ARTICLE

Disease trajectories and mortality among individuals diagnosed with depression: a community-based cohort study in UK Biobank

Xin Han [1,2] · Can Hou[1,2] · Huazhen Yang[1,2] · Wenwen Chen[3] · Zhiye Ying[1,2] · Yao Hu[1,2] · Yajing Sun[1,2] · Yuanyuan Qu[1,2] · Lei Yang[4] · Unnur A. Valdimarsdóttir[5,6,7] · Wei Zhang [1,2,8] · Haomin Yang [9] · Fang Fang[10] · Huan Song [1,2,5]

Abstract

Patients with depression are at increased risk for a range of comorbid diseases, with, however, unclear explanations. In this large community-based cohort study of the UK Biobank, 24,130 patients diagnosed with depression were compared to 120,366 matched individuals without such a diagnosis. Follow-up was conducted from 6 months after the index date until death or the end of 2019, for the occurrence of 470 medical conditions and 16 specific causes of death. The median age at the time of the depression diagnosis was 62.0 years, and most of the patients were female (63.63%). During a median follow-up of 4.94 years, 129 medical conditions were found to be significantly associated with a prior diagnosis of depression, based on adjusted Cox regression models. Using disease trajectory network analysis to visualize the magnitude of disease–disease associations and the temporal order of the associated medical conditions, we identified three main affected disease clusters after depression (i.e., cardiometabolic diseases, chronic inflammatory diseases, and diseases related to tobacco abuse), which were further linked to a wider range of other conditions. In addition, we also identified three depression-mortality trajectories leading to death due to cardiovascular disease, respiratory system disease and malignant neoplasm. In conclusion, an inpatient diagnosis of depression in later life is associated with three distinct network-based clusters of medical conditions, indicating alterations in the cardiometabolic system, chronic status of inflammation, and tobacco abuse as key pathways to a wide range of other conditions downstream. If replicated, these pathways may constitute promising targets for the health promotion among depression patients.

Introduction

Depression is a common psychiatric disorder affecting about 15–18% of the general population worldwide [1, 2]. According to the World Health Organization (WHO), depression has been

---

These authors contributed equally: Xin Han and Can Hou

✉ Haomin Yang
haomin.yang@ki.se

✉ Huan Song
songhuan@wchscu.cn

1   West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China

2   Medical Big Data Center, Sichuan University, Chengdu, China

3   Division of Nephrology, Kidney Research Institute, State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu, China

4   Department of Anesthesiology, West China Hospital, Sichuan University, Chengdu, China

5   Center of Public Health Sciences, Faculty of Medicine, University of Iceland, Reykjavík, Iceland

6   Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

7   Department of Epidemiology, Harvard T H Chan School of Public Health, Boston, MA, USA

8   Mental Health Center, West China Hospital of Sichuan University, Chengdu, China

9   Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, Fuzhou, China

10 Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden

ranked as the third contributor to the overall global burden of disease, as well as the most common cause of disease burden among females in low-, middle-, and high-income countries since 2004 [3]. In 2017, depression was ranked as a leading cause of disability worldwide by the WHO, with a prevalence of >427 million people of all ages globally [4].

In addition to psychological symptoms, individuals with depression have been reported to experience subsequently increased risks of various somatic diseases [5–13]. Longitudinal studies have reported elevated risks of asthma [5, 6, 11], diabetes [7, 13], cardiovascular disease [8, 12], Parkinson's disease [9], and dementia [10] among individuals diagnosed with depression. More recently, an increased risk of up to 30 different medical conditions and mortality following a mood disorder diagnosis (including depression) has been reported using data from the national register of Denmark [14–16]. Furthermore, a causal link has been suggested between depression and multiple disease outcomes through Mendelian Randomization analyses of the UK Biobank participants [17] and other populations [18, 19]. Hence, interventions to prevent future health declines are crucial for patients with depression, in addition to the psychotherapy and pharmacotherapy to address depressive symptoms. A comprehensive analysis of disease trajectories after a diagnosis of depression might assist the identification of key pathways linking depression with subsequent somatic problems.

Disease trajectory analysis has recently been proposed as a novel approach to explore disease progression over time [20]. Through visualizing the networks of emerged diseases—studying the magnitude of disease–disease associations and the temporal order of such associations, the analysis provides a foundation for examining causal relationships and sequential patterns of multiple morbidities. To date, disease trajectory analysis has been applied successfully in studies on the identification of sequential diseases, in general, or after a diagnosis of a specific medical condition, such as breast cancer [21, 22]. However, applications of this approach to explore patterns of disease networks after a specific psychiatric disorder are scarce.

With a focus on medical conditions associated with a prior diagnosis of depression, we aimed to clarify the disease trajectory network after depression using UK Biobank data. Because the clustering of sequential and interacting diseases may imply shared etiologies or disease mechanisms, our investigation can help identify the key affected pathways leading to the deterioration in general health after depression.

Methods and materials

Study design

Details of the UK Biobank study design are described elsewhere [23]. In brief, this prospective cohort study enrolled 502,507 participants from ~9.2 million people across the UK who were invited to join the cohort between 2006 and 2010. The participants' age was 40 to 69 years at recruitment, and the majority was of white ethnicity. Baseline assessments, including a computer-assisted interview using a touch-screen questionnaire, a physical examination and collection of biological samples, were conducted with all participants in 22 assessment centers. Health-related outcomes of the participants were obtained through regular linkages to multiple national databases. Inpatient hospital data on participants in England, Scotland and Wales were updated using their respective databases: the Hospital Episode Statistics database, the Scottish Morbidity Record and the Patient Episode Database [24]. The inpatient hospital data were deemed to cover all the UK Biobank participants since 1997 [24]. Mortality data on the UK Biobank participants in England and Wales were obtained from the National Health Service (NHS) Digital, whereas the mortality data for the participants in Scotland were obtained from the NHS Central Register [24].

