

Interaction of depression and unhealthy diets on the risk of cardiovascular diseases and all-cause mortality in Chinese population:

A PURE cohort substudy

(Supplementary Materials)

Supplementary Section S1A: PURE Study Participant Selection Methodology

as Excerpted from Teo et al. Am Heart J. 2009 Jul; 158 (1):1-7

Selection of Communities

Within each country, urban and rural communities were selected based on broad guidelines (see Guidelines for Selection of Countries, Communities, Households, and Individuals Recruited to PURE). A common definition for “community” that is applicable globally is difficult to establish. In PURE, a community was defined as a group of people who have common characteristics and reside in a defined geographic area. A city or large town was not usually considered a single community, rather communities from low-, middle-, and high-income areas were selected from sections of the city and the community area defined according to a geographical measure (e.g., a set of contiguous postal code areas or a group of streets or a village). The primary sampling unit for rural areas in many countries was the village. The reason for inclusion of both urban and rural communities is that for many countries, urban and rural environments exhibit distinct characteristics in social and physical environment, and hence, by sampling both, we ensured considerable variation in societal factors across PURE communities.

The number of communities selected in each country varied, with the aim to recruit communities with substantial heterogeneity in social and economic circumstances balanced against the capacity of local investigators to maintain follow-up. In some countries (e.g., India, China, Canada, and Colombia), communities from several states/provinces were included to capture regional diversity, in policy, socioeconomic status, culture, and physical environment.

Selections of Households and Individuals

Within each community, sampling was designed to achieve a broadly representative sample of that community of adults aged between 35 and 70 years (see Table below for general recommendations for countries in PURE). The choice of sampling frame within each center was based on both “representativeness” and feasibility of long-term follow-up, following broad study guidelines. Once a community was identified, where possible, common and standardized approaches were applied to the enumeration of households, identification of individuals, recruitment procedures, and data collection.

The method of approaching households differed between regions. For example, in rural areas of India and China, a community announcement was made to the village through contact of a community leader, followed by in-person door-to-door visits of all households. Households were eligible if at least 1 member of the household was between the ages of 35 and 70 years and the household members intended to continue living in their current home for a further 4 years.

For each approach, at least 3 attempts at contact were made. All individuals within these households between 35 and 70 years providing written informed consent were enrolled. When an eligible

household or eligible individual in a household refused to participate, demographics and self-reported data about CVD risk factors, education, and history of CVD, cancers and deaths in the households within the 2 previous years were recorded.

To ensure standardization and high data quality, we used a comprehensive operation manual, training workshops, DVDs, regular communication with study personnel and standardized report forms. We entered all data in a customized database programmed with range and consistency checks, which was transmitted, electronically to the Population Health Research Institute in Hamilton (Ontario, Canada) where further quality checks were implemented.

Guidelines for Selection of Countries, Communities, Households, and Individuals Recruited to PURE

Countries
1. High-income countries, middle-income countries, and low-income countries, with the bulk of the recruitment from low- and middle-income regions.
2. Committed local investigators with experience in recruiting for population studies.
Communities
1. Select both urban and rural communities. Use the national definition of the country to determine urban and rural communities.
2. Select rural communities that are isolated (distance of >50 km or lack easy access to commuter transportation) from urban centers. However, consider ability to process bloods samples, e.g., villages in rural developing countries should be within 45-min drive of an appropriate facility.
3. Define community to a geographical area, e.g., using postal codes, catchment area of health service/clinics, census tracts, areas bordered by specific streets or natural borders such as a river bank.
4. Consider feasibility for long-term follow-up, e.g., for urban communities, choose sites that have a stable population such as residential colonies related to specific work sites in developing countries. In rural areas, choose villages that have a stable population. Villages at greater distance from urban centers are less susceptible to large migration to urban centers.
5. Enlist a community organization to facilitate contact with the community, eg, in urban areas, large employers (government and private), insurance companies, clubs, religious organizations, clinic or hospital service regions. In rural areas, local authorities such as priests or community elders, hospital or clinic, village leader, or local politician.
Individual
1. Broadly representative sampling of adults 35 to 70 years within each community unit.
2. Consider feasibility for long-term follow-up when formulating community sampling framework, e.g., small percentage random samples of large communities may be more difficult to follow-up because they are dispersed by distance. In rural areas of developing countries that are not connected by telephone, it may be better to sample entire community (i.e., door-to-door systematic sampling).

3. The method of approach of households/individuals may differ between sites. In MIC and HIC, mail, followed up by phone contact may be the practical first means of contact. In LIC, direct household contact through household visits may be the most appropriate means of first contact.
4. Once recruited, all individuals are invited to a study clinic to complete standardized questionnaires and have a standardized set of measurements.

Specific recruitment strategies for China are summarized below.

Selection of Provinces: In China, we selected several provinces to capture regional diversity, in policy, socioeconomic status, culture, and physical environment. The choice and number of provinces selected in PURE China reflects a balance between involving a large number of communities in provinces at different economic levels, with substantial heterogeneity in social and economic circumstances and policies, and the feasibility of centres to successfully achieve long-term follow-up. Thus, PURE China study included sites in which investigators are committed to collecting good-quality data over the planned 10-year follow-up period and did not aim for a strict proportionate sampling of the entire country. We included 12 provinces of mainland China to capture regional diversity, policy, socioeconomic status, culture, and physical environment. Based on the national criteria for coordinated regional development according to their socioeconomic development status at the time of inclusion, these provinces were categorized into Eastern (Beijing, Jiangsu, Liaoning, Shandong), Central (Shanxi, Jiangxi), and Western regions (Shaanxi, Sichuan, Yunnan, Inner Mongolia, Qinghai, Xinjiang).

