Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study

Mika Kivimäki, G David Batty, Jaana Pentti, Martin J Shipley, Pyry N Sipilä, Solja T Nyberg, Sakari B Suominen, Tuula Oksanen, Sari Stenholm, Marianna Virtanen, Michael G Marmot, Archana Singh-Manoux, Eric J Brunner, Joni V Lindbohm, Jane E Ferrie, Jussi Vahtera

Summary

Background Socioeconomic disadvantage is a risk factor for many diseases. We characterised cascades of these conditions by using a data-driven approach to examine the association between socioeconomic status and temporal sequences in the development of 56 common diseases and health conditions.

Methods In this multi-cohort study, we used data from two Finnish prospective cohort studies: the Health and Social Support study and the Finnish Public Sector study. Our pooled prospective primary analysis data comprised 109 246 Finnish adults aged 17–77 years at study entry. We captured socioeconomic status using area deprivation and education at baseline (1998–2013). Participants were followed up for health conditions diagnosed according to the WHO International Classification of Diseases until 2016 using linkage to national health records. We tested the generalisability of our findings with an independent UK cohort study—the Whitehall II study (9838 people, baseline in 1997, follow-up to 2017)—using a further socioeconomic status indicator, occupational position.

Findings During 1 110 831 person-years at risk, we recorded 245 573 hospitalisations in the Finnish cohorts; the corresponding numbers in the UK study were 60 946 hospitalisations in 186 572 person-years. Across the three socioeconomic position indicators and after adjustment for lifestyle factors, compared with more advantaged groups, low socioeconomic status was associated with increased risk for 18 (32·1%) of the 56 conditions. 16 diseases formed a cascade of inter-related health conditions with a hazard ratio greater than 5. This sequence began with psychiatric disorders, substance abuse, and self-harm, which were associated with later liver and renal diseases, ischaemic heart disease, cerebral infarction, chronic obstructive bronchitis, lung cancer, and dementia.

Interpretation Our findings highlight the importance of mental health and behavioural problems in setting in motion the development of a range of socioeconomically patterned physical illnesses. Policy and health-care practice addressing psychological health issues in social context and early in the life course could be effective strategies for reducing health inequalities.

Funding UK Medical Research Council, US National Institute on Aging, NordForsk, British Heart Foundation, Academy of Finland, and Helsinki Institute of Life Science.

Introduction

Socioeconomic status, which captures social circumstances across the life course, is a powerful predictor of ill health. Studies have found increased morbidity and disability in individuals who are socioeconomically disadvantaged[1–3] and the disease burden in this group is increasing with population ageing.[4–6] However, to our knowledge, a comprehensive overview of the associations between socioeconomically patterned mental and physical health conditions is lacking.

Previous investigations have explored the relationship between socioeconomic status and multimorbidity.[3,7–14] These findings showed that having two or more diseases and developing multimorbidity was more common in people with low socioeconomic status. A limitation of these studies is the relatively restricted range of

morbidities investigated (communicable diseases are typically not included) and a failure to capture the temporal sequence between specific diseases. Considering temporality in disease onset could yield new insights into the cascades of health conditions that characterise morbidity in people with socioeconomic disadvantage.

To address these limitations, we examined the development of mental and physical health conditions among individuals with low and high socioeconomic status to determine temporal sequence and inter-relationships in the emergence of socioeconomically patterned conditions. We used a range of disease endpoints, adopting a data-driven approach, as we were not aware of any previous evidence-based test of hypotheses on disease cascades that characterised morbidity in people from different socioeconomic backgrounds.

Department of Epidemiology and Public Health
(Prof M Kivimäki FMedSci,
Prof G D Batty DSc,
M J Shipley MSc,
Prof A Singh-Manoux PhD
Prof E J Brunner PhD,
J E Ferrie PhD) **and Institute of Health Equity**
(Prof M G Marmot FBA),
University College London,
London, UK; Clinicum, Faculty
of Medicine, University of
Helsinki, Finland
(Prof M Kivimaki, J Pentti MSc,
P N Sipilä MD, S T Nyberg PhD,
J V Lindbohm MD); **School of Biological and Population Health Sciences, Oregon State University, Corvallis, OR, USA**
(Prof G D Batty); **Department of Public Health** (J Pentti,
Prof S B Suominen MD,
Prof S Stenholm PhD,
Prof J Vahtera MD) **and Centre for Population Health Research**
(J Pentti, Prof S Stenholm,
Prof J Vahtera), **University of Turku, Turku, Finland; School of Health and Education, University of Skövde, Skövde, Sweden** (Prof S B Suominen);
Turku University Hospital, Turku, Finland (J Pentti,
Prof S Stenholm, Prof J Vahtera);
Finnish Institute of Occupational Health, Helsinki, Finland (T Oksanen PhD);
School of Educational Sciences and Psychology, University of Eastern Finland, Joensuu, Finland (Prof M Virtanen PhD);
INSERM U1153, Epidemiology of Ageing and Neurodegenerative Diseases, Université de Paris, France
(Prof A Singh-Manoux); **and School of Community and Social Medicine, University of Bristol, Bristol, UK** (J E Ferrie)

Correspondence to:
Prof Mika Kivimäki, Department
of Epidemiology and Public
Health, University College
London, London WC1E 6BT, UK
m.kivimaki@ucl.ac.uk

Research in context

Evidence before this study

Low socioeconomic status, which captures multiple aspects of disadvantage, is a known risk factor for several diseases. We searched PubMed for research on low socioeconomic status and morbidity, without language or date restrictions, up to Feb 10, 2019, and identified thousands of studies using the search terms "socioeconomic" in combination with "cancer", "infection", "cardiovascular", "coronary heart disease", "stroke", and "psychiatric disorders". Studies on socioeconomic status in relation to other diseases, such as "diabetes", "endocrine disorder", "respiratory disease", "skin disease", "neurodegenerative disease", "dementia", and "digestive disease" were also very common. Few studies examined "multimorbidity" and we found no research on temporal sequences in mental and physical diseases across all bodily systems according to socioeconomic status.

Added value of this study

To facilitate a more comprehensive evaluation of morbidity associated with socioeconomic disadvantage, we combined individual-level data from two large cohort studies and examined low socioeconomic status as a risk factor for a range

of hospital-treated diseases. We determined temporal sequences in the emergence of diseases that were socioeconomically patterned. We repeated analyses in a third independent cohort. Across three indicators, low socioeconomic status was robustly associated with 18 (32·1%) of 56 specific diseases or health conditions, including 16 strongly interconnected conditions (hazard ratio >5 for each disease to be followed by another disease). This disease cascade started with psychiatric disorders, substance abuse, and self-harm and was followed later by diseases of the pancreas, liver, kidney, vascular and respiratory system, lung cancer, and dementia. Diabetes was associated with the cascade, but not with early psychiatric and behavioural disorders.

