Edinburgh Research Explorer

Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice

Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice

Lucy E Stirland,[1,2] Laura González-Saavedra,[3] Donncha S Mullin,[2,4] Craig W Ritchie,[1,2] Graciela Muniz-Terrera,[1,2] Tom C Russ[1,5,6]

For numbered affiliations see end of the article.

Correspondence to: L E Stirland
l.stirland@ed.ac.uk
(or @stirlandia on Twitter;
ORCID 0000-0002-5678-4583)

ABSTRACT

OBJECTIVES
To identify and summarise existing indices for measuring multimorbidity beyond disease counts, to establish which indices include mental health comorbidities or outcomes, and to develop recommendations based on applicability, performance, and usage.

DESIGN
Systematic review.

DATA SOURCES
Seven medical research databases (Medline, Web of Science Core Collection, Cochrane Library, Embase, PsycINFO, Scopus, and CINAHL Plus) from inception to October 2018 and bibliographies and citations of relevant papers. Searches were limited to English language publications.

ELIGIBILITY CRITERIA FOR STUDY SELECTION
Original articles describing a new multimorbidity index including more information than disease counts and not focusing on comorbidity associated with one specific disease. Studies were of adults based in the community or at population level.

RESULTS
Among 7128 search results, 5560 unique titles were identified. After screening against eligibility criteria the review finally included 35 papers. As index components, 25 indices used conditions (weighted or in combination with other parameters), five used diagnostic categories, four used drug use, and one used physiological measures. Predicted outcomes included mortality (18 indices), healthcare use or costs (13), hospital admission (13), and health related quality of life (7). 29 indices considered some aspect of mental health, with most including it as a comorbidity. 12 indices are recommended for use.

CONCLUSIONS
35 multimorbidity indices are available, with differing components and outcomes. Researchers and clinicians should examine existing indices for suitability before creating new ones.

SYSTEMATIC REVIEW REGISTRATION
PROSPERO CRD42017074211.

Introduction

Multimorbidity, usually defined as the coexistence of two or more chronic conditions within an individual, is important for patient outcomes and healthcare costs. Because its prevalence is rising as populations age, multimorbidity is attracting increasing attention from the research community worldwide.[1] More than 2800 publications on multimorbidity appeared between 1900 and 2016, 80% of which were published after 2010.[2] As no universally agreed measure or list of diseases exists to define multimorbidity, numerous indices have been developed. These might be designed, for example, to quantify multimorbidity as a covariate in other analyses, for mortality prediction or for risk adjustment. Previous systematic reviews identified multiple indices, but no searches have been done of indices since 2009.[3 4]

Multimorbidity research most often refers to a count of chronic conditions.[4] This method does not reflect patients' experience or the effects of different combinations or severity of diseases.[5] Some indices, however, combine disease counts with severity measures, physiological factors, or demographic items, thereby allowing a more holistic quantification of illness burden.

The coexistence of both physical and mental illness within multimorbidity is prevalent.[6] A 2018 report identifying priorities for multimorbidity research highlighted the need for more work into this coexistence.[1] Researchers exploring multimorbidity will therefore increasingly need to account for mental disorders. As previous reviews of multimorbidity indices have not covered mental health in depth we identified and summarised all community based multimorbidity measures that include more than simple disease counts, paying particular attention to mental health. This review should help guide clinicians and researchers to select an appropriate index according to their requirements.

WHAT IS ALREADY KNOWN ON THIS TOPIC

It is common for people to have two or more co-occurring chronic conditions (multimorbidity)

Researchers and clinicians use many different indices to measure multimorbidity

WHAT THIS STUDY ADDS

At least 35 objective measures of multimorbidity are available for people living in the community and each of these uses different variables to generate a score or index, linked with various or no outcome measures

No specific index investigates the interplay between mental and physical multimorbidity, although this is dealt with in a variety of ways across the measures

The recommendations in this study should guide researchers to find a suitable index for their purposes

Methods

No single accepted term describes the methods of measuring multimorbidity. In this review we use the term "index" to refer to any method of quantifying disease burden or predicting specific outcomes that includes more than a count of conditions. This could be by weighting conditions (for example, by allocating a score to each), adding other elements, or examining other variables such as drug or physiological parameters.

Search

To capture all relevant publications we conducted a broad search. We included a variety of terms for multimorbidity derived from previous systematic reviews on this topic[3 4] and other literature discussing terminology in this area of research.[7] We developed the search strategy iteratively with the support of an Academic Support Librarian. The final search terms are listed in the appendix (appendix eTable 1) and include multimorbidity, comorbidity, polypathology, polymorbidity, pluripathology, multi-condition, and multiple chronic conditions. The search was restricted to adults older than 18 years and to English language publications.

Eligibility criteria

We planned to summarise reports of novel indices and were primarily interested in the original report of each index. Therefore we excluded papers that either used existing indices or measured multimorbidity using only disease counts. In the initial screening process we included only the original form of each scale and not adaptations or updates; these were found later. Records that were not original research papers, such as conference abstracts, letters, and systematic reviews, were excluded. We defined multimorbidity as multiple co-occurring conditions without reference to a specific disease, so excluded papers were those that focused on comorbidities of an index disease or on comorbidities within one disease area (such as the coexistence of several psychiatric conditions). As most people with multimorbidity are adults living in the community and managed in primary care, we excluded articles about children, animals, or people admitted to hospital or living in residential care. We included studies of hospital inpatients when the primary focus was follow-up after discharge (for example, mortality one year later). As resources were limited, we excluded papers when full texts were not available in English.

Information sources

On 19 October 2018 we searched Medline, Web of Science Core Collection, Cochrane Library, Embase, PsycINFO, Scopus, and CINAHL Plus from inception onwards.

Study selection

Two authors used Covidence software independently to screen titles against exclusion criteria and the subsequent abstracts against the same criteria.[8] We then extracted the full texts of relevant abstracts for further screening. Any disagreements at the title, abstract, and full text stages were resolved by discussion, and a third author mediated unresolved conflicts. We excluded papers that referred to an existing index, but listed the indices that were used when excluding them at the abstract stage. We found the original papers describing these indices and returned them to the title screening stage. Additional relevant titles were found by reviewing previous systematic reviews on this topic, searching the bibliographies of included full text papers, and tracking their citations using Google Scholar. Emerging relevant titles were added to the screening process.

Usage, updates, and validation

After the list of included papers had been finalised, we searched their citations on Google Scholar for updates, revisions, or adaptations as well as validation papers. When original indices were adapted and validated numerous times, we listed the original performance and principal adaptations. We did not include adaptations where the original index was translated into another language with no other changes made. To assess the popularity of each index, we took the total number of citations for each original paper from Google Scholar on 7 September 2019, as a proxy for use. We then calculated the number of citations for each whole year since publication. To retain awareness of the context of their initial design and aims, we summarised index updates separately from the original papers.

