Effect on life expectancy of temporal sequence in a multimorbidity cluster of psychosis, diabetes, and congestive heart failure among 1·7 million individuals in Wales with 20-year follow-up: a retrospective cohort study using linked data

Rhiannon K Owen, Jane Lyons, Ashley Akbari, Bruce Guthrie, Utkarsh Agrawal, Daniel C Alexander, Amaya Azcoaga-Lorenzo, Anthony J Brookes, Spiros Denaxas, Carol Dezateux, Adeniyi Francis Fagbamigbe, Gill Harper, Paul D W Kirk, Eda Bilici Özyiğit, Sylvia Richardson, Sophie Staniszewska, Colin McCowan, Ronan A Lyons, Keith R Abrams

Summary

Background To inform targeted public health strategies, it is crucial to understand how coexisting diseases develop over time and their associated impacts on patient outcomes and health-care resources. This study aimed to examine how psychosis, diabetes, and congestive heart failure, in a cluster of physical–mental health multimorbidity, develop and coexist over time, and to assess the associated effects of different temporal sequences of these diseases on life expectancy in Wales.

Methods In this retrospective cohort study, we used population-scale, individual-level, anonymised, linked, demographic, administrative, and electronic health record data from the Wales Multimorbidity e-Cohort. We included data on all individuals aged 25 years and older who were living in Wales on Jan 1, 2000 (the start of follow-up), with follow-up continuing until Dec 31, 2019, first break in Welsh residency, or death. Multistate models were applied to these data to model trajectories of disease in multimorbidity and their associated effect on all-cause mortality, accounting for competing risks. Life expectancy was calculated as the restricted mean survival time (bound by the maximum follow-up of 20 years) for each of the transitions from the health states to death. Cox regression models were used to estimate baseline hazards for transitions between health states, adjusted for sex, age, and area-level deprivation (Welsh Index of Multiple Deprivation [WIMD] quintile).

Findings Our analyses included data for 1 675 585 individuals (811 393 [48·4%] men and 864 192 [51·6%] women) with a median age of 51·0 years (IQR 37·0–65·0) at cohort entry. The order of disease acquisition in cases of multimorbidity had an important and complex association with patient life expectancy. Individuals who developed diabetes, psychosis, and congestive heart failure, in that order (DPC), had reduced life expectancy compared with people who developed the same three conditions in a different order: for a 50-year-old man in the third quintile of the WIMD (on which we based our main analyses to allow comparability), DPC was associated with a loss in life expectancy of 13·23 years (SD 0·80) compared with the general otherwise healthy or otherwise diseased population. Congestive heart failure as a single condition was associated with mean a loss in life expectancy of 12·38 years (0·00), and with a loss of 12·95 years (0·06) when preceded by psychosis and 13·45 years (0·13) when followed by psychosis. Findings were robust in people of older ages, more deprived populations, and women, except that the trajectory of psychosis, congestive heart failure, and diabetes was associated with higher mortality in women than men. Within 5 years of an initial diagnosis of diabetes, the risk of developing psychosis or congestive heart failure, or both, was increased.

Interpretation The order in which individuals develop psychosis, diabetes, and congestive heart failure as combinations of conditions can substantially affect life expectancy. Multistate models offer a flexible framework to assess temporal sequences of diseases and allow identification of periods of increased risk of developing subsequent conditions and death.

Funding Health Data Research UK.

Lancet Public Health 2023; 8: e535–45

Population Data Science, Health Data Research, Swansea University Medical School, Faculty of Medicine, Health and Life Science, Swansea University, Swansea, UK (R K Owen PhD, J Lyons MSc, A Akbari MSc, Prof R A Lyons MD); **Advanced Care Research Centre, Usher Institute, University of Edinburgh, Edinburgh, UK** (Prof B Guthrie PhD); **Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK** (U Agrawal PhD); **Centre for Medical Image Computing, Department of Computer Science, Faculty of Engineering Sciences, University College London, London, UK** (Prof D C Alexander PhD, E B Özyiğit PhD); **School of Medicine, University of St Andrews, St Andrews, UK** (A Azcoaga-Lorenzo PhD, A F Fagbamigbe PhD, Prof C McCowan PhD); **Hospital Rey Juan Carlos, Instituto de Investigación Sanitaria Fundación Jimenez Diaz, Madrid, Spain** (A Azcoaga-Lorenzo); **Department of Genetics, University of Leicester, Leicester, UK** (Prof A J Brookes PhD); **Institute of Health Informatics, University College London, London, UK** (Prof S Denaxas PhD); **Clinical Effectiveness Group, Wolfson Institute of Population Health, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK** (Prof Carol Dezateux MD, Gill Harper PhD); **MRC**

Introduction

Multimorbidity is defined as the coexistence of two or more long-term health conditions in an individual.[1] It is associated with decreased quality of life, functional decline, polypharmacy, increased treatment burden, increased health-care use (including emergency hospital admissions),[2] and increased mortality.[3,4]

The burden placed on a patient and their health service providers is exacerbated with increasing numbers of coexisting conditions.[5] The health and social care costs

Biostatistics Unit, University of Cambridge, Cambridge, UK (P D W Kirk PhD, Prof S Richardson PhD); **Cambridge Institute of Therapeutic Immunology and Infectious Disease, University of Cambridge, Cambridge, UK** (P D W Kirk); **Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK** (Prof S Staniszewska DPhil); **Department of Statistics, University of Warwick, Coventry, UK** (Prof K R Abrams PhD); **Centre for Health Economics, University of York, York, UK** (Prof K R Abrams)

Correspondence to: Dr Rhiannon K Owen, Population Data Science, Health Data Research, Swansea University Medical School, Faculty of Medicine, Health and Life Science, Swansea University, Swansea SA2 8PP, UK
r.k.owen@swansea.ac.uk

Research in context

Evidence before this study

The epidemiology of multimorbidity has previously concentrated on the static clustering of diseases with respect to time, rather than their dynamic evolution or ordered trajectory over the life course. A systematic literature review and comparison of approaches used to model trajectories of disease in multimorbidity was published in 2021. We searched PubMed, Google Scholar, and relevant citations to identify articles published after this literature search (from May 1, 2020, to April 30, 2022) using two key search terms (multimorbidity AND trajectory). Several studies have used state transition modelling approaches to describe trajectories of disease in multimorbidity. However, none of these approaches have explored the sequence of disease occurrence within multimorbidity clusters and the effect of those sequences on patient-relevant outcomes, such as mortality.

Added value of this study

To our knowledge, this is the first study to explore the order of diseases within clusters of multimorbidity and the associated effect on mortality. The application of multistate models to physical–mental health multimorbidity (psychosis, diabetes, and congestive heart failure) with use of population-scale, individual-level, anonymised, linked, routinely collected, electronic health record data sources will help inform patients, clinicians, and health-care decision makers on the appropriate identification and management of patient care in this multimorbidity cluster, and thus have the potential to improve patient-relevant outcomes in an under-studied population.

Implications of all the available evidence

The novel application of multistate models in the context of evolving multimorbidity allows a holistic assessment of the effect of disease sequence on patient outcomes, such as mortality, together with associated implications for future research prioritisation, study design, and policy in multimorbidity.

associated with multimorbidity exceed the expected cost of managing the individual component conditions alone, with costs approximately doubling for every additional chronic condition.[6] Over the last 35 years, the prevalence of multimorbidity has increased worldwide.[7] In the UK, it is estimated that more than 25% of the adult population has two or more long-term conditions.[8,9] This estimate increases to 65% for those older than 65 years, and almost 82% for those aged 85 years or older.[8]

Multimorbidity is a major health concern worldwide as global populations live longer, with vast implications for health service providers such as the National Health Service and the Department of Health and Social Care in the UK.[10] Health-care decision making has previously focused on developing recommendations for disease-specific guidelines in single conditions.[11] However, standardised care for each chronic condition in isolation can be inappropriate for individuals with multimorbidity.[12] Targeted approaches to prevention or delay of multimorbidity are advocated as being a crucial component of future health-care planning.[13] To efficiently implement targeted approaches, we must first understand how commonly coexisting diseases develop over time and their associated burden on patients and health-care resources.