In the present study, based on the UK Biobank data, we excluded 14 individuals who withdrew their informed consent forms, leaving 502,493 individuals (229,115 males and 273,378 females) for the analysis. Individuals with depression were included in the exposed cohort on the date of their diagnosis of depression. We then randomly selected up to five unexposed individuals from the study population who were free of depression on the diagnosis date of the index patient (i.e., the index date of the unexposed individual) and individually matched to the index patient by birth year, sex, and Townsend deprivation index (transformed to decile variable) [25], using the method of incidence density sampling. Townsend deprivation index was assigned to each individual based on postcode location with a higher index score indicating a higher degree of deprivation. In a sub-analysis, the Charlson comorbidity index (CCI) on the index date was calculated for each participant as their baseline comorbidity level [26], based on UK Biobank inpatient hospital data. To minimize the influence of reverse causality and surveillance bias [27] (Supplementary Fig. 1), follow-up for all participants started from 6 months after the index date and continued until death, loss to follow-up or the end of study (December 31, 2019), whichever occurred first.

Ascertainment of depression

We defined patients with depression as individuals admitted for inpatient hospital care with a diagnosis (either primary or secondary) of major depression (according to the International Classification of Diseases, Tenth Revision [ICD-10]: F32 and F33), based on UK Biobank inpatient hospital data between January 1, 1997 and December 31, 2019. The

upper boundary was set to eliminate the potential influence of the Coronavirus Disease 2019 (COVID-19) outbreak on disease occurrence [28]. The diagnoses of depression, from the Hospital Episode Statistics database in England and the equivalent databases in Scotland and Wales, have been validated against detailed clinical evaluations, demonstrating good positive predictive value (75%) [29, 30]. As we focused on the identification of severe depression based on inpatient diagnoses, the unexposed individuals were defined as those who had no inpatient care records concerning depression at the time of the selection.

Diagnoses of other medical conditions

Diagnoses of other medical conditions were retrieved from the main and secondary diagnoses in the UK Biobank inpatient hospital data, which were also documented using the ICD-10. We excluded diagnoses related to pregnancy, childbirth, perinatal conditions, and unclassified symptoms or signs, and restricted the analyses to Chapters 1–14 and 19–20 of the ICD-10. We used the 3-digit ICD-10 codes for medical condition identification and combined conditions with clinical or biological similarities (Supplementary Data 1). For conditions identified in multiple records of the same individual, only the first record and the date of the first hospital visit were used as the date of diagnosis. The diagnoses of medical conditions in the UK inpatient hospital data have been validated, yielding an overall median diagnostic accuracy of 80.3% (interquartile range, 63.3–94.1%) comparing routinely collected data sets with case notes [31].

Death

The cause of death of an individual was defined as the underlying cause of death recorded in the UK Biobank mortality data. The underlying cause of death was also coded in accordance with the ICD-10, and the causes were classified into 16 categories, mainly based on the ICD-10 chapters (Supplementary Data 1) [32].

Statistical analysis

Trajectory analyses

We used a previously proposed method to identify disease trajectories following the diagnosis of depression [20, 21]. The method includes three interrelated steps (Supplementary Fig. 2): the first step was a phenome-wide association analysis (PheWAS) using conditional Cox regression, with the aim of investigating the risks of 470 medical conditions among individuals with depression compared to a matched group of unexposed individuals. To ensure statistical power,

we limited the analyses to medical conditions that occurred in at least 250 patients with depression (~1%). Only diseases with a $p <$ the Bonferroni corrected threshold and a hazard ratio (HR) > 1 were considered in the second step of the analysis. In the second step, for all possible disease 1 (D1) and disease 2 (D2) pairs with a temporal order experienced by at least 125 depression patients (250/2), binomial tests were conducted to investigate whether more depression patients (>50%) had D2 diagnosed after D1 among those with both D1 and D2 diagnoses. D1 → D2 disease pairs with a binomial test $p <$ the Bonferroni corrected threshold were then included in the third step of the analysis, where nested case-control study design and conditional Logistic regression were used to confirm and assess the magnitude of the associations. A detailed description is available in Supplementary Data 2.

To investigate disease pairs leading to death (e.g., D1 → D2 → Death), we repeated steps 1–3 of the above disease trajectory analyses (Supplementary Fig. 3). We restricted the analyses to D1 → D2 → Death pairs that were experienced by at least 20 depression patients.

The networks of the disease trajectories were formed by combining disease pairs with overlapping diseases (e.g., disease pairs D1 → D2 and D2 → D3 with overlapping D2 were combined in the trajectory D1 → D2 → D3).

Subgroup and sensitivity analyses

To investigate whether the disease trajectories after depression differed by sex, we performed separate analyses for females and males. Given the reduced sample size, we adjusted the thresholds of these analyses, by multiplying the original thresholds (250 for the step 1 analysis, 125 for the step 2 analysis and 20 for the analysis investigating the disease pairs leading to death) with the proportions of these subgroups in the total population of individuals with depression (i.e., females: 65% and males: 35%). To explore whether the impact of depression on other diseases changed over time, we performed separate PheWAS (i.e., step 1) for the different follow-up periods within five years of the index date and beyond. We further stratified the PheWAS analysis by season of cohort entry (January–March, April–June, July–September, and October–December) to detect any effect of seasonal variation on the disease associations analyzed.

As depression may be secondary to a preceding somatic disease and a large proportion of the patients with depression were hospitalized with a main diagnosis of somatic diseases (15,602/24,130, 64.66%) which might lead to subsequent diseases independent of depression, we repeated the PheWAS after excluding individuals with a history of any disease that belonged to the same category of the outcome disease (retrieved from the inpatient hospital data), as

Fig. 1 Flow chart of study population selection and main analysis steps. This figure shows the study design, the inclusion and exclusion process of study population selection, and main analysis steps.
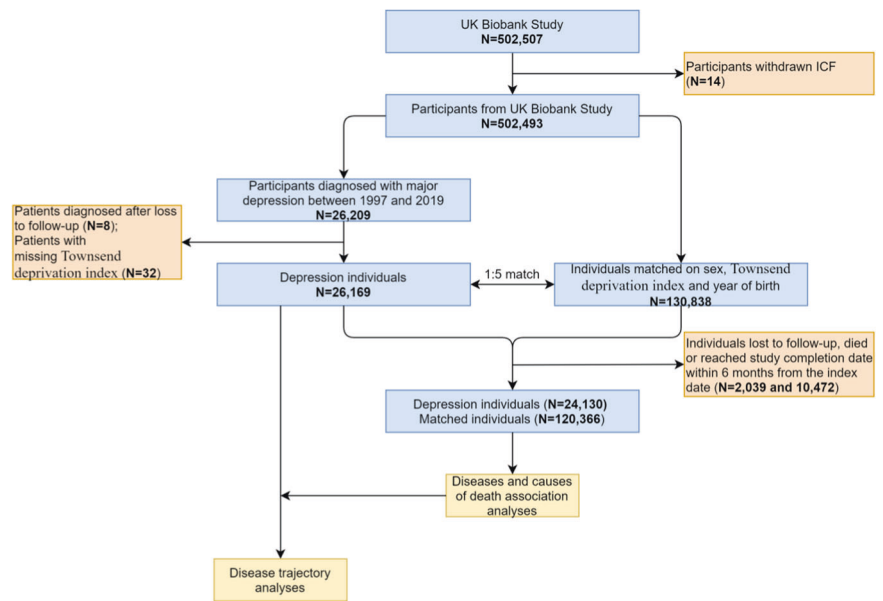
Table 1 Basic characteristics of the depression individuals and their matched individuals without depression.