Data collection process

We created a data extraction tool containing specific elements of interest for each original index. This tool took account of previous reviews on this topic as well as additional information relating to mental health. We used a broad definition of mental health, comprising any mental disorder, including mood disorders, dementia, delirium, and addictions as well as relevant symptoms. Many papers describe validation of the indices used so details on the size and demographic distribution of the populations under test was important.

Two authors independently extracted data from all full texts. We compared the consistency of extracted items and resolved any differences by discussion and reference to the original paper, with a third author available in case of substantially differing data extraction.

Data items

The variables of interest during data extraction were first author, year of publication, and name of index; original purpose of the index; characteristics of the population under test, including type of data source (eg, cohort study), location, number of participants, sex and age distribution, and mean number of concurrent medical conditions (when given); components included in the index; weighting method (if any) and details of model for its development; outcome measures; information and resources required to apply the index; internal validation or comparison

with another index (if applicable); external validation and performance compared with another index (if applicable); and inclusion of mental health (either in comorbidities or as outcomes).

Risk of reporting bias in original studies

We assessed the risk of bias of study design and reporting and aimed to develop overall recommendations for index choice. The Cochrane Collaboration advises against scales that generate total numerical scores, preferring emphasis on individual papers' performance on each criterion.[9] After our search date the Prediction model Risk Of Bias ASsessment Tool (PROBAST) was published. It focuses on risk of bias and applicability in prediction model studies.[10] As our search was not restricted to prediction models, it was not appropriate to apply this tool to every paper. We therefore developed our own list of criteria having referred to resources available for assessing clinical prediction indices.[11-13] Our assessment aimed to deal with risk of selection, observer, and funding bias. The list contained 10 questions on the population tested, description of the index, statistical methods, validity, and funding. The assessment tool is available in the appendix (appendix eTable 2), including division into domains. We also included an overall impression of the papers' risk of bias based on the Scottish Intercollegiate Guidelines Network standard, which was scored separately to the criteria rather than in an additive fashion.[14] Two reviewers independently assessed each paper and resolved disagreements by discussion.

When choosing an index, its predictive ability and its use elsewhere are important. We generated overall recommendations taking into account the generalisability of participants, selection and clinical relevance of index components, outcome measurement, risk of reporting bias, and model evaluation. These were separated into three main categories: recommended, potentially useful (usually when indices were applicable to certain situations), and not recommended.

Synthesis of results

We anticipated finding a wide variety of indices covering diverse outcome measures and therefore planned to summarise these narratively. Because of the range of outcomes included we did not expect to be able to perform meta-analysis. We listed performance and validity statistics as reported by the original papers or validation studies alongside each other, for comparison.

We did not expect to find indices specifically designed for measuring physical multimorbidity in relation to any aspect of mental health. Therefore, for separate narrative synthesis we planned to seek those indices that mentioned mental illnesses or symptoms, either as comorbidities or as outcomes.

Patient and public involvement

An early draft of this paper was discussed with a lay contributor who has personal experience of multimorbidity. We incorporated her comments into the text—for example, noting in the introduction that the number of conditions a person has might not reflect their experience of health, and in the discussion her suggestions about outcomes that could be important to patients. The lay contributor also commented on a lay summary of the paper (see appendix page 43), which we amended accordingly.

Results

Study selection

The searches yielded 7128 results. A search of bibliographies and citations identified 48 additional relevant titles, and a further 15 titles were added from the list of indices mentioned in excluded abstracts. The total number of titles was therefore 7191. Duplicates were removed using EndNote X8 and Covidence software,[8 15] leaving 5560 unique records for screening (fig 1).[16] Overall, 5236 titles were excluded at the screening stage, leaving 324 abstracts for eligibility assessment. Of these, 86 full texts were assessed and 35 papers finally included.

Study characteristics

Twenty articles originated from the United States[17-36]; three from Australia[37-39]; two each from the United Kingdom,[40 41] Taiwan,[42 43] and Italy[44 45]; and one each from Canada,[46] Spain,[47] Germany,[48] New Zealand,[49] Norway,[50] and India.[51] They were published between 1968 and 2017, with 15 (43%) published since the last systematic review on this topic in 2009.[17-20 40 42-51] The mean number of participants included in the derivation populations of indices developed after 2009 was 356 906, compared with 75 491 before 2009. The newer indices primarily required access to medical records in 11 (73%) cases,[18-20 40 42-45 47-49] and the remainder (4, 27%) self-report[17 46 50 51]; 10 (50%) indices before 2009 primarily used medical records[21 23 26 29 32-36 39] and 10 (50%) used self-report.[22 24 25 27 28 30 31 37 38 41]

The majority of papers described one final multimorbidity index, even if they tested various models in development, and four papers concluded with more than one index or measure.[23 25 30 46] For consistency, when articles were summarised and their quality assessed we considered each paper as a whole and noted when more than one index existed. We did not comment on models that used only unweighted disease counts, in keeping with our overall exclusion criteria.

Index components

Four indices primarily used weighted drug counts to quantify multimorbidity,[33 39 43 45] five used diagnostic groups or clusters,[23 26 34 36 48] and 25 included counts of diseases. Of these, 21 were weighted[17 19 20 22 24 27-29 31 32 35 37 38 40 41 44 46 47 49 50 51] and nine incorporated other parameters, demographic or otherwise.[18 19 22 25 28 30 42 47 51] Four papers used a combination of weighting and additional variables.[19 22 28 47] One index used physiological measurements to diagnose multimorbidity.[21] When
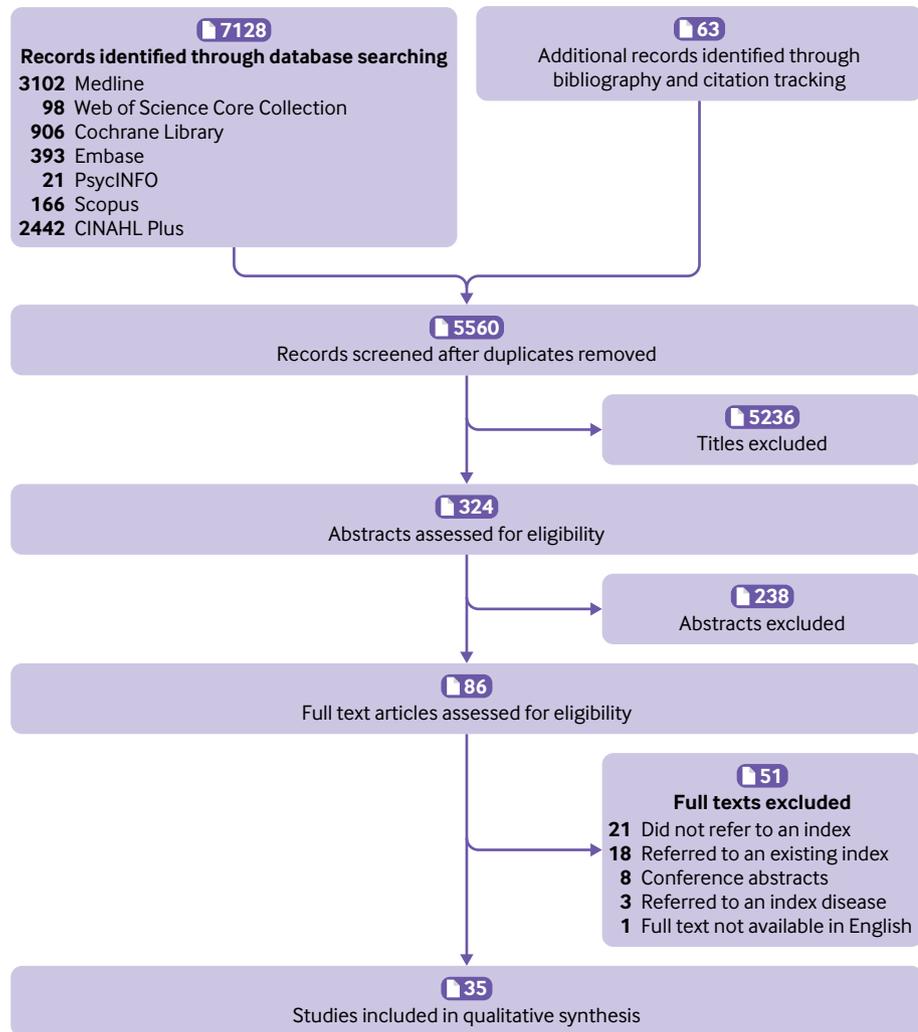
Fig 1 | Screening process according to PRISMA