Multimorbidity can encompass many different combinations of long-term conditions, and these combinations can cluster in a concordant or discordant manner.[14] To date, the epidemiology of multimorbidity has concentrated on the static clustering of diseases in cross-sectional analyses, with less attention paid to trajectories of disease onset and the sequence in which conditions develop.[15] Understanding the order of disease onset within a cluster of multimorbidity and the associated effect on patient outcomes allows the identification of disease trajectories with an increased risk of death, as well as periods of increased risk of developing further conditions within the cluster. Such an approach could facilitate health-care resource planning, including identification of optimal periods for screening and appropriate populations to target for preventive strategies, contributing to earlier diagnosis and management. Increasing clinicians' awareness of the risks of individuals developing new conditions could contribute to earlier identification and intervention, and thus improve patient outcomes. In physical–mental health multimorbidity, it is important to understand the potential synergistic effects of combinations of physical and mental health conditions, together with the importance of disease sequence on patient outcomes, which can help to inform care management and contribute to earlier diagnosis. Following a data-driven supervised cluster analysis, a cluster of three conditions in physical–mental health multimorbidity (psychosis, diabetes, and congestive heart failure) that commonly coexist in different ordered sequences and are associated with high all-cause mortality were identified, and further supported by clinical opinion and the current literature.[16–18] The objective of this study was to examine how diseases within a cluster develop over time and the associated effects of different disease trajectories on patient outcomes in Wales. More generally, our work paves the way for associating cross-sectional clustering with a time-resolved analysis of the sequence of events and their associated effects on patient outcomes.

Methods

Study design

We conducted a population-scale retrospective cohort study using the Wales Multimorbidity e-Cohort (WMC) held within the Secure Anonymised Information Linkage (SAIL) Databank.[19] The SAIL Databank's independent Information Governance Review Panel approved this research (SAIL Project 0911) and the development of the WMC.[20] Data acquired into the SAIL Databank must meet required legal and governance requirements, which can include consent as the appropriate legal mechanism. We worked with the SAIL Consumer Panel and the independent Information Governance Review Panel, which includes members of the public, to develop the WMC and study protocol for this work. In accordance with the Health Research Authority guidance, ethical approval is not required for projects involving anonymised data. However, the independent Information Governance Review Panel reviews all projects using the SAIL Databank prior to approval with consideration given to potential ethical concerns.

To identify common clusters of physical–mental health multimorbidity associated with a high risk of death, a supervised cluster analysis using Bayesian profile regression[21] was applied to the WMC. Of the multimorbid individuals within the WMC, 6% had developed two or more of psychosis, diabetes, and congestive heart failure. Of those, 2% developed all three conditions in various orders, allowing inferences to be drawn regarding the effects of different disease trajectories on patient outcomes. Following consensus discussion with clinical experts involved in the Measuring and Understanding Multimorbidity using Routine Data in the UK project through which the current study was funded, psychosis, diabetes, and congestive heart failure were selected from the common clusters of diseases identified by the supervised cluster analysis, and were further supported by the literature as being associated with a high risk of death when occurring together.[16–18]

Data sources and setting

The WMC has linked demographic, mortality, and electronic health record data sources through anonymised linkage fields at the individual level. Demographic data including sex assigned at birth, week of birth, and geographical variables such as Lower layer Super Output Area (LSOA) 2001 at cohort inception and LSOA 2011 at cohort end were obtained from the Welsh Demographic Service Dataset, which holds administrative information on the population of Wales registered with Welsh general practices. Mortality data were obtained from the Annual District Death Extract from the Office for National Statistics (ONS), which holds information on the dates and causes of death for all Welsh residents (including those who died outside of Wales). Electronic health record data were obtained from the Welsh Longitudinal General Practice dataset and the Patient Episode Database for Wales. Diagnoses and dates of diagnoses were obtained from these data sources for individuals registered to general practices providing data to the SAIL Databank, using Elixhauser comorbidity index phenotyping to identify diagnoses.[22] The Elixhauser comorbidity ICD-10 code list[23] was used to capture inpatient diagnoses and the Elixhauser comorbidity Read code list[24] was used to capture diagnoses recorded in primary care. Full code lists are provided in the appendix (pp 2–26). Elixhauser comorbidity index phenotyping was used because it includes more chronic conditions than the Charlson comorbidity index, and there are established code lists across both primary care (including Read codes) and secondary care (including ICD-10 codes) settings, whereas the Cambridge multimorbidity score has predominantly been designed and validated in primary care settings only. Full information regarding curation of the data in WMC and associated methods has been published previously.[20]

See Online for appendix

Participants

We included all individuals aged 25 years and older who were living in Wales on Jan 1, 2000 (the start of follow-up), with follow-up continuing until Dec 31, 2019, first break in Welsh residency, or death. Individuals had a minimum of 1 day and a maximum of 20 years of follow-up. To ensure that we captured all possible electronic health record data, including primary care data from Jan 1, 2000, we restricted the population to individuals registered with a general practice providing data to the SAIL Databank. These practices currently include 83·4% of general practices across all health boards in Wales, encompassing 83·7% of all individuals in the WMC. All individuals registered with a general practice providing data to the Welsh Longitudinal General Practice in SAIL can be linked to Welsh Demographic Service Dataset, Patient Episode Database for Wales, and Annual District Death Extract.

Outcome measures

The primary outcome measure was all-cause mortality identified using the Annual District Death Extract from the ONS. Secondary outcome measures included health state occupancy and transition rates between health states for psychosis, diabetes, and congestive heart failure. Health states were defined as the first diagnosis of psychosis, diabetes, congestive heart failure, or their combinations based on the Elixhauser comorbidity index phenotype definitions for diagnoses recorded in primary[24] and secondary care.[23] Diabetes included a diagnosis of either uncomplicated or complicated diabetes, including type 1 and type 2 diabetes, with the date of first diabetes diagnosis of any type being recorded. For psychosis, Elixhauser comorbidity index phenotyping includes symptomatic psychosis and clinical diagnoses (appendix pp 2–9).

