5-5-2021

Systematic review on the instruments used for measuring the association of the level of multimorbidity and clinically important outcomes

Eng Sing Lee
National Healthcare Group Polyclinics, Singapore

Hui Li Koh
National Healthcare Group Polyclinics, Singapore

Elaine Qiao Ying Ho
National University of Singapore

Sok Huang Teo
National Healthcare Group Polyclinics, Singapore

Fang Yan Wong
National Healthcare Group Polyclinics, Singapore

See next page for additional authors

Citation of this paper:
Lee ES, Koh HL, Ho EQ, et al Systematic review on the instruments used for measuring the association of the level of multimorbidity and clinically important outcomes BMJ Open 2021;11:e041219. doi: 10.1136/bmjopen-2020-041219

BMJ Open

Systematic review on the instruments used for measuring the association of the level of multimorbidity and clinically important outcomes

Eng Sing Lee ,[1,2] Hui Li Koh,[1] Elaine Qiao-Ying Ho,[3] Sok Huang Teo,[1] Fang Yan Wong,[1] Bridget L Ryan,[4,5] Martin Fortin,[6] Moira Stewart[5]

To cite: Lee ES, Koh HL, Ho EQ-Y, *et al*. Systematic review on the instruments used for measuring the association of the level of multimorbidity and clinically important outcomes. *BMJ Open* 2021;**11**:e041219. doi:10.1136/bmjopen-2020-041219

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-041219).

Received 02 June 2020
Revised 13 April 2021
Accepted 19 April 2021

For numbered affiliations see end of article.

Correspondence to
Dr Eng Sing Lee;
Eng_Sing_LEE@nhgp.com.sg

ABSTRACT

Objectives There are multiple instruments for measuring multimorbidity. The main objective of this systematic review was to provide a list of instruments that are suitable for use in studies aiming to measure the association of a specific outcome with different levels of multimorbidity as the main independent variable in community-dwelling individuals. The secondary objective was to provide details of the requirements, strengths and limitations of these instruments, and the chosen outcomes.

Methods We conducted the review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (PROSPERO registration number: CRD42018105297). We searched MEDLINE, Embase and CINAHL electronic databases published in English and manually searched the *Journal of Comorbidity* between 1 January 2010 and 23 October 2020 inclusive. Studies also had to select adult patients from primary care or general population and had at least one specified outcome variable. Two authors screened the titles, abstracts and full texts independently. Disagreements were resolved with a third author. The modified Newcastle-Ottawa Scale was used for quality assessment.

Results Ninety-six studies were identified, with 69 of them rated to have a low risk of bias. In total, 33 unique instruments were described. Disease Count and weighted indices like Charlson Comorbidity Index were commonly used. Other approaches included pharmaceutical-based instruments. Disease Count was the common instrument used for measuring all three essential core outcomes of multimorbidity research: mortality, mental health and quality of life. There was a rise in the development of novel weighted indices by using prognostic models. The data obtained for measuring multimorbidity were from sources including medical records, patient self-reports and large administrative databases.

Conclusions We listed the details of 33 instruments for measuring the level of multimorbidity as a resource for investigators interested in the measurement of multimorbidity for its association with or prediction of a specific outcome.

Strengths and limitations of this study

- This review builds on Huntley *et al*'s 2012 review article and provides an updated, comprehensive list of instruments that measure levels of multimorbidity in community-dwelling individuals.
- A thorough literature search of three major electronic databases was conducted with the involvement of a health science librarian.
- The review is reported based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.
- This review excluded non-English language articles and grey literature.

