



Article

# Towards a Parsimonious Pathway Model of Modifiable and Mediating Risk Factors Leading to Diabetes Risk

Yi-Ching Lynn Ho <sup>1,2,\*</sup>, Vivian Shu Yi Lee <sup>1,†</sup>, Moon-Ho Ringo Ho <sup>3</sup>, Gladis Jing Lin <sup>1,2</sup> and Julian Thumboo <sup>1,4,5</sup>

- <sup>1</sup> Office of Regional Health, Singapore Health Services, 167 Jalan Bukit Merah, Singapore 150167, Singapore; vivian.lee.s.y@singhealth.com.sg (V.S.Y.L.); gladis.lin.jing@singhealth.com.sg (G.J.L.); julian.thumboo@singhealth.com.sg (J.T.)
  - <sup>2</sup> Programme in Health Services and Systems Research, Duke-NUS Medical School, 8 College Rd., Singapore 169857, Singapore
  - <sup>3</sup> School of Social Sciences, Nanyang Technological University, 48 Nanyang Ave., Singapore 639818, Singapore; H0mh@ntu.edu.sg
  - <sup>4</sup> Department of Rheumatology and Immunology, Singapore General Hospital, Outram Rd., Singapore 169608, Singapore
  - <sup>5</sup> Medicine Academic Clinical Programme, Duke-NUS Medical School, 8 College Rd., Singapore 169857, Singapore
- \* Correspondence: gmshycl@nus.edu.sg  
† Co-first authors.

**Abstract:** Modifiable risk factors are of interest for chronic disease prevention. Few studies have assessed the system of modifiable and mediating pathways leading to diabetes mellitus. We aimed to develop a pathway model for Diabetes Risk with modifiable Lifestyle Risk factors as the start point and Physiological Load as the mediator. As there are no standardised risk thresholds for lifestyle behaviour, we derived a weighted composite for Lifestyle Risk. Physiological Load was based on an index using clinical thresholds. Sociodemographics are non-modifiable risk factors and were specified as covariates. We used structural equation modeling to test the model, first using 2014/2015 data from the Indonesian Family Life Survey. Next, we fitted a smaller model with longitudinal data (2007/2008 to 2014/2015), given limited earlier data. Both models showed the indirect effects of Lifestyle Risk on Diabetes Risk via the mediator of Physiological Load, whereas the direct effect was only supported in the cross-sectional analysis. Specifying Lifestyle Risk as an observable, composite variable incorporates the cumulative effect of risk behaviour and differentiates this study from previous studies assessing it as a latent construct. The parsimonious model groups the multifarious risk factors and illustrates modifiable pathways that could be applied in chronic disease prevention efforts.

**Keywords:** diabetes; risk factors; modifiable; mediators; lifestyle; physiological; pathways; mediation analysis; structural equation modeling; Indonesia



**Citation:** Ho, Y.-C.L.; Lee, V.S.Y.; Ho, M.-H.R.; Lin, G.J.; Thumboo, J. Towards a Parsimonious Pathway Model of Modifiable and Mediating Risk Factors Leading to Diabetes Risk. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10907. <https://doi.org/10.3390/ijerph182010907>

Academic Editor: Paul B. Tchounwou

Received: 8 September 2021

Accepted: 12 October 2021

Published: 17 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Diabetes mellitus alone contributed to 1.6 million deaths worldwide in 2017, a figure that is estimated to double by 2040 [1]. In terms of global burden, China (89.5 million), India (67.8 million), the USA (30.7 million), Indonesia (21.0 million), and Mexico (13.1 million) have been identified as five countries with the most diabetes cases, together with the most deaths and the highest disability-adjusted life years (DALYs) due to diabetes [2]. Among low- and middle-income countries, the disease is growing rapidly with Type 2 diabetes accounting for the majority of the cases. Three-quarters of diabetics are living in these countries [3,4].

Public health efforts in prevention and early control are needed to stem the rise in prevalence, given the disease burden from the higher risks of developing health complications, disability, and premature death [2]. Consequently, the healthcare and economic costs are also set to grow [1–3]. Identifying the risk factors and understanding their pathways

towards disease development is thus crucial in aiding preventative and early intervention efforts [3–5].

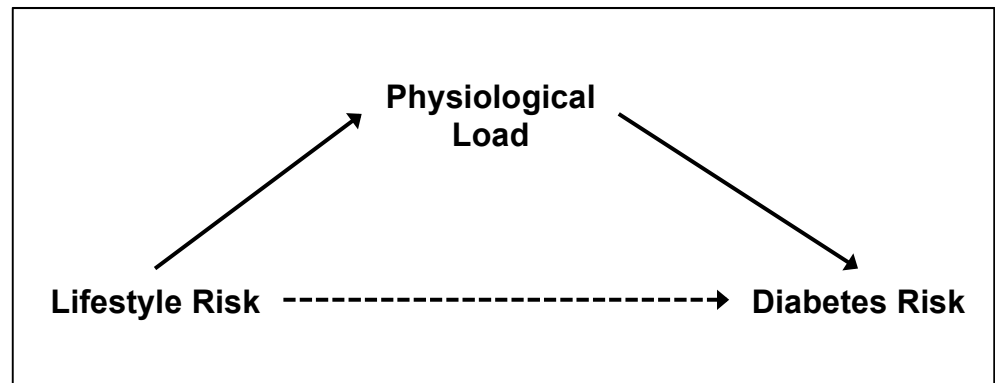
The development of Type 2 diabetes has been closely associated with multiple risk factors, from demographic factors e.g., age, sex, ethnicity, socioeconomic status, and educational level, to physiological factors and behaviour, e.g., hypertension, hyperlipidaemia, high body mass index (BMI), little sleep, pulse rate, c-reactive protein, unhealthy diets, physical inactivity, and tobacco use [4,6–15]. The risk factors may be classified according to whether they are modifiable, a good example being behaviour, or non-modifiable, such as demographic and genetic factors [11]. These risk factors can also be further differentiated in terms of their temporality: From the start point of a normal baseline, indicators of unfavourable metabolic and physiological changes, such as weight gain and increased blood pressure [11], can be regarded as downstream factors caused by preceding unhealthy lifestyle-related behaviour (e.g., a lack of physical activity, poor diet, smoking and lack of sleep) [16]. Obesity is known as the leading risk factor for Type 2 diabetes [17], but in order to prevent obesity itself, we need to go upstream to the risk factors for high BMI. Indeed physiological risk factors are often associated with lifestyle behavioural factors [18–20]. As a risk factor for diabetes, high BMI is associated with an unhealthy diet that includes high consumption of sodium and saturated fats [3,21,22].

As predictors of diabetes, these risk factors have often been investigated in terms of their direct relationship with diabetes outcomes [23] without distinguishing between the types of risk factors and potential dependence effects in a pathway system [16]; for example, how the risk factors may influence one other (both as independent and dependent variables) to impact diabetes risk. Furthermore, it is common to use categorical outcomes of diabetes status based on diagnostic thresholds, which neglects the continuum of diabetes risk and may also result in unbalanced samples of groups with and without diabetes [24,25]. This poses a problem for statistical analyses and requires compensating techniques (e.g., oversampling) [24].

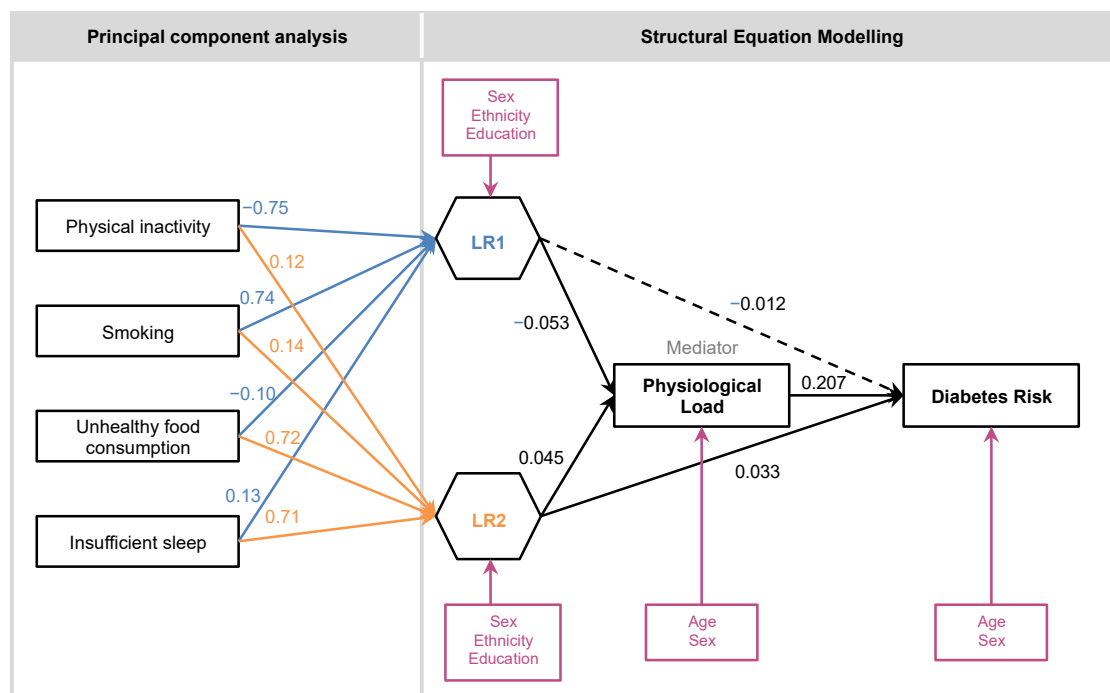
To address the challenges of multiple risk factors of different types and temporal effects, we used the approach to group risk factors in terms of their hypothesised effects, thus aiming to develop a parsimonious pathway model for diabetes risk, i.e., a simple model with minimal variables yet flexible enough to incorporate various classes of risk factors to explain how diabetes risk may be modulated by health behaviour changes. It is with this in mind that the modifiable lifestyle risk factors were set as the start point of the pathways (Figure 1). Lifestyle Risk variables form a natural grouping of independent variables concerned with health-related behaviour, such as the level of physical inactivity, smoking, the consumption frequency of unhealthy food, and insufficient sleep [4,6,8–10]. The second group of variables is termed Physiological Load, comprising clinical monitoring indicators associated with diabetes, such as body mass index (BMI), resting pulse rate (RPR), c-reactive protein (CRP), systolic (SBP), and diastolic (DBP) blood pressure [11,16,18–20]. Risk factors that cannot be modified, such as sociodemographics, will be specified as covariates of each main factor, rather than as main factors themselves. This strategy allows for the necessary accounting of the effects of the non-modifiable risk factors while focusing on the impact of the modifiable factors on diabetes risk. Lastly, we used the continuous variable of HbA1c as a marker of diabetes risk, which has the advantages of avoiding unbalanced diabetes outcome categories, gaining statistical power, and allowing a more informative interpretation of the relationships.

