Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies

Iris Szu-Szu Ho, Amaya Azcoaga-Lorenzo, Ashley Akbari, Corri Black, Jim Davies, Peter Hodgins, Kamlesh Khunti, Umesh Kadam, Ronan A Lyons, Colin McCowan, Stewart Mercer, Krishnarajah Nirantharakumar, Bruce Guthrie

Summary

Background A systematic understanding of how multimorbidity has been constructed and measured is unavailable. This review aimed to examine the definition and measurement of multimorbidity in peer-reviewed studies internationally.

Methods We systematically reviewed studies on multimorbidity, via a search of nine bibliographic databases (Ovid [PsycINFO, Embase, Global Health, and MEDLINE], Web of Science, the Cochrane Library, CINAHL Plus, Scopus, and ProQuest Dissertations & Theses Global), from inception to Jan 21, 2020. Reference lists and tracked citations of retrieved articles were hand-searched. Eligible studies were full-text articles measuring multimorbidity for any purpose in community, primary care, care home, or hospital populations receiving a non-specialist service. Abstracts, qualitative research, and case series were excluded. Two reviewers independently reviewed the retrieved studies with conflicts resolved by discussion or a third reviewer, and a single researcher extracted data from published papers. To assess our objectives of how multimorbidity has been measured and examine variation in the chronic conditions included (in terms of number and type), we used descriptive analysis (frequencies, cross-tabulation, and negative binomial regression) to summarise the characteristics of multimorbidity studies and measures (study setting, source of morbidity data, study population, primary study purpose, and multimorbidity measure type). This systematic review is registered with PROSPERO, CRD420201724090.

Findings 566 studies were included in our review, of which 206 (36·4%) did not report a reference definition for multimorbidity and 73 (12·9%) did not report the conditions their measure included. The number of conditions included in measures ranged from two to 285 (median 17 [IQR 11–23). 452 (79·9%) studies reported types of condition within a single multimorbidity measure; most included at least one cardiovascular condition (441 [97·6%] of 452 studies), metabolic and endocrine condition (440 [97·3%]), respiratory condition (422 [93·4%]), musculoskeletal condition (396 [87·6%]), or mental health condition (355 [78·5%]) in their measure of multimorbidity. Chronic infections (123 [27·2%]), haematological conditions (110 [24·3%]), ear, nose, and throat conditions (107 [23·7%]), skin conditions (70 [15·5%]), oral conditions (19 [4·2%]), and congenital conditions (14 [3·1%]) were uncommonly included. Only eight individual conditions were included by more than half of studies in the multimorbidity measure used (diabetes, stroke, cancer, chronic obstructive pulmonary disease, hypertension, coronary heart disease, chronic kidney disease, and heart failure), with individual mental health conditions under-represented. Of the 566 studies, 419 were rated to be of moderate risk of bias, 107 of high risk of bias, and 40 of low risk of bias according to the Effective Public Health Practice Project quality assessment tool.

Interpretation Measurement of multimorbidity is poorly reported and highly variable. Consistent reporting of measure definitions should be required by journals, and consensus studies are needed to define core and study-dependent conditions to include in measures of multimorbidity.

Funding Health Data Research UK.

Lancet Public Health 2021; 6: e587–97

Published Online June 21, 2021 https://doi.org/10.1016/S2468-2667(21)00107-9

Usher Institute, University of Edinburgh Medical School, Edinburgh, UK (I S-S Ho PhD, P Hodgins MBBS, Prof S Mercer PhD, Prof B Guthrie PhD); **University of St Andrews School of Medicine, Medical and Biological Sciences, St Andrews, UK** (A Azcoaga-Lorenzo PhD, Prof C McCowan PhD); **Institute of Life Science, Swansea University Medical School, Swansea, UK** (A Akbari MSc, Prof R A Lyons PhD); **School of Medicine, Medical Science and Nutrition, University of Aberdeen, Aberdeen, UK** (Prof C Black PhD); **Department of Computer Science, University of Oxford, Oxford, UK** (Prof J Davies PhD); **University of Leicester, Leicester General Hospital, Leicester, UK** (Prof K Khunti PhD, Prof U Kadam PhD); **Institute of Applied Health Research, University of Birmingham, Birmingham, UK** (K Nirantharakumar MD)

Correspondence to: Prof Bruce Guthrie, Usher Institute, University of Edinburgh Medical School, Edinburgh EH8 9AG, UK **bruce.guthrie@ed.ac.uk**

Introduction

Due to increasing longevity and improved survival from acute conditions, the number of people living with multiple long-term health conditions is rising.[1] In 2018, the UK Academy of Medical Sciences defined multimorbidity as the coexistence of two or more chronic health conditions in an individual, any one of which should be a long-term physical non-communicable disease, a mental health condition of long duration, or a long-term infectious disease.[2] Multimorbidity differs conceptually from comorbidity, for which the focus is on any additional conditions that people with a specified index condition also have.[3] Recognition is growing that existing health-care systems are seriously challenged by multimorbidity because they are largely designed to care for patients with single conditions, and multimorbidity is being increasingly researched.[4,5]

Despite the growing interest in multimorbidity, how it is defined and measured varies substantially. Several studies have examined the definition and measurement of multimorbidity.[6] Although there is agreement that multimorbidity is common and often burdensome, varying operational definitions of multimorbidity are recognised to

Research in context

Evidence before this study

To identify studies relevant to multimorbidity measurement, we used two key terms (multimorbidity AND measurement) to search Ovid (PsycINFO, Embase, Global Health, and MEDLINE), Web of Science, the Cochrane Library, CINAHL Plus, Scopus, and ProQuest Dissertations & Theses Global, from database inception to Jan 21, 2020, for English-language studies. Previous narrative systematic reviews examining multimorbidity measure design concluded that the number and type of conditions included in multimorbidity counts varies, but these reviews only included small numbers of studies with incomplete searching. Other reviews examined the performance of weighted indices in predicting one or more outcomes, concluding that the Charlson comorbidity index and Elixhauser's comorbidity index had high predictive performance for mortality, the health-related quality of life comorbidity index for quality of life, Bayliss' weighted index for physical functioning and self-rated health, and the Johns Hopkins Adjusted Clinical Groups system and medication-based indices for health-care use and costs. However, differences in predictive performance between measures (including between weighted indices and simple counts of conditions) were small.

Added value of this study

This systematic review is, to our knowledge, the largest to date, and systematically examines variation in the reporting and measurement of multimorbidity. The present analysis identifies considerable variation in how multimorbidity is measured in the research literature, particularly in relation to the number and type of conditions included in different studies. Reporting was often poor, with one in eight studies not reporting which conditions were included in their multimorbidity measure. We found large variation in the number of conditions included in multimorbidity measures (range 2 to 285 conditions), and which conditions were included. Only eight conditions were included by more than half of studies (diabetes, stroke, cancer, chronic obstructive pulmonary disease, hypertension, coronary heart disease, chronic kidney disease, and heart failure). Fewer than half of studies included any one mental health condition, and one quarter of studies did not include any mental health condition.