BACKGROUND

Multimorbidity is defined as the co-occurrence of two or more chronic medical conditions in an individual.[1] It is a growing public health challenge and accounts for most of the expenditures in the healthcare system.[2] The complex interactions of several coexisting diseases have profound implications on individuals[3 4] and their healthcare providers.[5 6]

There are multiple instruments for measuring multimorbidity and many of them do not usually specify the severity of individual conditions.[7] No gold standard multimorbidity measurement instrument exists and there is also no agreed categorisation of the available instruments. Sarfati[8 9] classified the various measurement instruments into four broad approaches. They are as follows: (1) by simple counts of individual conditions (ie, Disease Count), (2) by organ or system-based approaches, (3) by weighting conditions and combining them into indices and (4) by other miscellaneous approaches. Most of these measurements are used to measure the prevalence or patterns of multimorbidity. However, they can also be used to predict an outcome or to evaluate an intervention for a desired outcome. A set of core outcomes of multimorbidity (COSmm) was proposed

after consulting a panel of international experts in multimorbidity intervention studies using a Delphi process.[10] Core outcome sets represent the minimum that should be measured and reported in all clinical trials of multimorbidity.[11]

Huntley *et al*[12] published a systematic review in 2012 describing the instruments used to measure the morbidity burden in primary care and the general population. They found 17 different instruments from 194 articles. The most widely used instruments and those with the most significant evidence of validity were the Charlson Comorbidity Index (CCI), Disease Count and the Adjusted Clinical Groups (ACG) system.[12] However, this review was conducted in 2009 and multimorbidity research has increased exponentially since then.

The present review was to build on the review article by Huntley *et al*[12] in order to provide a current and comprehensive list of instruments that measure levels of multimorbidity for community-dwelling individuals. We used the term 'level of multimorbidity' to refer to the combined effects of multiple conditions on an individual. The main objective of this review was to list instruments for measuring the levels of multimorbidity. We specifically look for studies that measure the association of a clinically important outcome with different levels of multimorbidity as the main independent variable in community-dwelling individuals. Our second objective was to provide details of the requirements, strengths and limitations of these instruments, and the chosen outcomes in the studies so that clinicians and researchers can select or develop instruments that match their needs for predicting a specific outcome.

METHODS

A protocol for this systematic review (CRD42018105297) was published online on PROSPERO.[13] We searched MEDLINE, EMBASE, CINAHL and also manually searched the *Journal of Comorbidity* for potential studies. The medical subject headings and keywords used for the search are shown in online supplemental appendix 1.

We selected studies that included (1) adult patients from primary care or the general population as the majority of patients with multimorbidity are managed by primary care physicians[14]; (2) at least one specified outcome variable; and (3) published full-text articles from 1 January 2010 to 23 October 2020. Studies were excluded if they (1) selected patients from the hospital or nursing home only or patient data were drawn solely from the hospital or the nursing home; or (2) selected patients with an index condition; or (3) used level of multimorbidity as a covariate and not the main independent variable; or (4) were not written in English. We did not include a specific definition of multimorbidity instrument because, given a lack of consensus in the literature on the use of this term, we wanted to include a diverse range of studies on the above topic.

One reviewer (ESL) conducted a preliminary screen of titles and abstracts to exclude articles that were irrelevant. Abstracts of the remaining articles were screened independently by two reviewers (ESL and EQ-YH) according to the eligibility criteria. Disagreements were resolved through discussion until a consensus was reached. The full-text articles were then retrieved for the agreed list and independently assessed according to the eligibility criteria by the same reviewers. Disagreements were resolved through discussion with a third reviewer (TSH) until a consensus was reached. After agreement on the list of articles, the reference lists of included articles were hand-searched for additional eligible articles. We reported multimorbidity instruments that were described in all selected articles.

The risk of bias of the study design of selected articles was next appraised independently by three reviewers (ESL, EH and TSH) using the modified Newcastle-Ottawa Scale (NOS).[15 16] Each article was assessed under the three broad categories: (1) selection, (2) comparability and (3) outcome (online supplemental appendices 2 and 3).

We contacted the authors, as needed, for additional information or clarification up to three times spaced 1 week apart. We contacted 25 authors and 19 of them replied. Any disagreements on the risk of bias were resolved among the three reviewers through regular meetings. HLK and FYW were responsible for tracking and updating the final outcome of the risk of bias assessment.