Implications of all the available evidence

Low socioeconomic status is a risk factor for a range of disorders, including mental and behavioural problems, which seems to set in motion a lifelong cascade of physical diseases. These findings suggest that policy and health-care practice addressing psychological health issues in social context and early in the life course might be an effective strategy for reducing socioeconomic inequalities in health.

Methods

Study design and population

In this multi-cohort study, we used data from two Finnish prospective cohort studies: the Health and Social Support (HeSSup) study[15] and the Finnish Public Sector (FPS) study.[16] We tested the generalisability of our findings with an independent UK cohort study—the Whitehall II study.[17] Ethical approval for these three studies was obtained from local committees on the ethics of human research. The derivation of the analytical sample used in each of these studies is shown in figure 1 and the appendix (p 2).

In the HeSSup study, 21486 of the men and women who responded to the survey between June 7, 1998, and May 23, 2000, or Jan 7, and Aug 12, 2003, had no missing data on residential area deprivation, and were successfully linked electronically to national hospitalisation and mortality registers until Dec 31, 2012.[15] The FPS sample comprised 87760 men and women who responded to at least one of four surveys done between March 1, 2000, and June 30, 2002, March 1, 2004, and June 30, 2005, March 1, 2008, and Nov 30, 2009, and Dec 1, 2011, and Nov 30, 2013, and had data on residential area deprivation.[16] Study participants were linked to electronic health records until Dec 31, 2016.

For our replication analyses, we used data from the Whitehall II study, which comprises 9838 government workers who participated in clinical examinations between Sept 10, 1985, and March 29, 1988, had no missing data on occupational position and covariates, and were linked electronically to national hospitalisation and mortality registers from Jan 1, 1997, when these records achieved a high level of national coverage, to March 31, 2017.[18]

Assessment of socioeconomic status at baseline

To explore the consistency of our results, we used three different indicators of socioeconomic status in our analyses. In the Finnish studies, we derived a score for residential area deprivation, similar to that developed by Townsend and colleagues,[19] and a measure of educational attainment. The area deprivation score was obtained from Statistics Finland and is based on the proportion of adults with low education, the unemployment rate, and the proportion of people living in rented housing in each 250 m by 250 m grid area.[20] Higher scores on the continuous index denote greater deprivation. We categorised these data as follows: low socioeconomic status (an area deprivation score higher than national mean), intermediate socioeconomic status (deprivation score from national mean to 0·5 SD below), and high socioeconomic status (the remaining data).

Educational attainment, obtained from Statistics Finland via record linkage (for the FPS study) or from a survey (for the HeSSup study), was based on the following two categories: high (tertiary qualification, college or university) and low (all other qualifications, including none).

In our replication analysis, we indexed socioeconomic status by a third indicator, the British civil service occupational grade.[17] Broadly equivalent to the Registrar General's indicators of occupational social class,[21] this index of socioeconomic circumstances is related to salary, occupational prestige, level of responsibility at work, and future pension, and has three groups as follows: high (administrative occupations), intermediate (professional and executive occupations), and low (clerical and support occupations).

Assessment of lifestyle risk factors at baseline

Using predefined operationalisations, we chose the following baseline risk factors to determine the extent to which the associations between socioeconomic status and diseases were attributable to standard lifestyle factors: current smoking (yes *vs* no), risky alcohol use (consumption >210 g per week *vs* other), physical inactivity (yes *vs* no), and obesity (body-mass index ≥30 kg/m² *vs* other).

Follow-up for diseases, health conditions, and mortality

Participants from the HeSSup study[15] and FPS study[16] were linked by their unique identification number to national registries of hospital discharge information (recorded by the National Institute for Health and Welfare) and mortality (recorded by Statistics Finland). These electronic health records include cause and date of hospitalisation or mortality and their coverage (all hospital types, including private hospitals, and records cover emergencies) reflects the comprehensive nature of Finland's public health-care system. Additional information on site-specific cancers, diabetes, cardiovascular diseases (including hypertension), psychotic disorders, dementia, Parkinson's disease, multiple sclerosis, epilepsy, asthma, chronic obstructive bronchitis, inflammatory bowel disease, liver disease, rheumatoid arthritis, gout, and renal failure was available via record linkage to the National Cancer Registry and the Drug Reimbursement Register of the Social Insurance Institution of Finland.

Whitehall II study members were linked to the UK National Health Service (NHS) Hospital Episode Statistics (HES) database for hospital admissions and the NHS Central Registry for mortality. In studies of chronic diseases, the sensitivity and specificity of the HES database have been high.[18,22]

In all cohort studies, the diagnosis for incident disease was coded according to the WHO International Classification of Diseases Tenth Revision (ICD-10). We focused on fifteen ICD-10 disease chapters that concern infectious and parasitic diseases (A00-B99), neoplasms (C00-D48), diseases of the blood (D50-D89), endocrine, nutritional, and metabolic diseases (E00-E90), mental and behavioural disorders (F00-F99), diseases of the nervous system (G00-G99), the eye (H00-H59), the ear (H60-H95), the circulatory system (I00-I99), the respiratory system (J00-J99), the digestive system (K00-K93), the skin (L00-L99), the musculoskeletal system (M00-M99), and the genitourinary system (N00-N99), injuries and poisoning (S00-T98), and external causes (V01-Y98).

Statistical analysis

Linked records captured 1204 ICD codes, including 56 major diseases or health conditions used in this analysis (for a complete list see appendix pp 2–13). Our primary analysis included two steps as follows: examination of associations between socioeconomic status (the exposure) and first new onset of health conditions after baseline (outcome) and mapping of
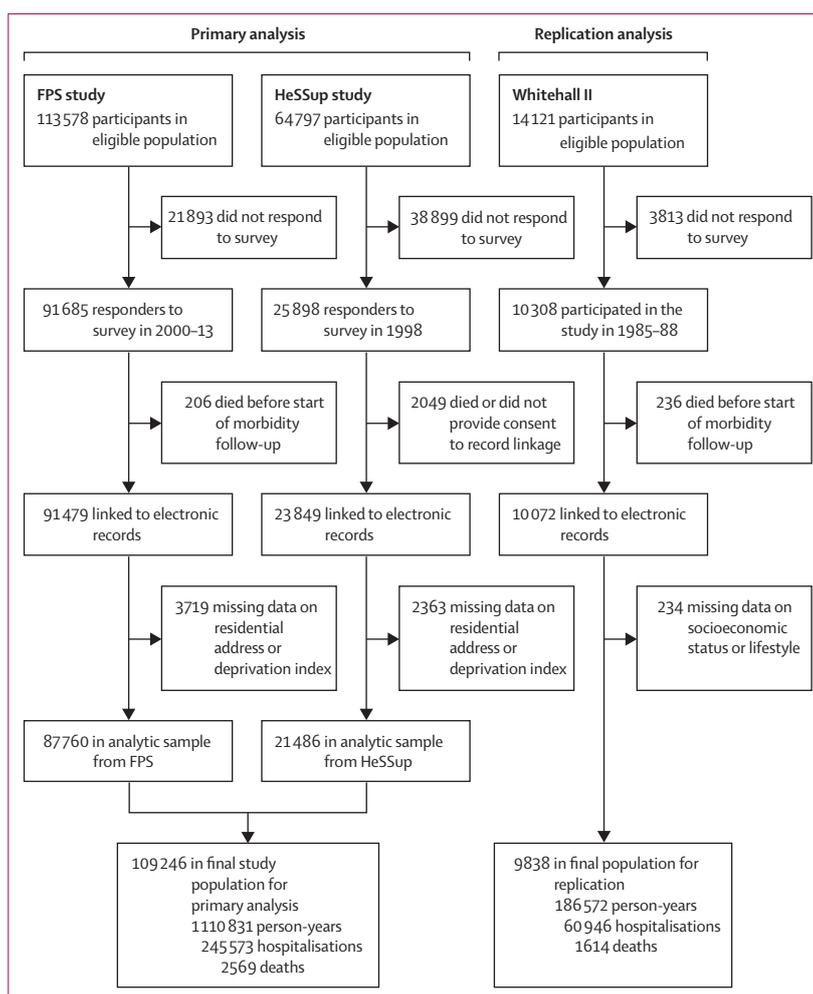
Figure 1: Selection of participants for primary and replication analyses
FPS=Finnish Public Sector. HeSSup=Health and Social Support.

