

Supplementary materials

Table S1 Features and reference levels listed in the logistic, random forest and XGBoost models

<b>Independent variables</b>	<b>Format</b>	<b>Classifications</b>
<b>Age</b>	Grouped	18- (reference)
		30-
		40-
		50-
		60-
		≥70
<b>Sex</b>	Grouped	Women (reference)
		Men
<b>BMI</b>	Grouped	Under/normal weight (reference)
		Overweight
		Obesity
<b>Waist circumference</b>	Continuous	Not applicable (NA)
<b>Cigarette smoking</b>	Grouped	Never (reference)
		Quit smoke
		Current smoke
<b>Alcohol drink</b>	Grouped	Never (reference)
		Ever
<b>Fiber-rich food intake</b>	Grouped	Never
		Occasionally
		Often (reference)
<b>Eating speed</b>	Grouped	Very fast
		A little bit fast
		Normal speed
		Slow (reference)
<b>Beverage preference</b>	Grouped	Boiled water (reference)
		Tea and coffee
		Sugary drinks
		Coffee
		Green tea
		Other kind of teas
<b>Breakfast frequency</b>	Grouped	Never
		1-2 times per week,
		3-5 times per week,
		Every day (reference)
<b>Dairy products intake</b>	Grouped	Never or less than 3 times per week
		3-5 times per week,
		6-7 times per week
		> 7 times per week (reference)
<b>The frequency of stay up late</b>	Grouped	Often
		Occasionally

		Never (reference)
<b>Daily cell phone use time</b>	Grouped	< 2h
		2-5 h
		More than 5h
<b>Hypertension</b>	Grouped	Yes
		No (reference)