	Depression individuals (N = 24,130)	Matched individuals without depression[c] (N = 120,366)
Age at index date, median (Q1–Q3), years	62.00 (54.00–69.00)	62.00 (54.00–69.00)
Sex (%)		
Female	15,354 (63.63%)	76,633 (63.67%)
Male	8,776 (36.37%)	43,733 (36.33%)
Townsend deprivation index[a], median (Q1–Q3)	−1.10 (−3.14–2.26)	−1.11 (−3.14–2.16)
Follow-up time[b], median (Q1–Q3), years	4.94 (2.18–8.63)	5.19 (2.38–8.82)

This table shows the basic characteristics of the depression individuals and their matched individuals without depression, including age at index date, sex, Townsend deprivation index and follow-up time (years).

Q1 lower quartile, *Q3* upper quartile.

[a]Townsend deprivation index was assigned to each individual based on their postcode location, a greater index score implies a greater degree of deprivation.

[b]Follow-up of the depression and matched individuals started 6 months after the index date to the date of death, date of loss to follow-up or December 31, 2019, whichever came first.

[c]At most five participants who were alive, retained in the cohort and free of depression at the corresponding index date were matched to each depression individual based on sex, Townsend deprivation index (transformed to decile variable) and year of birth.

indicated in Supplementary Data 1. In this analysis, we also additionally adjusted for the CCI on the index date to further control of the somatic comorbidities at baseline. For the same reason, we also conducted a sensitivity analysis by restricting to a subgroup of participants (both depression patients and matched unexposed individuals) without any inpatient care records during the 6 months prior to the index date. We further tested the robustness of our results in relation to the definition of depression by ascertaining depression solely through the primary diagnosis reported in the UK Biobank inpatient hospital data. All the statistical analyses were conducted using SciPy (version 1.4.1),

Statsmodels (version 0.11.1) and Lifelines (version 0.25.2) in Python 3.8. The p value for statistical significance was set to 0.05/number of analyses performed (Bonferroni corrections), to account for multiple testing.

Results

Based on the community-based cohort study of UK Biobank, 24,130 individuals with depression were identified, together with 120,366 age-, sex-, and Townsend deprivation index-matched unexposed individuals without such a

Fig. 2 Hazard ratios (HRs) of other medical conditions among depression individuals compared to matched individuals without depression. The *X* axis shows the disease categories according to ICD-10 codes A-N and S-Y. The *Y* axis shows the significant hazard ratios after Bonferroni correction of each medical condition when comparing depression individuals to individuals without depression. Details of the hazard ratios, number of cases, and 95% confidence intervals are listed in Supplementary Table 1.
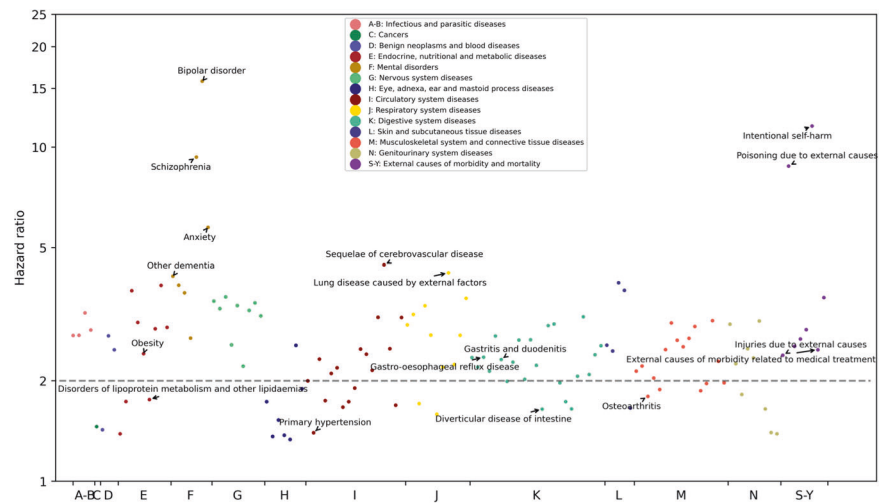
Table 2 Hazard ratios (HRs) with 95% confidence intervals (CIs) of different causes of death among depression individuals compared to matched individuals without depression.

Cause of death	Both sex (N = 24,130)		Males (N = 8776)		Females (N = 15,354)	
	N[a]	HR (95% CI)[b]	N[a]	HR (95% CI)[b]	N[a]	HR (95% CI)[b]
Malignant neoplasms	851	2.17 (2.01–2.35)	400	2.25 (2.01–2.53)	451	2.12 (1.90–2.36)
Cardiovascular diseases	469	3.00 (2.69–3.36)	296	3.06 (2.66–3.53)	173	2.96 (2.46–3.56)
Respiratory system diseases	314	4.65 (4.00–5.41)	150	3.82 (3.09–4.71)	164	5.89 (4.72–7.34)
Neural system diseases	160	3.92 (3.20–4.81)	83	4.74 (3.54–6.35)	77	3.32 (2.49–4.42)
Digestive system diseases	152	4.25 (3.43–5.28)	93	4.39 (3.33–5.80)	59	4.10 (2.90–5.78)
Unnatural causes	129	6.42 (4.95–8.33)	74	5.72 (4.11–7.95)	55	7.85 (5.12–12.04)
Mental disorder diseases	99	5.63 (4.24–7.49)	57	6.13 (4.18–8.97)	42	5.13 (3.35–7.86)
Endocrine system diseases	53	3.99 (2.79–5.71)	28	3.80 (2.34–6.18)	25	4.28 (2.52–7.29)
Genitourinary system diseases	22	2.71 (1.62–4.53)	12	2.25 (1.14–4.42)	10	3.66 (1.63–8.23)
Infectious and parasitic diseases	21	3.02 (1.76–5.17)	12	3.15 (1.54–6.43)	9	2.89 (1.27–6.53)

This table shows numbers and the HRs with 95% CIs of 10 different causes of death with statistically significant HRs after Bonferroni correction and $N \geq 20$ in all depression individuals, males and females.