	n	Median age, years*	Welsh Index of Multiple Deprivation					Sex	
			First quintile (most deprived)	Second quintile	Third quintile	Fourth quintile	Fifth quintile (least deprived)	Male	Female
OH/OD†	1 675 585	51·0 (37·0–65·0)	345 843 (20·6%)	337 683 (20·2%)	346 764 (20·7%)	306 130 (18·3%)	339 165 (20·2%)	811 393 (48·4%)	864 192 (51·6%)
P	24 990	53·9 (41·8–70·5)	7042 (28·2%)	5348 (21·4%)	4791 (19·2%)	3945 (15·8%)	3864 (15·5%)	11 486 (46·0%)	13 504 (54·0%)
PD	3346	49·7 (39·5–61·1)	1071 (32·0%)	733 (21·9%)	664 (19·8%)	444 (13·3%)	434 (13·0%)	1565 (46·8%)	1781 (53·2%)
PDC	276	63·4 (53·0–74·0)	100 (36·2%)	58 (21·0%)	56 (20·3%)	41 (14·9%)	21 (7·6%)	135 (48·9%)	141 (51·1%)
PC	1354	76·1 (65·5–84·0)	342 (25·3%)	300 (22·2%)	268 (19·8%)	234 (17·3%)	210 (15·5%)	549 (40·5%)	805 (59·5%)
PCD	115	68·7 (60·7–79·0)	31 (27·0%)	28 (24·3%)	19 (16·5%)	16 (13·9%)	21 (18·3%)	47 (40·9%)	68 (59·1%)
D	281 458	60·1 (49·4–70·2)	66 022 (23·5%)	59 457 (21·1%)	57 971 (20·6%)	47 574 (16·9%)	50 434 (17·9%)	145 113 (51·6%)	136 345 (48·4%)
DP	4209	55·2 (42·6–67·6)	1235 (29·3%)	921 (21·9%)	858 (20·4%)	624 (14·8%)	571 (13·6%)	1942 (46·1%)	2267 (53·9%)
DPC	397	68·9 (59·0–77·6)	115 (29·0%)	100 (25·2%)	73 (18·4%)	65 (16·4%)	44 (11·1%)	181 (45·6%)	216 (54·4%)
DC	39 457	75·8 (67·8–82·6)	9476 (24·0%)	8540 (21·6%)	8208 (20·8%)	6670 (16·9%)	6563 (16·6%)	21 473 (54·4%)	17 984 (45·6%)
DCP	320	73·9 (66·0–81·3)	93 (29·1%)	73 (22·8%)	58 (18·1%)	54 (16·9%)	42 (13·1%)	146 (45·6%)	174 (54·4%)
C	123 981	78·1 (69·4–84·9)	26 876 (21·7%)	26 070 (21·0%)	26 090 (21·0%)	22 266 (18·0%)	22 679 (18·3%)	60 428 (48·7%)	63 553 (51·3%)
CP	952	76·7 (68·0–83·9)	240 (25·2%)	208 (21·8%)	168 (17·6%)	160 (16·8%)	176 (18·5%)	370 (38·9%)	582 (61·1%)
CPD	80	69·6 (60·5–78·9)	23 (28·8%)	21 (26·3%)	10 (12·5%)	12 (15·0%)	14 (17·5%)	41 (51·3%)	39 (48·8%)
CD	11 377	75·1 (67·1–82·1)	2719 (23·9%)	2542 (22·3%)	2309 (20·3%)	1912 (16·8%)	1895 (16·7%)	6293 (55·3%)	5084 (44·7%)
CDP	116	75·4 (67·2–81·7)	31 (26·7%)	21 (18·1%)	21 (18·1%)	22 (19·0%)	21 (18·1%)	55 (47·4%)	61 (52·6%)
Death	424 180	80·2 (71·1–87·1)	96 249 (22·7%)	90 550 (21·3%)	89 120 (21·0%)	73 775 (17·4%)	74 486 (17·6%)	204 576 (48·2%)	219 604 (51·8%)

Data are n, median (IQR), or n (%). Diseases are referred to by their one-letter abbreviations, with the order of letters indicating the order in which the diseases were diagnosed within a trajectory. OH/OD=otherwise healthy or otherwise diseased. P=psychosis. D=diabetes. C=congestive heart failure. *Age at entry into the health state (or cohort entry, for the OH/OD row). †All individuals included in the cohort, who were in an OH/OD state at the start of follow-up.

Table 1: **Summary statistics for individuals in each health state**

Statistical analysis

Multistate models[25] were used to model trajectories of disease in multimorbidity and their associated effect on mortality, accounting for competing risks. The modelling framework for combinations of psychosis, diabetes, and congestive heart failure is illustrated in the appendix (p 27). All individuals start in the otherwise healthy or otherwise diseased health state at cohort inception, as it was not possible to capture diagnoses before cohort inception because of poor quality of reporting. Individuals can transition to death or other health states where a direct pathway exists. Death is included as an absorbing state. The model allows simultaneous estimation of each disease pathway, or disease trajectories, through estimation of all possible transitions between health states. The index condition refers to the first diagnosis received of either psychosis, diabetes, or congestive heart failure. Patients diagnosed with two or more of the three conditions on the same date do not inform transitions between diseases and, therefore, were excluded from the analysis (n=173). Time to diagnosis or time to death in years was used as the timescale for transition between health states. Separate baseline hazards were assumed for each of the transitions, and Cox proportional hazards regression models were fitted to adjust for covariate effects. All analyses were adjusted for age at entry into the health state with use of a time-varying covariate, sex, and area-level deprivation using quintiles according to the Welsh Index of Multiple Deprivation (WIMD; version 2011) at cohort inception. Complete data were obtained for age, sex, and area-level deprivation. Individuals were censored on the first date corresponding to a break in Welsh residency, end of follow-up period, or date of death when death is a competing risk. The proportional hazards assumption was assessed using the Schoenfeld test. Proportional hazards did not hold for age, sex, or area-level deprivation; therefore, we report restricted mean survival from entry into each health state as opposed to hazard ratios.[26] Restricted mean survival time represents the mean time free from an event (development of a condition of interest, or death) and is bound by the maximum follow-up of 20 years. Life expectancy was calculated as the restricted mean survival time for each of the transitions from the health states to death. Estimates of uncertainty were calculated via the Delta method using a second-order Taylor series expansion. The associated loss or gain in life expectancy for each of the health states was compared with the general population of people without a diagnosis of psychosis, diabetes, or congestive heart failure during follow-up. To make appropriate comparisons with respect to age, sex, and area-level deprivation, results are presented for a 50-year-old man in the third WIMD quintile. Results for older ages, women, and more deprived populations were explored in sensitivity analyses. Models were further adjusted for the time spent in the previous health state as a proxy measure of disease severity in further sensitivity analyses. All