temporal sequences of interconnected health conditions in analyses stratified by socioeconomic status.

First, having assessed the proportional hazards assumption (appendix pp 18–21), we examined associations between socioeconomic status and each of the 56 diseases in separate models using Cox proportional hazards regression. Follow-up continued until disease onset, death, or end of follow-up, whichever occurred first. Hazard ratios (HRs) computed for low socioeconomic status with high socioeconomic status as a reference were adjusted for the following potential confounding factors: age, sex, lifestyle factors (current smoking, heavy alcohol consumption, physical inactivity, and obesity), and cohort. In our analysis of socioeconomic status and new onset hospital-treated obesity, we did not control for baseline obesity. To identify diseases that were more common in participants with low socioeconomic status, related to socioeconomic differences that were likely to be meaningful for public health and unlikely to result from multiple testing,[23,24] only socioeconomic status–disease endpoint associations that

yielded predefined HRs equal to or greater than 1·2[23,24] and were statistically significant across the two different socioeconomic indicators (area deprivation and low education) were regarded as being sufficiently robust.

Second, we determined potential temporal sequences at recorded diagnosis of diseases that were robustly associated with socioeconomic status by testing prospective associations between all socioeconomically patterned disease pairs separately in individuals with low socioeconomic status and high socioeconomic status. We used Cox proportional hazards regression and determined temporal order in testing the associations between disease pairs based on the mean age at diagnosis; a disease with an earlier onset was treated as the predictor and a disease with a later onset as the outcome. Follow-up started at recorded diagnosis for the first disease and continued until the date of diagnosis for the next disease, death, or end of follow-up, whichever occurred first. We adjusted HRs and 95% CIs for age, sex, and study.

We constructed disease cascades from identified sequential interconnected disease pairs, starting from a single disease and continuing as far as interconnected disease pairs were available.[25] We considered diseases and disease pairs to be interconnected if the HR for the association between them exceeded an arbitrary threshold of 5 in the socioeconomic status group, irrespective of the indicator used to define the group. In sensitivity analyses, we used alternative HRs of greater than 2·5 and greater than 10 as criteria for interconnectedness.

For participants with high socioeconomic status, we analysed interconnected disease cascades, focusing on health conditions with HRs for indicators of socioeconomic status less than 1·0 (ie, diseases that were more common in high socioeconomic status groups than in low socioeconomic status groups). This more relaxed threshold was used because very few health conditions were more common in high socioeconomic status groups compared with low socioeconomic status groups.

To examine the generalisability of the findings from the Finnish cohorts in the primary analysis across geographical regions and health-care settings, we tested robust associations between socioeconomic status and health conditions in a replication analysis using the Whitehall II cohort.

All analyses were done using SAS version 9.4 and statistical code is provided in the appendix (pp 13–17).

Role of the funding source
The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. MK, JV, JP, and MJS had full access to all the data in the study. All authors had final responsibility for the decision to submit for publication.

Results
178 375 participants from the two cohorts of the primary analysis were eligible for inclusion (113 578 from the FPS study and 64 797 from the HeSSup study). 109 246 (61·2%) participants responded to the baseline questionnaire, were successfully linked to registers of socioeconomic status and health, and were included in the analytic sample (figure 1). 83 066 (76·0%) of 109 246 participants were women. The mean participant age was 44·3 years (SD 11·0) and the range was 17–77 years. According to area-based deprivation, 36 216 (33·2%) participants were in the low socioeconomic status group. 52 990 (48·5%) participants were in the low education group (appendix p 23).

During 1 110 831 person-years at risk, we recorded 245 573 hospitalisations in the 109 246 participants (figure 1). Compared with high socioeconomic status, low socioeconomic status was associated with an increased risk of 18 (32·1%) of the 56 health conditions for both indicators of socioeconomic status (HR ≥1·2; figure 2). By descending magnitude of association (ie, mean HR for the two indicators of socioeconomic status) these were self-harm, poisoning, psychotic disorders, arteriosclerosis, chronic obstructive bronchitis, lung cancer, dementia, obesity, disorders of substance abuse, pancreatitis, heart failure, anaemia, mood disorders, renal failure, diabetes, cerebral infarction, ischaemic heart disease, and disease of the liver. For disorders of substance abuse and ischaemic heart disease, the association was stronger in the first 3 years of follow-up than from year 3 onwards (appendix p 19). Minimally adjusted associations are also reported in the appendix (p 25).

Three further health conditions were associated with area deprivation, but not low education, eight health conditions with low education, but not area deprivation, and 23 with neither area deprivation nor low education (figure 2). Four health conditions were more common in groups with high socioeconomic status (melanoma, spontaneous abortion, hypertension in pregnancy, and breast cancer). Of these four conditions, only breast cancer was associated with both indicators of high socioeconomic status (ie, low area deprivation and high education).

Figure 3 shows interconnections between socioeconomically patterned health conditions and the mean age at diagnosis in participants with low socioeconomic status. The association between many disease pairs was stronger during the first 3 years of follow-up than from year 4 onwards (appendix p 22). 16 (88·9%) of 18 socioeconomically patterned health conditions were strongly interconnected as defined by HR greater than 5 for associations between disease pairs in the low socioeconomic status group, including both those with area deprivation and low education. Among participants

Figure 2: **Association between low versus high socioeconomic status and subsequent new onset diseases and health conditions by indicator of socioeconomic status**
HR=hazard ratio. *Adjusted for age, sex, cohort, lifestyle factors (smoking, heavy alcohol consumption, and physical inactivity), and obesity. †Statistically significant, HR of ≥1·2 for the association between both indicators of low socioeconomic status and disease.