Implications of all the available evidence

Multimorbidity has been increasingly researched in the past decade, but the literature is characterised by poor reporting and variation in the number and nature of conditions included in multimorbidity measurement. Reporting would be improved by studies clearly reporting which conditions they included in multimorbidity measurement and why conditions were chosen, and clinical code sets or clear definitions of how each individual condition was ascertained. A consensus is needed to establish a core condition set for multimorbidity that all studies should measure to facilitate comparison across studies, and guidance on principles for defining customised condition sets justified by purpose or context. Based on the present analysis, we recommend a potential core set of conditions to include in all measures, and other conditions that should be considered for inclusion deepening on context. Consensus studies are now needed to refine and inform this core set, to improve comparability and reproducibility of research in this field.

have led to heterogeneous estimates of multimorbidity prevalence and burden.[6–9] Operational definitions differ in terms of the number and types of conditions included, the cutoff number of conditions for defining when multimorbidity is present, whether conditions are simply counted or are weighted in relation to predefined outcomes, and the data sources and data collection methods used.[6,8,10–13] Therefore, no clear consensus exists on how to measure multimorbidity, nor on which conditions to include in a measure.[6,7,14] This lack of consensus makes comparison across research studies difficult because the underlying measure used by each study can be very different.

The aim of this systematic review was to examine how multimorbidity has been measured in peer-reviewed studies internationally, including which chronic conditions are included in measures and how these vary between studies.

Methods

Overview

For the **PROSPERO listing** see https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=172409

We systematically reviewed studies measuring multimorbidity, and examined the design and characteristics of multimorbidity measures used. The review protocol was registered with PROSPERO (CRD42020172409). This paper reports findings relating to the first two registered objectives, of how multimorbidity has been measured and variation in the chronic conditions included (in terms of type and number). Findings for the third registered objective, to identify factors associated with heterogeneity of estimated multimorbidity prevalence, were based on different analytical approaches and will be reported separately.

We followed the CoCoPop mnemonic (Condition, Context, and Population) for systematic reviews of observational studies to define eligibility criteria and conduct searches.[15] The condition focused on in the review was multimorbidity, for which we did not require alignment to any particular definition of multimorbidity but rather explored whether studies explicitly stated or justified their underlying definition of it. In terms of context and population, we examined studies in community, primary care, care home, or hospital settings. Other inclusion criteria were that studies had full text of the study report available and were reported in English. We excluded studies that did not measure multimorbidity (eg, were comorbidity studies), that focused on people within a specialist service such as critical care, or in which the morbidity measurement focused on acute conditions or

nursing conditions. Qualitative research and case series were also excluded.

Search strategy and selection criteria

The search strategy was developed in collaboration with a medical librarian (appendix p 3). Two sets of key terms (multimorbidity and measure) were combined with Boolean logic to search for relevant studies. Medical subject headings were used to capture concepts and maximise the number of studies retrieved. Searches were done in Ovid (PsycINFO, Embase, Global Health, and MEDLINE), Web of Science, the Cochrane Library, EBSCO (CINAHL Plus), Scopus, and ProQuest Dissertations & Theses Global, from inception to a last updated search on Jan 21, 2020. After the database searches, we hand-searched reference lists of retrieved articles and tracked citations.

Identified references were exported to EndNote (version X9) and Microsoft Excel 2016 for deduplication, and then imported to Covidence software for screening by two independent reviewers (IS-SH and PH) of titles, abstracts, and full-text articles against the eligibility criteria, with disagreement resolved by discussion and involvement of a third reviewer (BG) if necessary.

Data analysis

Data on the characteristics of the included studies were extracted by a single reviewer (IS-SH) with predesigned data extraction tables. Data were extracted from published reports only; authors were not contacted. The extracted data included authors, year of publication, study title, study purpose, method, country, mean age of study participants, reference definition of multimorbidity, type of multimorbidity measure (simple counts or weighted indices), data collection method or data source (self-report, or medical records or administrative databases), number of conditions included in the multimorbidity measure, and the actual conditions included. Due to the wide variation in labelling of conditions included, we categorised similar conditions into groups. An example of how we categorised conditions is our category of chronic pain, for which one or more of the studies variously counted neck pain, back pain, low back pain, chronic low back pain, pain, chronic pain, painful conditions, trigeminal neuralgia, or fibromyalgia.

In relation to the first registered review objective examining how multimorbidity has been measured, we used descriptive analyses to summarise the characteristics of multimorbidity studies and measures. Categorical data on multimorbidity study characteristics are presented as counts and percentages. To further explore how the choice of multimorbidity measure (simple *vs* weighted counts) was related to study purpose, setting, and data source, we cross-tabulated measure type by these variables, with statistical significance tested by Pearson's χ^2, considering a two-sided p value of 0·05 as the threshold for significance. If contingency table expected frequencies were less than 5, the p value was computed for a Monte Carlo test with 2000 duplicates.

In relation to the second registered review objective examining which conditions were included in multimorbidity measures, relationships between characteristics of multimorbidity studies and number of conditions included in the measure were examined with negative binomial generalised linear regression (the Poisson model was found to be overdispersed).[16] The study characteristics examined as predictors were data source, study setting, study population, and type of measure used (dichotomised into simple count and weighted measure). Since data were missing on both outcome and predictors, multiple imputation with 100 imputed datasets and 20 iterations was done with Bayesian logistic regression for binary variables, and negative binomial regression for count data.[17] Fraction of missing information (FMI) was computed to quantify loss of information due to missing status, and FMI was defined as small, moderate, or high as defined in White et al.[18] The imputed dataset was then used for analysis, with statistical significance of associations assumed when the confidence interval did not include 1. Univariate models were initially fitted with negative binomial model coefficients exponentiated to estimated incidence rate ratios (IRRs), followed by multivariate models with examination of plausible interaction terms, and an interaction between type of measure and data source retained in the model.

The conditions included in different studies were examined after grouping by International Classification of Diseases-10 chapter, partially modified to reflect the organisation of clinical care (for example, throat disorders were grouped with ear disorders rather than digestive disorders), and examined in terms of individual conditions. To assess whether choice of condition varied by where the study was done, for each condition we cross-tabulated whether it was included in multimorbiditiy measurement with two variables: the continent where the study was done, and the income level of the country where the study was done (defined by the World Bank,[19] and dichotomised into low-income and middle-income countries [LMICs] and high-income countries [HICs]). Statistical significance was examined with Pearson's χ^2 test. For continent, which has six categories, when χ^2 test indicated significant associations (at $p<0·05$), post-hoc analysis was done to identify which continent was significantly associated with inclusion of that condition.[20] To contextualise the observed choices of conditions in the reviewed studies, we extracted data on the burden of disability-adjusted life years (DALYs) and years of life lost (YLLs) from the Global Burden of Disease Study 2019 for the 44 chronic conditions with the highest DALY burden,[21] and examined how often these conditions were included in multimorbidity measurement.