	Transition to health state	Death	No event	Restricted mean survival*†, years	Estimated change in life expectancy†, years			
					vs OH/OD	*vs* P	*vs* D	*vs* C
OH/OD‡	1 675 585	257 016 (15·3%)	988 140 (59·0%)	19·23 (0·00)	..	2·29 (0·00)	−0·19 (0·00)	12·38 (0·00)
P	24 990/1 675 585 (1·5%)	6492 (26·0%)	13 798 (55·2%)	16·94 (0·00)	−2·29 (0·00)	..	−2·48 (0·00)	10·09 (0·00)
PD	3346/24 990 (13·4%)	702 (21·0%)	2368 (70·8%)	18·68 (0·01)	−0·55 (0·01)	1·74 (0·01)	−0·74 (0·01)	11·83 (0·01)
PDC	276/3346 (8·2%)	138 (50·0%)	138 (50·0%)	9·91 (0·87)	−9·32 (0·87)	−7·03 (0·87)	−9·51 (0·87)	3·06 (0·87)
PC	1354/24 990 (5·4%)	847 (62·6%)	392 (29·0%)	6·28 (0·06)	−12·95 (0·06)	−10·66 (0·06)	−13·14 (0·06)	−0·57 (0·06)
PCD	115/1354 (8·5%)	57 (49·6%)	58 (50·4%)	15·12 (0·53)	−4·11 (0·53)	−1·82 (0·53)	−4·30 (0·53)	8·27 (0·53)
D	281 458/1 675 585 (16·8%)	48 911 (17·4%)	188 881 (67·1%)	19·42 (0·00)	0·19 (0·00)	2·48 (0·00)	..	12·57 (0·00)
DP	4209/281 458 (1·5%)	1044 (24·8%)	2768 (65·8%)	18·22 (0·03)	−1·01 (0·03)	1·28 (0·03)	−1·20 (0·03)	11·37 (0·03)
DPC	397/4209 (9·4%)	224 (56·4%)	173 (43·6%)	6·00 (0·80)	−13·23 (0·80)	−10·94 (0·80)	−13·42 (0·80)	−0·85 (0·80)
DC	39 457/281 458 (14·0%)	22 096 (56·0%)	17 041 (43·2%)	10·25 (0·01)	−8·98 (0·01)	−6·69 (0·01)	−9·17 (0·01)	3·40 (0·01)
DCP	320/39 457 (0·8%)	189 (59·1%)	131 (40·9%)	9·55 (0·35)	−9·68 (0·35)	−7·39 (0·35)	−9·87 (0·35)	2·70 (0·35)
C	123 981/1 675 585 (7·4%)	79 126 (63·8%)	32 526 (26·2%)	6·85 (0·00)	−12·38 (0·00)	−10·09 (0·00)	−12·57 (0·00)	..
CP	952/123 981 (0·8%)	601 (63·1%)	271 (28·5%)	5·78 (0·13)	−13·45 (0·13)	−11·16 (0·13)	−13·64 (0·13)	−1·07 (0·13)
CPD	80/952 (8·4%)	51 (63·8%)	29 (36·3%)	16·77 (0·42)	−2·46 (0·42)	−0·17 (0·42)	−2·65 (0·42)	9·92 (0·42)
CD	11 377/123 981 (9·2%)	6605 (58·1%)	4656 (40·9%)	12·79 (0·01)	−6·44 (0·01)	−4·15 (0·01)	−6·63 (0·01)	5·94 (0·01)
CDP	116/11 377 (1·0%)	81 (69·8%)	35 (30·2%)	14·41 (0·39)	−4·82 (0·39)	−2·53 (0·39)	−5·01 (0·39)	7·56 (0·39)

Data are n (%) or mean (SD). Diseases are referred to by their one-letter abbreviations, with the order of letters indicating the order in which the diseases were diagnosed. OH/OD=otherwise healthy or otherwise diseased. P=psychosis. D=diabetes. C=congestive heart failure. *Mean time free of an event (development of a condition of interest, or death), bound by the maximum follow-up of 20 years. †Based on a 50-year-old man in the third quintile of the Welsh Index of Multiple Deprivation. ‡All individuals included in the cohort, who were in an OH/OD state at the start of follow-up.

Table 2: **Frequency of health state transitions and associated mortality outcomes**

analyses were done in the SAIL Databank trusted research environment using the mstate package[27] in R (version 4.1.1).

Role of the funding source

The funder of the study had no role in the data collection, data analysis, data interpretation, or writing of the report. The Health Data Research UK (HDRUK) National Implementation Multimorbidity Resource has a specified work package dedicated to patient and public involvement, including a representative panel from across the UK; this panel specifically discussed and directly fed into this work, including the research design, interpretation and presentation of findings, and dissemination.

Results

1 675 585 individuals were included in the analysis. The median age at cohort entry (Jan 1, 2000) was 51·0 years (IQR 37·0–65·0), 864 192 (51·6%) individuals were female, and 811 393 (48·4%) were male. As the index condition, 24 990 (1·5%) individuals developed psychosis, 281 458 (16·8%) developed diabetes, and 123 981 (7·4%) developed congestive heart failure during follow-up. Individuals diagnosed with congestive heart failure as the index condition (n=123 981) tended to be older compared with those who had psychosis or diabetes as the index condition at diagnosis, with a median age of 78·1 years (IQR 69·4–84·9). Among individuals who developed one or more of psychosis, diabetes, and congestive heart failure, people in the first quintile (most deprived) of the Welsh Index of Multiple Deprivation were over-represented. This pattern was particularly evident for individuals who developed all three conditions in the sequence psychosis, diabetes, congestive heart failure (PDC), among whom 100 (36·2%) of 276 individuals were categorised as being in the most deprived quintile (table 1).

6492 (26·0%) of 24 990 individuals who developed psychosis, 48 911 (17·4%) of 281 458 who developed diabetes, and 79 126 (63·8%) of 123 981 who developed congestive heart failure as a single diagnosis died during follow-up (table 2, figure 1). Of those who developed two conditions, less than 10% of individuals (ranging from 0·8% to 9·4%) developed a third condition. Disease trajectories with congestive heart failure as the index condition had the highest proportion of deaths (range 58–70%).

Development of congestive heart failure as a single diagnosis at 50 years of age in men in WIMD quintile 3 resulted in a mean loss in life expectancy of 12·38 years (SD 0·00) compared with the otherwise healthy or otherwise diseased population (ie, those without a diagnosis of psychosis, diabetes, or congestive heart failure during follow-up and with the same characteristics in terms of age at diagnosis, deprivation status, and sex; figure 2A, table 2). Estimated mean loss in life expectancy was 13·45 years (0·13) after a diagnosis of congestive heart failure followed by psychosis (CP), and 12·95 years (0·06) after a diagnosis of psychosis followed by congestive heart failure (PC) compared with the otherwise healthy or otherwise diseased state

