Impact of Behavioral Factors on Non-Communicative Diseases

Jack Rossi, Wenyu Yu, Tianran Zhang

Abstract

Non-communicable diseases such as diabetes, hypertension, and heart disease are known to be risk factors for communicable diseases such as COVID-19. Patients with these diseases are believed to be both at higher risk of mortality, and also higher risk of transmission to others. As a long-term pandemic mitigation strategy, we explored the behavioral factors related to emergence of non-communicable diseases. Our analysis considered respondents with one or more non-communicable diseases to be part of the NCD cohort, and respondents with 0 non-communicable diseases to be part of the non-NCD cohort. We found significant differences in diet, alcohol consumption, sleep quality, smoking and BMI between the NCD and non-NCD cohorts. In univariate and multivariate regressions on the outcome of 1 or more NCDs, we found that BMI, trouble sleeping, alcohol abstinence, smoking and fair diet habits were predictive of the NCD outcome. We also made NCD an ordinal variable to investigate the relationship between the number of NCDs patients suffer from and their lifestyles. We found that people with higher BMI, trouble in sleeping, alcohol abstinence, smoking habit and worse diet habits, tend to have more than one NCD. BMI, trouble sleeping, alcohol abstinence, smoking and fair diet habits are also significant indicator in predicting the severity of NCSs. These results show an association between behavioral factors and non-communicative diseases and suggest that behavioral interventions may be appropriate for limiting the long-term mortality from pandemic like COVID-19.

Introduction and Rationale

In the first quarter of 2020, reports of the factors associated with mortality in Severe Acute Respiratory Coronavirus Disease 2019 (COVID - 19) have emerged from the affected countries. While sample sizes are small and results tentative, it seems likely that mortality in sufferers of this disease is closely related to common comorbidities, including Cardiovascular Disease, diabetes, and hypertension (1). These medical problems are doubly debilitating in that their sufferers are more likely to suffer other disabilities (2), including sarcopenia, frailty, and skeletal muscle dysfunction (3). In many countries, insufficient testing is a dangerous fact of the situation at hand, especially in the short to medium term. In the absence of system capabilities to identify and isolate infected individuals, we identify the importance of instead protecting those who are at greatest risk of mortality, and those who have risk factors associated with the chronic diseases that often correspond to mortality. We also note that people with certain risk factors, being at higher risk for severe or deadly cases of COVID-19, are also more likely to further spread the disease through respiratory droplets. We will explore the lifestyle and behavioral risk factors that relate to diabetes, Cardiovascular Disease, and hypertension. Our hope is to understand the role of these factors in these non-communicable diseases so that future pandemics can be limited in severity.

According to the CDC, 34.2 million people have diabetes in the US, which accounts for 10.5% of the US population and may increase to 34.5% in the future. Given that patients with diabetes are more likely than others to suffer mortality from COVID-19, we decided that we were going to consider diabetes as a risk factor. A report indicated that lifestyle changes can reduce the risk of diabetes (4), which suggests there may be a similar reduction of the risk of death from COVID-19. Some literature suggests that vitamin C, an essential nutrient for bone health, may be related to the risk of contracting respiratory illness (5), so we will explore whether this nutrient has a role in mitigating risk factors such as diabetes. Further studies show that behavioral factors can play a role in hypertension and cardiovascular disease (6) (2), but no efforts to our knowledge have considered each of these non-communicable diseases in parallel.

The National Health and Nutrition Evaluation Survey (NHANES) provides sufficient data to explore the risk factors that impact the incidence of these diseases. We will build upon the current research by using the most recent data from the National Health and Nutrition Examination Survey (NHANES). We also propose several new covariates, including smoking, alcohol consumption, sleep habits, body weight, caffeine consumption and diet. We consider these covariates because they have either been considered previously, but not in a combined NCD manner (4), (6), or because they were present in the NHANES survey data, and subject or on-going behavioral public health initiatives (9).

Methods

Study Objectives

Our research seeks to identify risk factors associated with three common chronic disorders -- diabetes, cardiovascular disease and hypertension. We will test the following hypothesis: diabetes, cardiovascular disease and hypertension are associated with smoking, alcohol consumption, sleep habits, body weight, caffeine consumption and diet.