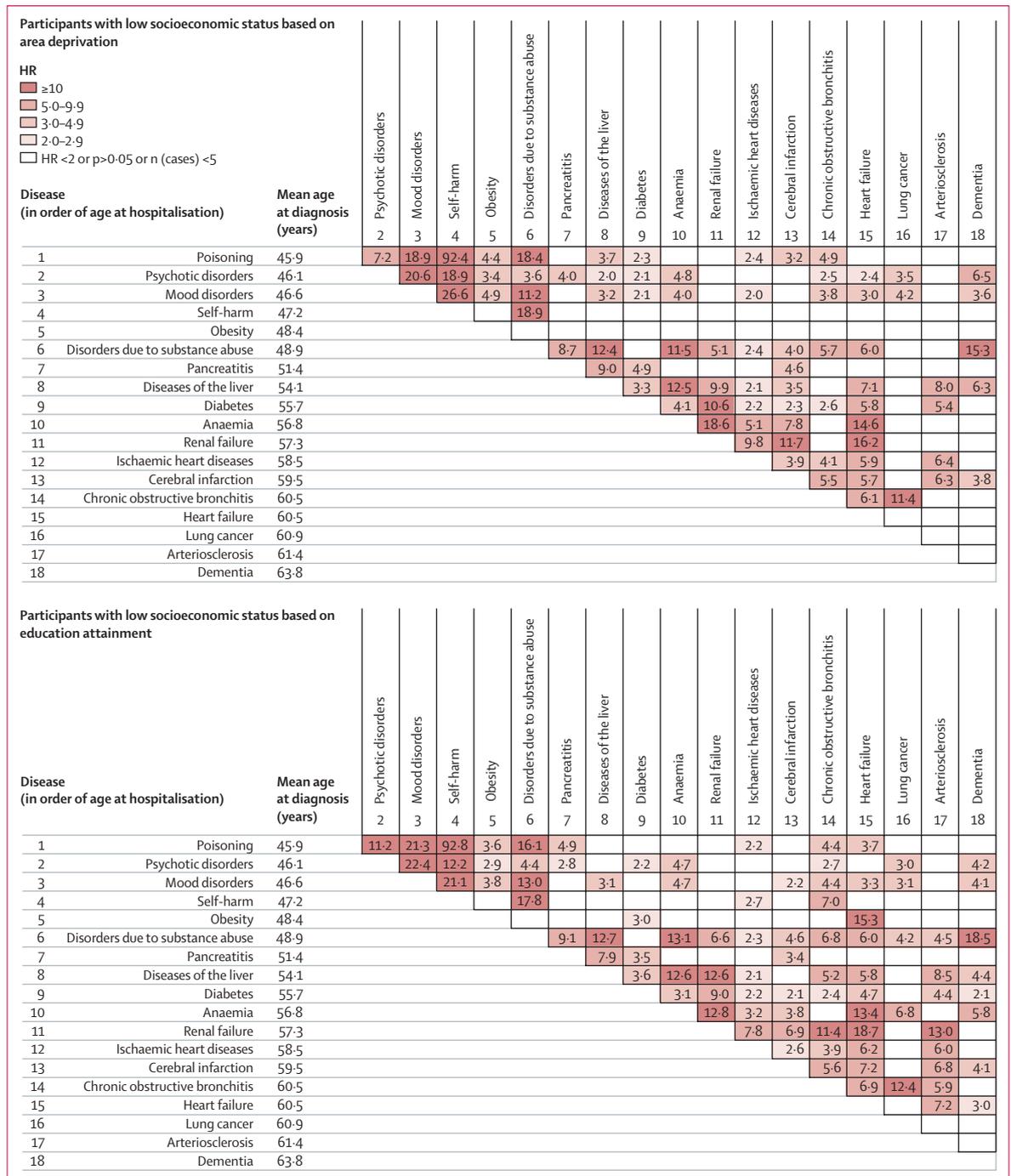|	Socioeconomic status (area deprivation)		Socioeconomic status (education)	
	New cases	HR (95% CI)*	New cases	HR (95% CI)*
Diseases associated with low socioeconomic status (both indicators)†				
Obesity	394	1·73 (1·34–2·22)	394	1·25 (1·01–1·54)
Self-harm	248	1·68 (1·23–2·28)	248	1·79 (1·37–2·34)
Poisoning	438	1·66 (1·31–2·10)	438	1·66 (1·34–2·05)
Psychotic disorders	599	1·64 (1·33–2·01)	596	1·52 (1·27–1·81)
Anaemia	358	1·53 (1·19–1·98)	357	1·27 (1·01–1·58)
Chronic obstructive bronchitis	454	1·50 (1·20–1·88)	454	1·56 (1·27–1·91)
Lung cancer	271	1·47 (1·10–1·95)	270	1·54 (1·18–2·01)
Disorders due to substance abuse	665	1·46 (1·21–1·77)	663	1·49 (1·25–1·78)
Dementia	452	1·45 (1·15–1·83)	451	1·52 (1·24–1·85)
Heart failure	466	1·42 (1·14–1·76)	466	1·41 (1·16–1·72)
Renal failure	313	1·41 (1·08–1·84)	312	1·26 (1·00–1·60)
Pancreatitis	291	1·40 (1·05–1·86)	291	1·54 (1·19–1·98)
Arteriosclerosis	208	1·39 (1·00–1·92)	208	1·78 (1·29–2·46)
Mood disorders	922	1·35 (1·15–1·60)	922	1·37 (1·19–1·58)
Diabetes	3206	1·28 (1·18–1·39)	3202	1·33 (1·23–1·43)
Cerebral infarction	765	1·23 (1·04–1·46)	763	1·28 (1·10–1·50)
Diseases of the liver	747	1·21 (1·02–1·44)	745	1·20 (1·03–1·39)
Ischaemic heart diseases	2309	1·20 (1·09–1·33)	2305	1·27 (1·16–1·38)
Diseases associated with area deprivation but not low education				
Intracerebral haemorrhage	182	1·95 (1·35–2·83)	182	0·92 (0·68–1·25)
Epilepsy	499	1·41 (1·14–1·75)	498	1·12 (0·93–1·35)
Viral infections	418	1·30 (1·02–1·65)	418	1·04 (0·85–1·27)
Diseases associated with low education but not area deprivation				
Kidney cancer	129	1·37 (0·90–2·10)	129	1·64 (1·13–2·38)
Soft tissue disorders	4826	1·09 (1·02–1·17)	4822	1·54 (1·45–1·64)
Gout	199	1·25 (0·89–1·76)	199	1·41 (1·05–1·90)
Osteoarthritis	5623	1·14 (1·07–1·21)	5616	1·35 (1·28–1·43)
Neurotic disorders	303	1·34 (1·00–1·79)	303	1·33 (1·04–1·70)
Deep vein thrombosis	431	1·20 (0·95–1·51)	429	1·26 (1·03–1·54)
Hypertension	4172	1·02 (0·95–1·10)	4167	1·24 (1·16–1·32)
Road accidents	380	1·14 (0·89–1·46)	380	1·23 (1·00–1·52)
Diseases not associated with socioeconomic status				
Diabetes in pregnancy	403	1·27 (0·96–1·67)	403	0·98 (0·79–1·21)
Pulmonary embolism	411	1·21 (0·96–1·54)	411	1·14 (0·93–1·39)
Infections and eczema	495	1·21 (0·97–1·50)	495	1·12 (0·92–1·35)
Brain cancer	111	1·19 (0·75–1·88)	111	1·21 (0·82–1·80)
Parkinson's disease	221	1·13 (0·83–1·56)	221	0·90 (0·69–1·19)
Bacterial infections	2674	1·12 (1·02–1·22)	2672	1·14 (1·05–1·24)
Inflammatory bowel disease	770	1·11 (0·94–1·33)	768	0·98 (0·84–1·13)
Sleep disorders	2702	1·11 (1·01–1·21)	2701	1·04 (0·96–1·13)
Back pain	592	1·08 (0·89–1·32)	591	1·18 (1·00–1·41)
Headaches	353	1·08 (0·83–1·41)	353	1·01 (0·81–1·27)
Transient ischaemic attack	748	1·08 (0·91–1·28)	747	1·03 (0·89–1·20)
Asthma	2393	1·08 (0·98–1·19)	2389	1·06 (0·98–1·16)
Appendicitis	1090	1·08 (0·93–1·25)	1090	1·00 (0·89–1·14)
Influenza and pneumonia	2246	1·06 (0·96–1·18)	2245	1·12 (1·03–1·23)
Falls	2282	1·05 (0·95–1·16)	2282	1·15 (1·06–1·25)
Arrhythmias	3748	1·02 (0·94–1·10)	3747	0·90 (0·84–0·96)
Injury	9374	1·00 (0·95–1·05)	9368	1·10 (1·05–1·15)
Sciatica	1426	0·99 (0·87–1·12)	1423	1·12 (1·01–1·26)
Leukaemia and lymphoma	522	0·96 (0·78–1·19)	522	1·06 (0·89–1·27)
Rheumatoid arthritis and related disorders	1873	0·95 (0·85–1·06)	1873	1·08 (0·98–1·19)
Colorectal cancer	488	0·91 (0·74–1·14)	488	1·03 (0·86–1·24)
Multiple sclerosis	222	0·89 (0·64–1·23)	222	0·83 (0·62–1·10)
Prostate cancer	497	0·83 (0·67–1·04)	496	0·84 (0·69–1·01)
Diseases associated with high socioeconomic status				
Melanoma	1236	0·88 (0·76–1·01)	1236	0·82 (0·73–0·92)
Spontaneous abortion	365	1·21 (0·90–1·61)	365	0·73 (0·58–0·91)
Hypertension in pregnancy	368	0·94 (0·70–1·24)	367	0·65 (0·52–0·82)
Breast cancer	2168	0·85 (0·77–0·94)	2166	0·85 (0·78–0·93)