The Effective Public Health Practice Project (EPHPP) quality assessment tool for quantitative studies was used to assess risk of bias and quality of the included studies.[22] The EPHPP tool evaluates selection bias, study design,

See **Online** for appendix

For the **Covidence software platform** see https://www.covidence.org/

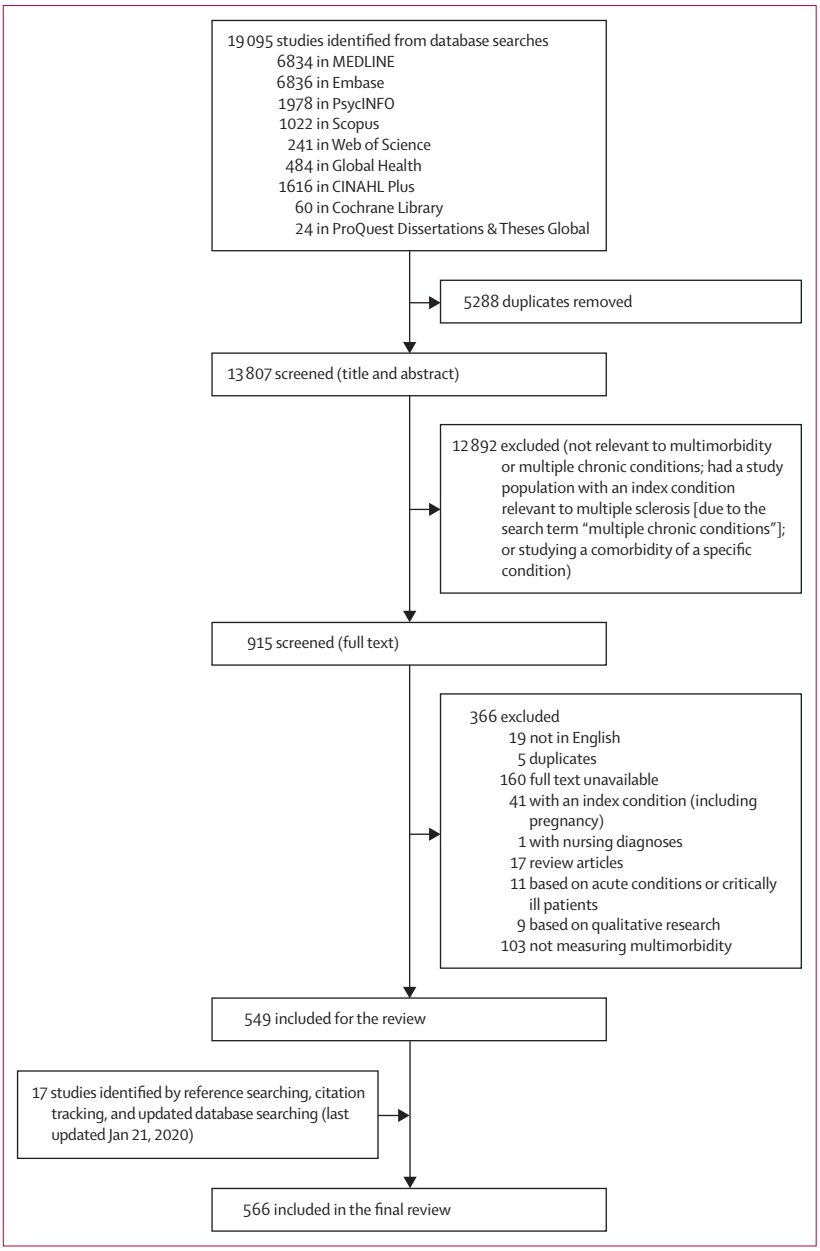

Figure 1: **Study selection for systematic review**

confounders, blinding, data collection method, and withdrawals and dropouts. We also evaluated two other types of bias: publication bias (for which high risk of bias was judged present if the same data were reported variably in different papers), and conflict of interest (for which risk of bias was rated unclear if no conflict of interest statement was present). Each domain was rated as low, moderate, high, or unclear risk of bias. The overall risk of bias for each study was classed as low, moderate, or high based on the most frequent component score. We also categorised each study in terms of clarity of reporting of multimorbidity measurement and definition. Clear reporting of multimorbidity was judged to be present if a study had reported the reference definition for measuring multimorbidity and listed all health conditions included in the multimorbidity measure. Associations between overall risk of bias and clarity of reporting were examined post hoc with cross-tabulation and Pearson χ^2 test.

All statistical tests were done with R (version 4.0.1). Where quantitative analyses were not suitable, a narrative overview was given.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit for publication.

Results

We identified 13 807 articles from our database search. Study selection and reasons for exclusion are summarised in figure 1. Overall, 566 studies were included in our systematic review (table 1, appendix pp 5–71). Most studies were done in Europe (209 [36·9%]), followed by North America (190 [33·6%]). The majority of studies were community-based (320 [56·5%]) or primary care-based (126 [22·3%]). Data on multimorbidity were mainly collected by patient self-report (312 [55·1%] studies) or medical records and administrative databases (240 [42·4%]). Eight (1·4%) studies did not report data source. Of 525 studies that examined a single measure of multimorbidity, 73 (13·9%) had incomplete reporting, with missing status highest for the count variable of number of conditions included in the multimorbidity measure (56 [10·7%] studies), and the binary variables of measure type (28 [5·3%] studies; with those including more than one measure type treated as missing]), and data source (eight [1·5%] studies).

The most common stated purposes of measuring multimorbidity in the included studies were examination of its association with various outcomes (209 [36·9%] studies) and examination of measure performance (117 [20·7%]). Other purposes were identification of multimorbidity patterns (62 [11·0%]) and examination of multimorbidity trajectories over time (19 [3·4%]; table 1, appendix pp 5–71). Multimorbidity was most commonly examined in studies of adults of any age (240 [42·4%]), in studies of older adults (216 [38·2%]), and in studies of middle-aged and older adults (80 [14·1%]; table 1, appendix pp 5–71).

Almost half of the studies (268 [47·3%]) defined multimorbidity as the presence of two or more chronic conditions, with other studies defining multimorbidity as multiple chronic conditions (56 [9·9%]), three or more chronic conditions (28 [4·9%]), or five or more chronic conditions (four [0·7%]). 206 (36·4%) studies did not report a reference definition (table 1, appendix pp 5–71). Three studies defined severe or complex multimorbidity (according to various terms) in addition to overall multimorbidity, with the cutoff point for severe defined as four or more chronic conditions in

	Number of studies (n=566)
Study setting	
Community	320 (56·5%)
Primary care	126 (22·3%)
Hospital	104 (18·4%)
Care home	14 (2·5%)
Primary care and hospitals	2 (0·4%)
Data source	
Patient self-report or interviews	312 (55·1%)
Medical records and administrative databases	240 (42·4%)
Patient self-report plus medical records and administrative databases	6 (1·1%)
Not reported*	8 (1·4%)
Study population†	
All adults (age ≥15 years)	240 (42·4%)
Only children (age <18 years)	5 (0·9%)
Young and middle-aged adults (age 18–64 years)	10 (1·8%)
Middle-aged adults (age 40–65 years)	8 (1·4%)
Middle-aged and older adults (age ≥40 years)	80 (14·1%)
Older adults (age ≥60 years)	216 (38·2%)
All ages	7 (1·2%)
Primary study purpose	
Association of multimorbidity with outcome	209 (36·9%)
Association of risk factors with multimorbidity	75 (13·3%)
Patterns or clusters of multimorbidity diseases	62 (11·0%)
Trajectory of multimorbidity	19 (3·4%)
Examination of multimorbidity measure performance‡	117 (20·7%)
Study of populations with multimorbidity	68 (12·0%)
Prevalence or burden of multimorbidity	16 (2·8%)
Definition (threshold) for multimorbidity	
Not reported	206 (36·4%)
Two or more chronic conditions	268 (47·3%)
Three or more chronic conditions	28 (4·9%)
Five or more chronic conditions	4 (0·7%)
Multiple chronic conditions§	56 (9·9%)
Patient-defined weighted cumulative score	4 (0·7%)
(Table 1 continues in next column)	

