

SUPPLEMENTAL MATERIALS**Trajectories of Early to Mid-life Adulthood BMI and Incident Diabetes: The China Health and Nutrition Survey**

Jiali Lv ^{1*}, Bingbing Fan ^{1*}, Mengke Wei ¹, Guangshuai Zhou ¹, Alim Dayimu ¹, Zhenyu Wu ²,
Chang Su ^{3#}, Tao Zhang ^{1#}

¹ Department of Biostatistics, School of Public Health, Shandong University, Jinan, China

² Department of Biostatistics, School of Public Health, Fudan University, Shanghai, China

³ National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention, Beijing, China

* Equally contributed

Running Title: Early to Mid-life Adulthood BMI Trajectories and Diabetes

Correspondence & Reprints:

Tao Zhang, MD, PhD.

Department of Biostatistics, School of Public Health, Shandong University

PO Box 100, 44 Wenhuxi Road, Jinan 250012, China.

Email: taozhang@sdu.edu.cn Tel: 0531-88382140

and Chang Su, MD, PhD.

National Institute for Nutrition and Health, Chinese Center for Disease Control and

Prevention, 155 Changba Road, Beijing 102206, China.

Email: suchang@nih.chinacdc.cn Tel: 0531-88382140

The authors do not have any conflict of interest.

Trajectory Analysis

Repeated trajectory analyses were performed to identify the latent classes by changing the number of groups from 2 to 5, with the same starting values calculated from the 1-group model. The shapes and optimal number of groups were determined by the following criteria: 1) Bayesian information criterion (BIC) decreased at least 20; 2) high mean posterior class membership probabilities (> 0.65); 3) high mean posterior probabilities (> 0.7). Estimation of latent class models was performed with *lcmm* (version 1.7.8) package in R (version 3.4.3). To avoid convergence towards local maxima, LCGMM models with 2 or 5 classes were performed for several times with different sets of random starting value based on 1-class model. Finally, the best fitting model based on above criteria was quadratic trajectories of three groups, and the final model was described as:

$$BMI_{ij} |_{c_i=g} = (v_{0g} + u_{0ig}) + (v_{1g} + u_{1ig})age + (v_{2g} + u_{2ig})age^2 + \varepsilon_{ij}$$

where $v = (v_{0g}, v_{1g}, v_{2g})$ is a vector of fixed effect parameters in the group “g”, $u = (u_{0ig}, u_{1ig}, u_{2ig})$ is a vector of random effect parameters of the individual “i” in the group “g”, ε_{ij} is an unknown error term.

Supplement Table S1. Baseline characteristics of participants included and excluded

Variable	Excluded	Included	<i>P</i> -value
N	26059	7289	
Age, year	31.2 (23.5)	31.4 (6.8)	<0.001
Male, n (%)	12478 (47.9)	3516 (48.2)	0.5934
BMI, kg/m ²	20.8 (4.8)	21.8 (2.7)	<0.001
Smoker, n (%) †	4616 (27.3)	1104 (31.7)	<0.001
Drinker, n (%) †	990 (6.2)	271 (8.8)	<0.001

BMI=body mass index. † Rate was calculated after removing missing value

Supplement Table S2. Descriptive data of baseline and follow-up characteristics by incident hyperglycemia at follow-up

Variable	Total (n=7289)	Normoglycemia (n=6318)	Hyperglycemia (n=971)	P-value
Baseline				
Age, year	31.4 (6.8)	31.8 (6.9)	28.9 (5.7)	< 0.001
BMI, kg/m ²	21.8 (2.7)	21.7 (2.6)	22.5 (3.1)	< 0.001
Smoker, n (%) †	1104 (31.7)	869 (30.6)	235 (36.3)	0.005
Drinker, n (%) †	271 (8.8)	208 (8.3)	63 (10.8)	0.065
Follow-up				
Age, year	42.8 (6.4)	42.8 (6.6)	42.5 (5.2)	<0.001
BMI, kg/m ²	23.1 (3.1)	23.0 (3.0)	24.3 (3.4)	<0.001
Smoker, n (%) †	2431 (33.4)	2085 (33.1)	346 (35.6)	0.118
Drinker, n (%) †	959 (13.8)	825 (13.6)	134 (14.7)	0.403
FPG, mmol/L	5.5 (1.4)	4.8 (0.4)	5.8 (1.5)	<0.001
Hb1Ac, %	5.7 (0.9)	5.2 (0.4)	5.9 (0.9)	<0.001
Follow-up years	11.3 (5.3)	11.0 (5.2)	13.6 (5.3)	<0.001