diagnoses were required, 14 indices took these from medical records[18 19 20 23 26 29 32 35 36 40 42 44 48 49] and 15 from self-report.[17 22 24 25 27 28 30 31 36 37 38 41 46 50 51] Figure 2 is a spider diagram of the papers, displayed according to their broad category of index components.

Outcome measures

The number of outcomes assessed in each paper ranged from none to seven. Eighteen studies measured mortality as an outcome,[19 21 22 25 28 29 33 35-40 42 44 45 47 49] 13 aimed to predict hospital admissions,[18 23 28 31 33 37-39 42-45 49] 10 measured general healthcare use,[18 19 23 25 27 31 33 34 38 46] seven measured independence with daily activities or disability (with or without physical functioning),[19 21 24 31 33 38 41] seven measured health related quality of life,[20 25 27 37 38 46 48] five measured overall self-reported health,[24 33 41 46 50] five measured healthcare costs,[23 26 27 30 44] and four measured drug use.[19 23 27 46] Mental health was a specific outcome in three papers,[24 33 41] with a further six including mental health aspects of established tools (eg, the mental component score of SF-36).[20 25 27 37 38 48] Adherence to screening programmes,[18] specific physiological parameters,[18] and care quality indicators[19] were assessed by one tool each. Ten (29%) papers used cross sectional data to derive their index weighting but measured longitudinal outcomes.[19 21 22 25 33 36 38 39 42 43] Table 1 lists the papers according to their original outcomes and index components.

Applicability

The applicability of each index depends on study design and intended usage. Table 1 summarises index components and outcome variables. Most of the original papers (27, 77%) contained sufficient information for readers to use the index, usually with lists of included conditions with or without weighting. Of the remainder, access to additional free resources was needed in four,[30 43 45 47] information from proprietary scales was needed in one,[51] and two were missing information that would allow the index to be applied.[18 31] One index was only available as proprietary software.[34 52]

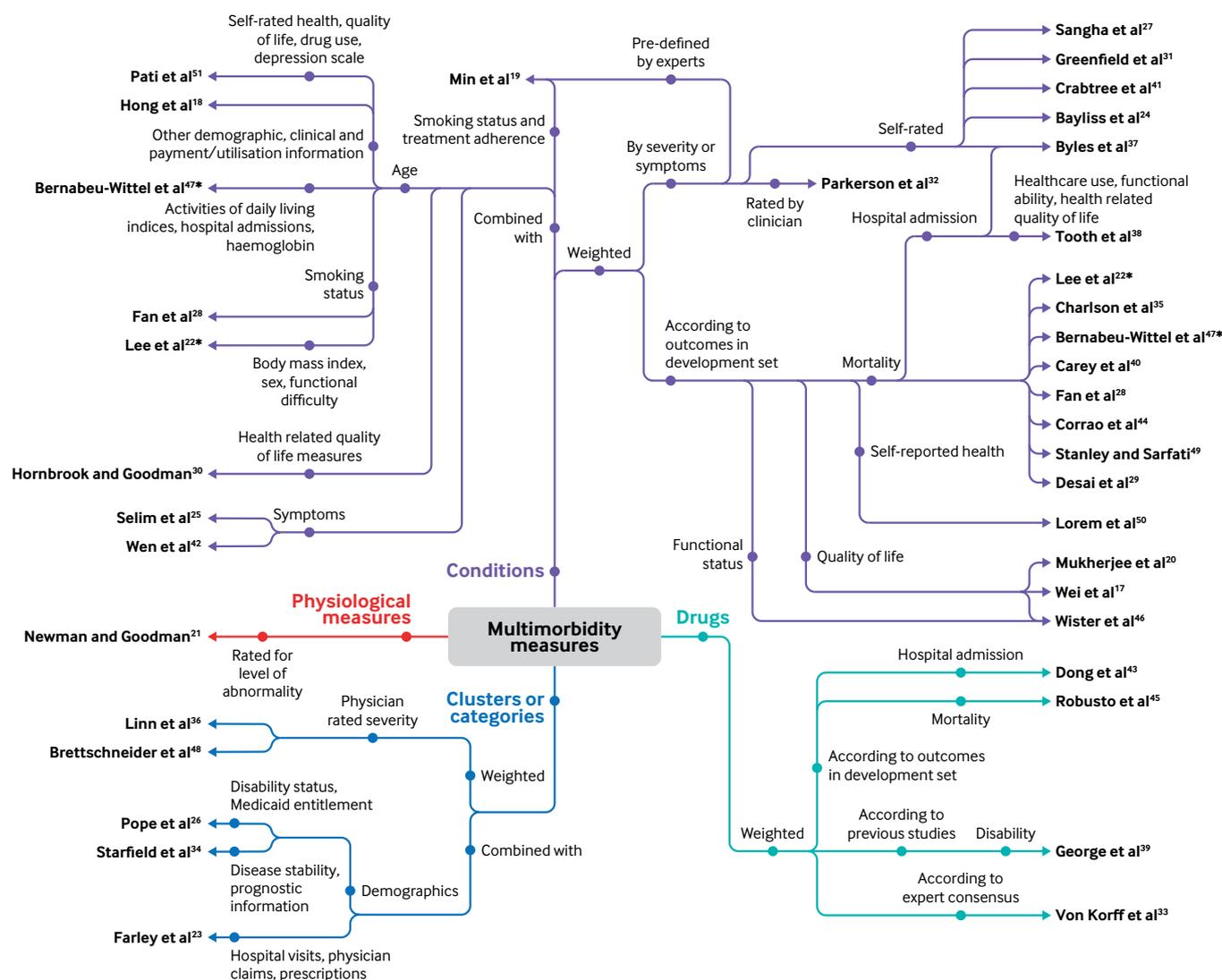Some of the indices, although designed to measure multimorbidity, were developed in cohorts of people