Participants with low socioeconomic status based on area deprivation

HR
- ≥10
- 5·0–9·9
- 3·0–4·9
- 2·0–2·9
- HR <2 or p>0·05 or n (cases) <5

Disease (in order of age at hospitalisation)	Mean age at diagnosis (years)	Psychotic disorders (2)	Mood disorders (3)	Self-harm (4)	Obesity (5)	Disorders due to substance abuse (6)	Pancreatitis (7)	Diseases of the liver (8)	Diabetes (9)	Anaemia (10)	Renal failure (11)	Ischaemic heart diseases (12)	Cerebral infarction (13)	Chronic obstructive bronchitis (14)	Heart failure (15)	Lung cancer (16)	Arteriosclerosis (17)	Dementia (18)
1 Poisoning	45·9	7·2	18·9	92·4	4·4	18·4		3·7	2·3			2·4	3·2	4·9				
2 Psychotic disorders	46·1		20·6	18·9	3·4	3·6	4·0	2·0	2·1	4·8				2·5	2·4	3·5		6·5
3 Mood disorders	46·6			26·6	4·9	11·2		3·2	2·1	4·0		2·0		3·8	3·0	4·2		3·6
4 Self-harm	47·2					18·9												
5 Obesity	48·4																	
6 Disorders due to substance abuse	48·9						8·7	12·4		11·5	5·1	2·4	4·0	5·7	6·0			15·3
7 Pancreatitis	51·4							9·0	4·9				4·6					
8 Diseases of the liver	54·1								3·3	12·5	9·9	2·1	3·5			7·1	8·0	6·3
9 Diabetes	55·7									4·1	10·6	2·2	2·3	2·6	5·8		5·4	
10 Anaemia	56·8										18·6	5·1	7·8		14·6			
11 Renal failure	57·3											9·8	11·7		16·2			
12 Ischaemic heart diseases	58·5												3·9	4·1	5·9		6·4	
13 Cerebral infarction	59·5														5·5	5·7	6·3	3·8
14 Chronic obstructive bronchitis	60·5														6·1	11·4		
15 Heart failure	60·5																	
16 Lung cancer	60·9																	
17 Arteriosclerosis	61·4																	
18 Dementia	63·8																	

Participants with low socioeconomic status based on education attainment

Disease (in order of age at hospitalisation)	Mean age at diagnosis (years)	Psychotic disorders (2)	Mood disorders (3)	Self-harm (4)	Obesity (5)	Disorders due to substance abuse (6)	Pancreatitis (7)	Diseases of the liver (8)	Diabetes (9)	Anaemia (10)	Renal failure (11)	Ischaemic heart diseases (12)	Cerebral infarction (13)	Chronic obstructive bronchitis (14)	Heart failure (15)	Lung cancer (16)	Arteriosclerosis (17)	Dementia (18)
1 Poisoning	45·9	11·2	21·3	92·8	3·6	16·1	4·9					2·2		4·4	3·7			
2 Psychotic disorders	46·1		22·4	12·2	2·9	4·4	2·8		2·2	4·7				2·7		3·0		4·2
3 Mood disorders	46·6			21·1	3·8	13·0		3·1		4·7			2·2	4·4		3·3	3·1	4·1
4 Self-harm	47·2					17·8						2·7		7·0				
5 Obesity	48·4							3·0							15·3			
6 Disorders due to substance abuse	48·9						9·1	12·7		13·1	6·6	2·3	4·6	6·8	6·0	4·2	4·5	18·5
7 Pancreatitis	51·4							7·9	3·5				3·4					
8 Diseases of the liver	54·1								3·6	12·6	12·6	2·1		5·2	5·8		8·5	4·4
9 Diabetes	55·7									3·1	9·0	2·2	2·1	2·4	4·7		4·4	2·1
10 Anaemia	56·8										12·8	3·2	3·8		13·4	6·8		5·8
11 Renal failure	57·3											7·8	6·9	11·4	18·7		13·0	
12 Ischaemic heart diseases	58·5												2·6		3·9	6·2	6·0	
13 Cerebral infarction	59·5														5·6	7·2	6·8	4·1
14 Chronic obstructive bronchitis	60·5														6·9	12·4	5·9	
15 Heart failure	60·5																7·2	3·0
16 Lung cancer	60·9																	
17 Arteriosclerosis	61·4																	
18 Dementia	63·8																	