Selection of Communities: Within each province, urban and rural communities were selected using pre-specified criteria to recruit participants with an almost 1:1 urban-to-rural proportion. A total of 115 communities were included in the PURE China study.

Selection of Households and Individuals: Within each community, households were approached by a community announcement through contact of the community organization, followed by in-person door-to-door visits of all households. Households were eligible if at least 1 member of the household was between the ages of 35 and 70 years and the household members intended to continue living in their current home for a further 4 years. All individuals within these households between 35 and 70 years providing written informed consent were enrolled.

Supplementary Section S1B: Data collection

Detailed information on data collection in PURE study has been published previously (Teo et al. Am Heart J. 2009 Jul; 158 (1):1-7). Baseline information was collected using standardized questionnaires and physical measurements. The Adult Questionnaire for adult participants collects data on the 9 INTERHEART risk factors (lipids, smoking, hypertension, diabetes, abdominal obesity, psychosocial factors [stress and symptoms of depression], consumption of fruits, and vegetables, consumption of alcohol, and regular physical activity). The physical examination includes 2 measures of resting blood pressure (sitting), anthropometric measures (weight, height, waist, and hip), spirometry (forced expiratory volume in 1 second, forced expiratory vital capacity), electrocardiogram, and hand grip strength (three measurements from both hands). A 10-mL fasting blood sample is collected from all consenting participants. Blood samples are centrifuged and transferred to centralized long-term storage in secure -70°C freezers or large -180°C liquid nitrogen tanks for future biochemical and genetic testing. In addition, detailed information on physical activity and diet is collected using the International Physical Activity Questionnaire (IPAQ) and country-specific semiquantitative Food Frequency Questionnaires (FFQs). Data on PURE study households are collected using the Household Questionnaire obtains information from the household head on house structure, amenities, access to water, and sanitation.

Quality of data collection is maintained through the use of standardized protocols and centralized training. Key staff from each centre attend an initial training session (standardization across regions is provided by Project Office staff attending these training sessions) and in turn, train local staff using centrally created manuals and training videos. These staff are trained, tested on “mock” subjects, and certified. If the variation between the staff compared to the coordinator is unacceptable, the staff member is retrained. All data are entered at each site electronically into a customized database programmed with range and consistency checks and transmitted electronically to the Project Office in Hamilton, Ontario, Canada.

Supplementary Section S1C: Standardized Event Definitions in PURE

Prospective Follow-up for Cardiovascular Events and Mortality: Information on specific events (e.g. death, myocardial infarction, stroke, heart failure, cancer, hospitalizations, new diabetes, injury, tuberculosis, human immunodeficiency viral infections, malaria, pneumonia, asthma, chronic obstructive pulmonary disease) were obtained from participants or their family members (events were reported by the participants if alive or by a relative if the individual had died). Because the PURE study involves urban and rural areas from middle- and low-income countries, supporting documents to confirm the cause of death and/or event varied in degrees of completion and availability. In most of middle- and low-income countries there was no central system of death or event registration. Therefore, information was obtained about prior medical illness and medically certified cause of death where available, and, second, best available information was captured from reliable sources in those instances where medical information was not available in order to be able to arrive at a probable diagnosis or cause of death. Event documentation was based on information from household interviews and medical records, death certificates and other sources. Verbal autopsies were also used to ascertain cause of death in addition to medical records which were reviewed by a health professional. This approach has been used in several studies conducted in middle- and low-income countries.

The following definitions and codes were used to classify CVD events and deaths in the study:

CVD Definitions and Coding

FATAL EVENTS

Cardiovascular Death – Definitions

01.00 DEATH DUE TO CARDIOVASCULAR EVENTS

01.10 Sudden unexpected Cardiovascular Death (SCVD)

Without evidence of other cause of death, death that occurred suddenly and unexpectedly (examples: witnessed collapse, persons resuscitated from cardiac arrest who later died) or persons seen alive less than 12 hours prior to discovery of death (example persons found dead in his/her bed).

01.10 SCVD is either definite, probable or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.11: Definite	One of the following in persons with: <ul style="list-style-type: none">• known cardiovascular disease, or• diabetes with an additional risk factor such as hypertension, smoking, dyslipidemia, micro albuminuria, serum creatinine 50% above upper limit of normal, or• 3 of the above risk factors, or• 2 of the above risk factors in men aged 60 and more and women aged 65 and more	

01.12: Probable	One of the following in persons with: <ul style="list-style-type: none">• diabetes, or• 2 of the above risk factors in men aged less than 60 and in women less than 65, or• one of the above risk factor in men aged 60 and more and in women aged 65 and more, or• typical of chest pain or sudden severe dyspnea of less than 20-minute duration preceding the event	No ICD-10 Code
01.13: Possible	In persons without risk factor	

For SCVD, the patient was well or had a stable CVD (example stable angina) when last seen alive. The event of a sudden death occurring during the hospitalization of MI is considered a fatal MI and not sudden death.

01.3 Fatal Myocardial Infarction (MI)

Symptoms of Myocardial Infarction:

Typical symptoms or suggestive symptoms of MI according to physician are characterized by severe anterior chest pain as tightness, crushing, burning, lasting at least 20 minutes, occurring at rest, or on exertion, that may radiate to the arms or neck or jaw and may be associated with dyspnea, diaphoresis and nausea. However, death associated with nausea and vomiting with or without chest pain not due to another cause may be considered as possible MI if ECG and cardiac markers are not done. These symptoms may have occurred the last month before death.