	Number of studies (n=566)
(Continued from previous column)	
Type of multimorbidity measure	
Simple count	376 (66·4%)
Weighted index of conditions	149 (26·3%)
Weighted index of medications	6 (1·1%)
Both simple count and weighted index	27 (4·8%)
Unclear¶	8 (1·4%)
Number of conditions included in measure	
Not reported	56 (9·9%)
2–10	104 (18·4%)
11–20	224 (39·6%)
21–30	68 (12·0%)
31–40	33 (5·8%)
41–50	15 (2·7%)
>50	25 (4·4%)
Not applicable (studies that examined multiple measures)	41 (7·2%)

Percentages might not add to 100% due to rounding. Complete definitions of variables and age groups across individual studies are provided in the appendix (p 4). *None of the studies reported a reason for not reporting data source. †Population age ranges were according to individual study definitions (appendix p 4). ‡Studies that examined the performance of multimorbidity measures (including development, validation, or adaptation); compared the predictive performance for different measures in relation to outcomes; examined different multimorbidity definitions and how these affected estimates of multimorbidity prevalence; or examined the difference in measure performance between self-report and medical records or administrative databases. §Implicitly two or more chronic conditions, but not stated numerically. ¶The same eight studies that did not report data source.

Table 1: **Study characteristics**

two studies[23,24] and five or more chronic conditions in one study.[25]

376 (66·4%) studies used a simple count of conditions to measure multimorbidity, 155 (27·4%) used weighted indices, 27 (4·8%) used both types of measures together, and eight (1·4%) did not report the type of measure used (table 1, appendix pp 5–71). In studies that used a weighted index, 149 used a disease-based index, and six used a medication-based index. The most commonly used disease-based indices were various versions of the Charlson comorbidity index (CCI; 64 studies) and the cumulative illness rating scale-geriatric (CIRS-G; 25 studies). Medication-based indices comprised the chronic disease score (two studies[26,27]), the Rx-risk comorbidity index (two studies[28,29]), the medication-based disease burden index (MDBI; one study[30]), and a medicines comorbidity index (one study[31]). Common weighting methods for indices were: use of regression models to calculate weights for each condition on the basis of association with a particular outcome (eg, the CCI); rating of condition severity on the basis of prespecified thresholds of severity (eg, the CIRS-G[32] and the weighted index developed by Bayliss and colleagues[33]); and assignment of weights on the basis of existing disease burden literature (eg, the MDBI[30]).

Weighted multimorbidity measures were marginally more likely to be used in studies collecting data from medical records or administrative databases (118 [53·9%] of 219 studies with non-missing data for type of multimorbidity measure), whereas simple counts of conditions were more commonly used in studies that were based on self-report data (271 [88·6%] of 306 studies; χ^2=109·33, p<0·0001; appendix p 72). The choice of measure varied depending on study purpose, with weighted measures primarily used to examine measure performance (92 [84·4%] of 109 studies), whereas simple counts were more commonly used for all other purposes (χ^2=208·22, p=0·0005). Similarly, the choice of measure varied by study setting, with simple counts more commonly used in

	Number of conditions: mean (SD)	Unadjusted IRR (95% CI)	Adjusted IRR (95% CI)	Fraction of missing information
Data source				
Patient self-report	16 (11)	1 (ref)	1 (ref)	1 (ref)
Databases	32 (41)	1·8 (1·6–2·0)*	2·2 (1·8–2·6)*	0·12
Setting				
Community	19 (20)	1 (ref)	1 (ref)	1 (ref)
Primary care	26 (27)	1·3 (1·1–1·5)*	0·9 (0·8–1·1)	0·11
Hospital	27 (34)	1·5 (1·2–1·7)*	1·1 (0·9–1·3)	0·10
Care home	43 (77)	2·3 (1·5–3·4)*	1·4 (1·0–2·2)*	0·16
Population				
All adults	25 (32)	1 (ref)	1 (ref)	1 (ref)
Only children	6 (1)	0·3 (0·1–0·9)*	0·4 (0·1–1·1)	0·53
Young and middle-aged adults	12 (4)	0·5 (0·3–0·9)*	0·6 (0·4–1·0)	0·19
Middle-aged adults	14 (11)	0·6 (0·3–1·2)	0·6 (0·3–1·0)	0·30
Middle-aged and older adults	24 (40)	0·9 (0·8–1·2)	1·0 (0·9–1·2)	0·07
Older adults	19 (12)	0·8 (0·7–0·9)*	0·9 (0·7–1·0)	0·11
All ages	28 (11)	1·2 (0·7–2·0)	1·3 (0·8–2·1)	0·09
Type of measure				
Simple count	21 (30)	1 (ref)	1 (ref)	1 (ref)
Weighted measure	25 (17)	1·2 (1·0–1·4)	1·7 (1·3–2·2)*	0·04
Interaction				
Weighted measure: databases	NA	NA	0·4 (0·3–0·5)*	0·09

IRR=incidence rate ratio. NA=not applicable. *p<0·05.

Table 2: **Association between number of conditions included in the multimorbidity measure and study characteristics**

community settings (269 [85·9%] of 313 studies) and primary care (81 [73·6%] of 110 studies), whereas weighted measures were predominantly used in hospital settings (73 [77·7%] of 94 studies; χ^2=147·66, p=0·0005).

56 (9·9%) of the 566 studies did not report how many chronic conditions were included in their multimorbidity measure, and 41 (7·2%) compared the performance of multiple measures in relation to outcomes (table 1, appendix pp 5–71). For the 469 (82·9%) remaining studies examining a single measure, the number of conditions included in the measure ranged from two to 285 (median 17 [IQR 11–23]; appendix p 73).

Associations between multimorbidity study characteristics and the number of conditions included in measures of multimorbidity are shown in table 2. In univariate analysis, studies measuring multimorbidity on the basis of database sources included more conditions in measures of multimorbidity than studies based on patient self-report, and community-based studies included fewer conditions in measures of multimorbidity than studies in primary care, hospital, and care home settings. Studies that included older adults included more conditions in measures of multimorbidity than those restricted to children or younger and middle-aged adults. In the multivariate model, studies of database sources were estimated to include more than twice as many conditions in measures of multimorbidity than self-report studies (IRR 2·2 [95% CI 1·8–2·6]). Study setting and study population were not significantly associated with the number of conditions included in measures in the adjusted model. With the exception of five studies measuring multimorbidity in children in which FMI was 0·53 (high), FMI ranged from 0·04–0·30 (small to moderate) indicating that multiple imputation uncertainty was acceptable.