Patient and public involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or to interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

RESULTS

The number of included studies was 96, of which 69 were assessed to have low risk of bias. A summary of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart is depicted in figure 1. Forty-eight studies selected participants from the general population and the other 48 studies selected participants from primary care. Most of the studies in this systematic review were from Europe and North America with very few Asian studies. There were 44 cohort studies, of which 36 were assessed to have low risk of bias, and 52 cross-sectional studies, of which 33 were assessed to have low risk of bias. We found 33 unique instruments from the 96 studies. The instruments were categorised according to Sarfati[8 9] into (1) simple counts of individual conditions; (2) organ or system-based approaches; (3) conditions that have been weighted and combined into indices; and (4) other approaches. A total of 150 outcomes were reported from all the studies. No studies were excluded for an outcome that was not deemed to be clinically important. Online

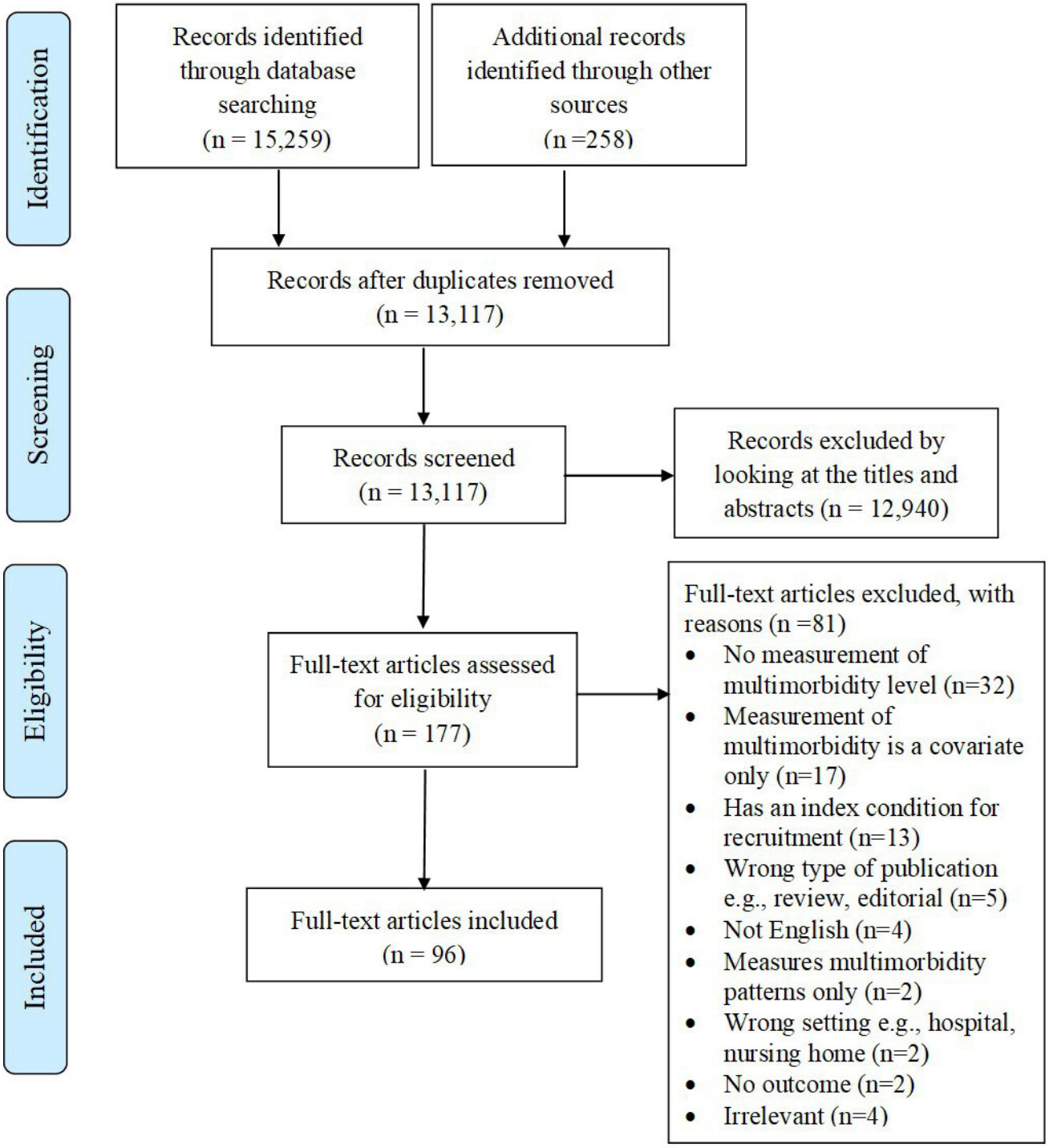

Figure 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

supplemental appendices 4 and 5 summarise the risk of bias assessment of each study. Table 1 provides a summary of the study design, population source, age group, multimorbidity measurements, outcome measures and risk of bias assessment of all the studies.

Table 2 summarises the 33 instruments that were identified from all the studies. Table 3 provides a summary of multimorbidity instruments and their associations with the outcomes measured from all the included studies.

Simple counts of individual conditions