We aimed to examine the direct and indirect effects of two groups of modifiable risk factors on Diabetes Risk, namely “Lifestyle Risk” and “Physiological Load”. We hypothesised that Lifestyle Risk would have a positive but indirect effect on Diabetes Risk through the mediator of Physiological Load (Figure 1), as Lifestyle Risk is postulated to be upstream of Physiological Load. Using data from the Indonesian Family Life Survey (IFLS), which had five survey waves from 1993 to 2015, we performed structural equation modeling (SEM) on two main models (Figures 2 and 3). As the hypothesis concerns temporal effects, longitudinal data would be important. However, as only the latest survey

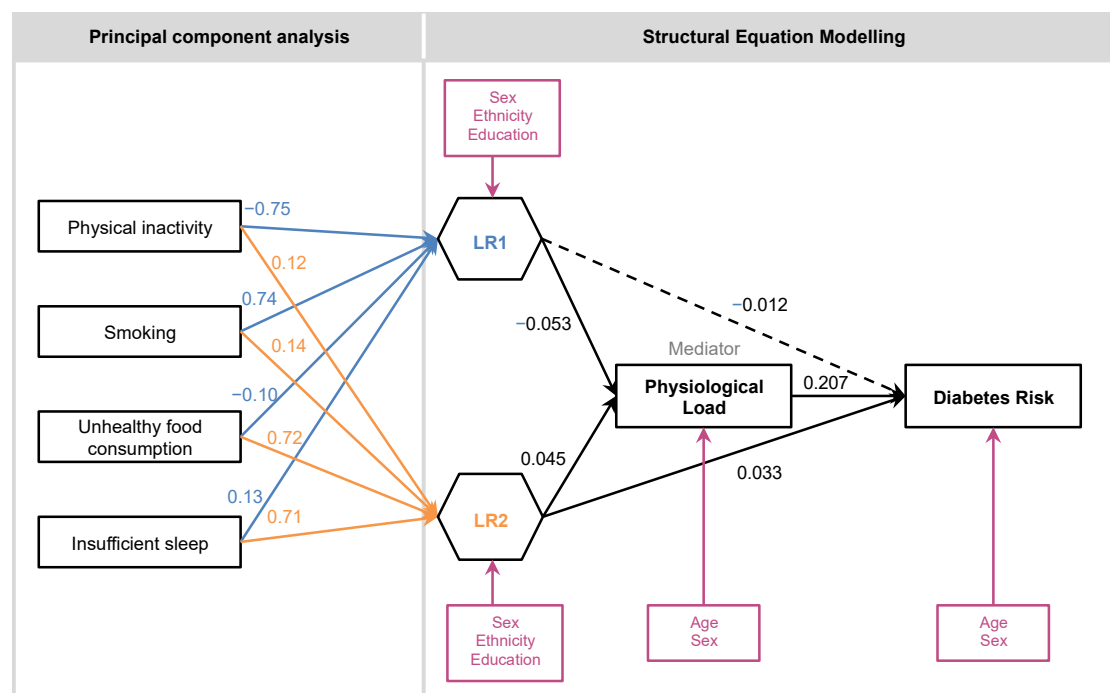
wave (IFLS5) had all the indicators of interest, we first tested a comprehensive model (Model A) with cross-sectional data from IFLS5. This model includes all four Lifestyle Risk factors of interest. We then specified a second model (Model B) to allow the testing of longitudinal data (IFLS4 and IFLS5), but with only two of the Lifestyle Risk factors, as sleep and food intake were not collected in the earlier wave. We account for sociodemographic risk factors as covariates in these analyses.



**Figure 1.** Illustration of the hypothesized pathways between Lifestyle Risk, Physiological Load and Diabetes Risk. Lifestyle Risk has a positive but indirect effect on Diabetes Risk through the mediator of Physiological Load. Note: Straight line with one arrowhead denotes positive and direct effect; dashed line with one arrowhead denotes positive but indirect effect.



**Figure 2.** Model A: Factor loadings of the structural equation model (SEM) and the two principal components of Lifestyle Risk (LR1, LR2). Physiological Load was demonstrated to be a mediator of Lifestyle Risk to Diabetes Risk. There was also a significant direct path from Lifestyle Risk (LR2) to Diabetes Risk. Only statistically significant sociodemographic covariates (pink boxes) were included in the figure. Note: A straight line with one arrowhead denotes a significant effect; a dashed line with one arrowhead denotes an insignificant effect. Boxes indicate observed variables; hexagons indicate composite variables. Principal Component Analysis (PCA) loadings were rounded to 2 decimal places; Pathway loadings in SEM model were rounded to 3 decimal places.



**Figure 3.** Model B: Factor loadings of the structural equation model (SEM) and the two principal components of Lifestyle Risk (LR1, LR2). Physiological Load was demonstrated to be a mediator of Lifestyle Risk (LR2) to Diabetes Risk. Only statistically significant sociodemographic covariates (pink boxes) were included in the figure. Note: A straight line with one arrowhead denotes a significant effect; a dashed line with one arrowhead denotes an insignificant effect. Boxes indicate observed variables; hexagons indicate composite variables. Principal Component Analysis (PCA) loadings were rounded to 2 decimal places; Pathway loadings in SEM model were rounded to 3 decimal places.

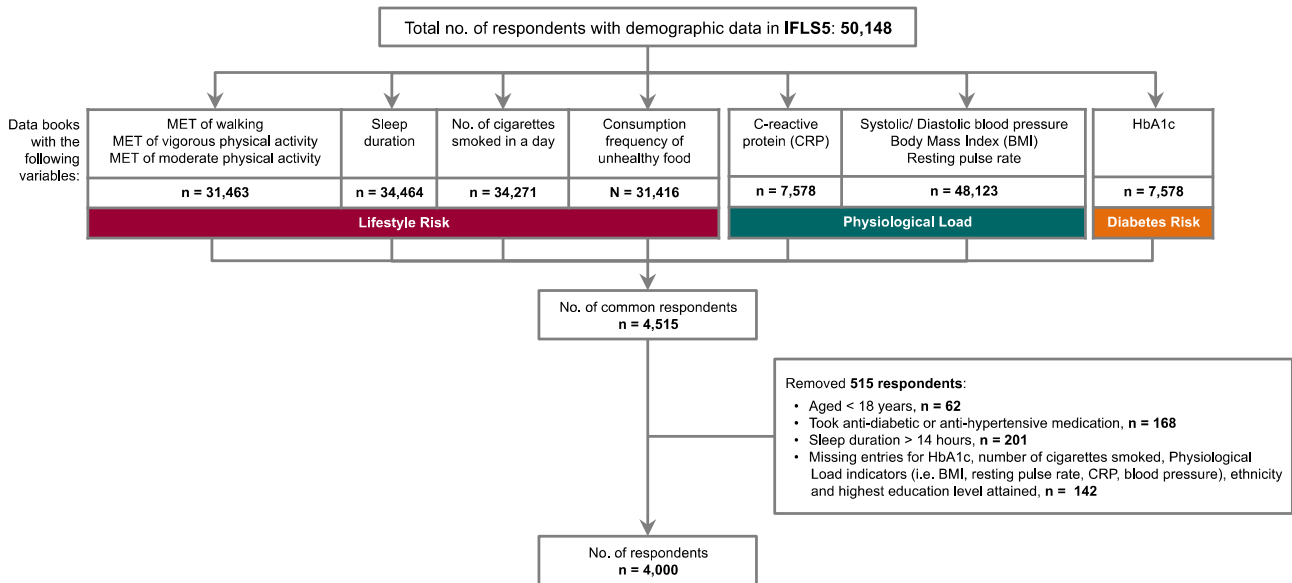
## 2. Methods

### 2.1. Study Design and Respondents

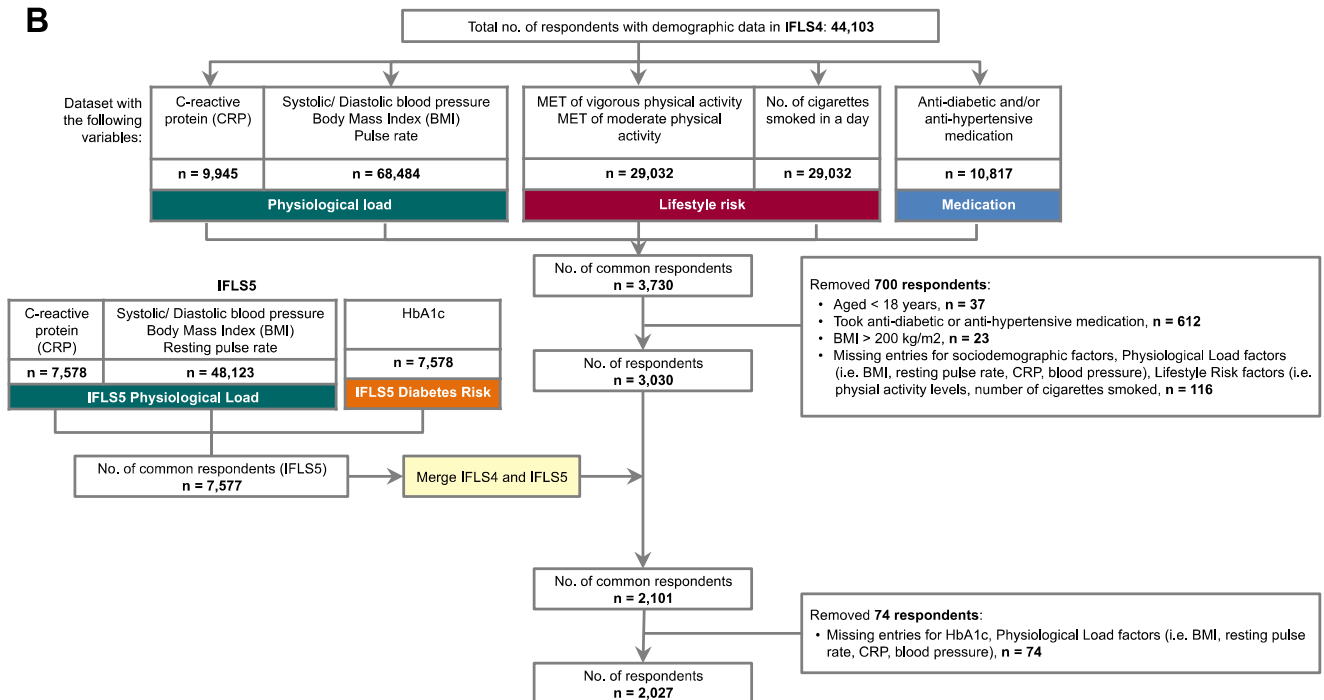
In this retrospective cohort study, we used publicly available data from the Indonesian Family Life Survey (IFLS), which has been organised by the RAND Corporation in collaboration with Lembaga Demografi, the University of Indonesia, Center for Population and Policy Studies, the University of Gadjah Mada, the University of California, Los Angeles, and SurveyMETER. A stratified random sampling scheme on the provinces was adopted and the resulting sample included 13 of 27 provinces in Indonesia, representative of 83% of the population. Households that were randomly selected and participated in the first survey in 1993 were followed up in subsequent waves. Sampling and survey methods have been discussed in detail elsewhere [26].