Study Population

In all, 1136 respondents are included in our study. The study population is people within 15-80 years old who responded to questionnaires in the National Health and Nutrition Examination Survey in 2017 and 2018. Patients were excluded if they had missing values of the variables and three diseases.

Statistical Analysis

After extracting all the data from NHANES, we tried 2 different approaches in creating the “NCD” variable:
a) Create “NCD” with binary outcome indicating one subject has at least one of the three diseases or not.

b) Create “NCD” with ordinal outcome values 0 to 3, indicating the number of diseases out of 3 one subject suffers from.

Within each approach, we evaluated the lifestyle and behavioral characteristics with descriptive statistics. For numerical covariates (Alcohol, vitamin C, caffeine consumption, SLD012, smoking, Weight, bmi), mean ± SD and median (range) were calculated. Two-sample t-tests were carried on comparing the mean differences of these covariates between disease and non-disease groups. ANOVA tests were conducted to compare their mean differences among different severity groups. We also used boxplots to compare the differences in distribution for each covariate among different outcome groups. For categorical covariates (Diet, Alcohol frequency, Sleeping), differences in proportions of each category for each variable across experimental groups were examined using chi-square test. We also drew Mosaic plots to visually compare the differences in proportion for each covariate among different outcome groups.

In our first approach with binary outcomes, to determine the relationship between each variable and NCDs, we performed univariate logistic regressions. Multivariate logistic regression was followed with only the significant predictors included assessing the association of behaviors and outcome variables using stepwise selection.

In the second approach considering the ordinal outcome, we performed a univariate proportional odds model to determine the association between each covariate and NCD outcome. The multivariate proportional odds model was followed with only the significant predictors included measuring the association of behaviors and outcome using stepwise selection.

The strength of each variable was measured using chi-square test. Odds ratio and 95% confidence intervals are given for these variables. The linear relationship, normality, and homoscedasticity between predictor and outcome variables were examined with residual plots. The PH assumptions were tested in the second approach. Multicollinearity was also tested. To maintain interpretability, we chose not to use any transformations though slightly violation to normality and linearity. All statistical testing was two-sided, and a P value of $\leq .05$ was considered statistically significant. R version 3.6.1 was used for all statistical analyses.

Results

Descriptive characteristics

Table 1 and 2 show the behavioral factors characteristics by the binary and ordinal NCD outcome groups.

	TRUE (N=833)	**FALSE (N=301)**	**P-value**
Diet			
Excellent	76 (9.1%)	38 (12.6%)	0.00578

	TRUE (N=833)	FALSE (N=301)	P-value
Very good	156 (18.7%)	74 (24.6%)	
Good	319 (38.3%)	119 (39.5%)	
Fair	212 (25.5%)	53 (17.6%)	
Poor	70 (8.4%)	17 (5.6%)	
Alcohol_freq			
Mean (SD)	2.85 (13.4)	3.09 (4.49)	0.648
Median [Min, Max]	1.00 [0, 365]	2.00 [0, 60.0]	
Alcohol			
Yes	266 (31.9%)	71 (23.6%)	0.00825
No	567 (68.1%)	230 (76.4%)	
Vitamin C			
Mean (SD)	70.7 (80.7)	75.2 (94.6)	0.465
Median [Min, Max]	41.7 [0, 921]	42.8 [0, 745]	
Caffeine			
Mean (SD)	207 (258)	232 (360)	0.264
Median [Min, Max]	144 [0, 2950]	156 [0, 4530]	
Sleep			
Mean (SD)	7.64 (1.68)	7.57 (1.38)	0.488
Median [Min, Max]	8.00 [2.00, 13.5]	7.50 [3.00, 11.5]	
Trouble sleep			
Yes	341 (40.9%)	69 (22.9%)	<0.001
No	492 (59.1%)	232 (77.1%)	
Smoking age			
Mean (SD)	18.2 (7.15)	18.8 (6.60)	0.251
Median [Min, Max]	17.0 [0, 58.0]	18.0 [0, 54.0]	
Weight			
Mean (SD)	189 (43.9)	171 (35.2)	<0.001
Median [Min, Max]	182 [78.0, 386]	165 [101, 300]	
BMI			
Mean (SD)	28.7 (7.29)	26.2 (5.99)	<0.001
Median [Min, Max]	28.2 [0.000956, 64.2]	25.8 [0.00105, 53.7]	