Disease Count was based on the total number of all the conditions an individual had, usually from a prespecified list of chronic conditions. It was used in 59 out of the 96 studies (61.5%). Disease Count was reported to be associated with activity limitations, continuity of care, disability, healthcare cost, healthcare utilisation, medications, mental disorders, mortality, general health, physical function, quality of life and self-rated health (table 3).

Table 1 Summary of included studies

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Agborsangaya *et al* (2013)[41]	CS	GP	≥18	DC	HRQoL	Good
Bähler *et al* (2015)[42]	CS	GP	≥65	DC-ATC classification system	Total number of consultations	Good
Barile *et al* (2013)[43]	Cohort	GP	≥65	DC	ADL limitations, physically unhealthy days, mentally unhealthy days	Good
Barile *et al* (2012)[44]	CS	GP	≥65	DC	Physical HRQoL, mental HRQoL	Good
Barnett *et al* (2012)[45]	CS	PC	≥0	DC	Presence of mental health disorder	Good
Biehl *et al* (2016)[46]	Cohort	PC	≥65	ERA, CCI	Presence of critical illness	Good
Boeckxstaens *et al* (2015a)[17]	CS	PC	≥80	DC, CCI, CIRS	Disability (measured by ADL), frailty (five components)	Poor
Boeckxstaens *et al* (2015b)[18]	Cohort	PC	≥80	DC, mCCI, CIRS	Mortality at 3 years, hospitalisation at 3 years, functional decline at 19 months (ADL, physical, mental decline)	Fair
Brilleman *et al* (2014)[47]	Cohort	PC	≥18	QOF count, CCI, EDC count, ACG, RUB	Primary healthcare cost The EDC performed best followed by the QOF and ACG	Good
Brilleman and Salisbury (2013)[48]	Cohort	PC	≥18	QOF count, CCI, EDC count, ACG, RUB, prescribed drugs count	Mortality: The CCI was the best performing measure followed by the number of prescribed drugs. Number of primary care consultations (3-year period): The number of prescribed drugs had the greatest predictive validity followed by the ACG-based measures (ACG, EDC count and RUB).	Good
Caballer-Tarazona *et al* (2019)[49]	CS	GP	≥0	CRG	Expenditure of integrated healthcare (hospital, primary healthcare (PHC) and pharmaceutical prescription)	Poor