We analysed data from Wave 4 (“IFLS4”, conducted from 2007–2008) and Wave 5 (“IFLS5”, conducted from 2014–2015), given that the indicators of interest were found in these two waves. Both waves had individual-level data on anthropometric measurements, morbidity indicators, healthcare utilization, health behaviours, employment, and household expenditure [26], though specific data on food frequency, sleep duration, and glycosylated haemoglobin (HbA1c) were not available in IFLS4. Figure 4 illustrates the data sampling for the two main models. To reduce the possible confounding effect of medication in this cross-sectional study, respondents who self-reported to be on anti-diabetic or anti-hypertensive medication were excluded. The final study sample for Model A included 4000 respondents aged 18 and above (12.9% of total 31,102 adult respondents in IFLS5). For Model B, it included 2027 respondents aged 18 and above (11.7% of the total 17,396 adult respondents present in IFLS4 and 5).

**A**



**B**



**Figure 4.** Flow charts of the study sampling for Model A (A) and Model B (B). Note: The number of respondents removed based on specific exclusion criteria do not add up to the total number removed, due to respondents who met multiple exclusion criteria.

## 2.2. Sociodemographic Factors

Sociodemographic variables are important determinants in chronic disease and should be accounted for [27,28]. Age, sex, ethnicity, and highest educational level were specified as covariates separately for each of the main factors (Lifestyle Risk, Physiological Load, and Diabetes Risk) to account for confounding effects at any level.

Ethnicity was grouped into 3 categories: Javanese; Sundanese; and Others. Besides the Javanese and Sundanese, the other ethnicities each made up less than 6% of the sample population and thus were combined, similar to other studies [29,30].

The highest level of education attained was grouped into four categories following previous studies [9,29,31]: no education (unschooled); elementary (grade school, kindergarten, Islamic elementary school, adult education A); high school (general junior high, vocational junior high, Islamic Junior high, adult education B, general senior high, vocational senior high, Islamic senior high school, adult education C, Pesantren boarding school); and college/university (open university, Diploma, University).

### 2.3. Estimation of Lifestyle Risk

Four unhealthy lifestyle behaviours (physical inactivity, smoking, consumption frequency of unhealthy food, and insufficient sleep) were used as indicators of Lifestyle Risk (though in Model B, only physical inactivity and smoking were available for testing).

The level of physical activity in the last 7 days was assessed through the International Physical Activity Questionnaire (IPAQ). If respondents reported engaging in walking or any moderate or vigorous physical activities, the total duration of activities was converted to Metabolic Equivalent of Task (MET)-hours per week using the IPAQ recommended formula [32]. The amount of physical activity was reverse-coded and used as a measure of “physical inactivity”, which would contribute to Lifestyle Risk. The negatively phrased physical inactivity was used to maintain a consistent interpretation of Lifestyle Risk indicators.

The number of cigarettes smoked was used as an indicator for smoking intensity. This indicator was collected by questionnaires in IFLS4 and 5, which included questions on smoking status, whether the respondent has ever smoked self-rolled cigarettes, manufactured cigarettes, or cigars, followed by questions on smoking intensity, how many cigarettes/cigars smoked currently or before quitting. Respondents who reported “No” for questions on smoking status would have a smoking intensity of zero.

Unhealthy foods are highly processed food items that are higher in saturated fat, sugar, and sodium, such as instant noodles, sweet snacks, fried snacks, fast food, or soft drinks [33]. High consumption of such foods has been shown to exacerbate the risk of chronic diseases [34]. Therefore, for the purpose of this study, the consumption frequency of unhealthy foods in the past 7 days was used as an indicator of Lifestyle Risk. Unhealthy foods refer to instant noodles, sweet snacks, fried snacks, fast food, or soft drinks. (Food Frequency Section of IFLS5). The total score was calculated by summing the consumption frequencies of individual unhealthy food items in the past week with saturation at a score of seven, as the focus was on the consumption frequency of any type of unhealthy food within a week, therefore, the range of scores was 0–7.

Sleep duration was determined by two self-reported questions in IFLS5, the time the respondent slept the night before and the time the respondent woke up the day before the survey. The time respondent slept was subtracted by the time respondent woke up to compute sleep duration. Sleep duration was then reverse-coded and used as a measure of insufficient sleep, i.e., shorter sleep duration corresponded to more insufficient sleep. The negatively phrased insufficient sleep was used to maintain a consistent interpretation of Lifestyle Risk indicators.

To specify an appropriate structural equation model, we first assessed the suitability of Lifestyle Risk to be a latent construct within the model. A statistical check for an underlying latent construct requires its effect indicators to be positively correlated with one another [35]. However, we found weak correlation results for the indicators of Lifestyle Risk (Spearman’s rank  $r < |0.4|$ ), showing that it would not be appropriate to create a Lifestyle Risk latent variable. See Table S1. The Lifestyle Risk indicators would be better identified as a composite variable because it does not violate this conceptual check [35]. Furthermore, the composite or cumulative effect of these high-risk lifestyle behaviours is associated with higher health risks [36]. As a composite score of high-risk lifestyle behaviours, this means



that Lifestyle Risk can be estimated, rather than it being an underlying, unobservable construct.

As there are no standardized thresholds for risk levels in lifestyle behaviours, we used principal component analysis (PCA) to derive the weighted composites of the Lifestyle Risk indicators. The Lifestyle Risk indicators were included as continuous variables in the PCA models, and the principal components (i.e., weighted composites) were orthogonally transformed using Varimax rotation and underwent Kaiser normalization to achieve a structure with independent components for greater interpretability. Principal components with eigenvalue > 1, scree test, and parallel analysis were considered in determining the number of components to retain [37–39].

#### 2.4. Estimation of Physiological Load

In this study, Physiological Load is conceptualised as an estimate of the cumulative physiological burden or stress on the body system, similar to the concept of allostatic load or “cumulative biological risk” [40]. Allostatic load describes the physiological consequences of the body’s attempts to adapt to chronic stressors, which may result in dysregulation spreading among multiple body systems, potentially combining to increase disease risk [40,41]. Summary measures are typically used to characterise allostatic load across the cardiovascular, metabolic, immune, nervous, and hormonal systems [40,42]. As this study does not focus on the nervous or hormonal systems, we have used the term, “Physiological Load”, based on five routine clinical monitoring indicators (body mass index, resting pulse rate, c-reactive protein, systolic and diastolic blood pressure) from IFLS4 and 5. Each of these Physiological Load indicators has been shown to be associated with diabetes [12,43–48].

The body mass index (BMI) was computed by taking the weight (kilograms) divided by the height (metres) squared. Resting pulse rate (RPR) and blood pressure were averaged from three measurements on alternate arms while respondents were seated upright. CRP concentrations were derived from finger prick dried blood samples and measured by a high-sensitivity CRP enzyme-linked immunosorbent assay (ELISA) [49]. CRP plasma equivalent values were used.

Prior to creating a summary measure of Physiological Load, we performed a conceptual check, similar to the one done for Lifestyle Risk, in order to ascertain suitability for specification as a latent construct. Except for SBP and DBP, which are closely related, low correlations, Spearman rank  $r < |0.4|$ , were found between the rest of the Physiological Load indicators, indicating that a latent variable would not be suitable [35] (See Table S1). We then proceeded to create a summary index of Physiological Load by using the method similar to allostatic load summary measurements: for each biomarker, a score of one is given for values beyond a clinical threshold reflecting high risk, with a score of zero given if otherwise [50]. These were then added up to form a non-weighted summary index (range of 0 to 5) of the Physiological Load for each respondent. High risk was defined as: BMI of  $\geq 25 \text{ kg/m}^2$  [51], RPR of  $\geq 90 \text{ bpm}$  [52], SBP  $\geq 140 \text{ mmHg}$  and DBP  $\geq 90 \text{ mmHg}$  [53]. CRP  $\geq 30 \text{ mg/L}$  (3 mg/dL) [54].

#### 2.5. Estimation of Diabetes Risk

As the primary outcome variable, the risk of diabetes was estimated with the level of glycosylated haemoglobin (HbA1c), a surrogate biomarker of average glycemic control over the previous three months. HbA1c was assayed using dried blood samples from the IFLS respondents using a validated protocol [49,55] and was available only in IFLS5.

#### 2.6. Structural Equation Modelling

Categorical variables (e.g., sex, ethnicity) were summarised through counts and percentages, while all the other variables were continuous and were summarised through minimum, maximum, median, and interquartile ranges (IQR).

Mediation analysis via SEM [56] was used to test the hypothesis that Lifestyle Risk has a positive but indirect effect on Diabetes Risk through Physiological Load mediators. In the models, weighted composites representing Lifestyle Risk were considered independent (exogenous) variables, Physiological Load was a mediator and HbA1c was the marker of the outcome variable of Diabetes Risk. Sociodemographic variables were controlled for as covariates of Lifestyle Risk, Physiological Load, and HbA1c [35].

Two main models were tested. Model A was a comprehensive model, encompassing all indicators of interest (Figure 2). We first tested the comprehensive model with cross-sectional data from IFLS5, as only IFLS5 had all indicators of interest. We then created a smaller model (Model B, Figure 3) that allowed us to do a longitudinal analysis with IFLS4 data. This model was similar to Model A, except for the Lifestyle Risk indicators of the frequency of unhealthy food consumption and sleep duration, both of which were unavailable in IFLS4. Physiological Load was modelled as the mean of Physiological Load from IFLS4 and 5, in order to obtain a value that would be temporally midway between the surveys of IFLS4 and 5. The continuous outcome variable of Diabetes Risk was represented by HbA1c values.

We used the maximum likelihood procedure that provides correction to estimates and standard errors, and a mean- and variance-adjusted chi-square test statistic robust to non-normality [57]. Standardized estimates with 95% confidence intervals (CI) and *p*-values were reported. Model fit was assessed with the root mean squared error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean squared residual (SRMR). The following values indicate acceptable model fits: RMSEA < 0.08, CFI and TLI > 0.8, and SRMR < 0.08 [58].