BMI=body mass index, FPG = fasting plasma glucose, Hb1Ac = Hemoglobin A1c. † Rate was calculated after removing missing value

Supplement Table S3. The distribution of age of onset for hyperglycemia group by gender

Age	Total	Male	Female
N	971	497	474
20~30	19 (2.0%)	12 (2.4%)	7 (1.5%)
30~40	268 (2.8%)	144 (29.0%)	124 (26.2%)
40~50	684 (70.4%)	341 (68.6%)	343 (72.3%)

Supplement Table S4. Latent Class Growth Mixture Models (LCGMM) results of model fitting process

No. Latent class	Polynomial degree	Log-Lik	BIC	% Participants per class	Mean posterior probabilities	% Posterior probabilities > 70%
1	Linear	-65101	130246	100	NA	NA
	Quadratic	-65052	130166	100	NA	NA
	Cubic	-65049	130179	100	NA	NA
2	Linear	-64654	129388	12.31/ 87.69	0.81/ 0.95	72.46/ 94.98
	Quadratic	-64472	129050	14.84/ 85.16	0.83/ 0.95	75.14/ 94.68
	Cubic	-64465	129064	14.89/ 85.11	0.83/ 0.95	76.04/ 94.65
3	Linear	-64509	129134	26.33/ 3.32/ 70.35	0.76/ 0.85/ 0.88	65.61/ 80.17/ 86.02
	Quadratic	-64311	128774	26.26/ 3.28/ 70.46	0.78/ 0.85/ 0.9	68.76/ 77.82/ 87.87
	Cubic	-64301	128788	26.51/ 3.72/ 69.78	0.77/ 0.84/ 0.89	66.98/ 75.28/ 87.28
4	Linear	-64452	129056	11.78/ 1.91/ 37.49/ 48.81	0.77/ 0.84/ 0.7/ 0.81	66.01/ 78.42/ 53.49/ 72.93
	Quadratic	-64214	128624	4.35/ 3.24/ 66.88/ 25.53	0.64/ 0.81/ 0.87/ 0.71	34.07/ 72.03/ 83.82/ 51.53
	Cubic	-64173	128586	3.35/ 5.53/ 24.26/ 66.87	0.82/ 0.64/ 0.72/ 0.86	73.36/ 33.5/ 52.88/ 82.81
5	Linear	-64358	128902	10.74/ 1.76/ 2.13/ 37.65/ 47.73	0.72/ 0.81/ 0.7/ 0.69/ 0.8	55.68/ 71.88/ 49.68/ 49.09/ 71.92
	Quadratic	-64131	128502	9.92/ 2.02/ 4.64/ 36.15/ 47.28	0.7/ 0.82/ 0.65/ 0.69/ 0.81	50.48/ 75.51/ 36.09/ 47.4/ 73.13
	Cubic	-64082	128458	9.96/ 2.22/ 6.8/ 48.57/ 32.45	0.68/ 0.83/ 0.65/ 0.79/ 0.66	47.66/ 78.4/ 37.5/ 67.94/ 42.16

No. Latent class: latent class number of the model; Log-Lik: the maximum Log-Likelihood; BIC: the Bayesian information Criterion;

% Participants per class: proportion of participants per class; The best fitting model is highlighted in bold characters. (NA: not applicable).

Supplement Table S5. Parameter estimates for the best fitting 3-class quadratic Latent Class Growth Mixture Model

	Intercept (se)*	Linear (se)	Quadratic (se)
Fixed effect			
Low--increasing	21.544 (0.1656)	0.0736 (0.0026)	0.0018 (0.0002)
Medium-increasing	not estimated†	0.1554 (0.0044)	-0.0052 (0.0005)
High-increasing	29.461 (0.2067)	0.1617 (0.0126)	-0.0098 (0.0011)
Random effects: variance-covariance matrix			
$\sigma_{int}^2 = 1.935$			
$\sigma_{linear\ slope}^2 = 0.008$			
$\sigma_{quadratic\ slope}^2 < 0.001$			
$\sigma_{error}^2 = 1.000$			

se=standard error

*: Intercept interpreted as the expected level of BMI in kg/m² at 37 years of age (centering to the mean age of the sample)

†: The se of the intercept of this group was not estimated.