[a]Number of the depression individuals died with the corresponding underlying causes.

[b]Hazard ratios (with 95% confidence intervals) of the causes of death in depression individuals compared to sex, Townsend deprivation index and year of birth matched individuals without depression.

diagnosis. Follow-up of all participants was conducted from 6 months after the index date until death or the end of 2019, for the occurrence of 470 medical conditions and 16 specific causes of death (Fig. 1). The participants' median age was 62.0 years on the index date, and most of them were females (63.63%). The median follow-up time was 4.94 years (Table 1).

Only 132, out of all 470 studied medical conditions, involved >250 patients with depression. Among these, 129 showed a statistically significant association with a prior diagnosis of depression (Supplementary Table 1). As expected, associations were found between depression and other psychiatric conditions, including bipolar disorder (hazard ratio [HR] = 15.77, 95% confidence intervals [CI], 13.18–18.87), intentional self-harm (HR = 11.57, 95% CI, 9.93–13.49), and schizophrenia (HR = 9.34, 95% CI, 7.75–11.24) (Fig. 2). Subsequent somatic diseases with the strongest associations

with depression were sequelae of cerebrovascular disease (4.45, 95% CI 3.79–5.22), pressure ulcer (3.94, 95% CI 3.45–4.49), and ulcer of lower limb (3.73, 95% CI 3.18–4.38).

For mortality, increased risks were present in 10 of 16 studied categories of causes of death, with the strongest association noted for death due to unnatural causes (HR = 6.42, 95% CI, 4.95–8.33) (Table 2).

Temporal disease trajectories in individuals with depression

In total, 110 disease pairs were established among all 129 depression-related medical conditions identified in step 1 (Supplementary Fig. 2). An overview of the disease trajectories subsequent to the diagnosis of depression is presented in the Supplementary Material (Supplementary
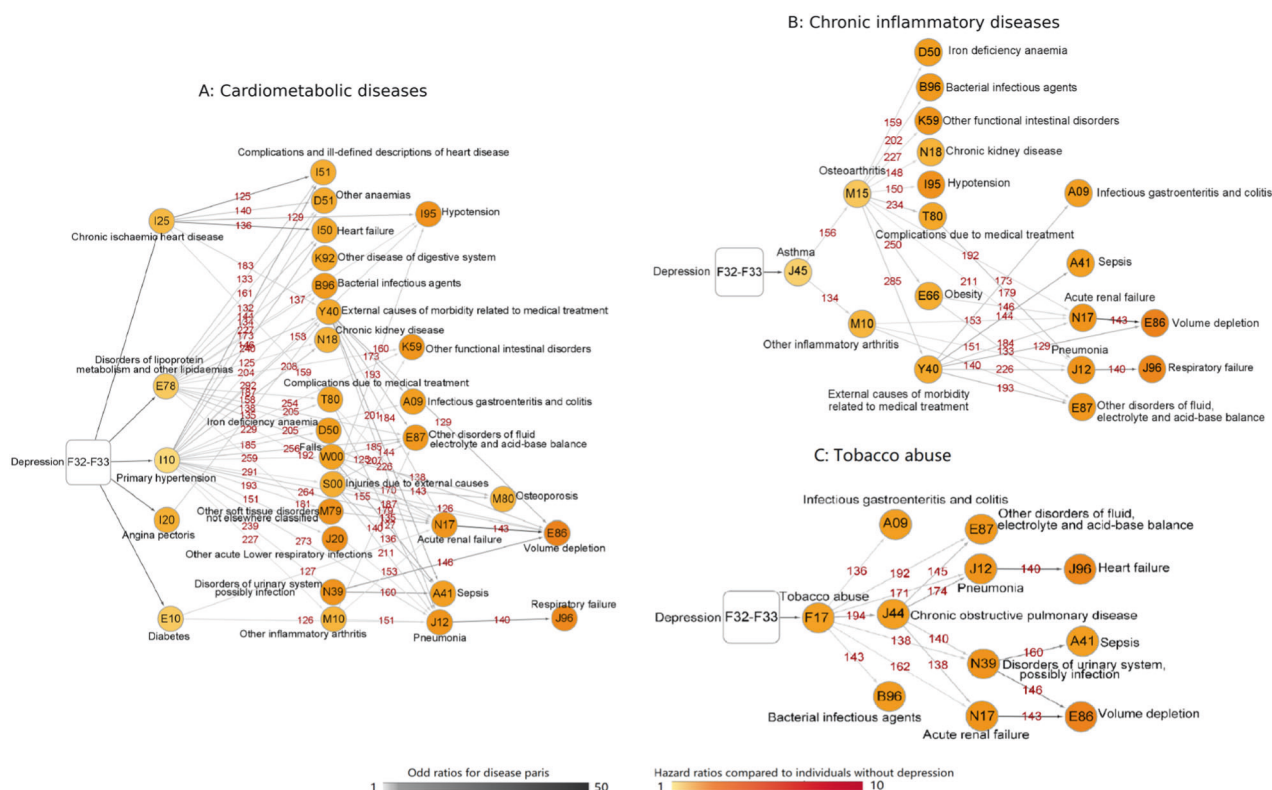
Fig. 3 Trajectories of three main disease clusters among depression individuals. This figure illustrates three following subgroups of disease trajectories identified in our analysis: A-Cardiometabolic diseases, B-Chronic inflammatory diseases, C-Tobacco abuse. The combined ICD-10 codes for the medical conditions are shown within the circle and round rectangle. The color of the circle represents the hazard ratios of this condition when comparing depression individuals to matched individuals without depression. The number above the arrow connecting two circles corresponds to the number of disease pairs among depression individuals. The color of the arrows indicates the odds ratio of the sequential association between the two medical conditions among depression individuals.