Data preparation and descriptive analyses were performed using STATA (version 14.0) software [59]. PCA and parallel analysis were performed on Rstudio (R version 4.0.2) software, using the *psych* (version 1.9.12) and *paran* (version 1.5.2) packages [60–62]. All mediation analyses were performed using Mplus 8 (version 1.5) software [63]. In all statistical analyses, *p* < 0.05 and 95% CI that do not include zero were considered statistically significant.

### 2.7. Post-Hoc Analyses

To investigate the possibility that the involvement of physical labour in respondents' primary jobs confounded physical inactivity levels of Lifestyle Risk composite variables, physical labour was assessed by responses to two self-reported questions, whether respondents' primary jobs involved physical effort or heavy lifting all the time, almost all the time, or most of the time. In this model (Model A2), involvement of physical labour in primary jobs was classified as "yes" or "no" based on their responses and adjusted for as a covariate of Lifestyle Risk.

To mitigate the weakness that the analysis for Model A used cross-sectional data, we compared the model fit results against an alternative model with reversed directional relationships (Model A3). The Bayesian information criterion (BIC) values were assessed, with the smaller BIC value indicating a better model fit.

## 3. Results

### 3.1. Sample Characteristics and Comparisons with the National Population

Table 1 presents a descriptive summary of demographic characteristics of the samples used in the analysis of Model A and Model B. It also includes summary statistics of the Lifestyle Risk indicators, Physiological Load indicators, and HbA1c levels.

The median age in the IFLS5 sample was 40 years (range 18–102 years) with 53.0% males. 59.5% of the respondents attained highest education level of high school and above (High school = 46.1%; College/university = 13.4%). This is similar to the Indonesian population, where 50.3% of the national population were males, the two largest ethnic groups were also Javanese and Sundanese, and 54.9% of the population attained an education level of at least high school and above [30,64–66].



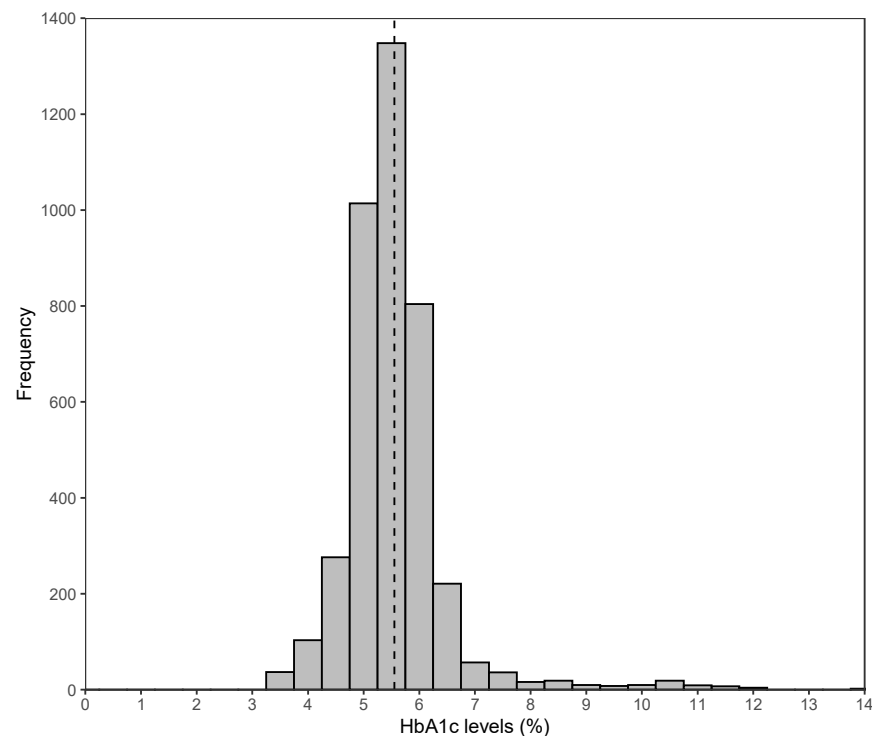
**Table 1.** Sociodemographic, lifestyle and physiological characteristics.

Characteristics	IFLS5 Sample		IFLS4 Sample			
	n (%) or Median (IQR)	Proportion of Sample beyond Clinical Threshold	n (%) or Median (IQR)	Proportion of Sample beyond Clinical Threshold		
<b>Age</b> , median (IQR)	40.00	(31.0, 56.0)	-	55.00	(49.0, 62.0)	-
<b>Sex</b> , n (%)						
Male	2078	(53.0%)	-	879	(43.4%)	-
Female	1854	(47.0%)	-	1148	(56.6%)	-
<b>Ethnicity</b> , n (%)						
Javanese	1853	(47.6%)	-	982	(48.4%)	-
Sundanese	425	(10.8%)	-	228	(11.2%)	-
Others	1654	(41.6%)	-	817	(40.3%)	-
<b>Highest Education level</b> , n (%)						
No education	235	(5.93%)	-	360	(17.8%)	-
Elementary	1360	(34.6%)	-	1139	(56.2%)	-
High school	1810	(46.1%)	-	424	(20.9%)	-
College/University	527	(13.4%)	-	104	(5.13%)	-
<b>Lifestyle Risk indicators</b> , median (IQR)						
Physical activity levels (MET h/w)	31.54	(8.25, 86.7)	-	48.96	(17.0, 113.0)	-
Smoking (cigarettes smoked in a day)	0.00	(0.00, 10.0)	-	0.00	(0.00, 6.00)	-
Consumption frequency of unhealthy food (days in a week)	5.00	(2.00, 7.00)	-	-	-	-
Sleep duration (h)	7.00	(5.67, 8.00)	-	-	-	-
<b>Physiological Load indicators</b> , median (IQR)						
BMI (kg/m <sup>2</sup> )	22.56	(19.9, 25.9)	30.7%	22.12	(19.7, 25.4)	26.9%
Resting pulse rate (bpm)	74.67	(68.0, 82.3)	9.58%	75.33	(69.3, 82.0)	10.2%
CRP (mg/L)	0.75	(0.263, 2.05)	0.375%	0.84	(0.350, 1.99)	0.00%
Systolic blood pressure (mmHg)	125.67	(116, 138)	22.9%	136.00	(124, 152)	34.9%
Diastolic blood pressure (mmHg)	78.00	(71.7, 85.7)	14.9%	80.17	(73.8, 87.5)	18.8%
<b>Diabetes Risk</b> , median (IQR)						
HbA1c level (% level)	5.45	(5.10, 5.85)	6.83%	-	-	-

Abbreviations: Indonesian Family Life Survey Wave 4/Wave 5 (IFLS4/IFLS5); Interquartile Range (IQR); Metabolic Equivalent of Task (MET); Body Mass Index (BMI); C-reactive protein (CRP); Glycosylated haemoglobin (HbA1c); Descriptive statistics are presented in n (%) for categorical variables and median (interquartile range) for continuous variables. All values with decimals were rounded off to 3 significant figures. Age statistics were calculated based on respondents' ages in year 2015 for IFLS5 and respondents' ages in year 2008 for IFLS4. Physical activity and sleep duration were later reverse-coded by multiplying with (−1) to maintain a consistent negative interpretation of all Lifestyle Risk indicators. Clinical thresholds for Physiological Load indicators were BMI ≥ 25 kg/m<sup>2</sup>; resting pulse rate ≥ 90 bpm; c-reactive protein (CRP) ≥ 30 mg/L; systolic blood pressure ≥ 140 mmHg; diastolic blood pressure ≥ 90 mmHg; HbA1c ≥ 6.50%.

The IFLS4 sample had some different characteristics. While the Javanese and Sundanese remained the two largest ethnic groups, the median age was 55 years (range 23–88 years) with 43.4% males, and the highest level of education for the majority was elementary school (56.2%). Differences in these factors between the IFLS4 sample and the national statistics can be attributed to the filtering of respondents that did not meet the inclusion criteria as outlined in the Methods (see Figure 4). Differences between the IFLS4 and IFLS5 samples could also be attributed to change in demographic progression across the 7–8 years between Waves 4 and 5, such as the increase in the proportion of those with higher education.

Figure 5 illustrates the distribution of the HbA1c level among the respondents. The median HbA1c level was 5.45% (range 3.50–14.0%). Applying the World Health Organisation (WHO) cut-off for indication of diabetes, i.e.,  $\text{HbA1c} \geq 6.50\%$  [67], 6.83% of our respondents were diabetic (see Table 1), which is very similar to the 6.90% national diabetes prevalence reported in 2013 Indonesia Basic Health Research survey (RISKESDAS) [68], indicating representativeness of our sample.



**Figure 5.** Distribution of HbA1c levels (%) among the IFLS5 sample (4000 respondents aged 18 and above). Note: Dashed line indicates the mean at 5.55%. The minimum was 3.50% and the maximum was 14.0%.

### 3.2. Model A

Model A was tested with IFLS5 data. Following PCA to determine the composite weights of Lifestyle Risk, we selected the first two components for Lifestyle Risk data upon inspection of the scree plot, applying the Kaiser-Guttman rule, and parallel analysis. The two components were termed LR1 and LR2. The variance explained by LR1 and LR2 was 28.27% and 26.63%, respectively. The total proportion of variance explained by the selected components was 54.9% for IFLS5 data.

Table 2 presents the retained components and their loadings. The strong loadings in LR1 were physical inactivity (−0.75) and smoking (0.74), while the strong loadings in LR2 were consumption frequency of unhealthy food (0.72) and insufficient sleep (0.71). All component loadings were positive, except for physical inactivity (−0.75) and the consumption frequency of unhealthy food (−0.10) in LR1.

**Table 2.** Component matrix of the 2-component principal component analysis (PCA).

<b>Lifestyle Risk Indicators (Model A)</b>	<b>LR1</b>	<b>LR2</b>
Physical inactivity	−0.75	0.12
Smoking	0.74	0.14
Consumption frequency of unhealthy food	−0.10	0.72
Insufficient sleep	0.13	0.71
<b>Lifestyle Risk Indicators (Model B)</b>	<b>LR1</b>	<b>LR2</b>
Physical inactivity	−0.75	0.66
Smoking	0.75	0.66

Abbreviations: Lifestyle Risk component 1 (LR1); Lifestyle Risk component 2 (LR2); Principal components were orthogonality transformed using varimax rotation. All values were rounded off to 2 decimal places.

Table 3 presents the direct and indirect relationships between Lifestyle Risk and Diabetes Risk (HbA1c). The model had good fits (RMSEA < 0.05; CFI and TLI > 0.95; SRMR < 0.08). Both components of Lifestyle Risk (LR1 and LR2) were shown to have effects on Diabetes Risk through the mediator of Physiological Load. LR1 did not have a direct effect, but there was an indirect effect on Diabetes Risk, through an inverse relationship with Physiological Load. LR2 had both positive direct and indirect effects on Diabetes Risk through Physiological Load. For the detailed results, including statistics for the sociodemographic covariates, refer to Table S2.