Figure 3: **Associations between pairs of diseases and health conditions and mean age at diagnosis in participants with low socioeconomic status**
Empty cells refer to an association with HR <2, a statistically non-significant association (p>0·05), or insufficient case numbers (less than five). HR=hazard ratio.

with high socioeconomic status, we observed no strong interconnections between the four common diseases for this group (appendix p 26).

Using mean age at recorded diagnosis, we were able to formulate a cascade of diseases in socioeconomically disadvantaged participants (figure 4). This cascade started with mental and behavioural disorders (psychiatric disorders, self-harm, and substance abuse) and was followed by pancreatitis, liver disease, anaemias, renal and heart failure, ischaemic heart disease, cerebral infarction, heart failure, arteriosclerosis, chronic obstructive bronchitis, lung cancer, or dementia. Diabetes was strongly connected
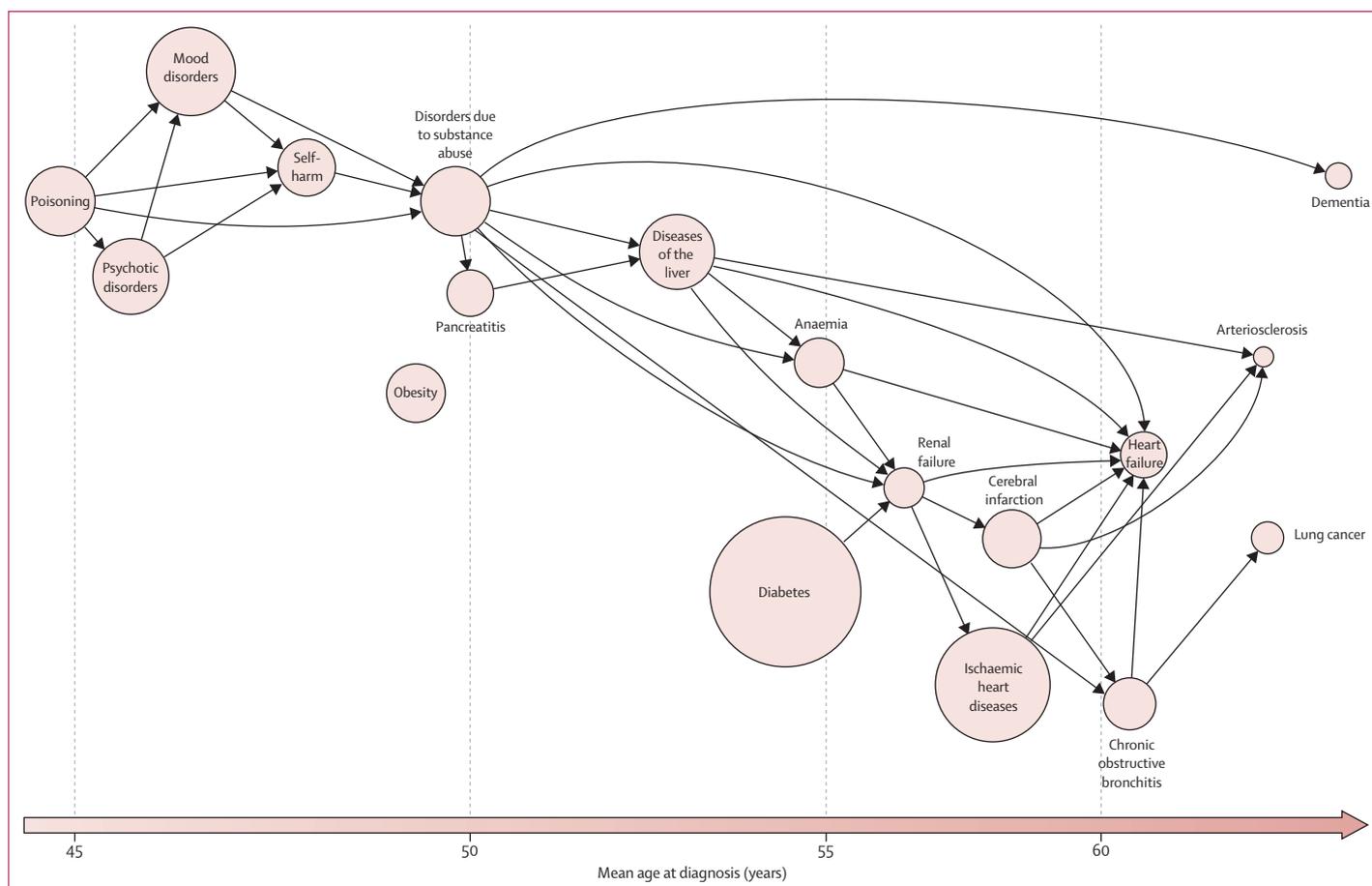
Figure 4: **Associations between diseases and health conditions in participants with low socioeconomic status**
Diseases and health conditions along the x-axis are in order of increasing mean age at diagnosis or hospitalisation. The size of plots is proportional to the incidence of disease or health condition.
Arrows link pairs of diseases or disease groups with HR >5 in low socioeconomic groups based on area deprivation and low education. HR=hazard ratio.

with this cascade via association with renal failure, whereas hospital-treated obesity was not associated with any of the diseases in the cascade. In groups with low socioeconomic status defined using only one indicator, we observed additional connections between health conditions, particularly among participants with low education (appendix pp 27–28).

When repeating our analysis of the 18 socioeconomically patterned health conditions among participants with low socioeconomic status using alternative thresholds (HRs 2·5 and 10) for connectedness (appendix p 29), the cascade of diseases starting from mental and behavioural disorders and including subsequent physical diseases remained apparent.

In a subsidiary analysis of bidirectional associations, several physical diseases were associated with subsequent mental ill health—eg, heart failure (HR 3·18, 95% CI 1·31–7·71), cerebral infarction (3·68, 1·81–7·45), chronic obstructive bronchitis (3·63, 1·50–8·82), pancreatitis (7·65, 3·94–14·85), and renal failure (5·38, 2·23–13·02) predicted later mood disorders.