Fatal myocardial infarction is either definite, probable or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.31: Definite	1. Autopsy demonstrating fresh myocardial infarction and/or recent coronary occlusion, or 2. ECG showing new and definite sign of MI (Minnesota code 1-1-1) or 3. Symptoms typical or atypical or inadequately described but attributed to cardiac origin lasting at least 20 minutes and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN 4. ECG with new ischemic changes (new ST elevation/depression or T wave inversion ≥ 2 mm) and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN	
01.32: Probable	1. ECG with sign of probable MI (Minnesota code 1-2-1), or	

	<ol style="list-style-type: none"> Typical symptoms lasting at least 20 minutes considered of cardiac origin, with only new ST-T changes (new ST elevation/depression or T wave inversion ≥ 1 but < 2mm) without documented increased cardiac markers or enzyme as in PURE definition 1.31 (above), or Increased cardiac enzymes as in PURE definition 1.31 (above) showing a typical pattern of MI as above without symptoms or significant ECG changes 	I21- I22
01.33: Possible	<ol style="list-style-type: none"> ECG with sign of possible MI (Minnesota code 1-3-1) or Typical symptoms or symptoms suggestive of MI according to the physician lasting at least 20 minutes without documented ECG or cardiac marker. 	

The Minnesota codes for MI is taken from Rose and Blackburn and published in their book “Evaluation Methods of Cardiovascular Disease WHO 1969”.

- Definite MI is Q/R ratio $\geq 1/3$ and Q duration ≥ 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-1-1)
- Probable MI is Q/R ratio $\geq 1/3$ and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-2-1)
- Possible MI is Q/R ratio between $1/5$ and $1/3$ and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-3-1)

01.40 Fatal Stroke

Fatal stroke is either definite or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.41: Definite	<p>Stroke death is defined as death within 30 days from an acute focal neurological deficit <i>diagnosed by a physician</i> and thought to be of vascular origin (without other cause such as brain tumor) with signs and symptoms lasting ≥ 24 hrs.</p> <p>Stroke death is also considered if death occurred within 24 hrs. of onset of persisting signs and symptoms, or if there is evidence of a recent stroke on autopsy.</p> <p>N.B.</p> <ul style="list-style-type: none"> In a subject with a stroke ≤ 30 days: If death occurred with a pneumonia due to possible aspiration, death will be considered to be due to stroke. In a subject with a stroke > 30 days: If death occurred with a pneumonia due to possible aspiration, the adjudicator will make a 	

	<p>decision according to his/her clinical judgment if death is related to stroke or not.</p> <ul style="list-style-type: none"> • Subarachnoid hemorrhage death manifested by sudden onset headache with/without focal signs and imaging (CT or MRI) evidence of bleeding primarily in the subarachnoid space is considered a fatal stroke in absence of trauma or brain tumor or malformation • Subdural hematoma death is not considered as a stroke death and may be related to previous trauma or other cause. 	I60- I64, I69
01.43: Possible	Death in a participant with a history of sudden onset of focal neurological deficit of one or more limbs, loss of vision or slurred speech lasting about 24 hours.	

01.50 Fatal Congestive Heart Failure

Fatal congestive heart failure is either definite or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.51: Definite	<p>The diagnosis of congestive heart failure may be an autopsy finding in absence of other cause or requires signs (rales, increased jugular venous pressure or ankle edema) or symptoms (nocturnal paroxysmal dyspnea, dyspnea at rest or ankle edema) of congestive heart failure and one or both of the following:</p> <ul style="list-style-type: none"> • radiological signs of pulmonary congestion, • treatment of heart failure with diuretics <p><i>If sudden death occurred in a patient with chronic severe heart failure, it should be adjudicated as fatal congestive heart failure.</i></p>	I50
01.52: Probable	Progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough in a person with known hypertension or a history of previous MI/angina or other heart disease	
01.53: Possible	Progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough	

01.60 Death Due to Other Cardiovascular Deaths (*other causes [1.10 to 1.50 above] having been excluded*)

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.61	Arterial rupture of aneurysm	I71- I72
01.62	Pulmonary embolism <i>NOTE: Death associated with pulmonary embolism occurring within 2 weeks after a fracture such as hip, femur should attributed to death due to injury. Refer to Injury, Section 6.0</i>	I26
01.63	Arrhythmic death (A-V block, sustained ventricular tachycardia in absence of other causes)	I44- I45, I47- I49
01.64	Death after invasive cardiovascular intervention: a perioperative death extending to 30 days after coronary or arterial surgical revascularization and to 7 days after a coronary or arterial percutaneous dilatation (angioplasty) with or without a stent or an invasive diagnostic procedure.	I97
01.65	Congenital heart disease	Q20-Q28
01.66	Heart valve disease (including rheumatic heart disease)	I01, I05- I09, I34- I37
01.67	Endocarditis	I33, I38
01.68	Myocarditis	I40
01.69	Tamponade (pericarditis)	I30, I31, I32
01.70	Other cardiovascular events (<i>Excluding 1.61 to 1.69 above</i>) <i>Valid ICD-10 codes would include the following:</i> <i>I11, I12, I13, I23, I24, I25, I27, I28, I42, I51, I52, I65-I68, I73, I74, I96, I98, I99 (Refer to ICD-10 Listing for associated definitions for each code)</i>	Any valid 'I' (Cardiovascular) ICD-10 code that can be classified as underlying cause of death, not specified above