Fig 2 | Spider diagram summarising index components. *Paper appears under more than one category

with a specific disease, and therefore this condition was not included in the list of comorbidities.[23 31] Most indices were based on diagnoses or drugs, but two indices required results of laboratory or other investigations.[21 47]

The provenance of papers will affect their applicability to other settings. For example, the majority of papers came from the US where the predominant health system is commercial and healthcare costs are of interest to insurers. Some of the indices were designed with a particular population in mind, such as the questionnaire for Indian primary care. This included diseases that are less prevalent in other geographical areas, such as filariasis and tuberculosis, which might limit generalisability.[51] Other original indices have domains that might be outdated. One example is the Charlson comorbidity index.[35] This index assigns the maximum weight of 6 points to AIDS, but the life expectancy for HIV/AIDS in high income countries has changed considerably since the index's publication in 1987.[53]

Although several papers mentioned outcomes as relevant to patient experience, only one clearly described involving patients in the study design, by developing their rating scale with focus groups.[37]

Summary of evidence

Appendix eTables 3 to 6 summarise all included papers according to their index's original purpose, components and outcome variables, and information used. The data source, location, and demographics of the population used to derive or test the measure are listed as they are relevant for context. Our overall recommendations are also included. In appendix eTable 3, the index components are weighted condition counts, in appendix eTable 4, the index components are conditions with additional parameters, in appendix eTable 5 they are weighted drug counts, and appendix eTable 6 comprises the remainder, including diagnostic groups and physiological measures.

Weighting

The majority of indices (n=29, 83%) included some form of component weighting. Conditions, diagnostic categories, and drugs were weighted by severity or symptoms, either self-reported or defined by clinicians, or according to their associated outcomes in a derivation cohort. Different methods were used for developing each index, and disparities existed in the level of methodological detail reported. The appendix

eTable 7 summarises the method for developing each model, the details provided, and baseline outcome reporting.

Inclusion of mental health

Twenty nine (83%) of the papers contained some measure of mental health or dementia, with 18 of these including mental health markers exclusively as comorbidities (including psychotropic drugs when

Table 1 | Studies classified by components and outcomes of original versions of indices

Original outcomes	Conditions		Drugs (all weighted)	Categories or clusters		Physiological measures
	Weighted	With additional parameters		Weighted	With additional parameters	
Mortality	Corrao et al[44]	Wen et al[42]	Robusto et al[45]	Linn* et al[36]		Newman* et al[21]
	Stanley and Sarfati[49]	Min et al[19]	George et al[39]			
	Min et al[19]	Bernabeu-Wittel et al[47]	Von Korff* et al[33]			
	Carey et al[40]	Lee* et al[22]				
	Bernabeu-Wittel et al[47]	Selim et al[25]				
	Tooth et al[38]	Fan et al[8]				
	Lee* et al[22]					
	Byles et al[37]					
	Fan et al[28]					
	Desai* et al[29]					
	Charlson* et al[35]					
Healthcare use and costs	Corrao et al[44]	Hong et al[18]	Von Korff* et al[33]	Pope* et al[26]	Farley et al[23]	
	Wister et al[46]	Min et al[19]		Starfield* et al[34]		
	Min et al[19]	Selim et al[25]				
	Tooth et al[38]	Hornbrook and Goodman[30]				
	Sangha* et al[27]					
	Greenfield* et al[31]					
Hospital admission	Corrao et al[44]	Wen et al[42]	Robusto et al[45]		Farley et al[23]	
	Stanley and Sarfati[49]	Hong et al[18]	Dong et al[43]			
	Byles et al[37]	Fan et al[28]	George et al[39]			
	Fan et al[28]		Von Korff* et al[33]			
	Greenfield* et al[31]					
Independence or disability	Min et al[19]	Min et al[19]	Von Korff* et al[33]			Newman* et al[21]
	Tooth et al[38]					
	Bayliss* et al[24]					
	Crabtree et al[41]					
	Greenfield* et al[31]					
Self-rated overall health	Lorem et al[50]		Von Korff* et al[33]			
	Wister et al[46]					
	Bayliss* et al[24]					
	Crabtree et al[41]					
Health related quality of life or life satisfaction	Wister et al[46]	Selim et al[25]		Brettschneider et al[48]		
	Mukherjee* et al[20]					
	Tooth et al[38]					
	Byles et al[37]					
	Sangha* et al[27]					
Drug use	Wister et al[46]	Min et al[19]			Farley et al[23]	
	Min et al[19]					
Physical functioning	Bayliss* et al[24]					Newman* et al[21]
	Greenfield* et al[31]					
Mental health, depression or anxiety	Bayliss* et al[24]		Von Korff* et al[33]			
	Crabtree et al[41]					
	Greenfield* et al[31]					
Specific physiological measures		Hong et al[18]				
Quality indicators or adherence to screening	Min et al[19]	Hong et al[18]				
		Min[19]				
No outcomes measured	Wei et al[17]	Pati et al[51]				
	Pati et al[51]					
	Parkerson* et al[32]					

*Studies have been subsequently updated or adapted.

relevant)[17-19 22 26 30 32 34-36 39 40 42 43 45 47 49 51] and three including mental health markers as an outcome only.[24 31 33] Seven measures included different aspects of mental health as both comorbidities and outcomes,[20 25 27 37 38 44 48] and one paper included anxiety and depression as both a comorbidity and an outcome.[41] In appendix eTable 8, we summarise whether each index dealt with aspects of mental health, as either comorbidities or outcomes, and how these were measured. Where papers discussed specific findings relating to multimorbidity and mental health, we present their conclusions.

Risk of reporting bias within studies
Using our quality assessment tool, we classified six papers as high quality with little or no risk of bias in reporting[22 24 27 32 37 49] and seven as low quality with moderate to high risk of bias.[23 28 34 35 36 41 42] The remaining 22 papers were of satisfactory quality. Of the five domains we assessed, the best reported were index description and funding source. Validity and statistical methods were the least well reported across all papers. Table 2 shows the scores for all papers across each domain. As we had agreed in advance to judge the overall impression without summing domain scores, the domain scores did not always lead to the same overall impression. For example, one study was given an overall impression of satisfactory with a total domain score of 6,[48] whereas another study scored 8 and was also deemed satisfactory.[51]

Risk of bias across studies
It was not possible to quantify publication bias owing to the variety of methods and outcomes used. It is likely that more unpublished methods of measuring multimorbidity exist and are used in clinical practice, especially tailored to specific patient populations or available clinical information.

