metabolic disorders, including T2DM. Abdominal obesity, which is characterised by the accumulation of excessive body fat in the abdominal region, is a potent risk factor for metabolic diseases. These conditions disrupt insulin activity in regulating glucose levels within the body, ultimately leading to elevated blood glucose levels [27, 28].

In addition, sleep duration also contributed substantially to the HLI, as a high proportion of teachers had less than 7 h of sleep daily. A teacher's occupation is well-known to be characterised by high levels of work stress and job strain, which can result in sleep deprivation [29]. This sleep deprivation has been directly linked to teachers' mental health, such as stress level [30]. Additionally, chronic sleep deprivation among teachers may lead to further stressful conditions, thereby elevating the cortisol levels in their bodies [31]. This may ultimately lead to chronic mental health issues thus leading to metabolic diseases. On the other hand, regular physical activity can prevent T2DM, as physical activity

Table 4 Factors associated with T2DM among teachers in CLUSTer cohort (training data, $n=8559$)

Main factor	Model 1		Model 2		Model 3	
	aOR	95% CI	aOR	95% CI	aOR	95% CI
Healthy lifestyle index (HLI)	0.53**	0.48, 0.60	0.53**	0.47, 0.59	0.41**	0.33, 0.50
R^2	26.8%		26.6%		26.8%	
AIC	3999.93		3993.18		3986.45	
AUC	0.837		0.836		0.837	
LR test	Model 1 vs Model 2 (df = 7) = 7.252, $p=0.403$					
	Model 1 vs Model 3 (df = 6) = 1.478, $p=0.961$					

Model 1 (saturated model) = adjusted for age, sex (male/female), ethnicity (Malay/Chinese/Indian/other races), education level (diploma and below/degree/master and above), marital status (single/married/divorced or widowed), family history of T2DM (no/yes), total cholesterol, triglycerides, low-density lipoprotein, high-density lipoprotein, systolic blood pressure, and diastolic blood pressure

Model 2 (reduced model) = adjusted for age, sex (male/female), ethnicity (Malay/Chinese/Indian/other races), family history of T2DM (no/yes), total cholesterol, triglycerides, high-density lipoprotein, systolic blood pressure, and diastolic blood pressure

Model 3 (reduced model + interaction term) = adjusted for age, sex (male/female), ethnicity (Malay/Chinese/Indian/other races), family history of T2DM (no/yes), total cholesterol, triglycerides, high-density lipoprotein, systolic blood pressure, diastolic blood pressure, and interaction term between healthy lifestyle index with family history of T2DM

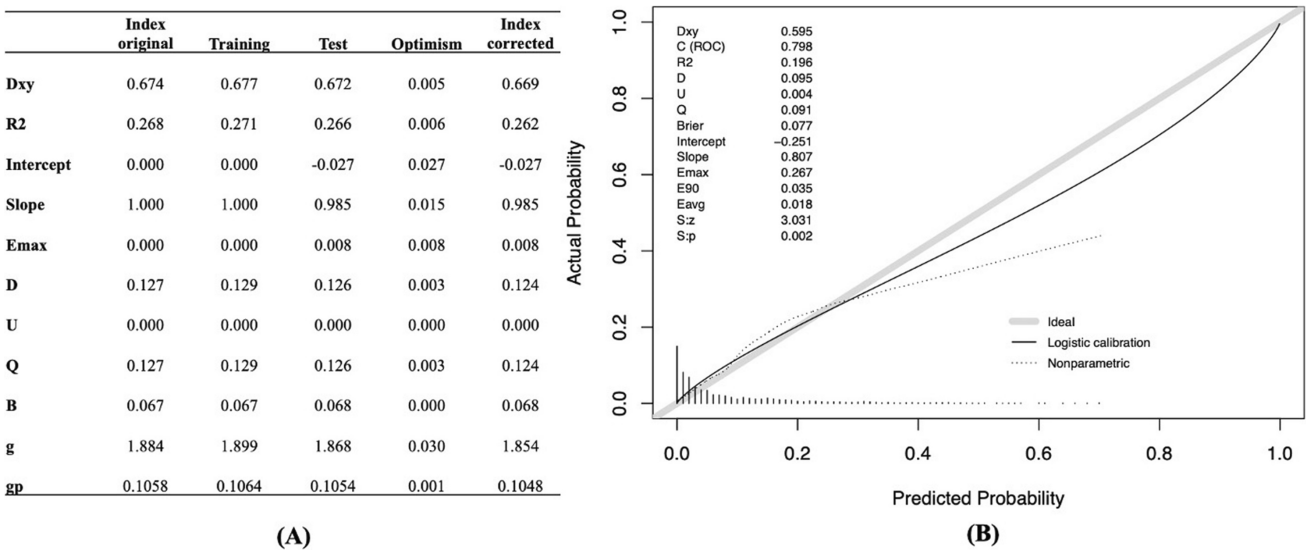
aOR adjusted odds ratio, *CI* confidence interval, R^2 percentage variant explained, *AIC* Akaike information criterion, *AUC* area under the curve, *LR test* likelihood ratio test