Supplement Table S6. Model-estimated levels and linear slopes of BMI in means (SD) by incident hyperglycaemia at follow-up

Age (year)	BMI Level (kg/m ²)			BMI Slope (kg/m ² /year)		
	NGLY	HGLY	<i>P</i> -value	NGLY	HGLY	<i>P</i> -value
20	21.529 (2.071)	22.207 (2.282)	<0.001	0.139 (0.222)	0.215 (0.241)	<0.001
21	21.667 (2.215)	22.419 (2.431)	<0.001	0.137 (0.212)	0.209 (0.230)	<0.001
22	21.803 (2.362)	22.625 (2.583)	<0.001	0.134 (0.203)	0.203 (0.219)	<0.001
23	21.936 (2.508)	22.825 (2.736)	<0.001	0.132 (0.193)	0.197 (0.209)	<0.001
24	22.066 (2.653)	23.019 (2.887)	<0.001	0.129 (0.183)	0.191 (0.198)	<0.001
25	22.195 (2.793)	23.207 (3.034)	<0.001	0.127 (0.174)	0.185 (0.188)	<0.001
26	22.320 (2.929)	23.390 (3.176)	<0.001	0.124 (0.164)	0.179 (0.177)	<0.001
27	22.443 (3.058)	23.566 (3.312)	<0.001	0.122 (0.155)	0.174 (0.167)	<0.001
28	22.564 (3.181)	23.737 (3.441)	<0.001	0.119 (0.145)	0.168 (0.156)	<0.001
29	22.682 (3.296)	23.901 (3.562)	<0.001	0.117 (0.136)	0.162 (0.146)	<0.001
30	22.798 (3.404)	24.060 (3.674)	<0.001	0.114 (0.127)	0.156 (0.136)	<0.001
31	22.911 (3.503)	24.213 (3.778)	<0.001	0.112 (0.118)	0.150 (0.126)	<0.001
32	23.022 (3.595)	24.360 (3.873)	<0.001	0.109 (0.109)	0.144 (0.116)	<0.001
33	23.130 (3.678)	24.501 (3.959)	<0.001	0.107 (0.100)	0.138 (0.107)	<0.001
34	23.236 (3.752)	24.636 (4.036)	<0.001	0.105 (0.092)	0.132 (0.097)	<0.001
35	23.339 (3.818)	24.766 (4.103)	<0.001	0.102 (0.084)	0.126 (0.089)	<0.001
36	23.440 (3.874)	24.889 (4.160)	<0.001	0.100 (0.076)	0.120 (0.081)	<0.001
37	23.538 (3.922)	25.006 (4.207)	<0.001	0.097 (0.070)	0.115 (0.073)	<0.001
38	23.634 (3.961)	25.118 (4.245)	<0.001	0.095 (0.064)	0.109 (0.067)	<0.001
39	23.727 (3.991)	25.224 (4.272)	<0.001	0.092 (0.059)	0.103 (0.062)	<0.001
40	23.818 (4.012)	25.324 (4.290)	<0.001	0.090 (0.056)	0.097 (0.059)	<0.001
41	23.907 (4.024)	25.417 (4.298)	<0.001	0.087 (0.055)	0.091 (0.058)	0.021
42	23.993 (4.027)	25.505 (4.296)	<0.001	0.085 (0.055)	0.085 (0.058)	0.492
43	24.076 (4.021)	25.588 (4.284)	<0.001	0.082 (0.058)	0.079 (0.061)	0.363
44	24.157 (4.007)	25.664 (4.262)	<0.001	0.080 (0.061)	0.073 (0.066)	0.018
45	24.235 (3.983)	25.734 (4.230)	<0.001	0.077 (0.067)	0.067 (0.072)	<0.001
46	24.311 (3.951)	25.799 (4.189)	<0.001	0.075 (0.073)	0.061 (0.079)	<0.001
47	24.385 (3.910)	25.857 (4.137)	<0.001	0.072 (0.080)	0.056 (0.087)	<0.001
48	24.456 (3.861)	25.910 (4.077)	<0.001	0.070 (0.087)	0.050 (0.096)	<0.001
49	24.525 (3.803)	25.957 (4.007)	<0.001	0.067 (0.096)	0.044 (0.105)	<0.001
50	24.591 (3.738)	25.997 (3.927)	<0.001	0.065 (0.104)	0.038 (0.114)	<0.001