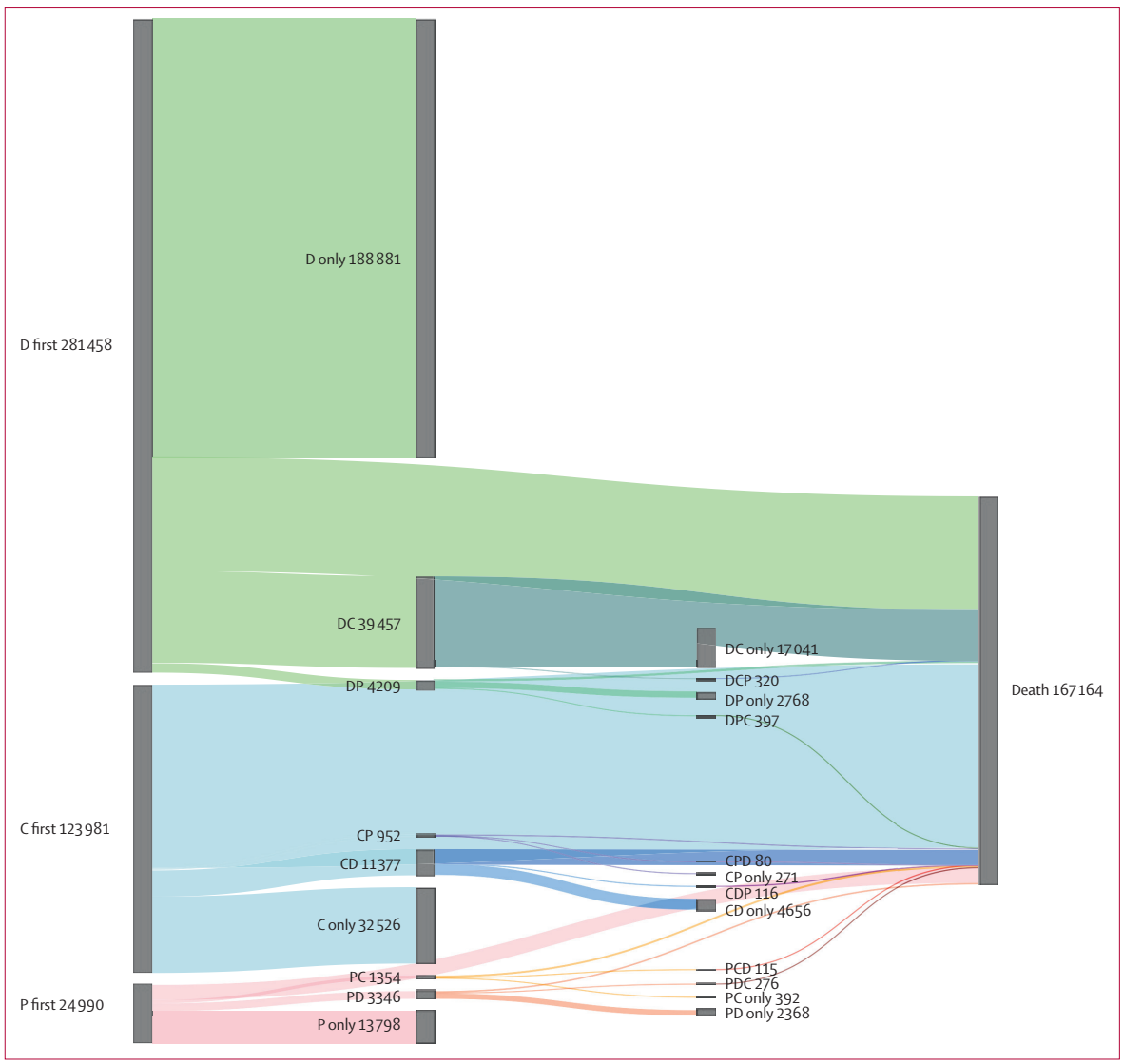

Figure 1: **Accrual and trajectory of diseases in a cluster of psychosis, diabetes, and congestive heart failure**
Diseases are referred to by their one-letter abbreviations, with the order of letters indicating the order in which the diseases were diagnosed within a trajectory. C=congestive heart failure. D=diabetes. P=psychosis.

(figure 2B, table 2). The development of congestive heart failure after psychosis resulted in a mean loss of 10·66 years (0·06) compared with the development of psychosis alone. As a combination of three diseases, development of diabetes, psychosis, and congestive heart failure (DPC) in this sequence resulted in the highest cumulative all-cause mortality (figure 2C), and the largest mean loss in life expectancy (13·23 years [0·8]) of any combination of the three diseases compared with an otherwise healthy or otherwise diseased population with the same characteristics (table 2). It is notable that congestive heart failure as a single condition and in combination with psychosis (PC and CP) had similar decreases in life expectancy to that of the DPC trajectory (table 2).

Individual hazard plots focusing on each of the transitions in the DPC trajectory are provided in the appendix (p 28). Individuals had an increased probability of developing psychosis following diabetes (DP) and congestive heart failure following psychosis (DPC) within the first 5 years of each subsequent diagnosis. Following the increased risk within the first 5 years, the probability of developing congestive heart failure in the DPC sequence drops before gradually increasing over time, reflecting the ageing process. The probability of death is also elevated within the first 5 years from the last diagnosis of congestive heart failure.

Diabetes as a single condition was associated with a slightly improved life expectancy (estimated mean change 0·19 years [SD 0·00]) compared with the

otherwise healthy or otherwise diseased population (table 2, figure 2A). Furthermore, psychosis and diabetes in combination were associated with an improved life expectancy compared with psychosis as a single condition (1·28 years [0·03] for DP and 1·74 [0·01] for psychosis followed by diabetes [PD] in a 50-year-old man in WIMD quintile 3; table 2). Cumulative hazard rates for all transitions between health states are provided in the appendix (p 29). Findings were robust in older ages (appendix p 30) and the most deprived populations (appendix p 31), and were robust in women for most trajectories (appendix p 32) except for the psychosis, congestive heart failure, and diabetes (PCD) trajectory, for which women had a noticeably higher probability of death (approximately 0·75) at 20 years compared with men (<0·6). Findings were also robust to adjustment for the time spent in the previous health state as a proxy measure for disease severity (appendix p 33).

Discussion

In this multimorbidity cluster of physical and mental health conditions, the sequence of diagnoses had an important effect on all-cause mortality. Development of diabetes, psychosis, and congestive heart failure in that order had a poorer prognosis than that of the same conditions developed in different orders. Comorbid psychosis and congestive heart failure as a combination of two diseases had a poorer prognosis than combinations of psychosis and diabetes or combinations of congestive heart failure and diabetes. Congestive heart failure had the worst prognosis as a single condition, with development at 50 years of age in men in WIMD quintile 3 associated with a reduction in life expectancy of 12·38 years compared with the general population without a diagnosis of psychosis, diabetes, or congestive heart failure during follow-up and with the same characteristics in terms of age at diagnosis, deprivation status, and sex.

Development of multimorbidity associated with DPC as an ordered sequence at 50 years of age resulted in a loss of 13·23 years in life expectancy compared with the general population. The losses in life expectancy for combinations of psychosis and congestive heart failure at 50 years of age (12·95 years for PC and 13·45 years for CP) and congestive heart failure alone (12·38 years) were similar to that of DPC. These findings would suggest that the key components resulting in an increased loss in life expectancy were attributable to the development of congestive heart failure or congestive heart failure and psychosis in combination. We further note that identification and therapeutic targets for congestive heart failure and psychosis might be beneficial in the first 5 years following the diagnosis of congestive heart failure. Identification of periods of increased risk of the development of subsequent conditions or death is essential to improve the management and care of individuals at risk of, and living with, multimorbidity.

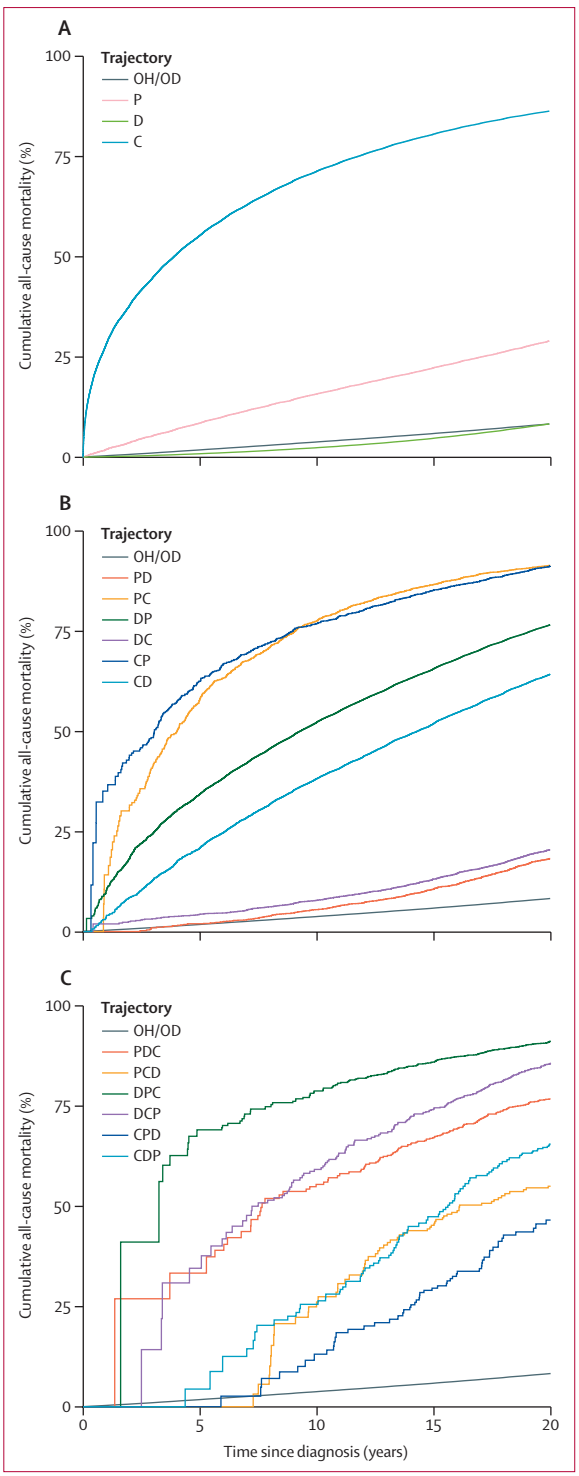

Figure 2: **Cumulative all-cause mortality following diagnosis with one (A), two (B), or three (C) conditions for a 50-year-old man in the third quintile of the Welsh Index of Multiple Deprivation**

Diseases are referred to by their one-letter abbreviations, with the order of letters indicating the order in which the diseases were diagnosed within a trajectory. C=congestive heart failure. D=diabetes. OH/OD=otherwise healthy or otherwise diseased. P=psychosis.