73 (12·9%) of the 566 studies did not report which health conditions were included in their measure, and 41 (7·2%) compared the predictive performance of multiple measures. In the remaining 452 (79·9%) studies, 67 conditions were included across the individual study measures of multimorbidity. Most of the 452 studies included at least one cardiovascular condition (441 [97·6%] studies), metabolic and endocrine condition (440 [97·3%]), respiratory condition (422 [93·4%]), musculoskeletal condition (396 [87·6%]), or mental health condition (355 [78·5%]) in the measure of multimorbidity used (figure 2). Conversely, chronic infections (123 [27·2%] studies), haematological conditions (110 [24·3%]), ear, nose, and throat conditions (107 [23·7%]), skin conditions (70 [15·5%]), oral conditions (19 [4·2%]), and congenital conditions (14 [3·1%]) were included in the morbidity measure by a minority of studies.

With respect to individual health conditions (appendix p 74), only eight conditions were included by more than half of studies in the multimorbidity measure used. Diabetes was included in 411 (90·9%) of 452 studies, with fewer studies including stroke (357 [79·0%]), cancer (349 [77·2%]), chronic obstructive pulmonary disease (322 [71·2%]), hypertension (316 [69·9%]), coronary heart disease (275 [60·8%]), chronic renal disease (240 [53·1%]), and heart failure (230 [50·9%]) in the multimorbidity measure. By contrast, health conditions included in less than 2% of studies were sleep apnoea (five [1·1%]), chronic fatigue syndrome (four [0·9%]), and attention deficit hyperactivity disorder (three [0·7%]).

Fewer than half of studies included any one mental health condition in their multimorbidity measure. Multimorbidity measures included both physical and mental health conditions in 331 (73·2%) studies, exclusively physical conditions in 115 (25·4%) studies, and exclusively mental health conditions in six (1·3%) studies. In studies including both physical and mental health conditions, measures usually only included a small number of mental health conditions. In all studies, the most commonly included mental health conditions were depression (219 [48·5%] studies), dementia (197 [43·6%]), and anxiety (98 [21·7%]). All other mental health conditions, including schizophrenia, alcohol and drug use disorders, and bipolar disorder, were included in less than 20% of studies.

Inclusion of a condition had no significant association with country level of income or continent of study, except for tuberculosis, on Pearson's χ^2 tests (data not shown). In our post-hoc analysis, we found that 23 (6·2%) of

371 studies in HICs included tuberculosis in their multimorbidity measurement, compared with 11 (16·4%) of 67 studies in LMICs (χ^2=6·91, p=0·0009; appendix p 75). By continent, significant associations were observed regarding inclusion of tuberculosis in the multimorbidity measurement, with our post-hoc test identifying Asia as having a significantly higher proportion of studies including tuberculosis (14 [20·6%] of 68 studies) than Europe (eight [4·9%] of 163 studies), North America (five [3·3%] of 150 studies), Australasia (five [15·6%] of 32 studies), Africa (one [33·3%] of three studies), and South America (one [4·5%] of 22 studies; χ^2=27·42, p=0·0002; appendix p 75).

We analysed how often conditions with the highest global burden of DALYs and YLLs in the Global Burden of Disease Study 2019[21] were included in multimorbidity measures among the 452 studies that specified chronic conditions (figure 3, appendix p 76). Of the 10 conditions with highest DALY burden, coverage in multimorbidity measures ranged from 12 (2·7%) studies to 411 (90·9%) studies. Among the 25 conditions with highest DALY burden, conditions included in less than 10% of multimorbidity measures were tuberculosis, malnutrition, gynaecological disorders, and oral disorders.

Of the 566 studies measuring multimorbidity, 107 were rated to be at high risk of bias, 419 at moderate risk of bias, and 40 at low risk of bias (appendix pp 77–118). In our post-hoc analysis, risk of bias was significantly associated with clarity of reporting of multimorbidity measurement; studies at low risk of bias were significantly less likely to have unclear reporting (four [10·0%] of 40 studies) than studies with moderate risk of bias (101 [24·1%] of 419 studies) and high risk of bias (82 [76·6%] of 107 studies; (χ^2=116·64, p<0·0001; appendix p 119).

Discussion

In this review, more than a third of studies did not report their definition of multimorbidity. When reported, the most common definition used was two or more chronic conditions, with small numbers of studies using three or more conditions or five or more conditions as cutoffs. One in ten studies did not report the number of conditions included in the multimorbidity measure, and one in eight did not report which conditions were included. The number of conditions included in multimorbidity measures showed large variation, ranging from two to 285 (median 17 [IQR 11–23]). Approximately 66% of studies used a simple count for multimorbidity measurement (for a range of study purposes), whereas nearly 30% of studies measured multimorbidity with weighted indices (mainly for predicting outcomes).

The 452 studies which reported the conditions included in their multimorbidity measure varied in terms of which conditions they counted. More than 90% of studies included at least one cardiovascular, metabolic and endocrine, or respiratory condition.