NGLY = Normoglycaemia; HGLY = hyperglycaemia

Supplement Table S7. Characteristics of participants among three classes by hyperglycemia

Variable	Normoglycemia (n=6318)				Hyperglycemia (n=971)			
	Low increasing	Medium increasing	High increasing	P- value	Low increasing	Medium increasing	High increasing	P- value
N	4594	1559	165		542	355	74	
Age, year	32.1 (7.0)	31.21 (6.7)	29.5 (6.4)	<0.001	28.9 (5.8)	28.9 (5.5)	29.3 (6.0)	0.844
Female, n (%)	2388 (52.0)	820 (52.6)	91 (55.2)	0.683	281 (51.8)	158 (44.5)	35 (47.3)	0.096
BMI, kg/m ²	20.8 (1.9)	23.7 (2.5)	27.8 (2.7)	<0.001	20.9 (1.9)	28.4 (3.2)	23.7 (2.4)	<0.001
Smoker, n (%) †	563 (29.7)	276 (33.3)	30 (26.5)	0.103	115 (34.7)	100 (39.4)	20 (31.7)	0.378
Drinker, n (%) †	122 (7.3)	75 (10.4)	11 (10.7)	0.028	27 (8.9)	7 (12.3)	29 (12.9)	0.314
Follow-up years	11.0 (5.3)	11.2 (5.2)	10.1 (4.4)	0.016	13.8 (5.4)	12.1 (5.3)	13.5 (5.2)	0.040
FPG*, mmol/L	4.8 (0.4)	4.8 (0.4)	4.87 (0.6)	0.691	5.6 (1.2)	6.4 (2.0)	5.9 (1.7)	<0.001
Hb1Ac*, %	5.2 (0.4)	5.2 (0.3)	5.4 (0.3)	0.337	5.7 (0.7)	6.1 (1.1)	6.3 (1.3)	<0.001

Data are means ± SD, or n (%). BMI = body mass index, FPG = fasting plasma glucose, Hb1Ac = Hemoglobin A1c. All characteristics are baseline characteristics unless otherwise indicated. *Follow-up information. † Rate was calculated after removing missing value

Supplement Table S8. Hazard ratios and 95% CIs of BMI trajectory groups for incident pre-diabetes and diabetes

Variable	Pre-diabetes			Diabetes		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Trajectory groups						
Low increasing	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Medium increasing	1.55 (1.33, 1.79)	1.29 (1.09, 1.52)	1.18 (0.94, 1.47)	2.19 (1.63, 2.94)	1.40 (1.00, 1.97)	1.52 (0.88, 2.61)
High increasing	2.91 (2.15, 3.95)	1.76 (1.22, 2.55)	1.39 (0.89, 2.15)	7.46 (4.93, 11.30)	2.45 (1.32, 4.56)	2.28 (0.93, 5.58)
Covariates						
Age*	-	1.05 (1.03, 1.07)	1.07 (1.05, 1.09)	-	1.15 (1.10, 1.19)	1.15 (1.10, 1.20)
Female	-	0.91 (0.79, 1.04)	0.89 (0.71, 1.12)	-	0.59 (0.44, 0.78)	0.34 (0.20, 0.57)
BMI*	-	1.08 (1.05, 1.12)	1.06 (1.02, 1.11)	-	1.20 (1.12, 1.28)	1.20 (1.09, 1.33)
Smoker	-	-	1.23 (0.97, 1.57)	-	-	0.69 (0.41, 1.14)
Drinker	-	-	0.90 (0.65, 1.25)	-	-	1.12 (0.64, 1.94)

Data are hazard ratios (95% CIs). Model 1 unadjusted for any covariates. Model 2 adjusted for baseline age, gender, BMI. Model 3 also adjusted for baseline smoking and alcohol drinking. Low increasing group is the referent group. CI = confidence interval; BMI = body mass index.