### 3.3. Model B

Model B was tested with longitudinal data from IFLS4-5. Following PCA to determine Lifestyle Risk, we also selected the first two components (LR1 and LR2), as they passed the Kaiser-Guttman rule, though parallel analysis recommended retaining only the first component. The positive loadings of the second component were also in line with our hypothesis. LR1 and LR2 explained 55.97% and 44.03% of variance, respectively. The loadings of LR1 were similar to that of Model A (Table 2), where the loadings for physical inactivity and smoking were −0.75 and 0.75, respectively. The loadings of LR2 were 0.66 for both physical inactivity and smoking.

The model had acceptable fits (RMSEA < 0.08; CFI > 0.95; SRMR < 0.08) (See Table 3). The first Lifestyle Risk component, LR1, did not have a significant effect on Physiological Load or Diabetes Risk. However, the second Lifestyle Risk component, LR2, had a positive indirect effect on Diabetes Risk through positive effects on Physiological Load. For detailed results, including statistics for the sociodemographic covariates, refer to Table S3.

### 3.4. Comparisons with Alternative Models

With reference to physical activity guidelines [69,70], at least 8.30 MET hours/week of physical activity is recommended. Compared to this, the medians for the IFLS4 and IFLS5 samples were higher by approximately 6 times (49.0 MET hours/week) and 4 times (31.5 MET hours/week) respectively (Table 1). Due to the surprisingly robust physical activity levels, we postulated that these levels could be confounded by jobs that involved physical labour. Indeed, individuals with physical labour in their primary jobs had significantly higher MET levels than individuals who did not,  $t(3998) = 13.1, p < 0.001$ , indicating that the jobs with manual labour contributed to this “lifestyle behaviour”. Therefore, the involvement of physical labour in respondents’ primary jobs was adjusted for as a covariate of Lifestyle Risk composite variables in a post-hoc analysis (Model A2). There was a positive and significant association between the involvement of physical labour and the Lifestyle Risk composite variable of LR1 (0.123,  $p < 0.001$ ), but not for LR2. Otherwise, the standardized estimates and significance of direct and indirect relationships in the model were similar to Model A, with model fit indices indicating good fit (RMSEA < 0.05; CFI > 0.95; TLI > 0.80; SRMR < 0.08) (see Table S4).

**Table 3.** Model A and Model B with standardized estimates, *p*-values and 95% confidence interval.

		Model A			Model B		
		Standardized Estimate	<i>p</i> -Value	95% CI	Standardized Estimate	<i>p</i> -Value	95% CI
Direct effects on Diabetes Risk	LR1	−0.012	0.461	(−0.045, 0.020)	0.007	0.817	(−0.054, 0.069)
	LR2	<b>0.033</b>	<0.050	(0.002, 0.064)	0.032	0.244	(−0.022, 0.086)
Indirect effects on Diabetes Risk via Physiological Load mediator	LR1	− <b>0.011</b>	<0.050	(−0.019, −0.004)	−0.003	0.633	(−0.013, 0.008)
	LR2	<b>0.009</b>	<0.050	(0.003, 0.016)	<b>0.010</b>	<0.050	(0.001, 0.020)
Model fit indices	RMSEA	0.014			0.069		
	CFI	1.000			0.988		
	TLI	0.989			0.601		
	SRMR	0.002			0.016		

Abbreviations: Lifestyle Risk component 1 (LR1); Lifestyle Risk component 2 (LR2); Root mean squared error of approximation (RMSEA); Comparative fit index (CFI); Tucker-Lewis index (TLI); Standardized root mean squared residual (SRMR); Significant estimates at  $p < 0.05$  are shown in bold. All values were rounded off to 3 decimal places. Model A (using IFLS5 data) adjusted for sociodemographic covariates: age, sex, ethnicity, and highest education level attained in IFLS5 (full model results in Table S2). Model B (using IFLS4 and 5 data) adjusted for sociodemographic covariates: age, sex, ethnicity, and highest education level attained in IFLS4 (full model results in Table S3). Statistically significant estimates (in bold) showed Physiological Load mediated the indirect effects of Lifestyle Risk (LR2) on Diabetes Risk. Model fit indices were within acceptable thresholds.

Given that Model A was based on cross-sectional data, we compared it to a model in which the pathways were reversed (Model A3). BIC values were lower for Model A (42,703.629) than Model A3 (42,709.770), indicating a better fit for Model A.

### 3.5. Sociodemographic Covariates

Results of the sociodemographic covariates can be found in Figures 2 and 3, and Tables S2 and S3. The significant covariates were mostly consistent between the models. In summary, increasing age was associated with increasing Physiological Load. Males were associated with poorer lifestyles, especially in terms of diet and sleep (LR2). Higher education was associated with higher Lifestyle Risk and Physiological Load.

## 4. Discussion

We have developed a general pathway model from the start point of modifiable lifestyle behaviour and have demonstrated how the behavioural components in Lifestyle Risk can affect Diabetes Risk via the mediating factor of Physiological Load. The overall results support our hypothesis that the effect of Lifestyle Risk on Diabetes Risk is likely to be indirect, and thus offers a stepwise perspective, whereby upstream and downstream modifiable factors could be modeled pathwise. Specifying Lifestyle Risk as an observable composite variable incorporates the cumulative effect of risk behaviour and differentiates this study from previous studies looking at it as a latent construct [16,71,72]. There was also the advantage of being able to assess causality using seven-year follow-up data, albeit only for Model B. From a disease prevention perspective, it helps to narrow our focus to an initial set of lifestyle risk factors, from which to monitor the progression of health risk towards more downstream physiological factors to disease.

The use of mediation analysis via SEM allowed us to simultaneously assess multiple pathways within a single model, in addition to accommodating a variable to be both independent and dependent (i.e., a mediator). This confers advantages over traditional regression, where multiple pathways in a single model need to be tested separately, resulting in potential problems with multiple comparisons [73]. To date, only a few studies have simultaneously analysed risk factors in diabetes in a multiple pathway system. Bardenheier et al., 2013 [16] performed the first study, using 10 variables with 27 hypothesised pathways in an SEM and found through best-fit iterations of the model that physical activity and poor diet were significant lifestyle factors (other lifestyle factors were not studied) that contributed to diabetes risk via large waist circumference, high blood pressure, triglycerides, and high-density lipoprotein (HDL). Subsequently several other studies applied similar models in their own population data [71,72] with varying results, but generally finding that physical activity and poor diet impact diabetes risk through separate mediators like BMI, blood pressure, HDL, and triglycerides. One difference between the earlier studies and this study is the use of latent constructs to model lifestyle behaviour in the earlier studies. As latent variables are meant to be unobservable constructs [35], a behaviour such as physical activity may not be suitable to be modeled as a latent construct, since it is actually observable. It would be necessary to evaluate the assumption for latent variables by testing if the effect indicators are correlated, a result that we did not find in our analysis (Table S1). We thus modelled Lifestyle Risk as a composite variable comprising linear, weighted combinations of risk from the uncorrelated lifestyle behaviours [35].

### 4.1. Components of Lifestyle Risk

In both models, two Lifestyle Risk components (LR1 and LR2) were derived from the data as complementary composites of unhealthy behaviour. PCA as a data-driven approach was used because there are currently no standardized thresholds for risk levels in lifestyle behaviours. The LR2 component encompassed a straightforward set of unhealthy lifestyle behaviours (all behaviour indicators had positive coefficients). As hypothesised, an increase in Lifestyle Risk goes on to increase Diabetes Risk through the mediator of Physiological Load. Physical activity and poor diet are both well established risk factors corresponding to the risk of chronic disease, in particular diabetes [74–76]. Smoking is

also another risk factor, showing a dose-response phenomenon with risk of diabetes [77], while short sleep duration is gaining attention as a factor involved in developing risk of diabetes through associations with BMI and blood pressure [78–80]. The path coefficients from LR2 to Physiological Load were relatively small (0.04–0.05), though the effect sizes were within the range of a study that used regressions to assess lifestyle variables and BMI with the same IFLS dataset [81]. While the indirect effect of LR2 on Diabetes Risk was clear in both models, the direct effect was uncertain, given that this effect was found in Model A but not in Model B. However the use of longitudinal data to test Model B lends weight to its results and aligns with studies showing the indirect effects of physical activity on the risk of developing diabetes through intermediate variables like BMI, but without detected direct effects [16,73,82].

The other lifestyle risk component, LR1, is an intriguing mix of behaviour found in both models: the composite weightings for physical inactivity and smoking are in opposite directions, i.e., being physically active while smoking contributes to LR1. This apparent contradiction in lifestyle may be explained by our post hoc analysis showing that a significant contributor to the high physical activity scores in the sample is involvement in physical labour as part of work, rather than recreational exercise. The effects of LR1 on Diabetes Risk were equivocal, as there was a significant indirect effect in Model A, but not in Model B. We lean towards the results of Model B, given the longitudinal testing. The factor of time may explain the different results, for example, it has been shown that the number of cigarettes smoked appears to have a negative correlation with the risk of high blood pressure, but when adjusted for life-course, the correlation turned positive [36]. In the longitudinal Model B, the negative effects of smoking may have over time counterbalanced the protective effects of physical activity, thus nullifying any overall effect of LR1 on Diabetes Risk.

#### 4.2. Physiological Load as a Mediator of Diabetes Risk

Physiological Load was found to be a significant mediator in both main models, with every unit increase in Physiological Load corresponding to an increase in HbA1c value by approximately 0.2 percentage points (based on the unstandardised estimates). (A unit increase in Physiological Load can be achieved by a cross in the clinical threshold for any of the five indicators, which are also used in routine clinical monitoring.) The pathway coefficients are within the ranges found in studies relating individual physiological markers to diabetes risk [16,71,72]. Furthermore, each of them have been associated with the effects of lifestyle behaviour [7,19,21,83–86], and thus may be considered intermediate markers of diabetes risk. The measure of Physiological Load is a subset of the allostatic load summary measure [40], as we focused on using typical clinical routine monitoring indicators in the metabolic and cardiovascular domains and did not include indicators of nervous and hormonal responses to chronic stress [40–42]. There is growing research that points to the utility of using grouped measures of physiological indicators to predict clinical risk, for example, the presence of adverse risk factors across multiple physiological systems strongly predicts morbidity and mortality [87]. A higher allostatic load has been found in patients with Type 2 diabetes [88,89] and correlates with higher glycated haemoglobin [89]. Physiological dysfunction can also spread across multiple physiological systems and combine to elevate disease risk [42,50]. Importantly, the grouped measures of physiological dysregulation appear to better predict morbidity and mortality risks as compared to individual risk indicators [50,90,91].