Usage, performance, and validation
As a proxy for usage, we calculated the number of annual citations for each paper. The number of citations for each year since publication ranged from three[41] to 949,[35] with a median of 8.8 (interquartile range 5.3-16.2). This information is listed alongside measures of the indices' performance at predicting outcomes and validation in appendix eTables 9 (indices without external validation) and 10 (externally validated indices). Sixteen original papers described designing indices within a derivation cohort and testing their ability to predict specific outcomes within a separate validation set.[20 22 28-30 33 35 37 38 40 43 44 45 47 49 50]

Fourteen original papers measured an index's performance at predicting outcomes[17 22 28 30-32 34 36-38 41 42 46 48] and 20 compared an index to an existing measure of multimorbidity.[18-21 23-27 29 33 35 39 40 43 44 45 47 49 50] Fourteen indices were validated elsewhere, of which 11 were compared with other indices[20 22 26 27 33-36 39 43 47] and three only measured ability to predict outcomes.[17 24 32] Among the indices that were externally validated, 11 were tested at predicting

different or additional outcomes to those in the original index design.[20 24 26 32-36 39 43 47] Some indices were tested against other indices that had been developed with different original outcomes—for example, the Charlson index where the outcome in question was admission to hospital[43] or health related quality of life.[20]

Updates and adaptations
Thirteen (37%) of the indices had updates or adaptations published, by either original or separate research teams. These revised versions included updated comorbidities or weights, focused on specific patient groups, or mapped a clinical index to codes for administrative data. Two of the indices are risk adjustment methods undergoing regular review and updates.[26 34] The relevant indices are listed in appendix eTable 11 alongside a summary of their adaptations and updates and any reported performance metrics. The older and widely used indices such as Charlson, Chronic Disease Score, and Cumulative Illness Rating Scale have been adapted and updated many times; we include the most cited versions. Most updated indices are broadly based on the aim and outcome measures of the original, with some exceptions.[57-60]

Of the indices that were not updated, in some cases this was because the original index was unsuccessful at predicting specific outcomes[37] or was not designed for use outside of the original study.[30]

Discussion
In this review we collated descriptions of 35 distinct multimorbidity indices. The papers were diverse in study design, intended purpose, and variables included. Similarities did, however, emerge, such as index components concentrating on conditions, diagnostic categories, drug classes, or physiological measures. Mortality was the most commonly studied outcome, with healthcare use, hospital admission, functional ability, and health related quality of life as other important groups. Those that measure mortality will be of most relevance to clinicians and researchers, whereas healthcare use and costs are more useful to healthcare providers and funders, particularly in predominantly private healthcare systems. For patients, the most relevant outcomes might be quality of life or self-reported health.

Strengths and weaknesses of this study
A major strength of this review was the use of an updated search in a rapidly expanding area of research and a focus on multimorbidity measures that specifically include mental health.

Although the medical subject heading (MeSH) term "comorbidity" has existed since 1990, a new MeSH term, "multimorbidity," was introduced in January 2018, after our search had been designed and pre-registered.[1 61] We found that the word "indexes" was used in some titles when we had used "indices" in our search terms. One paper was found by citation tracking and had apparently been missed during our search because we had omitted the term "score."[44]

Table 2 | Overall risk of reporting bias: domain scores

Reference	Sample selection (maximum++)	Index description (maximum ++)	Statistical methods (maximum ++)	Validity (maximum +++)	Funding source (maximum ++)	Overall quality
Corrao et al[44]	-	++	+	+	++	Satisfactory
Wen et al[42]	++	++	+	-	++	Low
Stanley and Sarfati[49]	++	++	+	+	++	High
Wei et al[17]	+	++	+	+	++	Satisfactory
Robusto et al[45]	+	++	+	+	++	Satisfactory
Lorem et al[50]	+	++	++	+	++	Satisfactory
Pati et al[51]	++	++	+	++	++	Satisfactory
Hong et al[18]	++	++	+	+	++	Satisfactory
Wister et al[46]	++	+	++	-	++	Satisfactory
Brettschneider et al[48]	++	++	+	-	++	Satisfactory
Min et al[19]	-	++	+	++	+	Satisfactory
Carey et al[40]	+	++	+	+	++	Satisfactory
Dong et al[43]	+	++	+	+	++	Satisfactory
Bernabeu-Wittel et al[47]	++	++	+	+	++	Satisfactory
Mukherjee et al[20]	++	++	+	+	++	Satisfactory
Newman and Goodman[21]	++	++	+	+	++	Satisfactory
Tooth et al[38]	+	++	+	+	++	Satisfactory
George et al[39]	-	++	+	++	+	Satisfactory
Lee et al[22]	++	++	+	+	++	High
Farley et al[23]	-	-	+	+	+	Low
Byles et al[37]	++	++	+	+	-	High
Bayliss et al[24]	++	++	++	+	++	High
Selim et al[25]	+	++	+	+	++	Satisfactory
Pope et al[26]	++	++	+	+	++	Satisfactory
Sangha et al[27]	++	++	+	+++	++	High
Fan et al[28]	+	+	+	+	++	Low
Desai et al[29]	++	++	+	+	++	Satisfactory
Crabtree et al[41]	-	++	+	+	++	Low
Hornbrook and Goodman[30]	+	+	+	+	++	Satisfactory
Greenfield et al[31]	++	++	+	+	++	Satisfactory
Parkerson et al[32]	++	++	+	++	+	High
Von Korff et al[33]	-	+	+	++	++	Satisfactory
Starfield et al[34]	-	++	-	+	++	Low
Charlson et al[35]	-	++	+	+	-	Low
Linn et al[36]	-	+	+	++	-	Low

However, we aimed to minimise the number of missed relevant papers by searching bibliographies, citations, and indices that had been mentioned in the abstracts we excluded. A limitation of this review is that we limited our search to full texts in the medical literature, excluding conference abstracts and other grey literature. This approach might have missed indices in clinical use that are either unpublished or based on guidelines from healthcare quality institutions.

In this review, we excluded papers that used simple counts of conditions to measure multimorbidity, focusing instead on indices. We chose to make this distinction because indices tend to include multiple parameters to quantify different aspects of multimorbidity and use sophisticated models to predict outcomes. However, as disease counts are the most commonly used method of measuring multimorbidity, their exclusion is a limitation of our work.[4] The ease of applying disease counts means they are frequently used and they are comparable between studies as long as the list of candidate conditions is clear.[62] One paper in this review reported that a count of diagnosis clusters was a better predictor of healthcare expenditure than more complex indices.[23] Other research has drawn similar conclusions.[63 64] Reviewing the use of disease counts is

outside the scope of this paper and has been discussed elsewhere.[65-67] Simple counts of drugs have also been shown to predict healthcare use and mortality,[68] and using these or disease counts are more practical than indices in many settings. For example, they are used in clinical care, as they do not require calculations or particular software, or in large population studies where the impact of each condition on individuals is unknown.[6]

As this review is aimed at those undertaking population or community research, we also excluded studies of people who had been admitted to hospital or live in residential care. This meant that several commonly used indices were not included in this review, such as the Elixhauser index.[69] We did include the Charlson, PROFUND, Self-Administered Comorbidity Questionnaire, and High-Risk Diagnoses for the Elderly Scale indices as although the studies recruited hospital patients, they did so when the patients were discharged, and the main period of interest was later, in the community setting.[27 29 35 47]