Continued

Table 1 Continued

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Carey *et al* (2013)[50]	Cohort	PC	≥60	Standard QOF, extended QOF, CCI (Khan)	Mortality (1-year period) The standard QOF score outperformed the CCI (Khan). The extended QOF score produced only a modest improvement in overall model performance.	Good
Chapman *et al* (2015)[51]	Cohort	GP	≥18	CCI, CCI-PSR	Mortality (5, 10, 15, 20, 25-year period) The CCI-PSR showed substantially better discrimination than the CCI.	Good
Charlson *et al* (2014)[20]	Cohort	GP	≥0	CCI	Healthcare cost, utilisation of services	Good
Chen *et al* (2011)[52]	CS	GP	≥18	DC	General health, mental distress, physical distress, activity limitations	Good
Chen *et al* (2018)[53]	CS	GP	≥45	DC	Health service utilisation	Poor
Chu *et al* (2018)[54]	CS	PC	≥40	DC, CIRS	Healthcare utilisation	Good
Clynes *et al* (2020)[55]	CS	GP	(Born in 1931–1939)	DC	Physical functioning	Poor
Crane *et al* (2010)[56]	Cohort	PC	≥60	ERA	Number of hospital visits, ED visits, hospital admissions, days hospitalised (1-year period)	Good
Crooks *et al* (2016)[57]	Cohort	PC	20–100	Comorbidity linked score, CCI, EI	Mortality (1-year period) The linked score had significantly improved discrimination and fit compared with the CCI and the Elixhauser Index	Good
Crooks *et al* (2015)[58]	Cohort	PC	≥20	CCI (Read), CCI (ICD-10), CCI (Read and ICD-10)	All-cause mortality (1–5 years) There was no large difference in the discrimination of the model for whichever codes that were used to derive the CCI.	Good
DiNapoli *et al* (2017)[19]	CS	PC	≥50	Organ systems with chronic disease	Presence of depressive or anxiety disorder	Good
Formiga *et al* (2013)[59]	Cohort	PC	85	CCI	Mortality (3-year period)	Good
Formiga *et al* (2011a)[60]	Cohort	GP	90 to 99	CCI	Mortality (5-year period)	Good

Continued

Table 1 Continued

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Formiga *et al* (2011b)[61]	CS	PC	85	CCI	Successful ageing	Good
Formiga *et al* (2016)[62]	Cohort	PC	85	CCI	Mortality (5-year period)	Good
Fraccaro *et al* (2016)[63]	Cohort	PC	≥18	CCI (Khan)	Mortality (1, 5, 10-year period), mortality (3, 6, 12-month period)	Good
Galenkamp *et al* (2011)[64]	CS	GP	57–98	DC	SRH	Good
Garin *et al* (2014)[65]	CS	GP	≥50	DC	QOL, disability	Good
Glynn *et al* (2011)[66]	CS	PC	>50	DC	Primary care consultations, hospital outpatient visits, hospital admissions, healthcare cost (all 1-year period)	Good
Gunn *et al* (2012)[67]	CS	PC	18–76	DC	Depressive symptoms (CES-D score)	Fair
Haas *et al* (2013)[21]	Cohort	PC	≥18	ACG, Minnesota Healthcare Home Tiering, HCC, ERA, CCC, CCI, hybrid model	Hospitalisation, ED visits, readmission within 30 days, healthcare expenditure (all 1-year period) The ACG model outperformed the other five models in all outcomes.	Good
Hanmer *et al* (2010)[68]	CS	GP	22 to 106	Additive model, minimum model, multiplicative model	Health utility (SF-6D)	Fair
Hu *et al* (2017)[69]	CS	PC	≥65	Age-adjusted CCI	Frequency of family physician visits	Fair
Hwang *et al* (2015)[23]	Cohort	GP	≥0	ACE-27, ACE-27 count	Healthcare expenditure The model, using year 1 data to determine if an individual would be classified into the persistent high-user group for the following 3 years, indicates a very high level of accuracy in predicting membership in a high-user group.	Good
Isaacs *et al* (2014)[70]	CS	PC	18–101	DC	Prescription costs	Poor
Jennings *et al* (2015)[71]	Cohort	PC	≥75	DC	Count of fall-related injuries in the 24 months after the date of screening	Fair
Jia *et al* (2018)[72]	Cohort	GP	≥65	DC	Quality-adjusted life years (QALY)	Poor
Jia and Lebetkin (2017)[73]	Cohort	GP	≥65	DC	Quality-adjusted life years (QALY)	Poor