** $p<0.01$

increases skeletal muscle mass and glucose demands and regulates glucose production from the liver cells [32].

This study has several strengths. Firstly, the HLI was developed specifically for teachers and possibly other academicians in higher educational institutions, like colleges and universities, who may have different lifestyles compared to other occupational groups. Secondly, the large number of teachers involved in the study were recruited from multiple representative states in Peninsular Malaysia, which increases the likelihood that this index can be generalised to all Malaysian teachers, as supported by the internal and external validations.

However, the study also had a few limitations. As this was a cross-sectional study, the causal relationship between the index and T2DM status among teachers could not be established. To establish the causal effect and validate the index, a follow-up study could be conducted from the CLUSTer cohort to test this HLI in response to the incidence of



	Index original	Training	Test	Optimism	Index corrected
Dxy	0.674	0.677	0.672	0.005	0.669
R2	0.268	0.271	0.266	0.006	0.262
Intercept	0.000	0.000	-0.027	0.027	-0.027
Slope	1.000	1.000	0.985	0.015	0.985
Emax	0.000	0.000	0.008	0.008	0.008
D	0.127	0.129	0.126	0.003	0.124
U	0.000	0.000	0.000	0.000	0.000
Q	0.127	0.129	0.126	0.003	0.124
B	0.067	0.067	0.068	0.000	0.068
g	1.884	1.899	1.868	0.030	1.854
gp	0.1058	0.1064	0.1054	0.001	0.1048

Fig. 1 Internal (training data, $n=8559$) (**A**) and external (validation data, $n=2853$) (**B**) validations of Model 3 on T2DM among teachers

Table 5 Logistic regression models on healthy lifestyle index (T1–T3) associated with other health status (IFG, hypertension, and dyslipidaemia) among teachers in CLUSTer cohort ($n = 11412$)

Health status	T1 (Ref) aOR (95% CI)	T2 aOR (95% CI)	T3 aOR (95% CI)	*p* for trend
Impaired fasting glucose (> 5.6 mmol/L)				
Model 1[a]	1.00	0.61 (0.50, 0.74)	0.30 (0.23, 0.38)	< 0.001
Model 2[b]	1.00	0.72 (0.60, 0.88)	0.44 (0.34, 0.57)	< 0.001
Hypertension (yes)				
Model 1[a]	1.00	0.59 (0.50, 0.70)	0.36 (0.29, 0.44)	< 0.001
Model 2[c]	1.00	0.78 (0.65, 0.93)	0.65 (0.52, 0.82)	< 0.001
Dyslipidaemia (yes)				
Model 1[a]	1.00	0.61 (0.53, 0.70)	0.51 (0.43, 0.60)	< 0.001
Model 2[d]	1.00	0.63 (0.55, 0.74)	0.58 (0.49, 0.69)	< 0.001

Ref reference, *aOR* adjusted odds ratio

[a]Adjusted for age, sex, and ethnicity

[b]Adjusted for age, sex, ethnicity, education level, marital status, family history of T2DM, systolic/diastolic blood pressure, total cholesterol, triglycerides, low-density lipoprotein, and high-density lipoprotein

[c,d]Adjusted for age, sex, ethnicity, education level, marital status, systolic/diastolic blood pressure, total cholesterol, triglycerides, low-density lipoprotein, and high-density lipoprotein

T2DM and other metabolic disorders such as hypertension and dyslipidaemia. Furthermore, there may be social desirability and recall biases as the information on lifestyle behaviours were self-reported. To minimise the recall bias, the researchers recruited enumerators to communicate with those teachers while answering the questions. Unfortunately, we are unable to rule out the social desirability bias; thus, the inputs from the respondents on high-risk behaviours (e.g. smoking and alcohol consumption) should be interpreted with caution.