Individuals developed congestive heart failure as a first disease later in life, with a median age of around 78 years, resulting in a higher proportion of deaths from health states with congestive heart failure as the index condition. However, having adjusted for age at entry into the health state, we observed a lower rate of deaths for combinations of three diseases in which congestive heart failure was the first condition in the sequence, compared with combinations with either diabetes or psychosis as the index condition. This finding might be a consequence of a survivor effect, in which individuals who developed subsequent disease after congestive heart failure, and congestive heart failure and psychosis in combination, must have survived long enough to accrue further disease, and thus these individuals might have a less severe type of congestive heart failure, resulting in improved survival.

Diabetes as a single condition was associated with a slightly improved life expectancy compared with that of the general population of otherwise healthy and otherwise diseased individuals. For example, a 50-year-old man in WIMD quintile 3 diagnosed with diabetes could expect to gain a mean 0·19 years compared with the general population, which might include individuals with other life-shortening conditions such as cancer. Such a small improvement in life expectancy might not be considered important at the individual level, but could be considered more important at the population level. Combinations of psychosis and diabetes in any order were associated with an improved life expectancy compared with that of psychosis alone. For example, 50-year-old men in WIMD quintile 3 diagnosed with a combination of psychosis and diabetes might gain a mean 1·28–1·74 years compared with those diagnosed with psychosis alone. This might be a result of increased monitoring and management of diabetes, such as through routine diabetes clinics, and consequently improved management of additional conditions such as psychosis. Additionally, evidence suggests that metformin, which is used to treat type 2 diabetes, might increase life expectancy by reducing the risk of death associated with various conditions—including diabetes, cardiovascular disease, cognitive decline, and cancer—which might explain our findings, although these alleged benefits of metformin remain a point of discussion in the literature.[28] Compared with diabetes as a single condition, the development of psychosis in combination with diabetes in either sequence, PD or DP, at 50 years of age in men in WIMD quintile 3 resulted in a mean loss in life expectancy of 0·74–1·20 years. These findings are congruent with those of a previous study by Han and colleagues, who evaluated the effect of multimorbidity in people with diabetes and severe mental illness.[29] In that study, individuals with severe mental illness appeared to have increased all-cause and cardiovascular disease-specific mortality compared with individuals with diabetes alone. Han and colleagues concluded that increased mortality might be attributable to the underdiagnosis of cardiovascular disease and resultant delays in treatment in individuals with diabetes and severe mental illness. These findings, among others reported from studies of cardiometabolic multimorbidity[30] and physical–mental health multimorbidity,[31,32] further highlight the need to identify screening opportunities and therapeutic targets in populations with multimorbidity. It has previously been suggested that the relationship between common psychotic conditions and excess cardiovascular mortality is partly attributable to the increased risk of modifiable coronary heart disease risk factors such as obesity, smoking, diabetes, and dyslipidaemia in people with psychosis. In addition, antipsychotic medication can worsen cardiometabolic risk factors and result in weight gain.[33] The shared aetiology of psychosis and coronary heart disease has previously been discussed and includes biological, behavioural, psychological, and genetic mechanisms.[34] Evidence also suggests potential pleiotropic effects through shared gene loci associated with cardiometabolic and mental disorders.[35]

Future work could apply the modelling framework described in this Article to alternative populations with available population-level linked electronic health records to assess the external validity of the findings. However, for a truly external validation, any analysis should also be done by an independent external research group.[36]

The use of multistate models offers a flexible framework in which to assess the effect of disease trajectories in multimorbidity, and thereby facilitate the identification of potential screening opportunities, therapeutic targets, and risk factors, together with associated implications for future research prioritisation, study design, and policy.[32] Multistate models could also be used as a means to evaluate the potential effect of future interventions in multimorbid populations by incorporating estimates of treatment effects from randomised controlled trials with use of simulation-based methods.

We have demonstrated the application of this modelling framework to physical–mental health multimorbidity using population-scale, individual-level, anonymised, linked, routinely collected, electronic health record data sources in Wales. Such an approach could be applied to any situation with individuals accumulating conditions over time—eg, in alternative combinations of disease clusters in multimorbidity and post-COVID-19 condition (also known as long COVID). Application of these models could be used to help inform patients, clinicians, and health-care decision makers on the appropriate identification and management of patient care, and thus has the potential to improve patient outcomes.

Transition modelling approaches have previously been used in multimorbidity research, with several studies specifically looking at state transition modelling.[37–41] Two studies have evaluated trajectories of multimorbidity

patterns in terms of transitions between clusters of multimorbidity over time.[41,42] However, to our knowledge, none of these approaches have explored the sequences of disease occurrence within multimorbidity clusters. Two studies have evaluated the order of diseases;[43,44] however, these studies focused on summary statistics and visual representation of disease trajectories, and did not evaluate the effects of disease sequence on patient outcomes. A systematic review and comparison of approaches used to model trajectories of diseases in multimorbidity research has also been reported.[15] Although multistate models have been used previously in multimorbidity research,[45,46] they have not been used to look at the evolution of diseases within a cluster of multimorbid conditions over time.

The use of multistate modelling approaches can have limitations in multimorbidity research. Increasing the number of clusters or the number of diseases within clusters will exponentially increase the complexity of the model, which could lead to computational difficulties, especially when applied to population-scale data. For this reason, the number and type of conditions included in the analysis requires careful consideration to ensure that we do not compromise the interpretability of the results by pooling conditions, or increase the granularity of conditions to the extent that we compromise computational efficiency or sample size.