Some of the indices have been in use for many years and have several adaptations. The Charlson index is the most widely known and warranted its own systematic review within a medical specialty (critical care)[70]; we

> **Box 1: Guide to selecting an index using this paper when designing a multimorbidity study**
> - Identify reasons for including multimorbidity—for example, is multimorbidity important because of its association with quality of life or mortality? This will inform which outcomes of original indices are relevant
> - Identify the exposure variables available (eg, diagnoses, drugs)
> - Identify the outcomes to be measured
> - Use figure 3 to choose a recommended index
> - Use the information in the appendix (eTables 9 and 10) to compare usage and performance of any suitable indices

have only briefly summarised its performance and the principal adaptations. More information is available in another systematic review on this topic.[71]

Some of the indices were specifically designed for use with administrative data. These might have scored more poorly on quality assessment as our tool focused on reporting and clinical applicability. Our search also included papers that compared different measurements of multimorbidity but that were not intended for clinical use.[46] We included these papers for completeness when they met our inclusion criteria. We aimed to find measures of multimorbidity, and our exclusion criteria included studies of comorbidity with a specific index disease. However, in two of the papers studied the indices had been developed in populations with one condition only, either hypertension[23] or type 2 diabetes.[31] We referred to our protocol and included these studies because their aim was to study multimorbidity rather than comorbidities of those conditions as index diseases. These papers are, however, less generalisable to the general population.

Our search was intentionally broad, using a wide range of search terms in multiple medical research databases. We included a variety of studies measuring multimorbidity from different perspectives, which is a strength over previous more specific reviews.[71]

We generated overall impressions of the risk of reporting bias and recommendations for index use,

to provide a guide for researchers when choosing an index. Samples in the included studies were, however, heterogeneous, and the indices had varied purposes and components. Therefore, our recommendations are subjective. Formally comparing the predictive ability of the indices is outside the scope of this work but has been comprehensively performed by other investigators.[68 72-74]

Comparison with previous literature
The two most recent similar systematic reviews to ours did not formally assess the quality of publications.[34] Fifteen (43%) of the indices we included were published after 2009 and therefore would not have been found by the searches in these previous reviews. This is out of proportion to the increase in multimorbidity publications since 2010, suggesting that many recent papers have used either older measures or disease counts.[2] A newer systematic review on this topic focused only on tools used on administrative data and conducted its search in 2012.[71] A systematic review of multimorbidity systematic reviews, published in 2018, focused on definitions and measurement.[5] However, it did not include a search for new indices, thereby also omitting the 15 papers published since 2010.

Recommendations for index selection
We suggest that to select an index for clinical or research use, clinicians or researchers should first consider their desired outcomes and the information available. Box 1 summarises the process of selecting an index using this review.

Appendix eTables 3-6 list our overall recommendations, divided broadly into "not recommended," "potentially useful," and "recommended," and figure 3 displays the 12 indices that we recommend according to their design. The 10 indices that are not recommended could be useful for other purposes, such as recording symptoms[41] or comparing models,[37] but our recommendations focus on those that are practical for designing multimorbidity research.
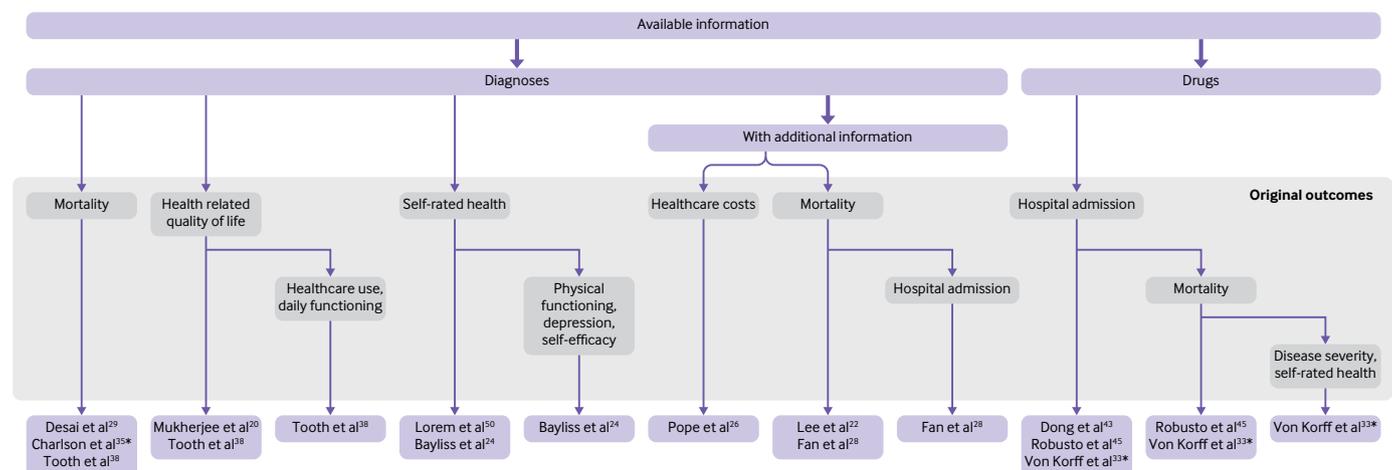

Fig 3 | Flowchart of recommended indices organised by components and original outcomes

Conclusions

At least 35 objective measures of multimorbidity are available for people living in the community, and each of these uses different variables to generate a score or index, linked with various or no outcome measures. We found no specific index for investigating the interplay between mental and physical multimorbidity, although this issue is dealt with in a variety of ways across the measures. The array of index components and outcomes means that a validated measure exists for many applications, including clinical, research, and cost prediction. It is important when choosing an index to consider its original purpose and the outcomes for which it is validated. Given the differing methodologies of multimorbidity research, it would not be appropriate to assume that a single index could definitively measure multimorbidity in all settings. However, with this research area at risk of saturation, we propose that anyone measuring multimorbidity should study existing indices before developing new ones.