Continued

Table 1 Continued

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Jindai *et al* (2016)[74]	CS	GP	≥65	DC	Functional limitations (ADL, IADL, leisure and social activities, lower-extremity mobility, general physical activities)	Good
Kim *et al* (2012)[75]	CS	GP	≥65	DC	Quality of life (EQ5D)	Poor
Kojima *et al* (2011)[76]	CS	PC	≥65	DC	Fall tendency	Poor
Kristensen *et al* (2014)[77]	CS	PC	>0	RUB	Fee-for-services expenditures	Good
Lapi *et al* (2015)[78]	CS	PC	≥15	HSMI	Total mean healthcare cost per year The HSMI explained 50.17% of the variation in costs	Good
Lawson *et al* (2013)[79]	CS	GP	≥20	DC	Preference-weighted HRQoL	Good
Lemke *et al* (2012)[80]	Cohort	GP	≥0	CCI, ACG	Inpatient hospitalisations ACG-based predictive model was superior to CCI model.	Good
Li *et al* (2016)[81]	CS	GP	16–68	DC	Health-related quality of life	Poor
Loprinzi *et al* (2016)[82]	CS	GP	60–85	DC	Cognitive function	Good
Macinko *et al* (2019)[83]	CS	GP	≥18	DC (categorical 2 and 3 or more) (self-reported)	Primary care experience (self-reported)	Good
Marengoni *et al* (2011)[84]	CS	GP	≥75 (baseline) ≥77 (follow-up)	DC	Disability	Good
McDaid *et al* (2013)[85]	CS	GP	≥50	DC	Disability, QoL, SRH	Good
Md Yusof *et al* (2010)[86]	Cohort	GP	64–85	CCI,	Mortality over 7 years	Fair
Milla-Perseguer *et al* (2019)[87]	CS	PC	≥18	CRG	Health-related quality of life (HRQL)—EQ-5D-3L	Good
Monterde *et al* (2020)[88]	Cohort	GP	≥18	Adjusted morbidity group (GMA), CCI, DC, CRG	Use of healthcare resources	Good
Muggah *et al* (2012)[89]	CS	GP	≥20	DC	Primary healthcare use	Poor
Mujica-Mota *et al* (2015)[90]	CS	PC	≥18	DC	Health-related quality of life (EQ5D)	Fair
Naessens *et al* (2011)[91]	CS	GP	18–64	DC	Healthcare cost	Poor
Østergaard and Foldager (2011)[92]	CS	PC	≥18	DC	Major depressive episode (measured by DSQ)	Poor

Continued

Table 1 Continued

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Palladino *et al* (2019)[93]	CS	GP	≥50	DC	Primary care use, reduced functional capacity, self-perceived health, hospital admissions, quality of life	Good
Pati *et al* (2019)[94]	CS	PC	≥18	Severity burden score (21 conditions)	Health-related quality of life (SF-12)	Good
Payne *et al* (2013)[95]	Cohort	PC	≥20	DC	Unplanned hospital admission, potentially preventable admission (all 1-year period)	Good
Payne *et al* (2014)[96]	Cohort	PC	≥20	DC	Unplanned hospital admissions (1-year period)	Good
Payne *et al* (2020)[97]	Cohort	PC	≥20	CCI, DC (37 read codes), Cambridge Multimorbidity Score	Mortality, unplanned inpatient hospital admission, primary care consultations	Good
Peters *et al* (2018)[98]	CS	PC	18–101	DC, DBIS	Quality of life	Fair
Quail *et al* (2011)[99]	Cohort	GP	≥20	DC, CCI (Quan), Elixhauser (Quan), number of different dispended drugs, CDS	Mortality (1-year period): Elixhauser (Quan) performed best followed by CCI. One or more hospitalisations; two or more hospitalisations: DC was the best performing measure	Good
Ranstad *et al* (2014)[100]	CS	GP	≥0	RUB	Registered active listing in primary care and all healthcare	Good
Reinke *et al* (2019)[101]	CS	PC	30–94	DC	Symptom burden (MSAS-SF), quality of life (Veterans RAND 12)	Good
Renne and Gobbens (2018)[102]	CS	PC	≥70	DC	Quality of life	Poor
Reyes *et al* (2014)[103]	Cohort	PC (men)	≥65	CCI	Hip fractures	Good
Ryu *et al* (2015)[104]	CS	PC	≥18	DC	Deficits of perceived general health, depressive symptoms	Good
Salisbury *et al* (2011)[105]	Cohort	PC	≥18	QOF count, EDC count	Primary care consultation rates, continuity of care (all 3-year period)	Good
Saver *et al* (2014)[106]	Cohort	GP	≥65	CCI (Romano)+Hypertension	Acute ACSH, chronic ACSH	Good