Table1. Descriptive characteristics of binary NCDs

	0 (N=301)	1 (N=418)	2 (N=347)	3 (N=68)	P-value
Diet					
Excellent	38 (12.6%)	47 (11.2%)	26 (7.5%)	3 (4.4%)	0.00186
Very good	74 (24.6%)	87 (20.8%)	54 (15.6%)	15 (22.1%)	
Good	119 (39.5%)	166 (39.7%)	128 (36.9%)	25 (36.8%)	
Fair	53 (17.6%)	89 (21.3%)	105 (30.3%)	18 (26.5%)	
Poor	17 (5.6%)	29 (6.9%)	34 (9.8%)	7 (10.3%)	
Alcohol					
Yes	71 (23.6%)	121 (28.9%)	116 (33.4%)	29 (42.6%)	0.0041
No	230 (76.4%)	297 (71.1%)	231 (66.6%)	39 (57.4%)	
Vitamin_C					
Mean (SD)	75.2 (94.6)	67.3 (83.1)	76.2 (79.6)	63.4 (69.4)	0.133
Median [Min, Max]	42.8 [0, 745]	39.9 [0, 921]	49.8 [0, 545]	33.0 [0.400, 331]	
Caffeine					
Mean (SD)	232 (360)	221 (277)	184 (224)	236 (293)	0.0911
Median [Min, Max]	156 [0, 4530]	160 [0, 2950]	132 [0, 1920]	153 [0, 1760]	
Sleep					
Mean (SD)	7.57 (1.38)	7.74 (1.66)	7.53 (1.68)	7.60 (1.82)	0.413
Median [Min, Max]	7.50 [3.00, 11.5]	8.00 [2.50, 13.5]	7.50 [2.00, 13.0]	8.00 [2.00, 11.0]	
Trouble_sleep					
Yes	69 (22.9%)	147 (35.2%)	155 (44.7%)	39 (57.4%)	<0.001
No	232 (77.1%)	271 (64.8%)	192 (55.3%)	29 (42.6%)	
Smoking_age					
Mean (SD)	18.8 (6.60)	18.4 (6.69)	18.1 (7.70)	17.6 (7.03)	0.0115
Median [Min, Max]	18.0 [0, 54.0]	18.0 [0, 49.0]	17.0 [0, 58.0]	16.0 [0, 45.0]	
BMI					
Mean (SD)	26.2 (5.99)	27.8 (6.63)	29.0 (7.45)	32.7 (8.89)	<0.001
Median [Min, Max]	25.8 [0.00105, 53.7]	27.5 [0.000956, 55.3]	28.3 [0.000984, 57.2]	32.4 [0.00102, 64.2]	

Table2. Descriptive characteristics of ordinal NCDs

We conducted two different approached corresponding to two types of outcome variable.

a) NCD with binary outcome

1) Hypothesis: Is there an association between each of the behavioral factors and the presence of at least one NCD?

We conducted two-sample t-test in evaluating the association between continuous factors and the outcome, and chi-square test in evaluating the association between categorical factors and the outcome. The results are shown in Table 1. We found that diet, alcohol, trouble sleeping, weight and BMI were significantly different between NCD and non-NCD groups. For diet, the majority of the NCD group had either a Fair or Good diet, while most of the non-NCD group had a Good, Very Good, or Excellent diet (chi-square test p-value = 0.006). For alcohol, the proportion of respondents who report drinking alcohol (31.9%) is higher in the NCD group than in the non-NCD group (23.6%) (chi-square test p-value = 0.0825). For trouble sleeping, the proportion who reported having trouble sleeping (40.9%) is higher in the NCD group than the non-NCD group

(22.9%) (chi-square test p-value < 0.001). For weight and BMI, the mean of them are significantly different between NCD and non-NCD groups (both have two-sample t-test p-values < 0.001).

2) Hypothesis: Are each of the covariate a predictor of NCD?