#### 4.3. Diabetes Risk

In this large sample, 6.83% of respondents had HbA1c values of 6.50% and above and can be classified diabetic, according to WHO criteria [67]. This statistic is very similar to the 6.90% found in a national health survey in Indonesia done around the same time [68], indicating the representativeness of our sample. We observed that 90% of survey respondents with HbA1c values of  $\geq 6.50\%$  did not report themselves as having received a diagnosis of diabetes (see Table S5). The incongruence of high HbA1c values with a low incidence of



awareness/reported diagnosis is a concern for public health efforts, notwithstanding the disease burden from known cases, being one of the top five countries with diabetes cases [2].

The availability of continuous HbA1c values in the IFLS data was well-suited for the performing of SEM and conferred an advantage over studies using categorical or binary diabetes outcomes, as unbalanced samples may require the employment of correction techniques, such as oversampling [24]. Besides yielding greater statistical power and precision, having a single continuous outcome reduces the number of parameters in the model, contributing to its parsimony [92].

#### 4.4. Applications and Limitations

As the purpose of the general model is to provide a simple pathway framework that groups distal (lifestyle) and proximal (physiological) factors, it can be applied to any analysis concerned with evaluating the relationship of lifestyle risks with chronic disease via the mediator of physiological risk. Specific composites of the distal or proximal factors can be determined by knowledge of the disease etiology, or driven by the dataset at hand, such as was done with the Lifestyle Risk indicators in this study.

Food frequency and sleep duration were lifestyle indicators important to our model, but as they were only found in IFLS Wave 5, the comprehensive model (Model A) was constrained to using the available cross-sectional data, which limited causal attribution. We assumed, similar to other studies (e.g., Bardenheier et al., 2013 [16]), that the reported lifestyle behaviour was habitual, built up over time, and thus preceded Physiological Load and Diabetes Risk. We mitigated the issue of cross-sectional data analysis in three ways. First, we excluded all the respondents who reported that they were taking medication for diabetes and hypertension, in order to avoid the confounding effect of medication on the Physiological Load indicators. Second, we generated a model with all pathways reversed to compare with our hypothesised model, in order to check that the hypothesised model was the better fit, which was shown to be the case. Third, we did a longitudinal analysis using a smaller model (Model B), without the Lifestyle Risk indicators of food and sleep. This longitudinal analysis supported the mediation effect found in the comprehensive model.

A potential confounding factor for the longitudinal analysis is the changing exposure to health policies and health promotion programs during the seven to eight years in between survey waves. The decentralized, district-level approach in Indonesian healthcare [93] precludes straightforward adjustments of effects to program exposure, since the IFLS respondents would not have had uniform exposure to programs, being from districts across Indonesia. The interpretation of the cross-sectional analysis for Model A remains unaffected.

The self-reports for lifestyle behaviour in the study pose a potential weakness, as the Lifestyle Risk indicators are subject to measurement errors and self-recall biases. As wearables and fitness trackers become more commonplace in the future, health behaviour data from such devices would form more objective sources for feeding into the model.

## 5. Conclusions

We have presented a general model illustrating modifiable pathways from Lifestyle Risk to Diabetes Risk via the mediating factor of Physiological Load and have tested it using large datasets from Wave 4 and Wave 5 of the Indonesian Family Life Survey. Non-modifiable sociodemographic covariates were accounted for in the model, while focusing on what is amenable for health outcomes. The model illustrates parsimonious and modifiable pathways that could be applied in public health efforts for diabetes or chronic disease prevention.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijerph182010907/s1>. Table S1: Spearman correlation matrix of Lifestyle Risk and Physiological Load indicators; Table S2: Model A—Full statistics of all direct and indirect relationships; Table S3: Model B—Full statistics of all direct and indirect relationships; Table S4: Model A2—To assess the inclusion of physical labour as a covariate; Table S5: Comparison of Diabetes prevalence determined by self-reports and HbA1c level.

**Author Contributions:** Conceptualization, Y.-C.L.H. and J.T.; Methodology, Y.-C.L.H., V.S.Y.L. and M.-H.R.H.; Validation, Y.-C.L.H., V.S.Y.L. and M.-H.R.H.; Formal analysis, Y.-C.L.H. and V.S.Y.L.; Data curation, V.S.Y.L.; Visualization, V.S.Y.L.; Supervision, J.T.; Writing—original draft preparation, Y.-C.L.H. and V.S.Y.L.; Writing—review and editing, Y.-C.L.H., V.S.Y.L., G.J.L., M.-H.R.H. and J.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** Y.-C.L.H. and V.S.Y.L. are funded by a grant, NMRC/CG/C027/2017.

**Institutional Review Board Statement:** The study was approved by the Centralised Institutional Review Board of Singapore Health Services (SingHealth) (reference number: 2020/2021).

**Informed Consent Statement:** The study did not collect data directly from patients and confined to secondary analysis of de-identified data made publicly available.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. This data can be found here: <https://www.rand.org/well-being/social-and-behavioral-policy/data/FLS/IFLS/access.html> (accessed on 25 August 2021).

**Acknowledgments:** We thank the RAND Corporation for providing access to the IFLS survey data.

**Conflicts of Interest:** The authors declare no conflict of interest.

### List of Abbreviations

95% CI	95% confidence intervals
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CFI	Comparative Fit Index
CRP	C-reactive protein
DALY	Disability-Adjusted Life Years
DBP	Diastolic blood pressure
HbA1c	Glycosylated haemoglobin
HDL	High-density lipoprotein
IFLS(4/5)	Indonesian Family Life Survey (Wave 4/Wave 5)
IPAQ	International Physical Activity Questionnaire
IQR	Interquartile range
LR1	Lifestyle Risk component 1
LR2	Lifestyle Risk component 2
MET	Metabolic Equivalent of Task
PCA	Principal component analysis
RMSEA	Root mean squared error of approximation
RPR	Resting pulse rate
SBP	Systolic blood pressure
SEM	Structural equation modelling
SRMR	Standardized root mean squared residual
TLI	Tucker-Lewis Index
WHO	World Health Organisation

### References

- World Health Organization. Projections of Mortality and Causes of Death, 2016 to 2060. Available online: [https://www.who.int/healthinfo/global\\_burden\\_disease/projections/en/](https://www.who.int/healthinfo/global_burden_disease/projections/en/) (accessed on 19 March 2021).
- Lin, X.; Xu, Y.; Pan, X.; Xu, J.; Ding, Y.; Sun, X.; Song, X.; Ren, Y.; Shan, P.F. Global, Regional, and National Burden and Trend of Diabetes in 195 Countries and Territories: An Analysis from 1990 to 2025. *Sci. Rep.* **2020**, *10*, 14790. [CrossRef] [PubMed]
- World Health Organization. Global Report on Diabetes. Available online: <https://www.who.int/publications/i/item/9789241565257> (accessed on 14 December 2020).
- International Diabetes Federation. *IDF Diabetes Atlas*, 7th ed.; International Diabetes Federation: Brussels, Belgium, 2015.
- Tuso, P. Prediabetes and Lifestyle Modification: Time to Prevent a Preventable Disease. *Perm. J.* **2014**, *18*, 88–93. [CrossRef]
- Monterrosa, A.E.; Haffner, S.M.; Stern, M.P.; Hazuda, H.P. Sex Difference in Lifestyle Factors Predictive of Diabetes in Mexican-Americans. *Diabetes Care* **1995**, *18*, 448–456. [CrossRef] [PubMed]
- Persson, P.G.; Carlsson, S.; Svanstrom, L.; Ostenson, C.G.; Efendic, S.; Grill, V. Cigarette Smoking, Oral Moist Snuff Use and Glucose Intolerance. *J. Intern. Med.* **2000**, *248*, 103–110. [CrossRef] [PubMed]