NON-FATAL EVENTS

Cardiovascular Events – Definitions

10.00 NON-FATAL CARDIOVASCULAR EVENTS

10.10 Non-Periprocedural Myocardial Infarction (MI)

MI is considered either definite, probable or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.11: Definite	<ol style="list-style-type: none">1. ECG showing new and definite sign of MI (Minnesota code 1-1-1) or2. Symptoms typical or atypical or inadequately described but attributed to cardiac origin lasting at least 20 minutes and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN3. ECG with new ischemic changes (new ST elevation/depression or T wave inversion ≥ 2 mm) and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN <p>Please note that increased markers may occur in trauma (CK, AST, myoglobin and CK MB to a lesser degree); renal insufficiency, heart failure, pulmonary embolism (troponin), cardioversion (all)</p>	I21-I22
10.12: Probable	<ol style="list-style-type: none">1. ECG with new and probable sign of MI (Minnesota code 1-2-1), or2. Typical symptoms lasting at least 20 minutes considered of cardiac origin, with only new ST-T changes (new ST elevation/depression or T wave inversion ≥ 1 but < 2mm) without documented increased cardiac markers as in PURE definition 10.11 (above), or3. Increased cardiac enzymes showing a typical pattern of MI as above without symptoms or significant ECG changes.	
10.13: Possible	<ol style="list-style-type: none">1. ECG with new and possible sign of MI (Minnesota code 1-3-1), or2. Typical symptoms lasting 20 minutes and more considered to be of cardiac origin without documented ECG or cardiac marker.	

10.20 Periprocedural Myocardial Infarction

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
------------------------------	------------	-------------------------------

10.21: Definite	<ol style="list-style-type: none"> ECG showing new and definite sign of MI (Minnesota code 1-1-1), or Increased cardiac markers within 48 hours of procedure: <ul style="list-style-type: none"> percutaneous coronary intervention: CKMB should be $\geq 5 \times$ ULN or troponin $\geq 5 \times$ above lower level of necrosis OR $> 20\%$ increase in cardiac markers if elevated at the beginning of the procedure in a patient with symptoms suggestive of myocardial ischemia Coronary surgery: Increased cardiac markers CKMB should be $\geq 10 \times$ ULN or troponin $\geq 10 \times$ above lower limit of necrosis. 	I21-I22
-----------------	---	---------

The Minnesota codes for MI is taken from Rose and Blackburn and published in their book “Evaluation Methods of Cardiovascular Disease WHO 1969”.

- Definite MI is Q/R ratio $\geq 1/3$ and Q duration ≥ 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-1-1)
- Probable MI is Q/R ratio $\geq 1/3$ and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-2-1)
- Possible MI is Q/R ratio between $1/5$ and $1/3$ and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-3-1)

10.30 Stroke/Transient Ischemic Attack (TIA)

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.31: Definite	<p>Stroke is defined as an acute focal neurological deficit <i>diagnosed by a physician</i> and thought to be of vascular origin (without other case such as brain tumor) with signs and symptoms lasting ≥ 24 hrs.</p> <p>N.B.</p> <ul style="list-style-type: none"> Subarachnoid hemorrhage manifested by sudden onset headache with/without focal signs and imaging (CT or MRI or lumbar puncture) showing evidence of bleeding primarily in the subarachnoid space is considered a stroke in absence of trauma or brain tumor or malformation Subdural hematoma is not considered as a stroke and may be related to previous trauma or other cause. 	I60-I64, I69
10.33: Possible	Stroke is possible if there is a history of sudden onset of focal neurological deficit of one or more limbs, loss of vision or slurred speech lasting about 24 hours or more	

10.34: TIA	The diagnosis of TIA requires the presence of acute focal neurological deficit thought to be of vascular origin with signs and symptoms lasting less than 24 hours	G45
------------	--	-----

10.40 Congestive Heart Failure

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.41: Definite	<p>The diagnosis of congestive heart failure requires signs (rales, increased jugular venous pressure or ankle edema) or symptoms (nocturnal paroxysmal dyspnea, dyspnea at rest or ankle edema) of congestive heart failure and one or both of the following:</p> <ul style="list-style-type: none"> • radiological signs of pulmonary congestion, • Treatment of heart failure with diuretics. 	I50
10.42: Probable	Progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough in a person with known hypertension or a history of previous MI/angina or other heart disease	
10.43: Possible	Congestive heart failure is considered possible when there is progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough	

Coding of Deaths in the PURE Study

Generic Event Codes	Description	Acceptable FATAL PURE Codes				Acceptable ICD 10 Code Ranges
		Definite	Probable	Possible	Codes for which Definite/ Probable/ Possible not defined	
1.00	Death Due to Cardiovascular Events					
1.10	Sudden unexpected cardiovascular death	1.11	1.12	1.13		No ICD Code
1.30	Fatal Myocardial Infarction	1.31	1.32	1.33		I21-I22
1.40	Fatal Stroke	1.41		1.43		I60-I64, I69
1.50	Fatal Congestive Heart Failure	1.51	1.52	1.53		I50
1.60	Other cardiovascular deaths					
-	- Arterial rupture of aneurysm				1.61	I71-I72
-	- Pulmonary Embolism				1.62	I26
-	- Arrhythmic death (A-V block, sustained ventricular tachycardia in absense of other causes)				1.63	I44-I45, I47-I49
-	- Death after invasive cardiovascular intervention				1.64	I97
-	- Congenital heart disease				1.65	Q20-Q28
-	- Heart valve disease (including RHD)				1.66	I01, I05-I09, I34-I37
-	- Endocarditis				1.67	I33, I38, I39
-	- Myocarditis				1.68	I40, I41
-	- Tamponade (Pericarditis)				1.69	I30-I32
-	- Other cardiovascular events				1.70	I00, I02, I10-I13, I15, I23-I25, I27, I28, I42, I43, I51, I52, I65-I68, I70, I73, I74, I77-I83, I85-I89, I95, I98, I99
2.00	Death due to most frequent infections					
2.10	Typhoid and Paratyphoid	2.11	2.12			A01
2.20	Diarrhoea and Gastroenteritis/Dysentery	2.21	2.22			A00, A02-A09
2.30	Pulmonary Tuberculosis	2.31	2.32	2.33		A15-A16, A19
-	Septicaemia				2.40	A40-A41
2.50	Viral Hepatitis	2.51	2.52			B15-B19
2.60	AIDS	2.61	2.62			B20-B24
2.70	Malaria	2.71	2.72			B50-B54
2.80	Covid-19	2.81	2.82			U07