A potential limitation of using routinely collected, electronic health record data is the appropriate coding of diagnoses, the time at which the condition is first recorded, and the resulting classification of multimorbidity. Busija and colleagues have discussed this potential issue in terms of clustering of diseases in multimorbidity research and the need for harmonised data and definitions.[47] HDRUK has taken steps to potentially improve the appropriate classification and harmonisation of human disease by developing a phenotype library. However, even with complete harmonisation across data resources, it is widely known that diagnostic codes in electronic health record data are not always accurate, and there could be missed diagnoses or delays between the development of the condition of interest and the resultant recording of the diagnosis that are unlikely to be captured. External factors affecting health-care use (eg, the COVID-19 pandemic) might exacerbate this issue. In our analysis, we excluded 16·3% of the WMC population who were registered with general practices that did not provide data to the SAIL Databank, and thus we were unable to access their complete diagnostic history in terms of diagnoses captured within primary care settings. Furthermore, we were unable to capture diagnoses before Jan 1, 2000, due to the poor quality of reporting; therefore, the otherwise healthy or otherwise diseased population might contain individuals diagnosed with psychosis, diabetes, or congestive heart failure before cohort inception, among other conditions. The otherwise healthy or otherwise diseased population might also include those with missed diagnoses during cohort follow-up, and thus the otherwise healthy or otherwise diseased population provides a reference group for general all-cause mortality. This situation illustrates a limitation of using linked electronic health records with a historically poor quality of reporting. In this study, we used the Elixhauser comorbidity index phenotyping classification to define conditions. This includes a broad range of definitions, particularly for psychosis, which includes both symptomatic psychosis and clinical diagnosis, which can make interpretation more difficult. The development of phenotype libraries and increasing use of electronic health records will improve the quality of reporting over time, increasing the potential of electronic health records for future research. A further potential limitation of using routinely collected, electronic health record data is the difficulty in capturing severity of disease and transient morbidity, such as capturing periods of psychotic episodes. We have assessed the potential effect of adjusting for disease severity by incorporating the time spent in the previous health state in the model as a proxy measure for disease severity and uncaptured conditions at cohort inception that might accelerate the development of the conditions of interest. However, further linkage of data to prescribing and dispensing data or routinely collected health observations (such as HbA1c in diabetes) at the individual level might help to validate clinical coding and provide an improved understanding of the effect of severity of disease and transient health conditions on long-term outcomes. For example, with access to continued prescribing and dispensing data, we would gain further clarity on the severity of disease through information on potential dose escalations or additional medications, as well as confirmation of the presence of disease. Subsequently, modelling techniques could be easily developed to incorporate the transient nature of these conditions in the analysis, in addition to incorporating repeated observations over time and their effects on patient outcomes, where data permit. A further limitation of the current study is the lack of data on ethnicity, and the lack of power to adjust for ethnicity in the analyses owing to a predominantly White population (93·8% in the 2021 Census).[48]

For the **HDRUK Phenotype Library** see https://phenotypes.healthdatagateway.org

In this Article, we focused on all-cause mortality and life expectancy as the primary outcome of interest. Using routinely collected electronic health record data sources, modelling approaches could be extended to also assess health-care and resource use. Increasingly, trials are being conducted in multimorbid populations, with a focus on alternative outcomes of importance to patients, clinicians, and health-care decision makers, such as independent living and quality-of-life measures.[49] As more data become available in multimorbid populations, further work could extend this modelling approach to assess the effects of trajectories of disease on patient-centred outcomes such as quality of life, independent living, or patient-reported

outcome measures. Extensions to the modelling approach using longitudinal biomarker data could further enable identification of subpopulations at particular increased risks of developing specific multimorbid conditions or of death.[50]

Patient and public involvement in health data science is still relatively new. From our discussions with public contributors, we identified clear potential for future public input into defining multimorbidities and interpreting trajectories of disease, as well as their implications for public contributors and communities, reviewing and critiquing the content of datasets, and the future curation of which variables are included in datasets. Moving from passive supplier of data to active collaborator in health data science offers much promise for the future and will strengthen our understanding of the community validity of such work.

In conclusion, we found that the sequence and timing of disease onset in a trio of long-term conditions associated with high mortality had an important and complex effect on all-cause mortality. Multistate models provide a flexible framework to analyse trajectories of disease development and their associated effects on patient outcomes, and thus allow assessment of potential risk factors, screening opportunities, and targeted interventions for health-care policy and decision making.

Contributors

RKO and KRA conceptualised the study. RKO did the analysis and drafted the manuscript. RKO, JL, AA, UA, DCA, AA-L, AJB, SD, CD, AFF, BG, GH, PDWK, EBÖ, SR, SS, CM, RAL, and KRA reviewed, edited, and approved the final version of the manuscript. RKO and JL accessed, curated, and verified the data underlying the study; due to SAIL data governance, full access to the underlying data was not possible for all authors.

Declaration of interests

RKO is a member of the National Institute for Health and Care Excellence (NICE) Technology Appraisal Committee, member of the NICE Decision Support Unit, and associate member of the NICE Technical Support Unit; has served as a paid consultant to the pharmaceutical industry, providing unrelated methodological advice; and reports teaching fees from the Association of British Pharmaceutical Industry (ABPI) and the University of Bristol. KRA is a member of the NICE Diagnostics Advisory Committee and the NICE Decision and Technical Support Units; is a National Institute for Health Research (NIHR) Senior Investigator Emeritus (NF-SI-0512-10159); has served as a paid consultant, providing unrelated methodological and strategic advice, to the pharmaceutical and life sciences industry generally, as well as to the Department of Health and Social Care and NICE; has received unrelated research funding from ABPI, European Federation of Pharmaceutical Industries & Associations, Pfizer, Sanofi, and Swiss Precision Diagnostics; has received course fees from ABPI; and is a Partner and Director of Visible Analytics Limited, a health technology assessment consultancy company. All other authors declare no competing interests.

Data sharing

Access to the research ready datasets used for all analyses, figures, and tables (including those in the appendix) are available to the research community via a Trusted Research Environment upon approval by the SAIL Databank and the independent Information Governance Review Panel. All code used in the analyses are available on request from the corresponding author.

For the **SAIL Databank** see https://saildatabank.com

Acknowledgments

This work was supported by Health Data Research UK (HDRUK) Measuring and Understanding Multimorbidity using Routine Data in the UK (MUrMuRUK; award numbers HDR-9006 and CFC0110). HDRUK is funded by the UK Medical Research Council (MRC), Engineering and Physical Sciences Research Council, Economic and Social Research Council, NIHR (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation, and Wellcome Trust. This work also was co-funded by the MRC and NIHR (grant number MR/S027750/1). The work was supported by the Administrative Data Research (ADR) Wales programme of work, part of the Economic and Social Research Council (part of UK Research and Innovation) funded ADR UK (grant ES/S007393/1). RKO is supported by a Springboard award (SBF006\1122) funded by the Academy of Medical Sciences, Wellcome Trust, Government Department of Business, Energy and Industrial Strategy, British Heart Foundation, and Diabetes UK. SS is part funded by the NIHR Applied Research Collaboration West Midlands, the NIHR Health Protection Research Unit (HPRU) in Gastrointestinal Infections, and the NIHR HPRU in Genomics and Enabling Data.

References

1 Fortin M, Bravo G, Hudon C, Vanasse A, Lapointe L. Prevalence of multimorbidity among adults seen in family practice. *Ann Fam Med* 2005; **3:** 223–28.

2 Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. *BMJ* 2015; **350:** h176.

3 Menotti A, Mulder I, Nissinen A, Giampaoli S, Feskens EJ, Kromhout D. Prevalence of morbidity and multimorbidity in elderly male populations and their impact on 10-year all-cause mortality: the FINE study (Finland, Italy, Netherlands, Elderly). *J Clin Epidemiol* 2001; **54:** 680–86.

4 Nunes BP, Flores TR, Mielke GI, Thumé E, Facchini LA. Multimorbidity and mortality in older adults: a systematic review and meta-analysis. *Arch Gerontol Geriatr* 2016; **67:** 130–38.

5 Starfield B. Challenges to primary care from co- and multi-morbidity. *Prim Health Care Res Dev* 2011; **12:** 1–2.

6 Lehnert T, Heider D, Leicht H, et al. Review: health care utilization and costs of elderly persons with multiple chronic conditions. *Med Care Res Rev* 2011; **68:** 387–420.

7 Uijen AA, van de Lisdonk EH. Multimorbidity in primary care: prevalence and trend over the last 20 years. *Eur J Gen Pract* 2008; **14** (suppl 1)**:** 28–32.