8. Quatromoni, P.A.; Copenhafer, D.L.; D'Agostino, R.B.; Millen, B.E. Dietary Patterns Predict the Development of Overweight in Women. *J. Am. Diet. Assoc.* **2002**, *102*, 1239–1246. [CrossRef]
9. Amalia, B.; Cadogan, S.L.; Prabandari, Y.S.; Filippidis, F.T. Socio-Demographic Inequalities in Cigarette Smoking in Indonesia, 2007 to 2014. *Prev. Med.* **2019**, *123*, 27–33. [CrossRef]
10. Christiani, Y.; Byles, J.E.; Tavener, M.; Dugdale, P. Gender Inequalities in Noncommunicable Disease Risk Factors among Indonesian Urban Population. *Asia-Pacific J. Public Health* **2015**, *28*, 134–145. [CrossRef]
11. World Health Organization. Global Status Report on Noncommunicable Diseases 2010. Available online: [http://www.who.int/nmh/publications/ncd\\_report2010/en/](http://www.who.int/nmh/publications/ncd_report2010/en/) (accessed on 31 August 2020).
12. de Rekeneire, N.; Peila, R.; Ding, J.; Colbert, L.H.; Visser, M.; Shorr, R.I.; Kritchevsky, S.B.; Kuller, L.H.; Strotmeyer, E.S.; Schwartz, A.V.; et al. Diabetes, Hyperglycemia, and Inflammation in Older Individuals: The Health, Aging and Body Composition Study. *Diabetes Care* **2006**, *29*, 1902–1908. [CrossRef]
13. Tripathy, J.P.; Thakur, J.S.; Jeet, G.; Jain, S. Structural Equation Modeling to Identify the Risk Factors of Diabetes in the Adult Population of North India. *Trop. Med. Health* **2018**, *46*, 23. [CrossRef]
14. Lidfeldt, J.; Nerbrand, C.; Samsioe, G.; Agardh, C.D. Women Living Alone Have an Increased Risk to Develop Diabetes, Which Is Explained Mainly by Lifestyle Factors. *Diabetes Care* **2005**, *28*, 2531–2536. [CrossRef]
15. Ng, R.; Sutradhar, R.; Yao, Z.; Wodchis, W.P.; Rosella, L.C. Smoking, Drinking, Diet and Physical Activity - Modifiable Lifestyle Risk Factors and Their Associations with Age to First Chronic Disease. *Int. J. Epidemiol.* **2020**, *49*, 113–130. [CrossRef]
16. Bardenheier, B.H.; Bullard, K.M.; Caspersen, C.J.; Cheng, Y.J.; Gregg, E.W.; Geiss, L.S. Novel Use of Structural Equation Models to Examine Factors Associated with Prediabetes among Adults Aged 50 Years and Older: National Health and Nutrition Examination Survey 2001–2006. *Diabetes Care* **2013**, *36*, 2655–2662. [CrossRef]
17. Barnes, A.S. The Epidemic of Obesity and Diabetes: Trends and Treatments. *Texas Hear. Inst. J.* **2011**, *38*, 142–144.
18. Forrester, S.N.; Leoutsakos, J.-M.; Gallo, J.J.; Thorpe, R.J., Jr.; Seeman, T.E. Association between Allostatic Load and Health Behaviours: A Latent Class Approach. *J. Epidemiol. Community Health* **2019**, *73*, 340–345. [CrossRef]
19. Gay, J.L.; Salinas, J.J.; Buchner, D.M.; Mirza, S.; Kohl, H.W.; Fisher-Hoch, S.P.; McCormick, J.B. Meeting Physical Activity Guidelines Is Associated with Lower Allostatic Load and Inflammation in Mexican Americans. *J. Immigr. Minor. Health* **2015**, *17*, 574–581. [CrossRef]
20. World Health Organization. Global Atlas on Cardiovascular Disease Prevention and Control. Available online: [http://www.who.int/cardiovascular\\_diseases/publications/atlas\\_cvd/en/](http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/) (accessed on 9 September 2020).
21. Paradis, A.M.; Godin, G.; Pérusse, L.; Vohl, M.C. Associations between Dietary Patterns and Obesity Phenotypes. *Int. J. Obes.* **2009**, *33*, 1419–1426. [CrossRef] [PubMed]
22. Schulze, M.B.; Hu, F.B. Primary Prevention of Diabetes: What Can Be Done and How Much Can Be Prevented? *Annu. Rev. Public Health* **2005**, *26*, 445–467. [CrossRef] [PubMed]
23. Fogelholm, M. Physical Activity, Fitness and Fatness: Relations to Mortality, Morbidity and Disease Risk Factors. A Systematic Review. *Obes. Rev.* **2010**, *11*, 202–221. [CrossRef] [PubMed]
24. Alghamdi, M.; Al-Mallah, M.; Keteyian, S.; Brawner, C.; Ehrman, J.; Sakr, S. Predicting Diabetes Mellitus Using SMOTE and Ensemble Machine Learning Approach: The Henry Ford Exercise Testing (FIT) Project. *PLoS ONE* **2017**, *12*, e0179805. [CrossRef] [PubMed]
25. Joshi, R.D.; Dhakal, C.K. Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7346. [CrossRef]
26. Strauss, J.; Witoelar, F.; Sikoki, B. *The Fifth Wave of the Indonesia Family Life Survey: Overview and Field Report*; RAND Labor and Population; WR-1143/1-NIA/NICHD; RAND Corporation: Laguna Hills, CA, USA, 2016. [CrossRef]
27. Heo, M.; Kim, R.S.; Wylie-Rosett, J.; Allison, D.B.; Heymsfield, S.B.; Faith, M.S. Inverse Association between Fruit and Vegetable Intake and BMI Even after Controlling for Demographic, Socioeconomic and Lifestyle Factors. *Obes. Facts* **2011**, *4*, 449–455. [CrossRef]
28. Di Milia, L.; Vandelanotte, C.; Duncan, M.J. The Association between Short Sleep and Obesity after Controlling for Demographic, Lifestyle, Work and Health Related Factors. *Sleep Med.* **2013**, *14*, 319–323. [CrossRef]
29. Finkelstein, E.A.; Chay, J.; Bajpai, S. The Economic Burden of Self-Reported and Undiagnosed Cardiovascular Diseases and Diabetes on Indonesian Households. *PLoS ONE* **2014**, *9*, 99572. [CrossRef]
30. Ananta, A.; Arifin, E.N.; Hasbullah, M.S.; Handayani, N.B.; Pramono, W. *Demography of Indonesia's Ethnicity*; Institute of Southeast Asian Studies: Singapore, 2015.
31. Hussain, M.A.; Al Mamun, A.; Reid, C.; Huxley, R.R. Prevalence, Awareness, Treatment and Control of Hypertension in Indonesian Adults Aged  $\geq 40$  Years: Findings from the Indonesia Family Life Survey (IFLS). *PLoS ONE* **2016**, *11*, e0160922. [CrossRef]
32. Blasio, A.D.; Donato, F.D.; Mazzocco, C. Automatic Report of the International Physical Activity Questionnaire. Available online: <https://sites.google.com/site/theipaq/> (accessed on 18 May 2020).
33. Baker, P.; Friel, S. Processed Foods and the Nutrition Transition: Evidence from Asia. *Obes. Rev.* **2014**, *15*, 564–577. [CrossRef] [PubMed]

34. Micha, R.; Shulkin, M.L.; Peñalvo, J.L.; Khatibzadeh, S.; Singh, G.M.; Rao, M.; Fahimi, S.; Powles, J.; Mozaffarian, D. Etiologic Effects and Optimal Intakes of Foods and Nutrients for Risk of Cardiovascular Diseases and Diabetes: Systematic Reviews and Meta-Analyses from the Nutrition and Chronic Diseases Expert Group (NutriCoDE). *PLoS ONE* **2017**, *12*, e0175149. [[CrossRef](#)] [[PubMed](#)]
35. Bollen, K.A.; Bauldry, S. Three Cs in Measurement Models: Causal Indicators, Composite Indicators, and Covariates. *Psychol. Methods* **2011**, *16*, 265–284. [[CrossRef](#)] [[PubMed](#)]
36. Sotos-Prieto, M.; Bhupathiraju, S.N.; Falcón, L.M.; Gao, X.; Tucker, K.L.; Mattei, J. A Healthy Lifestyle Score Is Associated with Cardiometabolic and Neuroendocrine Risk Factors among Puerto Rican Adults. *J. Nutr.* **2015**, *145*, 1531–1540. [[CrossRef](#)]
37. Kaiser, H.F. The Application of Electronic Computers to Factor Analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [[CrossRef](#)]
38. Cattell, R.B. The Scree Test For The Number Of Factors. *Multivariate Behav. Res.* **1966**, *1*, 245–276. [[CrossRef](#)]
39. Franklin, S.B.; Gibson, D.J.; Robertson, P.A.; Pohlmann, J.T.; Fralish, J.S. Parallel Analysis: A Method for Determining Significant Principal Components. *J. Veg. Sci.* **1995**, *6*, 99–106. [[CrossRef](#)]
40. McEwen, B.S.; Stellar, E. Stress and the Individual: Mechanisms Leading to Disease. *Arch. Intern. Med.* **1993**, *153*, 2093–2101. [[CrossRef](#)] [[PubMed](#)]
41. McEwen, B.S. Stress, Adaptation, and Disease: Allostasis and Allostatic Load. *Ann. N. Y. Acad. Sci.* **1998**, *840*, 33–44. [[CrossRef](#)]
42. McEwen, B.S. Protective and Damaging Effects of Stress Mediators. *N. Engl. J. Med.* **1998**, *338*, 171–179. [[CrossRef](#)]
43. Bays, H.E.; Chapman, R.H.; Grandy, S. The Relationship of Body Mass Index to Diabetes Mellitus, Hypertension and Dyslipidaemia: Comparison of Data from Two National Surveys. *Int. J. Clin. Pract.* **2007**, *61*, 737–747. [[CrossRef](#)]
44. Aune, D.; ó Hartaigh, B.; Vatten, L.J. Resting Heart Rate and the Risk of Type 2 Diabetes: A Systematic Review and Dose-Response Meta-Analysis of Cohort Studies. *Nutr. Metab. Cardiovasc. Dis.* **2015**, *25*, 526–534. [[CrossRef](#)]
45. Hayashi, T.; Tsumura, K.; Suematsu, C.; Endo, G.; Fujii, S.; Okada, K. High Normal Blood Pressure, Hypertension, and the Risk of Type 2 Diabetes in Japanese Men: The Osaka Health Survey. *Diabetes Care* **1999**, *22*, 1683–1687. [[CrossRef](#)]
46. Kahn, S.E.; Hull, R.L.; Utzschneider, K.M. Mechanisms Linking Obesity to Insulin Resistance and Type 2 Diabetes. *Nature* **2006**, *444*, 840–846. [[CrossRef](#)]
47. Wang, X.; Bao, W.; Liu, J.; Ouyang, Y.Y.; Wang, D.; Rong, S.; Xiao, X.; Shan, Z.L.; Zhang, Y.; Yao, P.; et al. Inflammatory Markers and Risk of Type 2 Diabetes: A Systematic Review and Meta-Analysis. *Diabetes Care* **2013**, *36*, 166–175. [[CrossRef](#)]
48. Gress, T.W.; Nieto, F.J.; Shahar, E.; Wofford, M.R.; Brancati, F.L. Hypertension and Antihypertensive Therapy as Risk Factors for Type 2 Diabetes Mellitus. *N. Engl. J. Med.* **2000**, *342*, 905–912. [[CrossRef](#)] [[PubMed](#)]
49. Herningtyas, E.H.; Hu, P.; Edenfield, M.; Strauss, J.; Crimmins, E.; Witoelar, F.; Yuan, Z.; Jung, K.K.; Thomas, D.; Sikoki, B. *Dried Blood Spot User's Guide for the Indonesia Family Life Survey, Wave 5*; RAND Labor and Population; WR-1143/6-NIA/NICHD; RAND Corporation: Laguna Hills, CA, USA, 2017.
50. Slopen, N.; Non, A.; Williams, D.R.; Roberts, A.L.; Albert, M.A. Childhood Adversity, Adult Neighborhood Context, and Cumulative Biological Risk for Chronic Diseases in Adulthood. *Psychosom. Med.* **2014**, *76*, 481–489. [[CrossRef](#)]
51. Nishida, C.; Barba, C.; Cavalli-Sforza, T.; Cutter, J.; Deurenberg, P.; Darnton-Hill, I.; Deurenberg-Yap, M.; Gill, T.; James, P.; Ko, G.; et al. Appropriate Body-Mass Index for Asian Populations and Its Implications for Policy and Intervention Strategies. *Lancet* **2004**, *363*, 157–163. [[CrossRef](#)]
52. Seccareccia, F.; Pannozzo, F.; Dima, F.; Minoprio, A.; Menditto, A.; Lo Noce, C.; Giampaoli, S. Heart Rate as a Predictor of Mortality: The MATISS Project. *Am. J. Public Health* **2001**, *91*, 1258–1263. [[CrossRef](#)]
53. Chobanian, A.V.; Bakris, G.L.; Black, H.R.; Cushman, W.C.; Green, L.A.; Izzo, J.L.; Jones, D.W.; Materson, B.J.; Oparil, S.; Wright, J.T.; et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: The JNC 7 Report. *J. Am. Med. Assoc.* **2003**, *289*, 2560–2572. [[CrossRef](#)] [[PubMed](#)]
54. Ridker, P.M. C-Reactive Protein: A Simple Test to Help Predict Risk of Heart Attack and Stroke. *Circulation* **2003**, *108*, e81–e85. [[CrossRef](#)] [[PubMed](#)]
55. Hu, P.; Edenfield, M.; Potter, A.; Kale, V.; Risbud, A.; Williams, S.; Lee, J.; Bloom, D.E.; Crimmins, E.; Seeman, T. Validation and Modification of Dried Blood Spot-Based Glycosylated Hemoglobin Assay for the Longitudinal Aging Study in India. *Am. J. Hum. Biol.* **2015**, *27*, 579–581. [[CrossRef](#)]
56. Gunzler, D.; Chen, T.; Wu, P.; Zhang, H. Introduction to Mediation Analysis with Structural Equation Modeling. *Shanghai Arch. Psychiatry* **2013**, *25*, 390–394. [[CrossRef](#)] [[PubMed](#)]
57. Byrne, B.M. Choosing Structural Equation Modeling Computer Software: Snapshots of LISREL, EQS, AMOS, and Mplus. In *R. Hoyle, Handbook of Structural Equation Modeling*; The Guilford Press: New York, NY, USA, 2012.
58. West, S.G.; Taylor, A.B.; Wu, W. Model Fit and Model Selection in Structural Equation Modeling. In *R. Hoyle, Handbook of Structural Equation Modeling*; The Guilford Press: New York, NY, USA, 2012.
59. StataCorp. *Stata Statistical Software: Release 14*; StataCorp LP: College Station, TX, USA, 2015.
60. RStudio Team. *RStudio: Integrated Development Environment for R*; RStudio, PBC: Boston, MA, USA, 2020.
61. Revelle, W. *Psych: Procedures for Psychological, Psychometric, and Personality Research*, Northwestern University, Evanston, Illinois, USA, Version 1.9.12. Comprehensive R Archive Network (CRAN), Northwestern University: Evanston, IL, USA, 2019.
62. Dinno, A. *Paran: Horn's Test of Principal Components/Factors*, Version 1.5.2. Comprehensive R Archive Network (CRAN), Portland State University: Portland, OR, USA, 2018.
63. Muthén, L.K.; Muthén, B.O. *Mplus User's Guide*, 8th ed.; Muthén & Muthén: Los Angeles, CA, USA, 2017.