-	Other Infections				2.90	A17-A18, A20-A39, A42-A99, B00-B09, B25-B49, B55-B99
3.00	Death due to cancer					
-	Mouth				3.01	C00-C06, C10, C12, C14
-	Esophagus				3.02	C15
-	Stomach				3.03	C16
-	Small Intestine				3.04	C17
-	Large intestine including rectum				3.05	C18-C20
-	Pancreas				3.06	C25
-	Liver				3.07	C22, C24
-	Lung/Pleura				3.08	C33-C34, C38-C39, C45
-	Breast				3.09	C50
-	Prostate				3.11	C61
-	Head and Neck				3.13	C31-C32, C70
-	Skin				3.14	C43-C44
-	Multi-site				3.15	C97
-	Other, specify				3.16	C07-C09, C11, C13, C23, C26, C30, C37, C46-C49, C57-C58, C63, C68-C69, C74- C79, C96, D37-D45, D47, D48
-	Cervical				3.17	C53
-	Uterine/Ovarian				3.18	C54-C56
-	Vaginal/vulva				3.19	C51-C52
-	Kidney				3.20	C64, C65
-	Bladder				3.21	C66-C67
-	Anus				3.22	C21
-	Testis				3.23	C62
-	Penis				3.24	C60
-	Brain/spinal cord				3.25	C71-C72
-	Leukemia				3.26	C91-C95, D46
-	Lymphoma				3.27	C81-C86, C88
-	Multiple Myeloma				3.28	C90
-	Musculoskeletal (muscle, bones, tendons, ligaments, joints, cartilage)				3.29	C40-C41
-	Thyroid				3.30	C73
-	Unknown site				3.40	C80
4.00	4.00 Death due to diseases of the respiratory system					

4.10	Pneumonia	4.11	4.12	4.13		J12-J18
4.30	Asthma	4.31	4.32	4.33		J45-J46
4.40	COPD (Obstructive airways disease)	4.41	4.42	4.43		J44
-	Other respiratory diseases				4.90	J00-J11, J20-J43, J47-J99, Q30-Q34
5.00	Death related to pregnancy/delivery/puerperium					
-	Death related to pregnancy/delivery/ puerperium (Direct obstetrical causes)				5.00	O00-O99
6.00	Death due to Injury					
-	Death due to Injury				6.00	V01-V99, W00-W99, X00-X99, Y00-Y98
7.00	Death due to other causes					
-	Diseases of Nervous system				7.10	G00-G99, Q00-Q07
-	Diseases of Digestive System				7.20	K00-K93, Q39-Q45
-	Diseases of genito-urinary system				7.30	N00-N99, Q50-Q64
-	Other diseases, specify				7.50	D50-D89, E00-E90, F00-F99, H58, H59, H94, H95, L00-L99, M00-M99, Q35-Q38, Q65-Q99
8.00	Death due to unspecified/presumed cause					
-	Presumed cardiovascular death				8.10	No ICD Code
-	Presumed cancer death				8.20	No ICD Code
-	Unknown death				8.30	R99

Supplementary Section S1D: Interaction measurements

The interaction effect of depression and unhealthy diet on cardiovascular diseases (CVD) and mortality was measured by additive- and multiplicative-scale. In terms of the additive interaction, the attributable proportion due to interaction [AP; Eq. (1)], relative excess risk due to interaction [RERI; Eq. (2)], and synergy index [SI; Eq. (3)] were used. The AP is the proportion of the risk due to the interaction in the doubly exposed group (null hypothesis: AP = 0). When RERI is positive, it indicates increased risk due to the additive interaction (null hypothesis: RERI = 0). SI could be interpreted as the ratio of an increased risk due to both exposures to the sum of individual increased risks (null hypothesis: SI = 1). These were used to assess whether the risk due to having both exposure is greater than the sum of the risks due to each exposure.

$$AP_{Interaction} = \frac{HR_{Depression\&Diet} - HR_{Depression} - HR_{Diet} + 1}{HR_{Depression\&Diet}} \quad (1)$$

$$RERI_{Interaction} = HR_{Depression\&Diet} - HR_{Depression} - HR_{Diet} + 1 \quad (2)$$

$$SI_{Interaction} = \frac{HR_{Depression\&Diet} - 1}{HR_{Depression} + HR_{Diet} - 2} \quad (3)$$

The multiplicative-scale interaction [Eq. (4)] has been widely used to examine the interaction effects by identifying whether the risk due to having both diseases is greater than the product of the risks due to each disease alone (null hypothesis: multiplicative interaction = 1).

$$Mult_{Interaction} = \frac{HR_{Depression\&Diet}}{HR_{Depression} * HR_{Diet}} \quad (4)$$

In Eqs. (1–4), $HR_{Depression\&Diet}$ represents the hazard ratio (HR) for those who have both depression and unhealthy diet. $HR_{Depression}$ and HR_{Diet} represent the HR for those who have depression or unhealthy diet alone, respectively, compared with individuals who have neither of the conditions.