AUTHOR AFFILIATIONS

[1]Edinburgh Dementia Prevention, Centre for Clinical Brain Sciences, University of Edinburgh, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh EH10 5HF, UK

[2]Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

[3]NHS Lothian, St John's Hospital, Howden, Livingston, UK

[4]University of Malawi College of Medicine, Blantyre, Malawi

[5]Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK

[6]NHS Lothian, Royal Edinburgh Hospital, Edinburgh, UK

1 The Academy of Medical Sciences. Multimorbidity: a priority for global health research. 2018. https://acmedsci.ac.uk/file-download/39787360
2 Xu X, Mishra GD, Jones M. Mapping the global research landscape and knowledge gaps on multimorbidity: a bibliometric study. *J Glob Health* 2017;7:010414. doi:10.7189/jogh.07.010414
3 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. doi:10.1093/gerona/glq208
4 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. doi:10.1370/afm.1363
5 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;29:182-9. doi:10.1093/eurpub/cky098
6 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. doi:10.1016/S0140-6736(12)60240-2
7 Almirall J, Fortin M. The coexistence of terms to describe the presence of multiple concurrent diseases. *J Comorb* 2013;3:4-9. doi:10.15256/joc.2013.3.22
8 Covidence. Covidence Systematic Review Software. Covidence. 2016. https://www.covidence.org/
9 Higgins JP, Altman DG, Sterne JA. Assessing risk of bias in included studies, Chapter 8 in Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. www.handbook.cochrane.org
10 Wolff RF, Moons KGM, Riley RD, et al, PROBAST Group†. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 2019;170:51-8. doi:10.7326/M18-1376
11 Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014;11:e1001744. doi:10.1371/journal.pmed.1001744
12 Cook C, Brismée J-M, Pietrobon R, Sizer PJr, Hegedus E, Riddle DL. Development of a Quality Checklist Using Delphi Methods for Prescriptive Clinical Prediction Rules: the QUADCPR. *J Manipulative Physiol Ther* 2010;33:29-41. doi:10.1016/j.jmpt.2009.11.010
13 Critical Appraisal Skills Programme. CASP Clinical Prediction Rule Checklist. 2013. http://www.casp-uk.net/checklists
14 Scottish Intercollegiate Guidelines Network (SIGN). Methodology Checklist 3: Cohort Studies, Version 3.0. 2012. https://www.sign.ac.uk/checklists-and-notes.html
15 Clarivate Analytics. EndNote X8. 2016. https://endnote.com/
16 Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097. doi:10.1371/journal.pmed.1000097
17 Wei MY, Kawachi I, Okereke OI, Mukamal KJ. Diverse Cumulative Impact of Chronic Diseases on Physical Health-Related Quality of Life: Implications for a Measure of Multimorbidity. *Am J Epidemiol* 2016;184:357-65. doi:10.1093/aje/kwv456
18 Hong CS, Atlas SJ, Ashburner JM, et al. Evaluating a Model to Predict Primary Care Physician-Defined Complexity in a Large Academic Primary Care Practice-Based Research Network. *J Gen Intern Med* 2015;30:1741-7. doi:10.1007/s11606-015-3357-8
19 Min L, Wenger N, Walling AM, et al. When comorbidity, aging, and complexity of primary care meet: development and validation of the Geriatric CompleXity of Care Index. *J Am Geriatr Soc* 2013;61:542-50. doi:10.1111/jgs.12160
20 Mukherjee B, Ou H-T, Wang F, Erickson SR. A new comorbidity index: the health-related quality of life comorbidity index. *J Clin Epidemiol* 2011;64:309-19. doi:10.1016/j.jclinepi.2010.01.025
21 Newman AB, Boudreau RM, Naydeck BL, Fried LF, Harris TB. A physiologic index of comorbidity: relationship to mortality and disability. *J Gerontol A Biol Sci Med Sci* 2008;63:603-9. doi:10.1093/gerona/63.6.603

22 Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA* 2006;295:801-8. doi:10.1001/jama.295.7.801

23 Farley JF, Harley CR, Devine JW. A comparison of comorbidity measurements to predict healthcare expenditures. *Am J Manag Care* 2006;12:110-9.

24 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. doi:10.1186/1477-7525-3-51

25 Selim AJ, Fincke G, Ren XS, et al. Comorbidity assessments based on patient report: results from the Veterans Health Study. *J Ambul Care Manage* 2004;27:281-95. doi:10.1097/00004479-200407000-00011

26 Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev* 2004;25:119-41.

27 Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis Rheum* 2003;49:156-63. doi:10.1002/art.10993

28 Fan VS, Au D, Heagerty P, Deyo RA, McDonell MB, Fihn SD. Validation of case-mix measures derived from self-reports of diagnoses and health. *J Clin Epidemiol* 2002;55:371-80. doi:10.1016/S0895-4356(01)00493-0

29 Desai MM, Bogardus STJr, Williams CS, Vitagliano G, Inouye SK. Development and validation of a risk-adjustment index for older patients: the high-risk diagnoses for the elderly scale. *J Am Geriatr Soc* 2002;50:474-81. doi:10.1046/j.1532-5415.2002.50113.x

30 Hornbrook MC, Goodman MJ. Chronic disease, functional health status, and demographics: a multi-dimensional approach to risk adjustment. *Health Serv Res* 1996;31:283-307.

31 Greenfield S, Sullivan L, Dukes KA, Silliman R, D'Agostino R, Kaplan SH. Development and testing of a new measure of case mix for use in office practice. *Med Care* 1995;33(Suppl):AS47-55.

32 Parkerson GRJr, Broadhead WE, Tse C-KKJ. The Duke Severity of Illness Checklist (DUSOI) for measurement of severity and comorbidity. *J Clin Epidemiol* 1993;46:379-93. doi:10.1016/0895-4356(93)90153-R

33 Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 1992;45:197-203. doi:10.1016/0895-4356(92)90016-G

34 Starfield B, Weiner J, Mumford L, Steinwachs D. Ambulatory care groups: a categorization of diagnoses for research and management. *Health Serv Res* 1991;26:53-74.

35 Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373-83. doi:10.1016/0021-9681(87)90171-8

36 Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc* 1968;16:622-6. doi:10.1111/j.1532-5415.1968.tb02103.x

37 Byles JE, D'Este C, Parkinson L, O'Connell R, Treloar C. Single index of multimorbidity did not predict multiple outcomes. *J Clin Epidemiol* 2005;58:997-1005. doi:10.1016/j.jclinepi.2005.02.025

38 Tooth L, Hockey R, Byles J, Dobson A. Weighted multimorbidity indexes predicted mortality, health service use, and health-related quality of life in older women. *J Clin Epidemiol* 2008;61:151-9. doi:10.1016/j.jclinepi.2007.05.015

39 George J, Vuong T, Bailey MJ, Kong DC, Marriott JL, Stewart K. Development and validation of the medication-based disease burden index. *Ann Pharmacother* 2006;40:645-50. doi:10.1345/aph.1G204

40 Carey IM, Shah SM, Harris T, DeWilde S, Cook DG. A new simple primary care morbidity score predicted mortality and better explains between practice variations than the Charlson index. *J Clin Epidemiol* 2013;66:436-44. doi:10.1016/j.jclinepi.2012.10.012

41 Crabtree HL, Gray CS, Hildreth AJ, O'Connell JE, Brown J. The Comorbidity Symptom Scale: a combined disease inventory and assessment of symptom severity. *J Am Geriatr Soc* 2000;48:1674-8. doi:10.1111/j.1532-5415.2000.tb03882.x

42 Wen YC, Chen LK, Hsiao FY. Predicting mortality and hospitalization of older adults by the multimorbidity frailty index. *PLoS One* 2017;12:e0187825. doi:10.1371/journal.pone.0187825