64. United Nations Statistics Division. Population by Sex and Urban/Rural Residence. Available online: <http://data.un.org/Data.aspx?d=POP&f=tableCode%3A1> (accessed on 13 January 2021).
65. Suryadinata, L.; Arifin, E.N.; Ananta, A. *Indonesia's Population: Ethnicity and Religion in a Changing Political Landscape*; Institute of Southeast Asian Studies: Singapore, 2003.
66. Statistics Indonesia-Badan Pusat Statistik-BPS; National Population and Family Planning Board-BKKBN/Indonesia; Kementerian Kesehatan-Kemenkes-Ministry of Health/Indonesia; ICF International. *Indonesia Demographic and Health Survey 2012*; BPS, BKKBN, Kemenkes, and ICF International: Jakarta, Indonesia, 2013.
67. World Health Organization. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus: Abbreviated Report of a WHO Consultation (WHO/NMH/CHP/CPM/11.1). Available online: [https://www.who.int/diabetes/publications/diagnosis\\_diabetes2011/en/](https://www.who.int/diabetes/publications/diagnosis_diabetes2011/en/) (accessed on 8 January 2021).
68. The National Institute of Health Research and Development. *National Report on Basic Health Research, RISKESDAS*; Indonesia Ministry of Health: Jakarta, Indonesia, 2013.
69. World Health Organization. WHO Guidelines on Physical Activity and Sedentary Behaviour. Available online: <https://www.who.int/publications/i/item/9789240015128> (accessed on 8 January 2021).
70. Kaminsky, L.A.; Montoye, A.H.K. Physical Activity and Health: What Is the Best Dose? *Am. Heart Assoc.* **2014**, *3*. [[CrossRef](#)]
71. Tripathy, J.P.; Thakur, J.S.; Jeet, G.; Chawla, S.; Jain, S.; Pal, A.; Prasad, R.; Saran, R. Prevalence and Risk Factors of Diabetes in a Large Community-Based Study in North India: Results from a STEPS Survey in Punjab, India. *Diabetol. Metab. Syndr.* **2017**, *9*, 1–8. [[CrossRef](#)]
72. Roman-Urrestarazu, A.; Ali, F.M.H.; Reka, H.; Renwick, M.J.; Roman, G.D.; Mossialos, E. Structural Equation Model for Estimating Risk Factors in Type 2 Diabetes Mellitus in a Middle Eastern Setting: Evidence from the STEPS Qatar. *BMJ Open Diabetes Res. Care* **2016**, *4*, e000231. [[CrossRef](#)]
73. Miller, R.G. *Simultaneous Statistical Inference*, 2nd ed.; Springer Series in Statistics; Springer New York: New York, NY, USA, 1981. [[CrossRef](#)]
74. Aune, D.; Norat, T.; Leitzmann, M.; Tonstad, S.; Vatten, L.J. Physical Activity and the Risk of Type 2 Diabetes: A Systematic Review and Dose-Response Meta-Analysis. *Eur. J. Epidemiol.* **2015**, *30*, 529–542. [[CrossRef](#)]
75. Mills, K.T.; Stefanescu, A.; He, J. The Global Epidemiology of Hypertension. *Nat. Rev. Nephrol.* **2020**, *16*, 223–237. [[CrossRef](#)]
76. Bellou, V.; Belbasis, L.; Tzoulaki, I.; Evangelou, E. Risk Factors for Type 2 Diabetes Mellitus: An Exposure-Wide Umbrella Review of Meta-Analyses. *PLoS ONE* **2018**, *13*, e0194127. [[CrossRef](#)]
77. Willi, C.; Bodenmann, P.; Ghali, W.A.; Faris, P.D.; Cornuz, J. Active Smoking and the Risk of Type 2 Diabetes: A Systematic Review and Meta-Analysis. *J. Am. Med. Assoc.* **2007**, *298*, 2654–2664. [[CrossRef](#)] [[PubMed](#)]
78. Jarrin, D.C.; Alvaro, P.K.; Bouchard, M.A.; Jarrin, S.D.; Drake, C.L.; Morin, C.M. Insomnia and Hypertension: A Systematic Review. *Sleep Med. Rev.* **2018**, *41*, 3–38. [[CrossRef](#)] [[PubMed](#)]
79. Buxton, O.M.; Marcelli, E. Short and Long Sleep Are Positively Associated with Obesity, Diabetes, Hypertension, and Cardiovascular Disease among Adults in the United States. *Soc. Sci. Med.* **2010**, *71*, 1027–1036. [[CrossRef](#)] [[PubMed](#)]
80. Reutrakul, S.; Van Cauter, E. Sleep Influences on Obesity, Insulin Resistance, and Risk of Type 2 Diabetes. *Metabolism* **2018**, *84*, 56–66. [[CrossRef](#)]
81. Pengpid, S.; Peltzer, K. The Prevalence of Underweight, Overweight/Obesity and Their Related Lifestyle Factors in Indonesia, 2014–15. *AIMS Public Health* **2017**, *4*, 633–649. [[CrossRef](#)]
82. Seixas, A.A.; Vallon, J.; Barnes-Grant, A.; Butler, M.; Langford, A.T.; Grandner, M.A.; Schneeberger, A.R.; Huthchinson, J.; Zizi, F.; Jean-Louis, G. Mediating Effects of Body Mass Index, Physical Activity, and Emotional Distress on the Relationship between Short Sleep and Cardiovascular Disease. *Medicine (Baltim.)* **2018**, *97*, e11939. [[CrossRef](#)]
83. Fogelholm, M.; Kukkonen-Harjula, K. Does Physical Activity Prevent Weight Gain—A Systematic Review. *Obes. Rev.* **2000**, *1*, 95–111. [[CrossRef](#)]
84. Reimers, A.; Knapp, G.; Reimers, C.-D. Effects of Exercise on the Resting Heart Rate: A Systematic Review and Meta-Analysis of Interventional Studies. *J. Clin. Med.* **2018**, *7*, 503. [[CrossRef](#)]
85. Bakker, E.A.; Sui, X.; Brellenthin, A.G.; Lee, D. Physical Activity and Fitness for the Prevention of Hypertension. *Curr. Opin. Cardiol.* **2018**, *33*, 394–401. [[CrossRef](#)]
86. Kasapis, C.; Thompson, P.D. The Effects of Physical Activity on Serum C-Reactive Protein and Inflammatory Markers: A Systematic Review. *J. Am. Coll. Cardiol.* **2005**, *45*, 1563–1569. [[CrossRef](#)]
87. Juster, R.P.; McEwen, B.S.; Lupien, S.J. Allostatic Load Biomarkers of Chronic Stress and Impact on Health and Cognition. *Neurosci. Biobehav. Rev.* **2010**, *35*, 2–16. [[CrossRef](#)]
88. Steptoe, A.; Hackett, R.A.; Lazzarino, A.I.; Bostock, S.; La Marca, R.; Carvalho, L.A.; Hamer, M. Disruption of Multisystem Responses to Stress in Type 2 Diabetes: Investigating the Dynamics of Allostatic Load. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15693–15698. [[CrossRef](#)]
89. Carlsson, A.C.; Nixon Andreasson, A.; Wändell, P.E. Poor Self-Rated Health Is Not Associated with a High Total Allostatic Load in Type 2 Diabetic Patients - But High Blood Pressure Is. *Diabetes Metab.* **2011**, *37*, 446–451. [[CrossRef](#)]
90. Seeman, T.; Epel, E.; Gruenewald, T.; Karlamangla, A.; McEwen, B.S. Socio-Economic Differentials in Peripheral Biology: Cumulative Allostatic Load. *Ann. N. Y. Acad. Sci.* **2010**, *1186*, 223–239. [[CrossRef](#)] [[PubMed](#)]
91. Poulter, N. Global Risk of Cardiovascular Disease. *Heart* **2003**, *89* (Suppl. 2), ii2–ii5. [[CrossRef](#)] [[PubMed](#)]

92. Lasic, S.E. Why We Should Use Simpler Models If the Data Allow This: Relevance for ANOVA Designs in Experimental Biology. *BMC Physiol.* **2008**, *8*, 1–7. [[CrossRef](#)] [[PubMed](#)]
93. Soewondo, P.; Ferrario, A.; Tahapary, D.L. Challenges in Diabetes Management in Indonesia: A Literature Review. *Global. Health* **2013**, *9*, 1–17. [[CrossRef](#)]