Supplement Table S9. Linear correlation between model-estimated levels and linear slopes by age

Age	Total		Male		Female	
	r	P-value	r	P-value	r	P-value
20	0.64	<0.001	0.66	<0.001	0.62	<0.001
21	0.69	<0.001	0.71	<0.001	0.67	<0.001
22	0.73	<0.001	0.74	<0.001	0.71	<0.001
23	0.76	<0.001	0.77	<0.001	0.74	<0.001
24	0.78	<0.001	0.79	<0.001	0.76	<0.001
25	0.79	<0.001	0.81	<0.001	0.78	<0.001
26	0.81	<0.001	0.82	<0.001	0.79	<0.001
27	0.81	<0.001	0.83	<0.001	0.80	<0.001
28	0.82	<0.001	0.83	<0.001	0.81	<0.001
29	0.82	<0.001	0.83	<0.001	0.81	<0.001
30	0.82	<0.001	0.83	<0.001	0.81	<0.001
31	0.81	<0.001	0.83	<0.001	0.80	<0.001
32	0.80	<0.001	0.82	<0.001	0.79	<0.001
33	0.79	<0.001	0.80	<0.001	0.77	<0.001
34	0.76	<0.001	0.78	<0.001	0.75	<0.001
35	0.73	<0.001	0.75	<0.001	0.71	<0.001
36	0.68	<0.001	0.70	<0.001	0.67	<0.001
37	0.62	<0.001	0.64	<0.001	0.60	<0.001
38	0.53	<0.001	0.56	<0.001	0.52	<0.001
39	0.42	<0.001	0.44	<0.001	0.40	<0.001
40	0.28	<0.001	0.30	<0.001	0.27	<0.001
41	0.13	<0.001	0.14	<0.001	0.11	<0.001
42	-0.04	<0.001	-0.03	0.061	-0.05	0.004
43	-0.19	<0.001	-0.19	<0.001	-0.20	<0.001
44	-0.33	<0.001	-0.33	<0.001	-0.33	<0.001
45	-0.43	<0.001	-0.44	<0.001	-0.43	<0.001
46	-0.52	<0.001	-0.53	<0.001	-0.51	<0.001
47	-0.58	<0.001	-0.59	<0.001	-0.57	<0.001
48	-0.62	<0.001	-0.63	<0.001	-0.62	<0.001
49	-0.66	<0.001	-0.67	<0.001	-0.65	<0.001
50	-0.68	<0.001	-0.69	<0.001	-0.67	<0.001

Supplement Table S10. Odds ratios and 95% CIs for incident hyperglycemia

Quartiles	Total AUC [†]		Baseline AUC [†]		Incremental AUC [‡]	
	OR (95% CI)	<i>P</i> -value	OR (95% CI)	<i>P</i> -value	OR (95% CI)	<i>P</i> -value
Quartile 1	Reference		Reference		Reference	
Quartile 2	1.41 (1.05,1.89)	0.020	1.31 (0.98,1.74)	0.067	0.97 (0.71,1.33)	0.865
Quartile 3	1.63 (1.23,2.17)	<0.001	1.32 (0.99,1.77)	0.059	1.20 (0.85,1.68)	0.304
Quartile 4	2.25 (1.73,2.94)	<0.001	2.10 (1.62,2.76)	<0.001	1.31 (1.03,1.66)	0.026

Quartile 1 group is the reference.

[†] adjusted for age, sex, smoking and alcohol drinking.

[‡] adjusted for age, sex, smoking, alcohol drinking and baseline AUC. OR = odds ratio; CI = confidence interval.