Conclusion

The HLI was validated both internally and externally using a split dataset index construction. This index includes lifestyle-related measures including waist circumference, physical activity, sleep durations, and acute mental health (anxiety, depression, and stress). Notably, a higher HLI score is associated with a decreased association with T2DM. This finding highlights the importance of integrating multidimensional lifestyle modifications rather than focusing on a single aspect. This study contributes to the advancement of our understanding of complex lifestyle interactions and their significance in T2DM.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13410-024-01374-w.

Acknowledgments The authors thank all enumerators who assist in collecting data during the baseline CLUSTer cohort recruitment.

Author contribution YHN, FMM, NNH, and AB conceptualise the study. Material preparation and data collection were done by YHN. Statistical analyses were conducted by YHN. The first draft of the manuscript was written by YHN and was critically revised by FMM, NNH, and AB. All authors reviewed and edited the final draft of the manuscript.

Funding This study was financially supported by the Ministry of Higher Education—Fundamental Research Grant Scheme (FRGS/1/2020/SKK06/UM/01/1) and Universiti Malaya Faculty of Medicine Postgraduate Scholarship Scheme (FOMPSS).

Data availability The datasets for the research presented in the publication are available from the corresponding author on reasonable request.

Declarations

Ethical approval and consent to participate This study obtained ethical approval (MREC ID: 950.1) from the University Malaya Medical Centre (UMMC) Medical Research Ethics Committee. Hard copies of written consent were obtained from all participants.

Competing interests The authors declare no competing interests.

References

1. Bodai BI, Nakata TE, Wong WT, Clark DR, Lawenda S, Tsou C, et al. Lifestyle medicine: a brief review of its dramatic impact on health and survival. Perm J. 2018;22:17–025.
2. Piché ME, Tchernof A, Després JP. Obesity phenotypes, diabetes, and cardiovascular diseases. Circ Res. 2020;126(11):1477–500.
3. Vodovotz Y, Barnard N, Hu FB, Jakicic J, Lianov L, Loveland D, et al. Prioritized research for the prevention, treatment, and reversal of chronic disease: recommendations from the lifestyle medicine research summit. Front Med Lausanne. 2020;7:585744.
4. Andersson U, Cuervo-Cazurra A, Nielsen BB. From the editors: explaining interaction effects within and across levels of analysis. J Int Bus Stud. 2014;45(9):1063–71.
5. Li Y, Ley SH, Tobias DK, Chiuve SE, VanderWeele TJ, Rich-Edwards JW, et al. Birth weight and later life adherence to unhealthy lifestyles in predicting type 2 diabetes: prospective cohort study. BMJ. 2015;351:h3672.