Supplementary Section S2: Tables and Figures

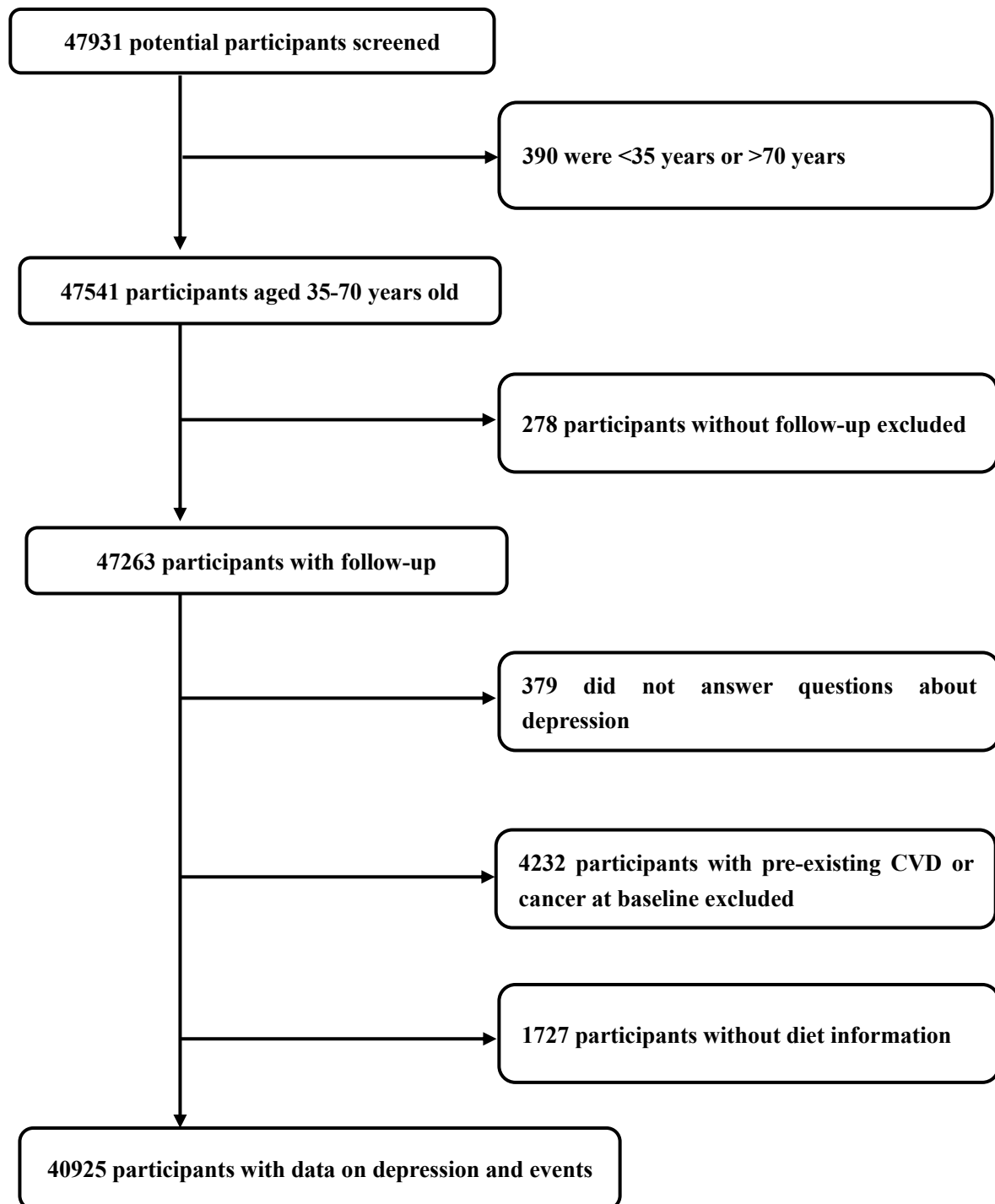


Figure S1. Flow chart of participant selection

Abbreviation: CVD, cardiovascular disease.