Fig. 4). Among the diseases ordered as D1, three major types were categorized according to the similarity of their underlying affected systems or their etiologies. Cluster 1 mainly included diseases of cardiometabolic system, where the disease tree thrived after the diagnoses of chronic ischemic heart disease, angina pectoris, primary hypertension, diabetes, and disorders of lipoprotein metabolism and other lipidaemias (Fig. 3a). Cluster 2 consisted of a group of diseases related to chronic inflammation, which first presented as the development of asthma, osteoarthritis, and other inflammatory arthritis (Fig. 3b). Unlike the other two clusters, Cluster 3 was a cluster of diseases and conditions related to tobacco abuse (Fig. 3c). All odd ratios (ORs) and numbers of cases of disease pairs in the trajectories are presented in the Supplementary Material (Supplementary Table 2).

Temporal disease trajectories leading to mortality following a diagnosis of depression

In total, 38 D1- > D2- > Death pairs were recognized based on the results of steps 1 and 2 of the analyses (Supplementary Fig. 3 and Supplementary Table 3), which constituted three networks of disease trajectories leading to deaths among depression patients (Fig. 4). The mortality trajectory of cardiovascular disease was associated with chronic ischemic heart disease, tobacco abuse, and injuries due to external causes. The mortality trajectory of diseases of the respiratory system was associated with primary hypertension, chronic obstructive pulmonary disease (COPD), and tobacco abuse, followed by pneumonia and respiratory failure. The mortality trajectory of malignant neoplasm was associated with several diseases, subsequent to anxiety, dorsalgia, tobacco abuse, and COPD.

Subgroup and sensitivity analyses

The separate PheWAS for females and males revealed largely similar subsequent diseases in relation to a prior diagnosis of depression, although the magnitude of the associations differed slightly. The females with depression had a higher relative risk for obesity, alcohol abuse, sleep disorder, heart failure, chronic ischemic heart disease, and chronic rheumatic heart disease. Male patients with depression had a higher relative risk for sepsis, malnutrition, anxiety, hypothyroid conditions, functional dyspepsia,

Fig. 4 Disease trajectories leading to mortality with different death causes among depression individuals. This figure shows the identified disease trajectories leading to CVDD (Cardiovascular disease death), RSDD (Respiratory system disease death), and MND (Malignant neoplasm death) among depression individuals. The death causes are shown within the octagons. The combined ICD-10 codes for the medical conditions are shown within the circle and round rectangle. The color of the circle and octagon represents the hazard ratios of this medical condition when comparing depression individuals to matched individuals without depression. The number above the arrow connecting two circles and circle with octagon corresponds to the number of disease pairs among depression individuals. The color of the arrows indicates the odds ratio of the sequential association between the two medical conditions among depression individuals.
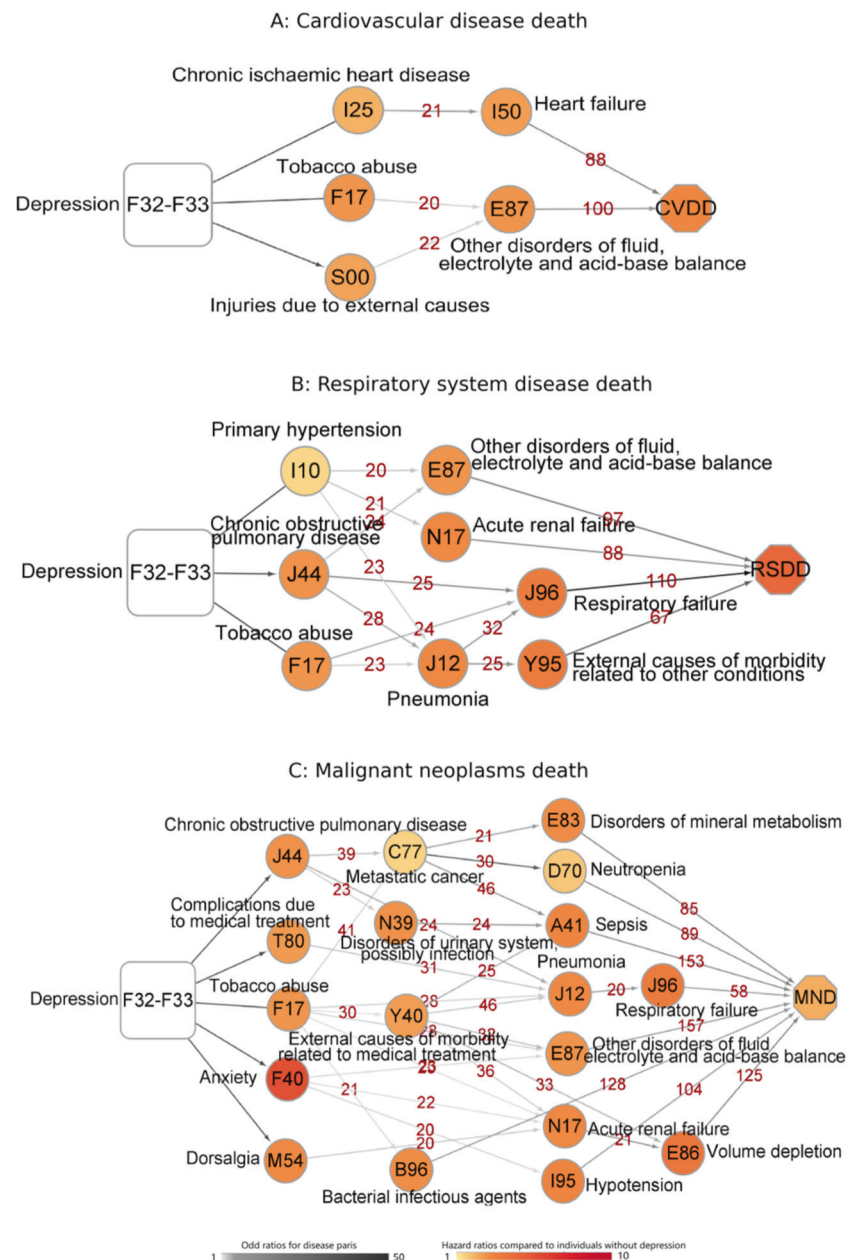
osteoporosis, and falls (Supplementary Table 1 shows the HRs and 95% CIs). In the disease trajectory analysis, female and male depression patients had several common trajectories (Supplementary Fig. 5, Supplementary Tables 4, 5). In addition, although both groups died from malignant neoplasms and respiratory system diseases, different progression pathways were found between the male and female patients (Supplementary Fig. 6, Supplementary Tables 6, 7). The subgroup PheWAS for the different follow-up periods revealed associations of almost identical diseases with depression both during the first 5 years of follow-up and thereafter (Supplementary Table 8). Similarly, stratified analysis by season rendered largely similar

estimates as those of the main analysis (Supplementary Table 9).