Continued

Table 1 Continued

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Shadmi *et al* (2011)[107]	CS	GP	≥18	ADG, CCI	Number of primary care physician visits, specialist visits, hospitalisation ADG explained the largest percent of variance or in healthcare resource use	Good
Sibley *et al* (2014)[108]	CS	GP	≥65	DC	Self-reported falls in the last 12 months	Poor
Stanley and Sarfati (2017)[109]	Cohort	PC	≥18	M3 Index, CCI, Elixhauser (van Walraven)	Mortality, overnight hospitalisation (all 1-year period) M3 Index outperformed both CCI and Elixhauser (van Walraven)	Good
St John *et al* (2014)[110]	Cohort	GP	≥65	DC (0–36 conditions)	Mortality in 5 years	Good
St John *et al* (2019)[111]	Cohort	GP	≥65	DC	Functional impairment in 5 years	Good
Streit *et al* (2014)[27]	Cohort	PC	50–80	CCI, DC	Quality of cardiovascular preventive care, quality of preventive care	Good
Sullivan *et al* (2012)[112]	CS	GP	≥18	DC	Preference-based HRQoL	Good
Takahashi *et al* (2011)[113]	Cohort	PC	>60	ERA	Mortality, nursing home placement (all 2-year period)	Good
Takahashi *et al* (2016)[114]	Cohort	PC	≥18	Minnesota Tiering (ACG), enhanced model	Hospitalisation/ED visits The enhanced model is better	Good
Tyack *et al* (2016)[115]	Cohort	PC	≥18	DC	Health-related quality of life	Fair
Ubalde-Lopez *et al* (2016)[24]	CS	GP	F (mean): 35.9, M (mean): 37.9	MDMS	Sickness absence episodes taken in last 2 years	Good
van den Bussche *et al* (2011)[116]	CS	PC	≥65	DC	Frequency of contacts with physicians, number of different ambulatory physicians contacted (all 1-year period)	Good
van Oostrom *et al* (2014)[117]	CS	PC	≥55	DC	Number of contacts with general practice, medications prescribed, referrals	Good
Vos *et al* (2013)[118]	CS	PC	70–74	DC	Self-rated health (SF-36)	Poor

Continued

Table 1 Continued

Author (Year)	Study design	Population source	Age	Multimorbidity measurement	Outcomes measured	Risk of bias
Wallace *et al* (2016a)[119]	Cohort	PC	≥70	Pra tool, modified Pra tool	Emergency hospital admission (1-year period) Both models demonstrated poor model discrimination	Good
Wallace *et al* (2016b)[120]	Cohort	PC	≥70	DC, Barnett conditions DC, CCI, prescribed drugs count, RxRisk-V	Emergency admission, functional decline (all 2-year period) All measures demonstrated poor discrimination	Good
Wei *et al* (2018)[121]	CS	GP	≥51	MWI	Subjective physical functioning, grip strength, gait speed, cognitive performance, ADL limitations, IADL limitations	Good
Wei *et al* (2019a)[122]	Cohort	GP	≥51	MWI	Physical functioning—SF-36, mortality	Good
Wei and Mukamal (2019b)[123]	Cohort	GP	≥51	MWI	Suicide mortality, health-related quality of life	Fair
Wei *et al* (2020a)[124]	Cohort	GP	≥51	MWI	Cognitive functioning	Good
Wei *et al* (2020b)[125]	Cohort	GP	≥51	MWI-ICD, DC, CCI, Elixhauser, health-related quality of life comorbidity index	Mortality, future physical functioning	Poor
Wei and Mukamal (2018)[28]	Cohort	GP	≥36	MWI, DC, CCI	Mortality (10-year period), future physical functioning MWI performed best in predicting mortality as compared with DC and CCI	Good
Wikman *et al* (2011)[126]	CS	GP	≥50	DC	QoL, affective well-being	Good
Wister *et al* (2015) [22]	CS	GP	≥65	MM additive scale, MM weighted by HUI3, MM weighted by ADL scale, MM weighted by HUI3 betas	Life satisfaction, perceived health status	Good