Table S1. Baseline information of participants grouped by depression and diet

Variables	Total		Healthier diet		Less healthy diet	
	Without depression (N=39973)	With depression (N=952)	Without depression (N=26635)	With depression (N=667)	Without depression (N=13338)	With depression (N=285)
W:R	0.4 (0.2-1.0)	0.4 (0.1-0.8)	0.5 (0.2-1.2)	0.4 (0.2-1.0)	0.3 (0.1-0.6)	0.2 (0.1-0.4)
P:S	0.7 (0.5-0.9)	0.7 (0.5-1.0)	0.8 (0.6-1.0)	0.8 (0.6-1.2)	0.5 (0.4-0.7)	0.5 (0.4-0.7)
total Fat %	17.6±7.7	17.6±8.4	18.2±7.5	18.0±8.0	16.4±8.0	16.6±9.2
sFat %	5.7±2.7	5.6±2.9	5.6±2.5	5.5±2.6	5.7±3.1	5.9±3.6
Protein %	15.3±2.9	14.9±3.1	15.5±2.8	15.1±3.0	14.8±3.1	14.6±3.4
CHO %	66.8±10.4	67.4±11.4	66.0±10.0	66.8±11.0	68.5±10.9	68.6±12.3
MUFA %	6.8±2.9	6.6±3.1	6.9±2.7	6.6±2.8	6.6±3.3	6.7±3.7
PUFA %	4.2±2.7	4.4±3.2	4.7±3.0	5.0±3.4	3.1±1.7	3.1±2.1
Vegetables Score	3.1±1.0	2.9±1.1	3.2±0.9	2.9±1.0	2.9±1.2	2.7±1.1
Fruits Score	1.9±1.8	2.2±2.2	2.3±2.0	2.6±2.3	1.2±1.1	1.3±1.3
Nuts Score	6.5±3.6	6.4±3.9	8.0±3.0	7.7±3.4	3.6±3.0	3.4±3.3
W:R Score	2.3±2.9	2.5±3.3	2.8±3.2	3.1±3.6	1.3±1.7	1.0±1.7
Fiber Score	6.0±3.7	6.3±3.5	6.8±3.4	7.0±3.2	4.2±3.7	4.7±3.5
P:S Score	6.9±2.2	7.2±2.3	7.5±2.1	7.8±2.1	5.8±1.9	5.6±2.0
Fried foods Score	9.7±1.1	9.8±1.0	9.7±1.0	9.8±0.9	9.7±1.2	9.7±1.1
AHEI Score	36.5±7.2	37.2±7.7	40.4±5.0	41.0±5.6	28.6±3.5	28.4±3.6

Abbreviation: W:R, Ratio of white to red meat; P:S, ratio of polyunsaturated to saturated fatty acids; sFat, saturated fat, CHO, carbohydrate; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; AHEI, Alternative Healthy Eating Index.

Table S2. Alcohol use information of participants grouped by depression and diet

Variables	Total		Healthier diet		Less healthy diet	
	Without depression (N=39973)	With depression (N=952)	Without depression (N=26635)	With depression (N=667)	Without depression (N=13338)	With depression (N=285)
Alcohol use ^a						
Never	29922 (76.1)	692 (73.8)	19881 (75.7)	486 (73.7)	10041 (76.8)	206 (73.8)
Former	1095 (2.8)	60 (6.4)	714 (2.7)	48 (7.3)	381 (2.9)	12 (4.3)
Current Low	6718 (17.1)	140 (14.9)	4588 (17.5)	98 (14.9)	2130 (16.3)	42 (15.1)
Current Moderate	1326 (3.4)	38 (4.1)	897 (3.4)	22 (3.3)	429 (3.3)	16 (5.7)
Current High	281 (0.7)	8 (0.9)	185 (0.7)	5 (0.8)	96 (0.7)	3 (1.1)
Wine ^b						
Current Low	6770 (17.9)	143 (16.0)	4630 (18.3)	99 (15.6)	2140 (17.0)	44 (16.8)
Current Moderate	18 (0.0)	0 (0.0)	12 (0.0)	0 (0.0)	6 (0.0)	0 (0.0)
Current High	13 (0.0)	0 (0.0)	9 (0.0)	0 (0.0)	4 (0.0)	0 (0.0)
Liquor<38% ^c						
Current Low	6915 (18.1)	147 (16.3)	4738 (18.6)	101 (15.9)	2177 (17.2)	46 (17.3)
Current Moderate	249 (0.7)	4 (0.4)	173 (0.7)	2 (0.3)	76 (0.6)	2 (0.8)
Current High	32 (0.1)	0 (0.0)	21 (0.1)	0 (0.0)	11 (0.1)	0 (0.0)
Liquor>38% ^d						
Current Low	7186 (18.6)	160 (17.3)	4905 (19.0)	110 (16.8)	2281 (17.7)	50 (18.5)
Current Moderate	385 (1.0)	13 (1.4)	249 (1.0)	10 (1.5)	136 (1.1)	3 (1.1)
Current High	69 (0.2)	0 (0.0)	42 (0.2)	0 (0.0)	27 (0.2)	0 (0.0)
Beer ^e						
Current Low	2493 (6.2)	81 (8.5)	1726 (6.5)	52 (7.8)	767 (5.8)	29 (10.2)
Current Moderate	441 (1.1)	7 (0.7)	282 (1.1)	5 (0.7)	159 (1.2)	2 (0.7)
Current High	95 (0.2)	6 (0.6)	67 (0.3)	3 (0.4)	28 (0.2)	3 (1.1)

Abbreviation: W:R, Ratio of white to red meat; P:S, ratio of polyunsaturated to saturated fatty acids; sFat, saturated fat, CHO, carbohydrate; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; AHEI, Alternative Healthy Eating Index.

^a Alcohol use: current low, alcohol drinks <1 time/day; currently moderate, alcohol drinks 1-3 times/day (men)/ 1-2 times/day(women); currently high, alcohol drinks >3 times/day (men)/ >2 times/day (women);

^b Wine: Same definition as above, and one time means 125ml intake;

^c Liquor<38%: Same definition as above, and one time means 125ml intake;

^d Liquor>38%: Same definition as above, and one time means 125ml intake;

^e Beer: Same definition as above, and one time means 375ml intake;