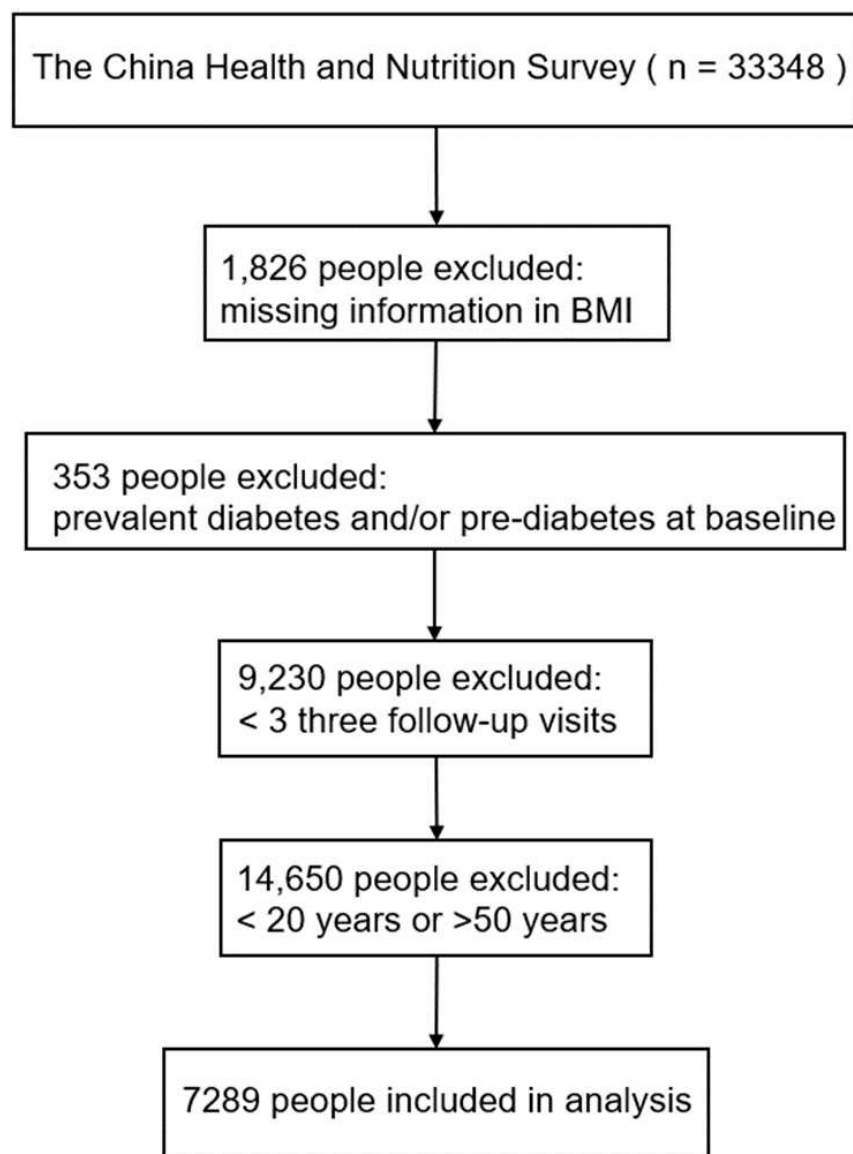


Figure S1. Flow chart of the study population selection

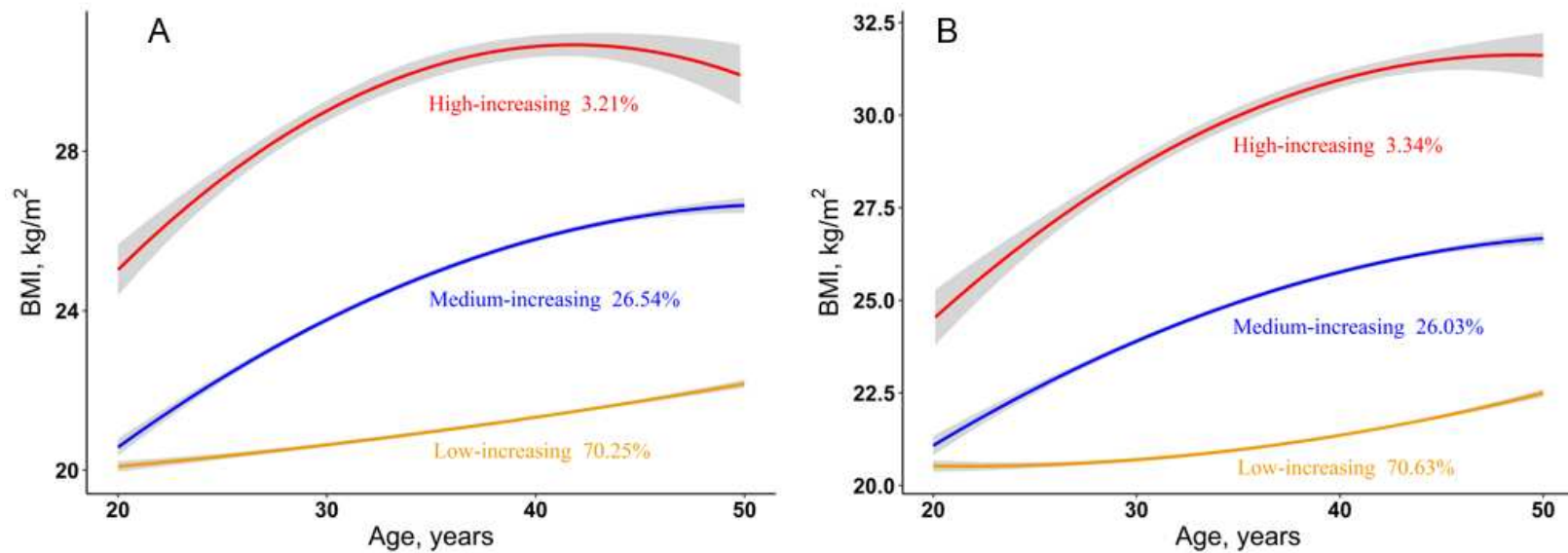


Figure S2. Predicted trajectories of BMI during young adulthood

The trajectories were shown in solid lines, and the 95% confidence intervals (CIs) were shown in shadow. The proportions in each trajectory were shown below solid lines. Fig.A for males, Fig.B for females.