6. Long GH, Johansson I, Rolandsson O, Wennberg P, Fhärm E, Weinehall L, et al. Healthy behaviours and 10-year incidence of diabetes: a population cohort study. Prev Med. 2015;71:121–7.
7. Lv J, Yu C, Guo Y, Bian Z, Yang L, Chen Y, et al. Adherence to a healthy lifestyle and the risk of type 2 diabetes in Chinese adults. Int J Epidemiol. 2017;46(5):1410–20.
8. Mozaffarian D, Kamineni A, Carnethon M, Djoussé L, Mukamal KJ, Siscovick D. Lifestyle risk factors and new-onset diabetes mellitus in older adults: the cardiovascular health study. Arch Intern Med. 2009;169(8):798–807.
9. Reis JP, Loria CM, Sorlie PD, Park Y, Hollenbeck A, Schatzkin A. Lifestyle factors and risk for new-onset diabetes: a population-based cohort study. Ann Intern Med. 2011;155(5):292–9.
10. Pau K, Ahmad A, Tang H-Y, Jusoh A, Perveen A, Tat K. Mental health and wellbeing of secondary school teachers in Malaysia. Int J Learn Teach Educ Res. 2022;21:50–70.
11. Mehta HB, Mehta V, Girman CJ, Adhikari D, Johnson ML. Regression coefficient-based scoring system should be used to assign weights to the risk index. J Clin Epidemiol. 2016;79:22–8.
12. Lamu AN, Gamst-Klaussen T, Olsen JA. Preference weighting of health state values: what difference does it make, and why? Value in Health. 2017;20(3):451–7.
13. Moy FM, Hoe VC, Hairi NN, Buckley B, Wark PA, Koh D, et al. Cohort study on clustering of lifestyle risk factors and understanding its association with stress on health and well-being among school teachers in Malaysia (CLUSTer)–a study protocol. BMC Public Health. 2014;14:611–9.
14. Ng YH, Moy FM, Hairi NN, Bulgiba A. Prevalence of type 2 diabetes mellitus and impaired fasting glucose, and their associated lifestyle factors among teachers in the CLUSTer cohort. PeerJ. 2024;12:e16678.
15. Zhang Z, Zhang H, Khanal MK. Development of scoring system for risk stratification in clinical medicine: a step-by-step tutorial. Ann Transl Med. 2017;5(21):436–44.
16. Mazerolle MJ. Model selection and multimodel inference using the AICcmodavg package. R Vignette. 2020. https://cran.r-project.org/web/packages/AICcmodavg/vignettes/AICcmodavg.pdf. Accessed 15 May 2023.
17. Harrell Jr FE. R Package 'rms'. 2021. https://cran.r-project.org/web/packages/rms/rms.pdf. Accessed 20 May 2023.
18. Sjoberg DD, Curry M, Hannum M, Whiting K, Zabor E. gtsummary: presentation-ready data summary and analytic result tables. 2020. https://education.rstudio.com/blog/2020/07/gtsummary/. Accessed 20 May 2023.
19. Wickham H, Wickham MH. Package tidyverse. Easily Install and Load the 'Tidyverse. 2017. https://tidyverse.tidyverse.org. Accessed 30 May 2023.
20. Lüdecke D. ggeffects: tidy data frames of marginal effects from regression models. JOSS. 2018;3(26):772–6.
21. Eertink JJ, Heymans MW, Zwezerijnen GJC, Zijlstra JM, de Vet HCW, Boellaard R. External validation: a simulation study to compare cross-validation versus holdout or external testing to assess the performance of clinical prediction models using PET data from DLBCL patients. EJNMMI Res. 2022;12(1):58–65.
22. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080.
23. Wilke CT, Pickard AS, Walton SM, Moock J, Kohlmann T, Lee TA. Statistical implications of utility weighted and equally weighted HRQL measures: an empirical study. Health Econ. 2010;19(1):101–10.
24. Monica SJ, John S. Risk of obesity among female school teachers and its associated health problems. Curr Res Nutr Food Sci. 2018;6(2):404–11.
25. Ojeogwu EA, Husin TMABR. Factors influencing overweight among teachers in south-south Nigerian training college and development of a self assessment checklist for overweight. Sch Int J Anat Physiol. 2020;3(1):11–5.
26 Rocha SV, Cardoso JP, Santos CAd, Munaro HLR, Vasconcelos LRC, Petroski EL. Overweight/obesity in teachers: prevalence and associated factors. Rev Bras Cineantropom Desempenho Hum. 2015;17:450–9.
27. Papaetis GS, Papakyriakou P, Panagiotou TN. Central obesity, type 2 diabetes and insulin: exploring a pathway full of thorns. Arch Med Sci. 2015;11(3):463–82.
28. Zatterale F, Longo M, Naderi J, Raciti GA, Desiderio A, Miele C, et al. Chronic adipose tissue inflammation linking obesity to insulin resistance and type 2 diabetes. Front Physiol. 2020;10:1067–86.
29. Schmidt LI, Steenbock LM, Sieverding M. Improving sleep among teachers: an implementation-intention intervention. Int J Behav Med. 2023;30(1):49–61.
30. Musa NA, Moy FM, Wong LP. Prevalence and factors associated with poor sleep quality among secondary school teachers in a developing country. Ind Health. 2018;3(56):407–18.
31. Sharma K, Akre S, Chakole S, Wanjari MB. Stress-induced diabetes: a review. Cureus. 2022;14(9):e29142.
32. Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, et al. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement. Diabetes Care. 2010;33(12):e147–67.