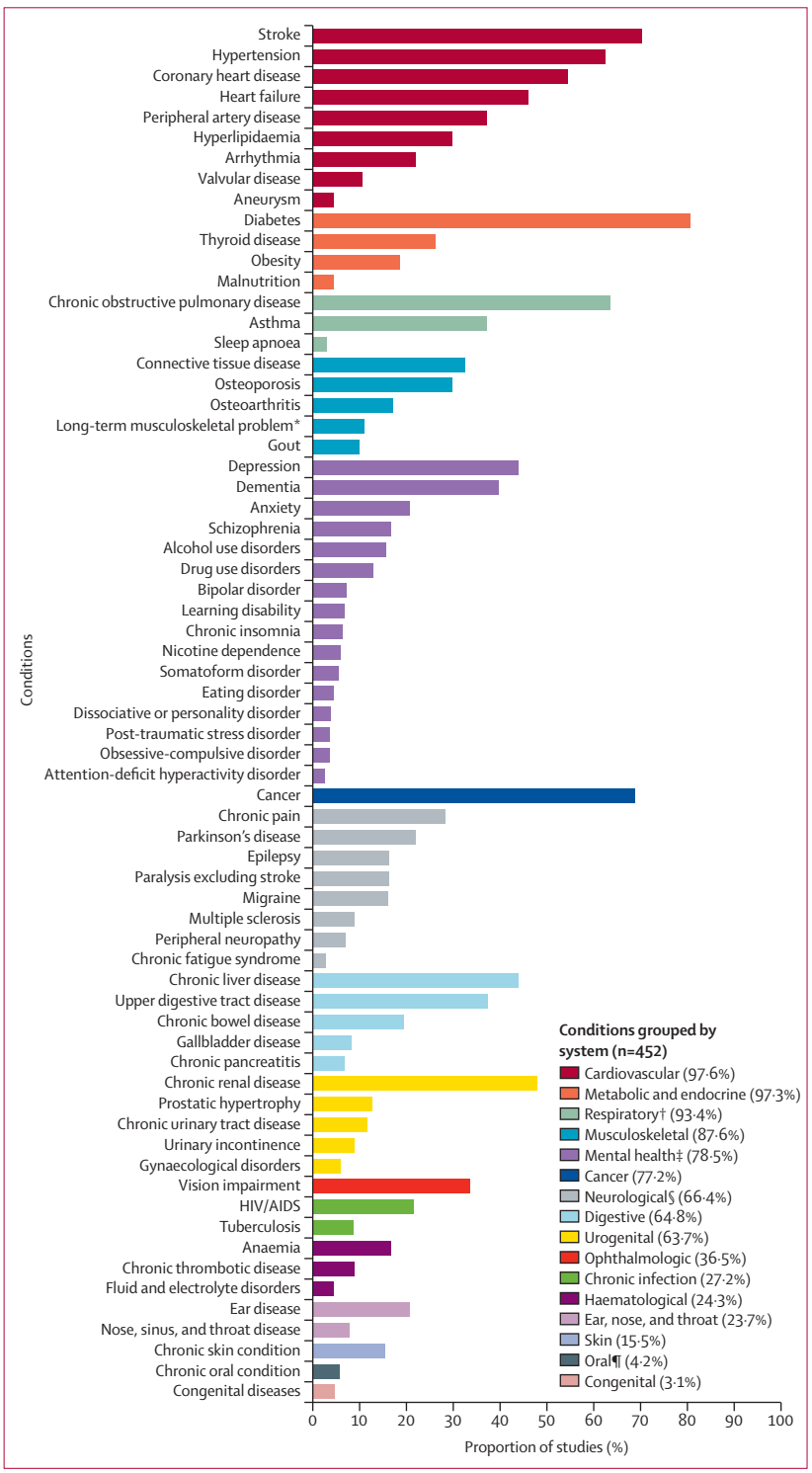

Figure 2: **Long-term conditions included in multimorbidity measures**
Conditions were grouped according to ICD-10 chapter with modification to reflect the organisation of clinical care. The 67 conditions are shown in descending order in the appendix (p 74). ICD=International Classification of Diseases. *Due to injury. †Sleep apnoea is classed as a nervous system disorder in ICD-10. ‡Learning disabilities include congenital conditions in which learning disability is a dominant feature (eg, Down syndrome). §Chronic pain and chronic fatigue syndrome are classed as symptoms or signs in ICD-10. ¶Throat conditions are classed as digestive system disorders in ICD-10. ||Oral conditions are classed as digestive system disorders in ICD-10.

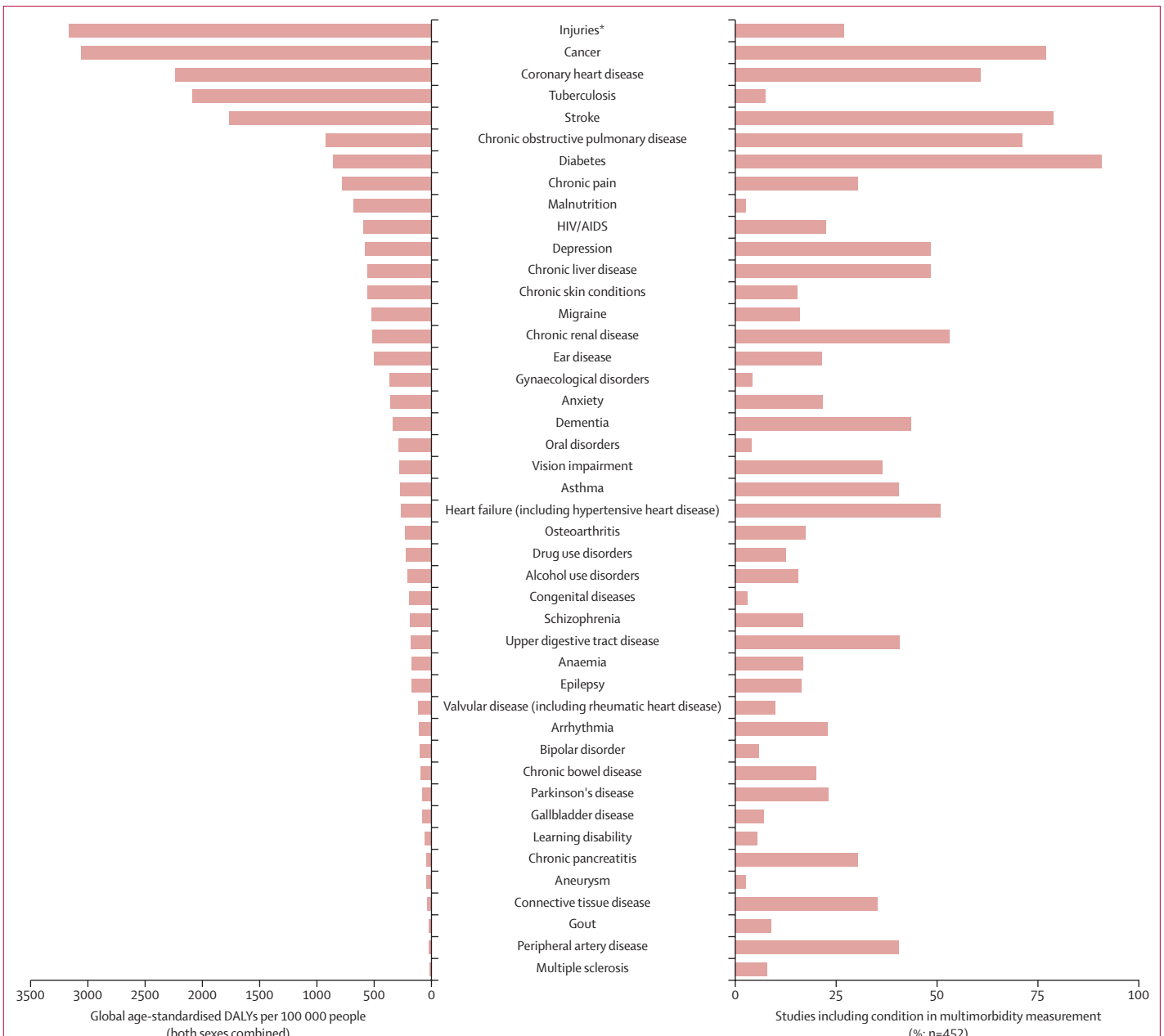

Figure 3: **DALYs of selected conditions and percentage of studies including each condition in multimorbidity measurement**
Conditions are ranked in descending order of attributable DALYs from the Global Burden of Disease Study 2019; data on DALYs were derived from the supplementary appendixes of Vos et al[21] (appendix p 75). DALYs comprise the estimates of years of life lost due to premature mortality and years lived with disability. DALYs=disability-adjusted life years. *Conditions relating to consequences of injury in this study are paralysis (from causes other than stroke; 75 [16·6%] of 452 studies) and long-term musculoskeletal impairment due to injury (47 [10·4%]).