The eligible population for the replication analysis was 14121 men and women from the UK Whitehall II study.[17] 10308 (73·0%) responded to the baseline survey and 9838 (69·7%) had no missing data on socioeconomic status or covariates and were successfully linked to electronic health records (figure 1; appendix p 24). In 186572 person-years at risk (mean follow-up 19·0 years), we recorded 60946 hospitalisations. All 18 associations between socioeconomic status and disease endpoints in the primary analysis were replicated (HRs ≥1·3; figure 5; minimally-adjusted HRs are presented in the appendix p 30). We found imprecision in the estimates for poisoning, self-harm, lung cancer, and arteriosclerosis as evidenced by the wide confidence intervals that included unity.

Discussion

In this study, we examined a range of mental and physical diseases and health conditions and found that low socioeconomic status was associated with 18 (32·1%) of the 56 diseases studied, independent of lifestyle factors and obesity and the indicator of socioeconomic status

	Patients	New cases	Age at diagnosis (years)		HR (95% CI)
Poisoning	9838	62	66·8		1·33 (0·60–2·99)
Psychotic disorders	9838	35	68·4		3·80 (1·35–10·70)
Neurotic disorders	9838	247	71·5		2·45 (1·64–3·66)
Mood disorders	9838	343	70·0		2·11 (1·49–2·99)
Self–harm	9838	31	66·4		2·09 (0·69–6·30)
Disorders due to substance abuse	9838	484	70·3		1·80 (1·35–2·39)
Pancreatitis	9838	56	67·0		2·41 (1·01–5·74)
Diseases of the liver	9838	248	69·2		1·62 (1·09–2·39)
Diabetes	9838	1076	69·2		1·62 (1·09–2·39)
Anaemia	9838	859	71·4		1·86 (1·49–2·32)
Renal failure	9838	775	74·7		1·89 (1·49–2·39)
Ischaemic heart disease	9838	1461	68·7		1·52 (1·29–1·80)
Cerebral infarction	9838	227	72·3		1·60 (1·04–2·46)
Chronic obstructive bronchitis	9838	508	72·8		1·75 (1·30–2·35)
Heart failure	9838	487	73·4		1·49 (1·11–1·99)
Lung cancer	9838	130	70·2		1·59 (0·87–2·90)
Arteriosclerosis	9838	71	72·6		1·58 (0·77–3·23)
Dementia	9838	346	76·4		1·56 (1·11–2·19)

0·5   1·0   2·0   4·0   4·0   16·0

Figure 5: **Associations between low versus high socioeconomic status (based on occupational position) with subsequent new-onset diseases and health conditions in Whitehall II replication cohort**
HRs, adjusted for age, sex, current smoking, heavy alcohol consumption, physical inactivity, and obesity. HR=hazard ratio.

used (area deprivation, education, or occupational position). 16 (88·9%) of these 18 socioeconomically patterned diseases were interconnected, directly or indirectly, with mental health problems and substance abuse, including conditions such as pancreatitis, liver, renal, cardiovascular, and cerebrovascular diseases, chronic obstructive bronchitis, lung cancer, and dementia. With a less stringent threshold for interconnectedness, the cascade from mental disorders to physical illness was replicated and comprised all 18 diseases. When a higher threshold for interconnectedness was set—a minimum HR of 10 between diseases—mental health problems and substance use remained strongly connected with diseases of the liver, the cardiovascular and cerebrovascular system, and dementias, the latter emphasising the importance of socioeconomic patterns in diseases related to the CNS.

Our findings are supported by several strands of evidence. The observed link between mental health and substance abuse, and between mental health and physical diseases, has been confirmed by meta-analyses measuring the impact of socioeconomically patterned adverse childhood experiences on mental disorders and chronic physical diseases in adulthood.[26] The morbidity trajectories identified in our study included several of these chronic physical diseases, such as liver, respiratory, and cardiovascular diseases. Studies have shown that mental disorders increase the risk of physical diseases, both communicable and non-communicable, via a higher tendency to commit risky behaviours, reduced self-care, and complications in help-seeking.[27,28] Additionally,

psychotropic medications have adverse effects on many aspects of physical health that accumulate over time.[27,28]

The link between mental and physical disorders could be strengthened by the bidirectional nature of this association. Although mental disorders are a risk factor for various physical diseases, the opposite direction of causality has also been observed. For example, socioeconomically patterned chronic conditions, such as cerebral infarction, heart failure, and chronic obstructive bronchitis, can increase the risk of mental disorders.[28] We confirmed these associations and found associations between pancreatitis and renal failure and subsequent mood disorders.

The range of health conditions in our study expands upon previous research on socioeconomic status and multimorbidity. A study of older Taiwanese people found low education to be associated with increasing trajectories of cardiovascular and chronic non-specific lung diseases[12] and a UK study reported a link between low occupational grade and an increased risk of developing cardiometabolic multimorbidity (ie, two or three of diabetes, myocardial infarction, and stroke).[3] A Canadian study identified an association between lower income and greater overall multimorbidity,[13] a German study of primary care patients identified an association between low education and a higher number of diagnoses, particularly cardiometabolic diseases,[14] and a study of Australian women found a relationship between low education and difficulties in managing income and increased self-reported multimorbidity in repeated questionnaire surveys.[11]

We adjusted the association between socioeconomic status and health conditions for lifestyle behavioural factors, such as self-reported heavy alcohol consumption, smoking, physical inactivity, and obesity. This approach is conservative, as these factors are both confounders and potentially part of the causal pathway from socioeconomic disadvantage to disease. Prospective life-course research supports socioeconomic disadvantage as an origin of unhealthy lifestyle behaviours and subsequent morbidity. In a cohort study of Finnish children and adolescents, for example, differences in risk factors between socioeconomic groups at the beginning of follow-up were small, but large differences emerged in the third decade of life.[20] In addition to risk behaviours, such as unhealthy diet, physical inactivity, and smoking, low socioeconomic status was associated with a poorer glycaemic profile in early adulthood and, like our findings, an excess prevalence of obesity, diabetes, fatty liver, and cardiovascular disease in middle age.[20]

Our findings have important implications for research and public health policy. The pattern of mental health problems and substance abuse preceding socioeconomically patterned physical diseases is not reflected in global strategies to prevent diseases. The WHO Sustainable Development Goals and the Global Action Plan for the Prevention and Control of Non-Communicable Diseases, for example, have their main focus on physical health;[6] the 2013–2020 WHO Global Plan for the Prevention and Control of Non-Communicable Diseases[29] and the Global Burden of Disease Collaboration do not include socioeconomic disadvantage as a modifiable risk factor.[30] Moreover, treatment of psychiatric disorders, physical disease, and substance abuse is often split between health-care and social services.[27,28] This approach is unlikely to be optimal for tackling problems with shared health determinants, including socioeconomic inequalities in morbidity. The 2019 *Lancet* Commission drew attention to the need to improve protection of physical health in people with psychiatric disorders;[27] our findings suggest this is particularly important for people living in socioeconomic disadvantage.