Table S3. Event rates and associations between diet and the outcomes

Variables	Healthy diet (N=27302)	Unhealthy diet (N=13623)
Primary outcomes		
Composite outcome^a		
Events N (%)	4317 (10.8)	122 (12.8)
Event rate/1000 person-years (95% CI)	10.2 (9.9, 10.5)	12.2 (10.0, 14.3)
Model 1	ref.	1.19 (1.12, 1.27)
Model 2	ref.	1.09 (1.02, 1.17)
Model 3	ref.	1.10 (1.03, 1.18)
Secondary outcomes		
Major CVD		
Events N (%)	3022 (7.6)	77 (8.1)
Event rate/1000 person-years (95% CI)	7.2 (6.9, 7.4)	7.7 (6.0, 9.4)
Model 1	ref.	1.16 (1.07, 1.26)
Model 2	ref.	1.09 (1.01, 1.18)
Model 3	ref.	1.10 (1.02, 1.20)
All-cause mortality		
Events N (%)	1998 (5.0)	68 (7.1)
Event rate/1000 person-years (95% CI)	4.6 (4.4, 4.8)	6.6 (5.0, 8.2)
Model 1	ref.	1.24 (1.13, 1.37)
Model 2	ref.	1.06 (0.96, 1.17)
Model 3	ref.	1.07 (0.97, 1.19)

^a Composite outcome was defined by the first of either a major cardiovascular event or death. In model 1, we adjusted for baseline age, gender and centre as the random effect. Model 2 was further adjusted for urban/rural region, physical activities, education level, wealth index, social isolation index and self-reported disabilities score. In model 3, we further adjusted baseline confounders, including current smoking, alcohol use, hypertension, diabetes, dyslipidemia, central obesity and use of statins.

Abbreviations: HR, hazard ratio; CI, confidence interval; CVD, cardiovascular disease.

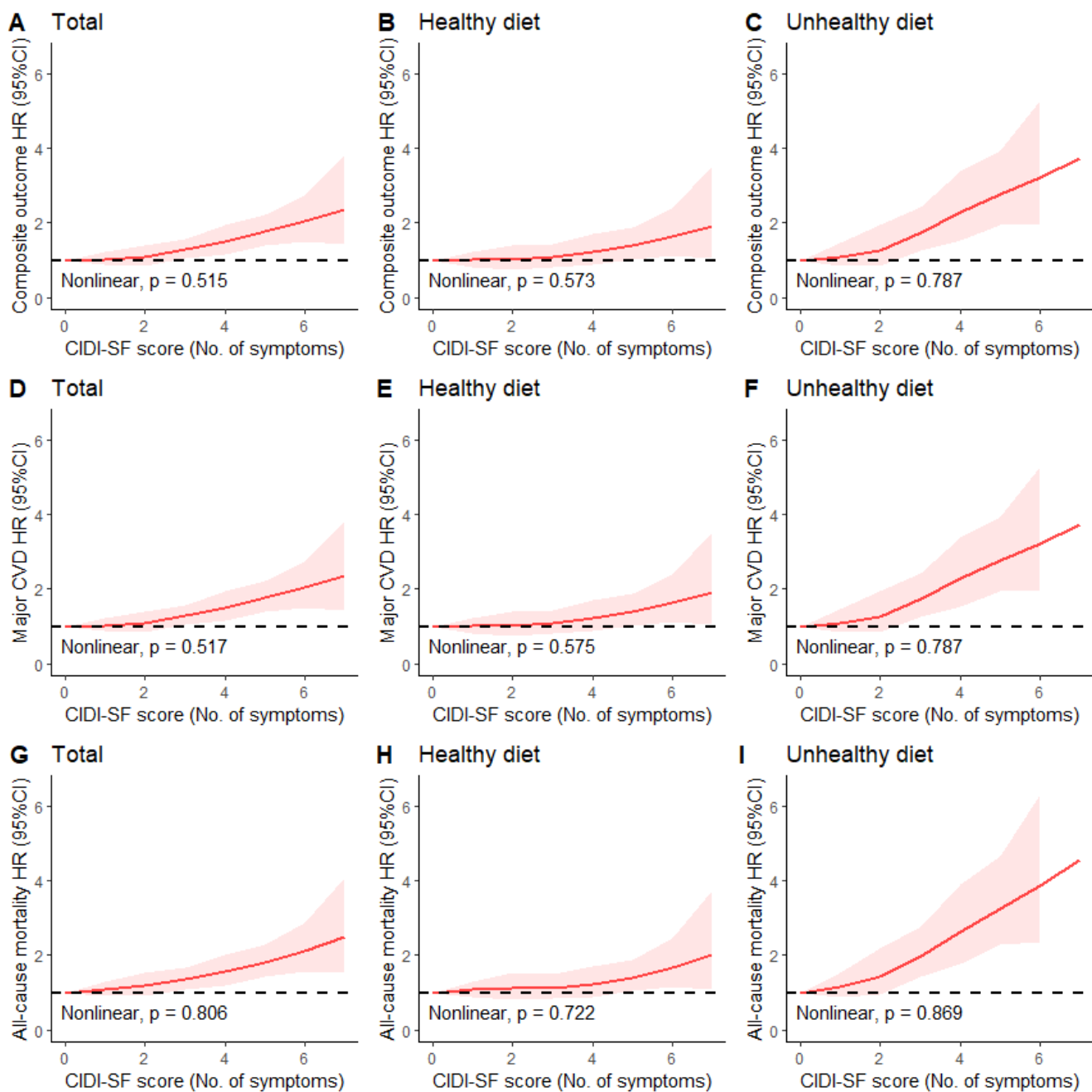


Figure S2. Restricted spline curve for the associations between depression and the primary outcomes by diet

The models were adjusted for age, gender, urban/rural regions, physical activities, education level, wealth index, social isolation index, self-reported disabilities score, current smoking, alcohol use, hypertension, diabetes, dyslipidemia, central obesity, use of statins and center.

Abbreviations: CI, confidence interval; CVD, cardiovascular disease; HR, hazard ratio.