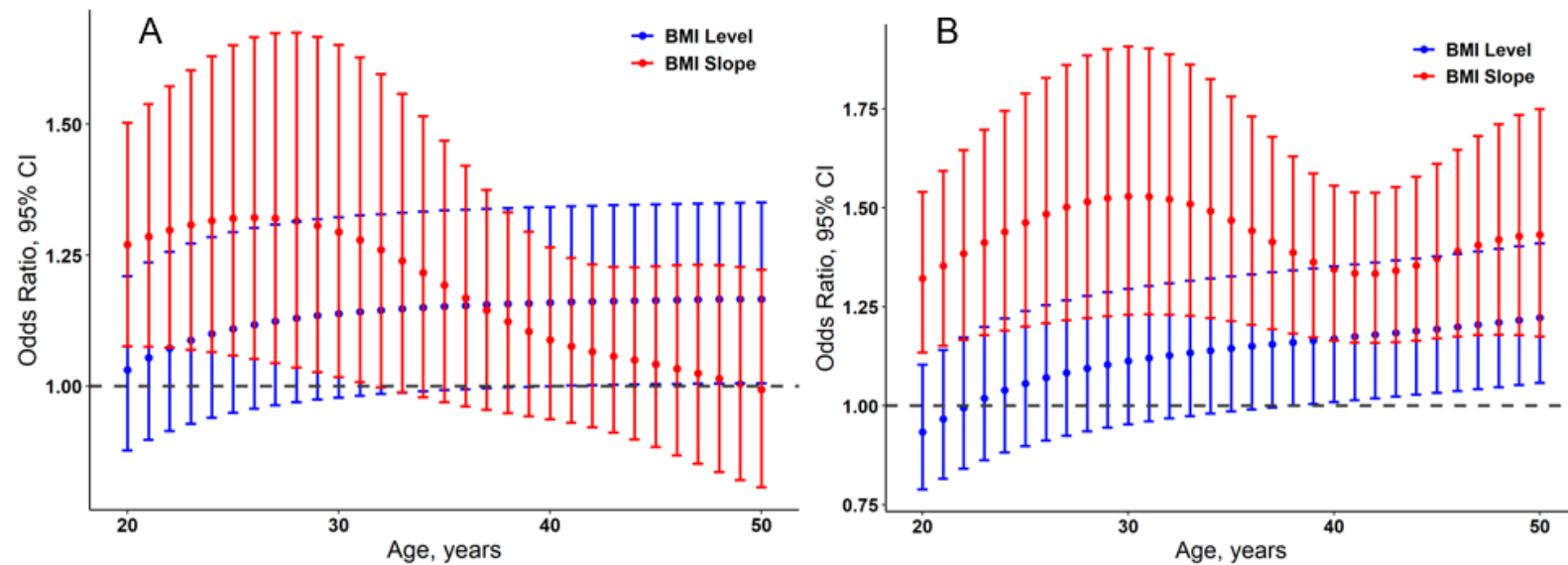


Figure S3. Standardized Odds ratios (ORs) and 95% confidence intervals (CIs) of model-estimated levels and level-adjusted linear slopes of BMI during young adulthood by age for incident hyperglycaemia, adjusted for age, gender, smoking, alcohol drinking. Fig.A for males, Fig.B for females.