This study has several limitations. The response to baseline assessment varied between 61% in the primary analysis and 70% in the replication analysis. Sample attrition might lead to an overestimation or underestimation of the true associations between socioeconomic status and health. We measured morbidity mainly using electronic health records, which covered hospital-treated diseases. For some conditions, such as asthma, diabetes, and hypertension, additional non-hospitalised cases were identified via linkage to records of eligibility for special reimbursement for medication. However, we will have inevitably omitted undiagnosed conditions and less severe cases that are largely dealt with in primary care (eg, obesity). Therefore, the observed interconnectedness between diseases reflected the temporal order of treatments rather than causal associations between health conditions. The age distribution of participants at study induction meant that we did not have data on children or very old people. The generalisability of our findings beyond Finland and the UK and to other health-care systems is also uncertain and requires testing.

However, by applying a data-driven approach to a wide set of diseases and health conditions, we refocus the field of socioeconomic inequality research from traditional analysis of specific diseases to the study of interconnected diseases. A large sample size, longitudinal design, minimal sample attrition after baseline because of follow-up via electronic health records, and the validation of our results across different indicators of socioeconomic status and health-care settings are further strengths.

In conclusion, by mapping morbidity from electronic health records we showed that low socioeconomic status is a risk factor for a spectrum of interconnected diseases and health conditions. Our analyses of interconnected diseases highlight the importance of mental health problems and substance abuse in the cascade of socioeconomically patterned physical illnesses. These findings suggest that policy and health-care practice addressing mental health issues in social context and early in the life course might be effective strategies for reducing health inequalities.

References

1   Dalstra JA, Kunst AE, Borrell C, et al. Socioeconomic differences in the prevalence of common chronic diseases: an overview of eight European countries. *Int J Epidemiol* 2005; **34:** 316–26.

2   Stringhini S, Carmeli C, Jokela M, et al. Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: multi-cohort population based study. *BMJ* 2018; **360:** k1046.

3   Singh-Manoux A, Fayosse A, Sabia S, et al. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of cardiometabolic multimorbidity and mortality: a cohort study. *PLoS Med* 2018; **15:** e1002571.

4   Mackenbach JP, Stirbu I, Roskam AJ, et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008; **358:** 2468–81.

5   WHO. Commission on Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. Geneva: World Health Organization, 2008.

6   UN. Transforming our world: the 2030 agenda for sustainable development. New York, NY: United Nations, 2015.

7   Frølich A, Ghith N, Schiøtz M, Jacobsen R, Stockmarr A. Multimorbidity, healthcare utilization and socioeconomic status: a register-based study in Denmark. *PLoS One* 2019; **14:** e0214183.

8   van den Akker M, Buntinx F, Metsemakers JF, Roos S, Knottnerus JA. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J Clin Epidemiol* 1998; **51:** 367–75.

9   Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012; **380:** 37–43.

10   Park B, Ock M, Lee HA, et al. Multimorbidity and health-related quality of life in Koreans aged 50 or older using KNHANES 2013–2014. *Health Qual Life Outcomes* 2018; **16:** 186.

11   Jackson CA, Dobson A, Tooth L, Mishra GD. Body mass index and socioeconomic position are associated with 9-year trajectories of multimorbidity: a population-based study. *Prev Med* 2015; **81:** 92–98.

12   Hsu HC. Trajectories of multimorbidity and impacts on successful aging. *Exp Gerontol* 2015; **66:** 32–38.

13   Canizares M, Hogg-Johnson S, Gignac MAM, Glazier RH, Badley EM. Increasing trajectories of multimorbidity over time: birth cohort differences and the role of changes in obesity and income. *J Gerontol B Psychol Sci Soc Sci* 2018; **73:** 1303–14.

14   Schäfer I, Hansen H, Schön G, et al. The influence of age, gender and socio-economic status on multimorbidity patterns in primary care. First results from the multicare cohort study. *BMC Health Serv Res* 2012; **12:** 89.

15   Korkeila K, Suominen S, Ahvenainen J, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol* 2001; **17:** 991–99.

16   Kivimäki M, Lawlor DA, Davey Smith G, et al. Socioeconomic position, co-occurrence of behavior-related risk factors, and coronary heart disease: the Finnish Public Sector study. *Am J Public Health* 2007; **97:** 874–79.

17   Marmot MG, Smith GD, Stansfeld S, et al. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991; **337:** 1387–93.

18   Kivimäki M, Batty GD, Singh-Manoux A, Britton A, Brunner EJ, Shipley MJ. Validity of cardiovascular disease event ascertainment using linkage to UK hospital records. *Epidemiology* 2017; **28:** 735–39.

19   Townsend P, Beattie A, Phillimore P. Health and deprivation: inequality and the North. London: Croom Helm, 1988.

20   Kivimaki M, Vahtera J, Tabák AG, et al. Neighbourhood socioeconomic disadvantage, risk factors, and diabetes from childhood to middle age in the Young Finns Study: a cohort study. *Lancet Public Health* 2018; **3:** e365–73.

21   Elovainio M, Ferrie JE, Singh-Manoux A, et al. Socioeconomic differences in cardiometabolic factors: social causation or health-related selection? Evidence from the Whitehall II Cohort Study, 1991-2004. *Am J Epidemiol* 2011; **174:** 779–89.

22   Sommerlad A, Perera G, Singh-Manoux A, Lewis G, Stewart R, Livingston G. Accuracy of general hospital dementia diagnoses in England: sensitivity, specificity, and predictors of diagnostic accuracy 2008-2016. *Alzheimers Dement* 2018; **14:** 933–43.

23   Olivier J, May WL, Bell ML. Relative effect sizes for measures of risk. *Commun Stat* 2017; **46:** 6774–81.

24   Siontis GC, Ioannidis JP. Risk factors and interventions with statistically significant tiny effects. *Int J Epidemiol* 2011; **40:** 1292–307.

25   Jensen AB, Moseley PL, Oprea TI, et al. Temporal disease trajectories condensed from population-wide registry data covering 6·2 million patients. *Nat Commun* 2014; **5:** 4022.

26   Hughes K, Bellis MA, Hardcastle KA, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *Lancet Public Health* 2017; **2:** e356–66.

27   Firth J, Siddiqi N, Koyanagi A, et al. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. *Lancet Psychiatry* 2019; **6:** 675–712.

28   Prince M, Patel V, Saxena S, et al. No health without mental health. *Lancet* 2007; **370:** 859–77.

29   WHO. Global action plan for the prevention and control of noncommunicable diseases 2013–2020. Geneva: World Health Organization, 2013.

30   GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; **392:** 1923–94.