Chronic infections, haematological conditions, ear, nose, and throat conditions, skin conditions, oral conditions, and congenital conditions were included by less than a third of studies. Only eight individual conditions were included by more than half of studies in the measure of multimorbidity used, all of which were physical health conditions (diabetes, stroke, malignancy, hypertension, chronic obstructive pulmonary disease, coronary heart disease, chronic renal disease, and heart failure).

Fewer than half of studies included any one mental health condition, and a quarter of studies did not include any mental health condition, which identifies a clear gap in the literature. Mental and physical multimorbidity has been well recognised to have collective effects on mortality, quality of life, disability, and patient activation in managing their own health care.[34–36] Therefore, the inclusion of mental health conditions, such as depression, dementia, anxiety, and schizophrenia is

required to properly understand the epidemiology and implications of multimorbidity.

Previous narrative systematic reviews on multimorbidity measurement have focused on the comparison of different weighted indices in predicting one or more outcomes, including mortality and health care utilisation.[10,37–39] Consistent with previous reviews,[13,39] in this review we found that studies commonly used a weighted index when the purpose was to predict outcomes, and Stirland and colleagues[39] in their recent systematic review provided recommendations on which weighted index to use for a particular outcome of interest. Although simple counts with a cutoff of two or more conditions to define multimorbidity are commonly used,[40] this approach has been criticised as being unable to identify people with the greatest needs.[41] However, in practice the predictive validity of weighted indices and a simple count often only differs slightly.[42,43] Beyond counting, a shift of focus has been suggested in recent years to improve understanding of the effects of particular disease clusters on clinical outcomes.[2,44]

We identified that a notable minority of studies did not report how multimorbidity was defined and measured. At a minimum, we believe that all studies of multimorbidity need to clearly report: their core multimorbidity definition, and whether measured by a simple count or a weighted index; which conditions are included in the measure; why these conditions were selected in relation to the purpose of the study; and how each condition has been defined including any clinical code sets used.

Clear reporting is essential, but our study also found large variation in the number and nature of the conditions included in multimorbidity measures. The optimal conditions to include are likely to vary somewhat by study context and purpose, but we believe there is value in identifying a core set of conditions that all studies should include, and factors which might then influence choice of other conditions. We recognise that no reason exists in principle to exclude any condition from morbidity counts, and that rare diseases are cumulatively common, but identifying rare conditions in routine data is problematic, and self-report studies are limited by concerns for participant burden when completing surveys, often leading to the use of general condition categories for participants to choose from.[45] Similarly, uncertainty exists about when to count condition subgroups separately (eg, myocardial infarction or angina) or within more heterogenous groups (eg, coronary artery disease). The 67 conditions we identified across 452 studies that reported type of conditions provide insight and, along with consideration of conditions with high DALYs or YLL, provide a potential core list to choose from, but we recognise that researchers might deviate from such a list depending on their purpose. Based on our analysis of conditions most commonly included in multimorbidity measures, and drawing on the recommendations of a previous review[10] and DALY burden of different conditions,[21] we propose a minimal core list of 20 conditions and considerations for choosing other conditions (figure 4), although consensus studies are needed to refine choices, and to underpin further research that is comparable and reproducible.

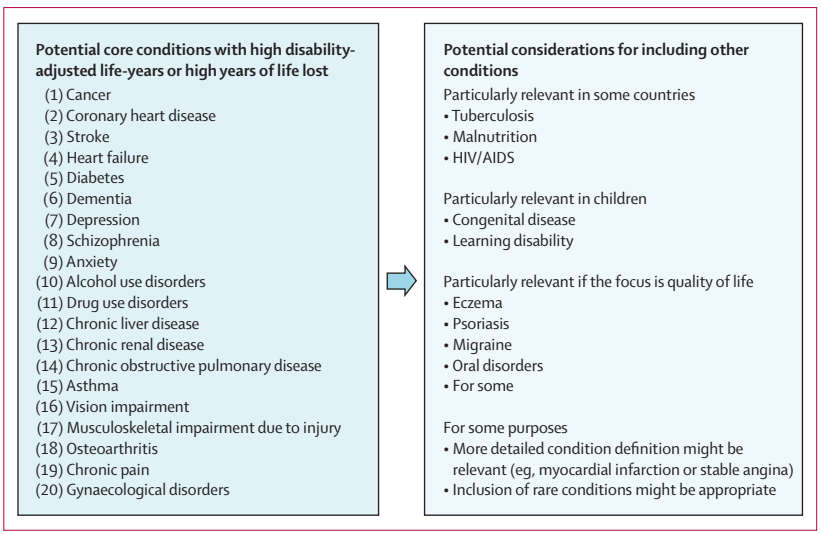

Figure 4: **A potential core list of 20 chronic conditions to include in measures of multimorbidity and considerations for choosing other conditions**

The key strengths of this review include comprehensive database searches, assessment of risk of bias, robust analysis, and the systematic examination of multimorbidity measurement in a large number of studies. A limitation is inconsistency in the labelling of relevant studies, meaning that not all relevant studies are likely to have been included, although we assume this had little effect on the overall conclusion that multimorbidity measurement is poorly reported and highly variable. There was also considerable variation in how different conditions were named or grouped by multimorbidity measures, which meant that for synthesis we had to group or combine some conditions that the underlying studies kept distinct. Restricting included studies to English might also mean that we have missed relevant studies in other countries, particularly low-income and middle-income countries. The possible consequence is that true heterogeneity of measurement is larger than we observed.

To conclude, this review found that many previous studies have not clearly reported their methods of measuring multimorbidity. Consistent reporting of measure definitions and condition definitions is required. When reported, measurement of multimorbidity was highly variable. A consensus is needed to establish a core condition set that all studies should measure (to facilitate comparison and synthesis across studies), and study purposes and contexts for which other condition sets will be appropriate.

Contributors
CMcC, KN, UK, KK, RAL, JD, CB, AA, AA-L, and SM were involved in conception of the work, acquisition of funding, and critically commenting on the manuscript. IS-SH and BG contributed to the design, analysis, and interpretation of data for the review. IS-SH and PH screened and reviewed retrieved studies, and IS-SH extracted the data. IS-SH and BG accessed and verified the underlying data. All authors had full access to all the data in the study and contributed to the edits of the manuscript. All authors approved the final draft and had final responsibility for the decision to submit for publication.

Declaration of interests
We declare no competing interests.

Data sharing
Study data are available on request to the corresponding author.

Acknowledgments
This work was supported by supported by Health Data Research UK (HDR-UK; grant number CFC0110) which receives its funding from the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (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 the Wellcome Trust. We would like to express our appreciation to members of the wider HDR-UK National Multimorbidity Resource for contribution to the conceptualisation of the study.