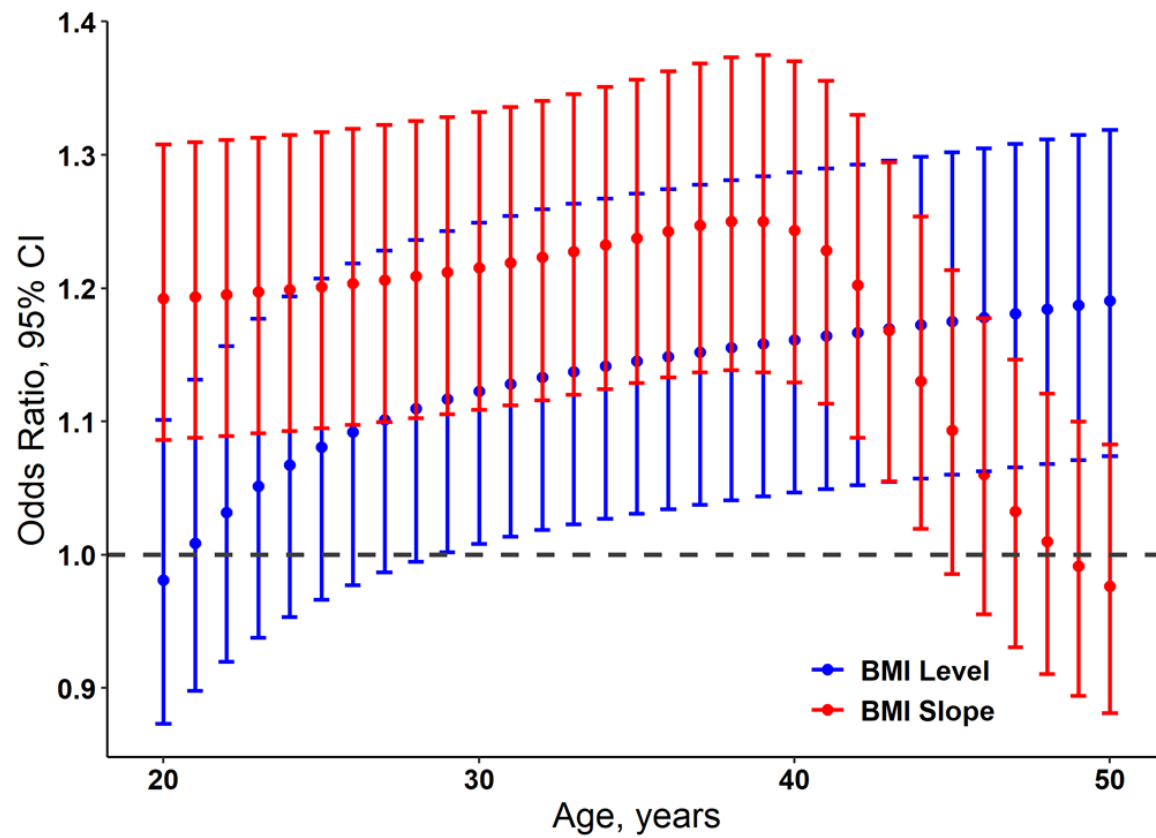


Figure S4. Standardized Odds ratios (ORs) and 95% confidence intervals (CIs) of model-estimated slopes and levels of BMI during young adulthood by age for incident hyperglycaemia, adjusted for age, gender, smoking, alcohol drinking.