In sensitivity analyses, we obtained largely similar PheWAS results when excluding depression patients who had a history of any disease that belonged to the same category of the outcome disease (Supplementary Table 10). Also, restricting the analyses to individuals without any inpatient record of a somatic disease during the 6 months before the index date did not change the results of the main analyses substantially (Supplementary Table 11). When using only the primary diagnoses ($n = 1920$, 7.96%) from the UK Biobank inpatient hospital data to identify depression, we found 141 out of 470 tested medical conditions to be

associated with a prior diagnosis of depression. Compared with the results of the main analysis, 111 (86.05%) significant medical conditions were identified in both analyses, although differential HRs were observed in 20 of these conditions (Supplementary Table 12).

Discussion

In this large community-based cohort study using the UK Biobank data, we found that individuals who received a diagnosis of depression through inpatient care were at a subsequently elevated risk of 129 medical conditions. Adding to the existing data that demonstrated similar temporal and causal links between depression and various health consequences, our analyses of the disease trajectories identified, for the first time, three major network-based clusters according to the similarities in the underlying affected systems or etiologies, revealing that alterations in cardiometabolic system, the chronic status of inflammation, together with behavior-related changes (i.e., tobacco abuse), may be key pathways linking depression with a wide range of other diseases and conditions downstream. In addition, we observed similar disease trajectories among males and females with depression. Given the recognizable decline in the general health of patients with depression, these findings call for future investigations of potential interventions to target these key pathways.

Although no previous study has addressed the disease trajectories after a diagnosis of depression, the results of our PheWAS gain support from Danish register-based studies, showing associations between depression and a wide range of psychiatric and somatic diseases and conditions [14, 16]. Notably, in addition to consistently reported associations with many common diseases, novel relationships between depression and subsequent individual diseases with relatively low incidence, such as sepsis, visual disturbances/blindness, and urolithiasis, were also unraveled in our analyses. More recently, a disease trajectory browser has been developed based on national data from inpatient wards, outpatient clinics, and emergency room visits between 1994 and 2018 in Denmark [22]. By searching for depression (ICD-10: F32) using the browser, some of the key diseases, such as cardiometabolic diseases, identified as occurring more-than-expected subsequent to depression in our study were found to occur more-than-expected prior to depression. These results should, however, not be compared directly because in our study we aimed to visualize physiological changes after depression by studying a cohort of patients with depression whereas the Danish work examined diseases occurring both before and after depression (and any other disease) by analyzing the entire general population of Denmark. Therefore, the inconsistent results

observed between these two studies do not necessarily invalidate each other. Moreover, using Mendelian Randomization, a previous study has demonstrated causal links between depression and 20, out of 42, phenotypically associated medical conditions [17]. Given that those conditions largely overlap with the key medical conditions identified in the three disease trajectory clusters in our study, results of such analysis corroborate our findings that there are key pathways linking depression to a general health decline.

In the present study, the visualized disease trajectory networks imply the cardiometabolic system may be impacted by the occurrence of depression, with initial presentations of chronic ischemic heart disease, angina pectoris, primary hypertension, diabetes, and disorders of lipoprotein metabolism and other lipidaemias. This is in line with the results of previous phenotypic studies linking depression with major types of cardiometabolic diseases, such as chronic ischemic heart disease [33], stroke [14, 34], heart failure [14], diabetes [7, 13], metabolic syndrome [35], and other non-specific heart diseases [36, 37]. Also, Mendelian Randomization analyses have suggested a causal association between depression and ischemic heart disease and lipid metabolism, independent of shared environmental and behavioral factors or medication use [17–19]. Other possible underlying mechanisms include sympathetic activation, hypothalamic–pituitary activation, and the chronic status of inflammation observed among individuals with depression, which may accelerate the development or progression of cardiometabolic diseases [38–44]. With further connections to a wide spectrum of diseases after the occurrence of cardiometabolic diseases in the disease trajectory networks, this pathway may be a promising target for the health promotion of depression patients.

We found an increased risk of asthma, followed by osteoarthritis and other inflammatory arthritis, after a depression diagnosis in the present study. Similar results were reported in a meta-analysis of six prospective studies with a total of 83,684 depression patients, indicating an increased risk (43%) of asthma after depression [11]. Likewise, a recent Mendelian randomization analysis provided evidence supporting a causal link between depression and asthma and osteoarthrosis [17]. As these diseases have key inflammatory mechanisms in common, the observed associations support the notion that depression may lead to exaggerated or prolonged inflammatory responses, as reported in the research literature on both animals and humans [45]. As a result, anti-inflammatory treatments may have the potential to prevent a general health decline after depression.

Behavior-related factors, mainly the tobacco abuse that often co-occurs and follows a diagnosis of depression, have been proposed as possible underlying mechanisms linking

depression with somatic problems [46–49]. Our analyses identified an array of diseases, including diseases of the respiratory system and infectious diseases, subsequent to tobacco abuse, underscoring the necessity for enhanced interventions for smoking cessation among patients with depression.

The major strengths of our study include the application of the disease trajectory analysis to visualize the associated disease networks after depression. By illustrating a whole picture of temporal disease progression patterns, this discovery-driven analysis complements association analysis in earlier studies often studying a single disease pair with a specific hypothesis. However, the nature of this analytic strategy prevents us from discovering novel disease pairs involving rare medical conditions. Additional targeted studies, are therefore, needed to examine rarer comorbidities of depression, as a complement to the present approach. Regardless, the disease-trajectory approach provides important additional data to the existing literature by clarifying the clusters and the temporal order of these subsequent diseases. The identification of common disease networks can on the other hand assist in the exploration of key physiological pathways, leading to a general health decline after depression. Other strengths include the use of a large community-based cohort, where complete data for diseases diagnoses were collected prospectively, through linkages with national health registers.