	OR	2.5%	97.5%	p_value
DietVery good	1.054	0.650	1.694	0.82898
DietGood	1.340	0.855	2.077	0.19466
DietFair	2.000	1.219	3.269	0.00578
DietPoor	2.059	1.080	4.050	0.03138
Alcohol_freq	0.998	0.987	1.013	0.761
AlcoholNo	0.658	0.483	0.888	0.00685
Vitamin_C	0.999	0.998	1.001	0.432
Caffeine	1.000	0.999	1.000	0.2
Sleep	1.027	0.946	1.115	0.5267
Trouble_sleepNo	0.429	0.315	0.578	<0.001
Smoking_age	0.990	0.972	1.008	0.269
Weight	1.012	1.008	1.015	< 0.001
BMI	1.054	1.034	1.076	<0.001

Table 3. Univariate regression of binary NCDs

We used univariate generalized linear regression model (GLM) with logit link function for this hypothesis. From the results shown in Table3, we found out that diet, alcohol, trouble sleeping, weight and BMI are independent predictors of NCDs. The univariate effect for each predictor is shown in fig1.

3) How strong is the association between these covariates and the presence of at least one NCD?

	OR	2.5%	97.5%	P-val
DietExcellent	Ref			NA
DietVery good	1.0064	0.6111	1.6442	0.9799
DietGood	1.2011	0.7547	1.8903	0.4330
DietFair	1.7663	1.0597	2.9321	0.0281
DietPoor	1.5279	0.779	3.0809	0.2249
AlcoholNo	0.7081	0.5142	0.9671	0.0320
Caffeine	0.9995	0.999	1	0.0546
Trouble_sleepNo	0.4385	0.3193	0.5963	<0.001
Weight	1.0116	1.0079	1.0156	<0.001

Table 4. Multivariate Logistic Regression of binary NCDs

Multivariate generalized linear regression model with logit link function was carried out for this question, and the results are shown in Table4. We found out that xxx are significant predictors for developing NCD. A fair diet has greater odds of NCDs than an excellent diet adjusted for other

covariates (OR 1.77, 95% CI 1.06 – 2.93, p = 0.028). Non-alcohol use has a less odds of NCDs than alcohol use adjusted for other covariates (OR 0.708, 95% CI 0.514-0.967, p = 0.032). Non-troubled sleep has less odds of NCDs than troubled sleep adjusted for other covariates (OR 0.439, 95% CI 0.319-0.596, p < 0.001). The higher BMI is, the higher odds of NCDs we have, with other covariates fixed (OR 0.439, 95% CI 0.319-0.596, p < 0.001). These covariates can interpret 68.17% of the presence of NCD, with the ROC curve is shown in fig2.

b) NCD with ordinary outcome

1) Hypothesis: Is there an association between each of the behavioral factors and the presence of at least one NCDs?

We conducted two-way Kruskal test in evaluating the association between continuous factors and the ordinal outcome, and chi-square test in evaluating the association between categorical factors and the ordinal outcome. The results are shown in Table 2. We found that diet, alcohol, trouble sleep, and BMI were significantly different among the ordinal NCD groups. For diet, increases in the number of NCDs are met with increases in the proportion of Poor diets reported and decreases in the number of Excellent diets reported (chi-square test p-value = 0.0186). For alcohol, the proportion of respondents drinking alcohol increases as the number of NCDs increases (chi-square test p-value = 0.0041). For trouble sleeping, the proportion of respondents having trouble sleeping increases as the number of NCDs increases (chi-square test p-value <0.001). In terms of weight and BMI, the mean of them are significantly different among the ordinal NCD groups, which are increasing by the increased NCDs (both with two-way Kruskal test p-values less than 0.001).

2) Hypothesis: Are each of the covariate an independent predictor of ordinal NCD?

We used univariate proportional odds model for this hypothesis. The estimated OR, 95% confidence interval, as well as PH assumption test p-value are all calculated. All factors are satisfied the PH assumption. We found out that diet, alcohol, trouble sleeping, weight and BMI are independent predictors of ordinal NCDs. The univariate effect for each predictor is shown in fig2.

3) How strong is the association between these covariates and the presence of at least one NCD?