8 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.

9 Cassell A, Edwards D, Harshfield A, et al. The epidemiology of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2018; **68:** e245–51.

10 MacMahon S. Multimorbidity: a priority for global health research. London, UK: The Academy of Medical Sciences, 2018.

11 Farmer C, Fenu E, O'Flynn N, Guthrie B. Clinical assessment and management of multimorbidity: summary of NICE guidance. *BMJ* 2016; **354:** i4843.

12 Salisbury C. Multimorbidity: redesigning health care for people who use it. *Lancet* 2012; **380:** 7–9.

13 Kingston A, Robinson L, Booth H, Knapp M, Jagger C. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. *Age Ageing* 2018; **47:** 374–80.

14 Schäfer I, von Leitner EC, Schön G, et al. Multimorbidity patterns in the elderly: a new approach of disease clustering identifies complex interrelations between chronic conditions. *PLoS One* 2010; **5:** e15941.

15 Cezard G, McHale CT, Sullivan F, Bowles JKF, Keenan K. Studying trajectories of multimorbidity: a systematic scoping review of longitudinal approaches and evidence. *BMJ Open* 2021; **11:** e048485.

16 Kilbourne AM, Morden NE, Austin K, et al. Excess heart-disease-related mortality in a national study of patients with mental disorders: identifying modifiable risk factors. *Gen Hosp Psychiatry* 2009; **31:** 555–63.

17 Ward M, Druss B. The epidemiology of diabetes in psychotic disorders. *Lancet Psychiatry* 2015; **2:** 431–51.

18 Kamalesh M. Heart failure in diabetes and related conditions. *J Card Fail* 2007; **13:** 861–73.

19 Ford DV, Jones KH, Verplancke JP, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Serv Res* 2009; **9:** 157.

20 Lyons J, Akbari A, Agrawal U, et al. Protocol for the development of the Wales Multimorbidity e-Cohort (WMC): data sources and methods to construct a population-based research platform to investigate multimorbidity. *BMJ Open* 2021; **11:** e047101.

21 Molitor J, Papathomas M, Jerrett M, Richardson S. Bayesian profile regression with an application to the National Survey of Children's Health. *Biostatistics* 2010; **11:** 484–98.

22 Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998; **36:** 8–27.

23 Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005; **43:** 1130–39.

24 Metcalfe D, Masters J, Delmestri A, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. *BMC Med Res Methodol* 2019; **19:** 115.

25 Andersen PK, Keiding N. Multi-state models for event history analysis. *Stat Methods Med Res* 2002; **11:** 91–115.

26 Stensrud MJ, Hernán MA. Why test for proportional hazards? *JAMA* 2020; **323:** 1401–02.

27 de Wreede LC, Fiocco M, Putter H. mstate: an R package for the analysis of competing risks and multi-state models. *J Stat Softw* 2011; **38:** 1–30.

28 Mohammed I, Hollenberg MD, Ding H, Triggle CR. A critical review of the evidence that metformin is a putative anti-aging drug that enhances healthspan and extends lifespan. *Front Endocrinol (Lausanne)* 2021; **12:** 718942.

29 Han L, Doran T, Holt RIG, et al. Impact of severe mental illness on healthcare use and health outcomes for people with type 2 diabetes: a longitudinal observational study in England. *Br J Gen Pract* 2021; **71:** e565–73.

30 Di Angelantonio E, Kaptoge S, Wormser D, et al. Association of cardiometabolic multimorbidity with mortality. *JAMA* 2015; **314:** 52–60.

31 Stubbs B, Koyanagi A, Veronese N, et al. Physical multimorbidity and psychosis: comprehensive cross sectional analysis including 242,952 people across 48 low- and middle-income countries. *BMC Med* 2016; **14:** 189.

32 Olsson E, Westman J, Sudic Hukic D, et al. Diabetes and glucose disturbances in patients with psychosis in Sweden. *BMJ Open Diabetes Res Care* 2015; **3:** e000120.

33 De Hert M, Dekker JM, Wood D, Kahl KG, Holt RI, Möller HJ. Cardiovascular disease and diabetes in people with severe mental illness position statement from the European Psychiatric Association (EPA), supported by the European Association for the Study of Diabetes (EASD) and the European Society of Cardiology (ESC). *Eur Psychiatry* 2009; **24:** 412–24.

34 De Hert M, Detraux J, Vancampfort D. The intriguing relationship between coronary heart disease and mental disorders. *Dialogues Clin Neurosci* 2018; **20:** 31–40.

35 Amare AT, Schubert KO, Klingler-Hoffmann M, Cohen-Woods S, Baune BT. The genetic overlap between mood disorders and cardiometabolic diseases: a systematic review of genome wide and candidate gene studies. *Transl Psychiatry* 2017; **7:** e1007.

36 Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol* 2014; **14:** 40.

37 Lindhagen L, Van Hemelrijck M, Robinson D, Stattin P, Garmo H. How to model temporal changes in comorbidity for cancer patients using prospective cohort data. *BMC Med Inform Decis Mak* 2015; **15:** 96.

38 Alaeddini A, Jaramillo CA, Faruqui SHA, Pugh MJ. Mining major transitions of chronic conditions in patients with multiple chronic conditions. *Methods Inf Med* 2017; **56:** 391–400.

39 Siriwardhana C, Lim E, Davis J, Chen JJ. Progression of diabetes, ischemic heart disease, and chronic kidney disease in a three chronic conditions multistate model. *BMC Public Health* 2018; **18:** 752.

40 Freisling H, Viallon V, Lennon H, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. *BMC Med* 2020; **18:** 5.

41 Violán C, Fernández-Bertolín S, Guisado-Clavero M, et al. Five-year trajectories of multimorbidity patterns in an elderly Mediterranean population using Hidden Markov Models. *Sci Rep* 2020; **10:** 16879.

42 Vetrano DL, Roso-Llorach A, Fernández S, et al. Twelve-year clinical trajectories of multimorbidity in a population of older adults. *Nat Commun* 2020; **11:** 3223.

43 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.

44 Ashworth M, Durbaba S, Whitney D, Crompton J, Wright M, Dodhia H. Journey to multimorbidity: longitudinal analysis exploring cardiovascular risk factors and sociodemographic determinants in an urban setting. *BMJ Open* 2019; **9:** e031649.

45 Bisquera A, Turner EB, Ledwaba-Chapman L, et al. Inequalities in developing multimorbidity over time: a population-based cohort study from an urban, multi-ethnic borough in the United Kingdom. *Lancet Reg Health Eur* 2021; **12:** 100247.

46 Chan MS, van den Hout A, Pujades-Rodriguez M, et al. Socio-economic inequalities in life expectancy of older adults with and without multimorbidity: a record linkage study of 1.1 million people in England. *Int J Epidemiol* 2019; **48:** 1340–51.

47 Busija L, Lim K, Szoeke C, Sanders KM, McCabe MP. Do replicable profiles of multimorbidity exist? Systematic review and synthesis. *Eur J Epidemiol* 2019; **34:** 1025–53.

48 Welsh Government. Ethnic group, national identity, language and religion in Wales (Census 2021). Nov 29, 2022. https://www.gov.wales/ethnic-group-national-identity-language-and-religion-wales-census-2021-html#110364 (accessed June 10, 2023).

49 Salisbury C, Man MS, Bower P, et al. Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach. *Lancet* 2018; **392:** 41–50.

50 Ferrer L, Rondeau V, Dignam J, Pickles T, Jacqmin-Gadda H, Proust-Lima C. Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer. *Stat Med* 2016; **35:** 3933–48.