Notable limitations include the lack of primary care data, which may have led to the exclusion of less severe medical conditions in our analyses. Furthermore, the completeness and accuracy of the observed disease trajectory networks may have been impacted the relatively limited follow-up period of the present study (i.e., the median follow-up time was 4.94 years after excluding the first 6 months of follow-up). Additional studies with longer follow-up are needed to investigate diseases with a later onset, such as neurodegenerative diseases, to obtain a more comprehensive picture of disease trajectories subsequent to depression. However, because the findings of the present study did not differ significantly during the first 5 years after the diagnosis of depression and beyond, such truncations of data are unlikely to have a significant influence on the validity of the present results. Notably, although we aimed to explore physiological changes subsequent to depression, the disease trajectory analysis does not, by itself, provide a sufficient basis for causal inference, while lends support to causal relationships indicated elsewhere [17–19]. Yet, the observed disease patterns might also be explained by the shared etiological factors between depression and these conditions, including genetics, lifestyle, and other environmental factors. Future exploration of disease trajectories by level of genetic susceptibility to depression could be valuable. Although similar estimates were obtained in the sensitivity analyses after excluding depression patients with a

pre-existing diagnosis of disease in the same category as the outcome disease and individuals with an inpatient record of a somatic disease during the 6 months prior to the index date (indicating a limited effect of pre-existing somatic comorbidities), we cannot rule out the possibility that these comorbidities contributed to the disease trajectories observed after depression. In addition, it is important to note that patients with depression in this study were identified through inpatient care (likely exhibiting a severe form of the disease) and had a median age of 62 years at the time of ascertainment, which is different from the median age of the diagnosis of depression in the general population. Further, separate analysis of patients with multiple episodes of depression was not feasible due to the limited number of cases. Therefore, disease trajectories after milder versions of depression, depression diagnosed at a younger age, or recurrent depression should be examined in future studies. Last, in a time sequential analysis, it is challenging to investigate how time-varying confounders, such as behavioral factors and life experience, for which we had little information about would have contributed to the observed disease networks.

In conclusion, this community-based study confirmed a general increase in the risk of many mental disorders, somatic diseases, and causes of death subsequent to severe depression. We identified three main disease clusters following an inpatient diagnosis of depression, which might imply that alterations in cardiometabolic diseases, chronic inflammatory diseases, and tobacco abuse may contribute significantly to the general health decline after depression. Investigations of potential interventions targeting these key pathways are therefore warranted.

Data availability

Data from the UK Biobank (http://www.ukbiobank.ac.uk/) are available to all researchers upon making an application. Part of this research was conducted using the UK Biobank Resource under Application 54803.

Code availability

All statistical analyses were conducted by the following software tools: Python v. 3.8, and Cytoscape Desktop v. 3.8.0. All codes associated with the current submission is available, and can be requested by contacting the corresponding authors.

analysis. XH, CH, HS, HY, and FF interpreted the data. XH, CH, FF, HY, and HS drafted the paper. All authors approved the final paper as submitted and agree to be accountable for all aspects of the work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Compliance with ethical standards

Ethics approval The UK Biobank has full ethical approval from the NHS National Research Ethics Service (reference number: 16/NW/0274), and this study was also approved by the biomedical research ethics committee of West China Hospital (reference number: 2019–1171).

Conflict of interest The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Kessler RC, Bromet EJ. The epidemiology of depression across cultures. Annu Rev Public Health. 2013;34:119–38.
2. Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, et al. Cross-national epidemiology of DSM-IV major depressive episode. BMC Med. 2011;9:90.
3. The global burden of disease: 2004 update. https://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/, 2004, Accessed Date Accessed 2004.
4. WHO. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018; 392:1789–858.
5. Akula M, Kulikova A, Khan DA, Brown ES. The relationship between asthma and depression in a community-based sample. J Asthma. 2018;55:1271–7.
6. Alonso J, de Jonge P, Lim CC, Aguilar-Gaxiola S, Bruffaerts R, Caldas-de-Almeida JM, et al. Association between mental disorders and subsequent adult onset asthma. J Psychiatr Res. 2014;59:179–88.
7. de Jonge P, Alonso J, Stein DJ, Kiejna A, Aguilar-Gaxiola S, Viana MC, et al. Associations between DSM-IV mental disorders and diabetes mellitus: a role for impulse control disorders and depression. Diabetologia. 2014;57:699–709.
8. Wium-Andersen MK, Wium-Andersen IK, Jorgensen MB, McGue M, Jorgensen TSH, Christensen K, et al. The association between depressive mood and ischemic heart disease: a twin study. Acta Psychiatr Scand. 2019;140:265–74.
9. Wang S, Mao S, Xiang D, Fang C. Association between depression and the subsequent risk of Parkinson's disease: a meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry. 2018;86:186–92.
10. Kessing LV. Depression and the risk for dementia. Curr Opin Psychiatry. 2012;25:457–61.
11. Gao YH, Zhao HS, Zhang FR, Gao Y, Shen P, Chen RC, et al. The relationship between depression and asthma: a meta-analysis of prospective studies. PLoS One. 2015;10:e0132424.
12. Seligman F, Nemeroff CB. The interface of depression and cardiovascular disease: therapeutic implications. Ann N. Y Acad Sci. 2015;1345:25–35.
13. Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. Diabetes Care. 2008;31:2383–90.
14. Momen NC, Plana-Ripoll O, Agerbo E, Benros ME, Borglum AD, Christensen MK, et al. Association between mental disorders and subsequent medical conditions. N. Engl J Med. 2020;382:1721–31.
15. Plana-Ripoll O, Pedersen CB, Agerbo E, Holtz Y, Erlangsen A, Canudas-Romo V, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. Lancet. 2019;394:1827–35.
16. Plana-Ripoll O, Pedersen CB, Holtz Y, Benros ME, Dalsgaard S, de Jonge P, et al. Exploring comorbidity within mental disorders among a danish national population. JAMA Psychiatry. 2019;76:259–70.
17. Mulugeta A, Zhou A, King C, Hypponen E. Association between major depressive disorder and multiple disease outcomes: a phenome-wide Mendelian randomisation study in the UK Biobank. Mol Psychiatry. 2020;25:1469–76.
18. Tang B, Yuan S, Xiong Y, He Q, Larsson SC. Major depressive disorder and cardiometabolic diseases: a bidirectional Mendelian randomisation study. Diabetologia. 2020;63:1305–11.
19. Li GH, Cheung CL, Chung AK, Cheung BM, Wong IC, Fok MLY, et al. Evaluation of bi-directional causal association between depression and cardiovascular diseases: a Mendelian randomization study. Psychol Med. 2020: 1–12. [Epub ahead of print].
20. Jensen AB, Moseley PL, Oprea TI, Ellesoe SG, Eriksson R, Schmock H, et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat Commun. 2014;5:4022.
21. Yang H, Pawitan Y, He W, Eriksson L, Holowko N, Hall P, et al. Disease trajectories and mortality among women diagnosed with breast cancer. Breast Cancer Res. 2019;21:95.
22. Siggaard T, Reguant R, Jorgensen IF, Haue AD, Lademann M, Aguayo-Orozco A, et al. Disease trajectory browser for exploring temporal, population-wide disease progression patterns in 7.2 million Danish patients. Nat Commun. 2020;11:4952.
23. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol. 2017;186:1026–34.
24. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779.
25. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North. London: Routledge; 1988.
26. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43:1130–9.
27. Chiolero A, Santschi V, Paccaud F. Public health surveillance with electronic medical records: at risk of surveillance bias and overdiagnosis. Eur J Public Health. 2013;23:350–1.
28. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoul-poor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19