References

1 WHO. Electronic tools: technical series on safer primary care. 2016. https://apps.who.int/iris/bitstream/handle/10665/252374/9789241511667-eng.pdf;jsessionid=1B003F23E58BBA44948F7862DEF30E9D?sequence=1 (accessed Jan 27, 2021).
2 The Academy of Medical Sciences. Multimorbidity: a priority for global health research. April, 2018. https://acmedsci.ac.uk/file-download/82222577 (accessed Jan 27, 2021).
3 Ricci-Cabello I, Violán C, Foguet-Boreu Q, Mounce LTA, Valderas JM. Impact of multi-morbidity on quality of healthcare and its implications for health policy, research and clinical practice. A scoping review. *Eur J Gen Pract* 2015; **21:** 192–202.
4 Hopman P, de Bruin SR, Forjaz MJ, et al. Effectiveness of comprehensive care programs for patients with multiple chronic conditions or frailty: a systematic literature review. *Health Policy* 2016; **120:** 818–32.
5 Giovannetti ER, Dy S, Leff B, et al. Performance measurement for people with multiple chronic conditions: conceptual model. *Am J Manag Care* 2013; **19:** e359–66.
6 Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: an overview of systematic reviews. *Ageing Res Rev* 2017; **37:** 53–68.
7 Lefèvre T, d'Ivernois JF, De Andrade V, Crozet C, Lombrail P, Gagnayre R. What do we mean by multimorbidity? An analysis of the literature on multimorbidity measures, associated factors, and impact on health services organization. *Rev Epidemiol Sante Publique* 2014; **62:** 305–14.
8 Harrison C, Britt H, Miller G, Henderson J. Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian general practice. *BMJ Open* 2014; **4:** e004694.
9 Violan C, Foguet-Boreu Q, Flores-Mateo G, et al. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PLoS One* 2014; **9:** e102149.
10 Diederichs C, Berger K, Bartels DB. The measurement of multiple chronic diseases—a systematic review on existing multimorbidity indices. *J Gerontol A Biol Sci Med Sci* 2011; **66:** 301–11.
11 van den Brink AM, Gerritsen DL, Voshaar RC, Koopmans RT. Residents with mental–physical multimorbidity living in long-term care facilities: prevalence and characteristics. A systematic review. *Int Psychogeriatr* 2013; **25:** 531–48.
12 Nowels MA, VanderWielen LM. Comorbidity indices: a call for the integration of physical and mental health. *Prim Health Care Res Dev* 2018; **19:** 96–98.
13 Nicholson K, Almirall J, Fortin M. The measurement of multimorbidity. *Health Psychol* 2019; **38:** 783–90.
14 Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012; **10:** 142–51.
15 Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. *Int J Evid-Based Healthc* 2015; **13:** 147–53.
16 Gardner W, Mulvey EP, Shaw EC. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychol Bull* 1995; **118:** 392–404.
17 van Buuren S. Package 'mice': multivariate imputation by chained equations. Jan 27, 2021. https://cran.r-project.org/web/packages/mice/mice.pdf (accessed Jan 27, 2021).
18 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011; **30:** 377–99.
19 World Bank Group. The world by income. 2018. https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html (accessed June 1, 2021).
20 Beasley TM, Schumacker RE. Multiple regression approach to analyzing contingency tables: post hoc and planned comparison procedures. *J Exp Educ* 1995; **64:** 79–93.
21 Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396:** 1204–22.
22 Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012; **18:** 12–18.
23 Prior A, Fenger-Grøn M, Larsen KK, et al. The association between perceived stress and mortality among people with multimorbidity: a prospective population-based cohort study. *Am J Epidemiol* 2016; **184:** 199–210.
24 Vassilaki M, Aakre JA, Mielke MM, et al. Multimorbidity and neuroimaging biomarkers among cognitively normal persons. *Neurology* 2016; **86:** 2077–84.
25 Chamberlain AM, Rutten LJF, Wilson PM, et al. Neighborhood socioeconomic disadvantage is associated with multimorbidity in a geographically-defined community. *BMC Public Health* 2020; **20:** 1–10.
26 MacKnight C, Rockwood K. Use of the chronic disease score to measure comorbidity in the Canadian Study of Health and Aging. *Int Psychogeriatr* 2001; **13** (suppl 1)**:** 137–42.
27 Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 1992; **45:** 197–203.
28 Kim S, Bennett K, Wallace E, Fahey T, Cahir C. Measuring medication adherence in older community-dwelling patients with multimorbidity. *Eur J Clin Pharmacol* 2018; **74:** 357–64.
29 Pratt NL, Kerr M, Barratt JD, et al. The validity of the Rx-risk comorbidity index using medicines mapped to the Anatomical Therapeutic Chemical (ATC) classification system. *BMJ Open* 2018; **8:** e021122.
30 George J, Vuong T, Bailey MJ, Kong DCM, Marriott JL, Stewart K. Development and validation of the medication-based disease burden index. *Ann Pharmacother* 2006; **40:** 645–50.
31 Narayan SW, Nishtala PS. Development and validation of a medicines comorbidity index for older people. *Eur J Clin Pharmacol* 2017; **73:** 1665–72.
32 Miller MD, Paradis CF, Houck PR, et al. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. *Psychiatry Res* 1992; **41:** 237–48.
33 Bayliss EA, Ellis JL, Steiner JF. Subjective assessments of comorbidity correlate with quality of life health outcomes: initial validation of a comorbidity assessment instrument. *Health Qual Life Outcomes* 2005; **3:** 51.
34 Scott KM, Von Korff M, Alonso J, et al. Mental-physical co-morbidity and its relationship with disability: results from the World Mental Health Surveys. *Psychol Med* 2009; **39:** 33–43.

35 Lynch CP, Gebregziabher M, Zhao Y, Hunt KJ, Egede LE. Impact of medical and psychiatric multi-morbidity on mortality in diabetes: emerging evidence. *BMC Endocr Disord* 2014; **14:** 68.

36 Blakemore A, Hann M, Howells K, et al. Patient activation in older people with long-term conditions and multimorbidity: correlates and change in a cohort study in the United Kingdom. *BMC Health Serv Res* 2016; **16:** 582–82.

37 Yurkovich M, Avina-Zubieta JA, Thomas J, Gorenchtein M, Lacaille D. A systematic review identifies valid comorbidity indices derived from administrative health data. *J Clin Epidemiol* 2015; **68:** 3–14.

38 Sharabiani MTA, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care* 2012; **50:** 1109–18.

39 Stirland LE, González-Saavedra L, Mullin DS, Ritchie CW, Muniz-Terrera G, Russ TC. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. *BMJ* 2020; **368:** m160.

40 Johnston MC, Crilly M, Black C, Prescott GJ, Mercer SW. Defining and measuring multimorbidity: a systematic review of systematic reviews. *Eur J Public Health* 2019; **29:** 182–89.

41 Mokraoui NM, Haggerty J, Almirall J, Fortin M. Prevalence of self-reported multimorbidity in the general population and in primary care practices: a cross-sectional study. *BMC Res Notes* 2016; **9:** 314.

42 Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med* 2012; **10:** 134–41.

43 Aubert CE, Schnipper JL, Roumet M, et al. Best definitions of multimorbidity to identify patients with high health care resource utilization. *Mayo Clin Proc Innov Qual Outcomes* 2020; **4:** 40–49.

44 Whitty CJM, MacEwen C, Goddard A, et al. Rising to the challenge of multimorbidity. *BMJ* 2020; **368:** l6964.

45 Fortin M, Hudon C, Haggerty J, Akker M, Almirall J. Prevalence estimates of multimorbidity: a comparative study of two sources. *BMC Health Serv Res* 2010; **10:** 111.