43 Dong YH, Chang CH, Shau WY, Kuo RN, Lai MS, Chan KA. Development and validation of a pharmacy-based comorbidity measure in a population-based automated health care database. *Pharmacotherapy* 2013;33:126-36. doi:10.1002/phar.1176

44 Corrao G, Rea F, Di Martino M, et al. Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. *BMJ Open* 2017;7:e019503. doi:10.1136/bmjopen-2017-019503

45 Robusto F, Lepore V, D'Ettorre A, et al. The Drug Derived Complexity Index (DDCI) Predicts Mortality, Unplanned Hospitalization and Hospital Readmissions at the Population Level. *PLoS One* 2016;11:e0149203. doi:10.1371/journal.pone.0149203

46 Wister AV, Levasseur M, Griffith LE, Fyffe I. Estimating multiple morbidity disease burden among older persons: a convergent construct validity study to discriminate among six chronic illness measures, CCHS 2008/09. *BMC Geriatr* 2015;15:12. doi:10.1186/s12877-015-0001-8

47 Bernabeu-Wittel M, Ollero-Baturone M, Moreno-Gaviño L, et al. Development of a new predictive model for polypathological patients. The PROFUND index. *Eur J Intern Med* 2011;22:311-7. doi:10.1016/j.ejim.2010.11.012

48 Brettschneider C, Leicht H, Bickel H, et al, MultiCare Study Group. Relative impact of multimorbid chronic conditions on health-related quality of life--results from the MultiCare Cohort Study. *PLoS One* 2013;8:e66742. doi:10.1371/journal.pone.0066742

49 Stanley J, Sarfati D. The new measuring multimorbidity index predicted mortality better than Charlson and Elixhauser indices among the general population. *J Clin Epidemiol* 2017;92:99-110. doi:10.1016/j.jclinepi.2017.08.005

50 Lorem GF, Schirmer H, Emaus N. Health impact index. Development and validation of a method for classifying comorbid disease measured against self-reported health. *PLoS One* 2016;11:e0148830. doi:10.1371/journal.pone.0148830

51 Pati S, Hussain MA, Swain S, et al. Development and Validation of a Questionnaire to Assess Multimorbidity in Primary Care: An Indian Experience. *Biomed Res Int* 2016;2016:6582487. doi:10.1155/2016/6582487

52 The Johns Hopkins University. The Johns Hopkins ACG® System. 2019. https://www.hopkinsacg.org/ (accessed 28 Mar 2019).

53 Antiretroviral Therapy Cohort Collaboration. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *Lancet* 2008;372:293-9. doi:10.1016/S0140-6736(08)61113-7

54 Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol* 2011;173:676-82. doi:10.1093/aje/kwq433

55 Fishman PA, Goodman MJ, Hornbrook MC, Meenan RT, Bachman DJ, O'Keeffe Rosetti MC. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. *Med Care* 2003;41:84-99. doi:10.1097/00005650-200301000-00011

56 Sanders JL, Minster RL, Barmada MM, et al. Heritability of and mortality prediction with a longevity phenotype: the healthy aging index. *J Gerontol A Biol Sci Med Sci* 2014;69:479-85. doi:10.1093/gerona/glt117

57 Sridharan S, Berdeprado J, Vilar E, Roberts J, Farrington K. A self-report comorbidity questionnaire for haemodialysis patients. *BMC Nephrol* 2014;15:134. doi:10.1186/1471-2369-15-134

58 Parkerson GRJr, Michener JL, Yarnall KS, Hammond WE. Duke Case-Mix System (DUMIX) for ambulatory health care. *J Clin Epidemiol* 1997;50:1385-94. doi:10.1016/S0895-4356(97)00217-5

59 McGregor JC, Perencevich EN, Furuno JP, et al. Comorbidity risk-adjustment measures were developed and validated for studies of antibiotic-resistant infections. *J Clin Epidemiol* 2006;59:1266-73. doi:10.1016/j.jclinepi.2006.01.016

60 Mistry R, Gokhman I, Bastani R, et al, UPBEAT Collaborative Group. Measuring medical burden using CIRS in older veterans enrolled in UPBEAT, a psychogeriatric treatment program: a pilot study. *J Gerontol A Biol Sci Med Sci* 2004;59:1068-75. doi:10.1093/gerona/59.10.M1068

61 National Center for Biotechnology Information. U.S. National Library of Medicine. Multimorbidity: MeSH Unique ID D000076322. 2018. https://www.ncbi.nlm.nih.gov/mesh/?term=multimorbidity (accessed 20 Jul 2018).

62 Nicholson K, Almirall J, Fortin M. The measurement of multimorbidity. *Health Psychol* 2019;38:783-90. doi:10.1037/hea0000739

63 Wallace E, McDowell R, Bennett K, Fahey T, Smith SM. Comparison of count-based multimorbidity measures in predicting emergency admission and functional decline in older community-dwelling adults: a prospective cohort study. *BMJ Open* 2016;6:e013089. doi:10.1136/bmjopen-2016-013089

64 Perkins AJ, Kroenke K, Unützer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. *J Clin Epidemiol* 2004;57:1040-8. doi:10.1016/j.jclinepi.2004.03.002

65 Willadsen TG, Bebe A, Køster-Rasmussen R, et al. The role of diseases, risk factors and symptoms in the definition of multimorbidity - a systematic review. *Scand J Prim Health Care* 2016;34:112-21. doi:10.3109/02813432.2016.1153242

66 Holzer BM, Siebenhuener K, Bopp M, Minder CE. Evidence-based design recommendations for prevalence studies on multimorbidity: improving comparability of estimates. *Popul Health Metr* 2017;15:9. doi:10.1186/s12963-017-0126-4

67 Fortin M, Stewart M, Poitras M-E, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward

a more uniform methodology. *Ann Fam Med* 2012;10:142-51. doi:10.1370/afm.1337

68 Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. *Fam Pract* 2013;30:172-8. doi:10.1093/fampra/cms060

69 Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27. doi:10.1097/00005650-199801000-00004

70 Needham DM, Scales DC, Laupacis A, Pronovost PJ. A systematic review of the Charlson comorbidity index using Canadian administrative databases: a perspective on risk adjustment in critical care research. *J Crit Care* 2005;20:12-9. doi:10.1016/j.jcrc.2004.09.007

71 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. doi:10.1016/j.jclinepi.2014.09.010

72 Mehta HB, Sura SD, Sharma M, Johnson ML, Riall TS. Comparative performance of diagnosis-based and prescription-based comorbidity scores to predict health-related quality of life. *Med Care* 2016;54:519-27. doi:10.1097/MLR.0000000000000517

73 Ng SK, Tawiah R, Sawyer M, Scuffham P. Patterns of multimorbid health conditions: a systematic review of analytical methods and comparison analysis. *Int J Epidemiol* 2018;47:1687-704. doi:10.1093/ije/dyy134

74 Sharabiani MTA, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care* 2012;50:1109-18. doi:10.1097/MLR.0b013e31825f64d0

Supplementary information: Additional methods and tables