pandemic: a systematic review and meta-analysis. Globalization health. 2020;16:1–11.

29. Davis KA, Sudlow CL, Hotopf M. Can mental health diagnoses in administrative data be used for research? A systematic review of the accuracy of routinely collected diagnoses. BMC Psychiatry. 2016;16:263.

30. Davis KAS, Bashford O, Jewell A, Shetty H, Stewart RJ, Sudlow CLM, et al. Using data linkage to electronic patient records to assess the validity of selected mental health diagnoses in English Hospital Episode Statistics (HES). PLoS ONE. 2018;13: e0195002.

31. Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, et al. Systematic review of discharge coding accuracy. J Public Health (Oxf). 2012;34:138–48.

32. International classification of diseases for mortality and morbidity statistics (10th Revision). https://www.who.int/classifications/icd/icdonlineversions/en/, 2019, Accessed Date Accessed 2019.

33. Smaardijk VR, Lodder P, Kop WJ, van Gennep B, Maas A, Mommersteeg PMC. Sex- and gender-stratified risks of psychological factors for incident ischemic heart disease: systematic review and meta-analysis. J Am Heart Assoc. 2019;8: e010859.

34. Swain NR, Lim CC, Levinson D, Fiestas F, de Girolamo G, Moskalewicz J, et al. Associations between DSM-IV mental disorders and subsequent non-fatal, self-reported stroke. J Psychosom Res. 2015;79:130–6.

35. Pan A, Keum N, Okereke OI, Sun Q, Kivimaki M, Rubin RR, et al. Bidirectional association between depression and metabolic syndrome: a systematic review and meta-analysis of epidemiological studies. Diabetes Care. 2012;35:1171–80.

36. Scott KM, de Jonge P, Alonso J, Viana MC, Liu Z, O'Neill S, et al. Associations between DSM-IV mental disorders and subsequent heart disease onset: beyond depression. Int J Cardiol. 2013;168:5293–9.

37. Martin DJ, Ul-Haq Z, Nicholl BI, Cullen B, Evans J, Gill JM, et al. Cardiometabolic disease and features of depression and bipolar disorder: population-based, cross-sectional study. Br J Psychiatry. 2016;208:343–51.

38. Miller AH, Maletic V, Raison CL. Inflammation and its discontents: the role of cytokines in the pathophysiology of major depression. Biol Psychiatry. 2009;65:732–41.

39. Heidt T, Sager HB, Courties G, Dutta P, Iwamoto Y, Zaltsman A, et al. Chronic variable stress activates hematopoietic stem cells. Nat Med. 2014;20:754–8.

40. Lanquillon S, Krieg JC, Bening-Abu-Shach U, Vedder H. Cytokine production and treatment response in major depressive disorder. Neuropsychopharmacology. 2000;22:370–9.

41. Kenis G, Maes M. Effects of antidepressants on the production of cytokines. Int J Neuropsychopharmacol. 2002;5:401–12.

42. Pollock BG, Laghrissi-Thode F, Wagner WR. Evaluation of platelet activation in depressed patients with ischemic heart disease after paroxetine or nortriptyline treatment. J Clin Psychopharmacol. 2000;20:137–40.

43. Musselman DL, Tomer A, Manatunga AK, Knight BT, Porter MR, Kasey S, et al. Exaggerated platelet reactivity in major depression. Am J Psychiatry. 1996;153:1313–7.

44. Kuijpers PMJC, Hamulyak K, Strik JJMH, Wellens HJJ, Honig A. Beta-thromboglobulin and platelet factor 4 levels in post-myocardial infarction patients with major depression. Psychiatry Res. 2002;109:207–10.

45. Kiecolt-Glaser JK, Derry HM, Fagundes CP. Inflammation: depression fans the flames and feasts on the heat. Am J Psychiatry. 2015;172:1075–91.

46. Boden JM, Fergusson DM. Alcohol and depression. Addiction. 2011;106:906–14.

47. Whooley MA, de Jonge P, Vittinghoff E, Otte C, Moos R, Carney RM, et al. Depressive symptoms, health behaviors, and risk of cardiovascular events in patients with coronary heart disease. JAMA. 2008;300:2379–88.

48. Weinberger AH, Kashan RS, Shpigel DM, Esan H, Taha F, Lee CJ, et al. Depression and cigarette smoking behavior: a critical review of population-based studies. Am J Drug Alcohol Abus. 2017;43:416–31.

49. Chaiton MO, Cohen JE, O'Loughlin J, Rehm J. A systematic review of longitudinal studies on the association between depression and smoking in adolescents. BMC Public Health. 2009;9:356.