	OR	2.5 %	97.5 %
DietExcellent	Ref		
DietVery good	1.148	0.76	1.738
DietGood	1.372	0.94	2.006
DietFair	2.005	1.338	3.01
DietPoor	1.759	1.041	2.975
AlcoholNo	0.682	0.538	0.865
Caffeine	1	0.999	1
Trouble_sleepNo	0.489	0.389	0.614
Smoking_age	0.988	0.972	1.003
BMI	1.054	1.037	1.072

Table 6. Multivariate Proportional Odds Model of ordinal NCDs

Multivariate proportional odds model was carried out for this question, and the results are shown in Table6. We found out that diet, alcohol, caffeine, sleeping, smoking and BMI are significant risk factors in developing more NCDs. As shown in Table 6, a worse diet would increase respondent's odds of occupying higher NCD ordinal groups. To be specific, compared to having an excellent diet, having a Very Good diet increased respondents' odds by a factor of 1.148, having a Good diet increased odds by a factor of 1.372, having a Fair diet increased odds by a factor of 2.005, and having a Poor diet by a factor of 1.759. Abstaining from alcohol and having no difficulty sleeping reduced the likelihood of occupying a higher NCD ordinal group by 0.682 and 0.489, respectively. Having an additional point of BMI increased the odds of occupying a higher NCD ordinal group by 1.054. Smoking age and caffeine consumption had little effect. The calculated AOC is 0.8054, meaning that these covariates can interpret 80.54% of the outcome.

Discussion

We believe these results to be preliminary indicators of the relationship between actionable, population-scale behaviors and the death toll of pandemics. However, there are several limitations to this study. First, we lack any primary data from the COVID-19 pandemic to make the direct link between behavioral factors and mortality via non-communicable diseases. We are also aware of several potentially confounding variables, such as the use of electronic cigarettes (8), some demographic variables of age, race, sex, and income. The survey also does not include geographic information which could reveal confounding environmental variables.

Group Member Contributions

Jack developed the abstract, rationale, introduction and discussion. He also helped with interpretation of the ordinal proportional odds model regression. Wenyu developed the method and

result part and helped with the R codes, and Tianran fit the univariate and multivariate models, and developed the statistical analysis part.

Citations

1. Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L., & Wei, Y. (2020). Articles Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan , China : a retrospective cohort study. *The Lancet*, *6736*(20), 1–9. https://doi.org/10.1016/S0140-6736(20)30566-3
2. Lisy, K. , Campbell, J. M. , Tufanaru, C. , Moola, S. , & Lockwood, C. . (2018). The prevalence of disability among people with cancer, cardiovascular disease, Cardiovascular Disease and/or diabetes:, a systematic review. International Journal of Evidence Based Healthcare, 16. https://www.ncbi.nlm.nih.gov/pubmed/29608458
3. Bone, A. E. , Hepgul, N. , Kon, S. , & Maddocks, M. . (2017). Sarcopenia and frailty in Cardiovascular Disease. Cardiovascular Disease, 14(1), 85-99.Wiener, R. C., https://www.ncbi.nlm.nih.gov/pubmed/27923981
4. N Engl J Med.(2002) Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1370926/
5. Roth, D. E., Abrams, S. A., Aloia, J., Bergeron, G., Bourassa, M. W., Brown, K. H., ... & Jefferds, M. E. (2018). Global prevalence and disease burden of vitamin D deficiency: a roadmap for action in low-and middle-income countries. https://www.ncbi.nlm.nih.gov/pubmed/30225965
6. Shapiro, D., & Goldstein, I. B. (1982). Behavioral Patterns as They Relate to Hypertension. *Arterial Hypertension*, 51–61. https://doi.org/10.1007/978-1-4612-5657-1_3
7. Campbell, T., Bacon, S. L., Dimsdale, J. E., & Carroll, D. (2012). Hypertension: A behavioral medicine perspective. International Journal of Hypertension, 2012, 2012–2014. https://doi.org/10.1155/2012/385690
8. Waters, C., Bhandari, R., Trickett Shockey, A. K., & Alshaarawy, O. (2020). The Association of Sleep Duration and the Use of Electronic Cigarettes, NHANES, 2015-2016. Sleep Disorders, 2020. https://www.ncbi.nlm.nih.gov/pubmed/32190389
9. National Healthy Sleep Awareness Project. http://sleepeducation.org/healthysleep

Appendix:

Fig1:

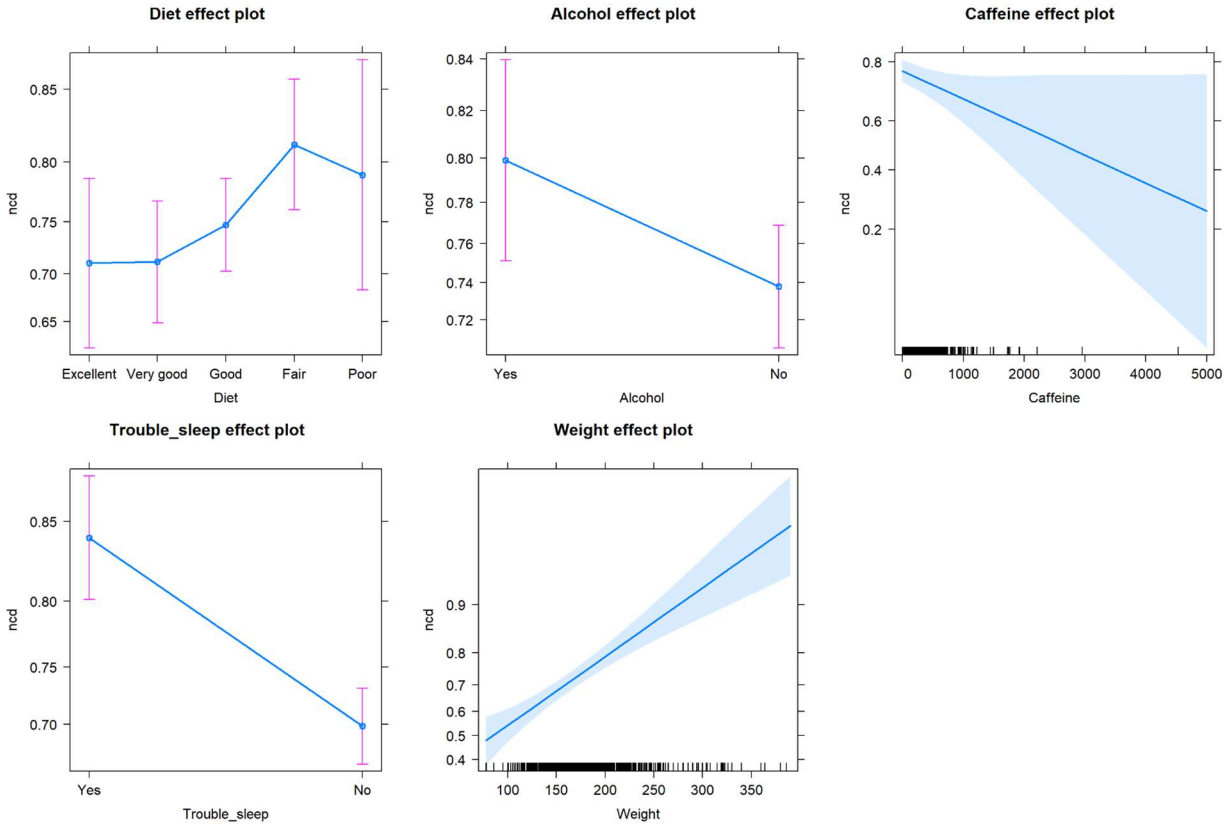

Fig 1. Effect plots of the final model by binary NCDs

Fig2:

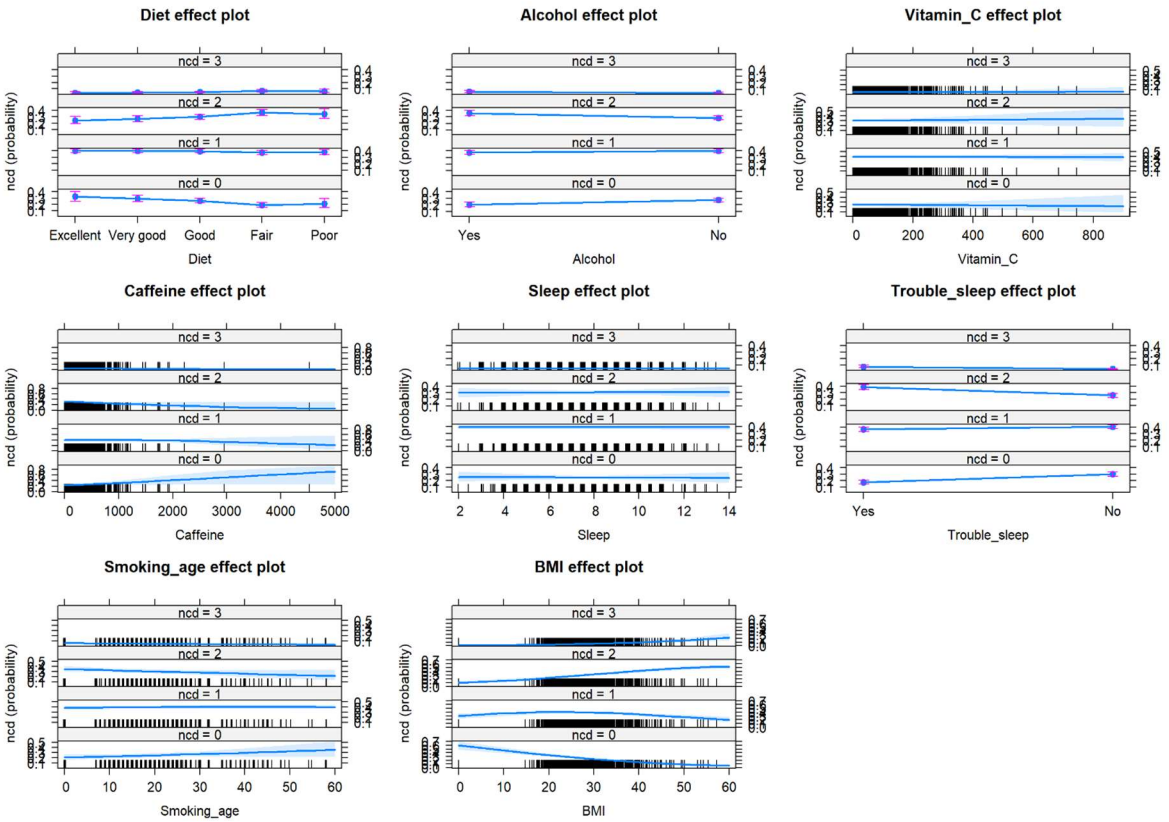

Fig 2. Effect plots of the final model by ordinal NCDs

Fig3:

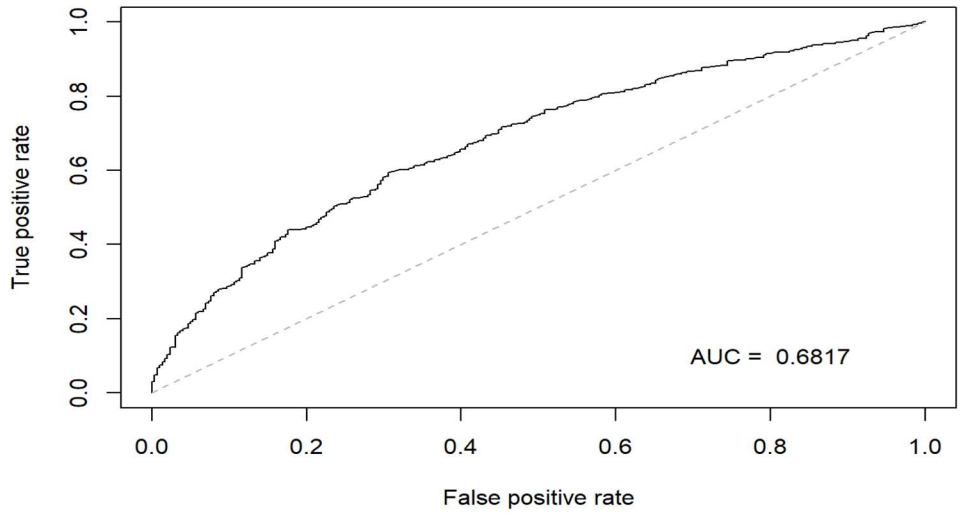

Fig 3. AOC curve of the final model by binary NCDs