ACE, Adult Comorbidity Evaluation; ACG, Adjusted Clinical Groups; ACSH, Ambulatory Care Sensitive Hospitalisation; ADG, Aggregated Diagnosis Groups; ADL, Activities of Daily Living; CCC, Chronic Condition Count; CCI, Charlson Comorbidity Index; CCI-PSR, Charlson Comorbidity Index-Psychosocial Risk; CDS, Chronic Disease Score; CIRS, Cumulative Illness Rating Scale; CRG, Clinical Risk Groups; CS, Cross-Sectional; DBIS, Disease Burden Impact Scale; DC, Disease Count (Unweighted); ED, Emergency Department; EDC, Expanded Diagnosis Clusters; EI, Elixhauser Index; ERA, Elder Risk Assessment; GP, General Population; HCC, Hierarchical Condition Categories; HRQoL, Health-Related Quality of Life; HSMI, Health Search Morbidity Index; HUI3, Health Utility Index; IADL, Instrumental Activities of Daily Living; ICD-10, International Classification of Diseases, Tenth Revision; mCCI, modified Charlson Comorbidity Index; MDMS, Multidimensional Multimorbidity Score; M3 Index, Multimorbidity Measure Index; MM, Multimorbidity; MWI, Multimorbidity-Weighted Index; PC, Primary Care; Pra tool, Probability of repeated admission risk prediction tool; QOF, Quality and Outcomes Framework; QoL, Quality of Life; RUB, Resource Utilisation Band; RxRisk-V, A Veterans Association adapted pharmacy-based case-mix instrument; SRH, Self-Rated Health.

Table 2 Description of instruments used for measurement of multimorbidity and the data sources and resources required

Category	Instrument	System/Condition based	Weightage; Scoring method	Data sources and resources required
A: Count of individual conditions				
A-1	DC	Condition (7–147)	Unweighted; condition count	ATC list of conditions, Elixhauser list of conditions, EMR, GP records, health service database, hospital discharge abstract, insurance claims or questionnaires— telephone, face-to-face, mailed surveys. Participant involvement required.
A-2	CCC[91]	Condition (6)	Unweighted; based on AHRQ's clinical classification software and number of conditions for each category	EMR
B: Organ or system-based approaches				
B-3	Organ systems with CDC	Organ system (17)	Unweighted sum of organ systems	EMR
B-4	CIRS[127 128]	Body systems (13)	1–5 (based on severity of the condition); different weightage for diseases	EMR
C: Weighted indices				
C-5	ACE[129]	Condition (27)	1–3 (based on severity of most severe condition); highest score of single item	Insurance claims' database
C-6	Cambridge MM Score[97]	Condition (20)	Weighted based on three different outcomes—primary care consultation, unplanned admission and mortality	EMR linked to mortality, hospital admission and socioeconomic deprivation
C-7	CCI[20]	Condition (19)	1–6 (based on impact on 1-year mortality (RR)—original); sum of weighted conditions	Administrative database, EMR, medical chart review, or interviews or postal questionnaire where participant involvement is required
C-8	CLS[57]	Condition (98)	Based on impact for mortality (HR); sum of beta coefficients of each category	Linked patients' records of all primary care events, hospital admissions and causes of death.
C-9	DBIS[130 131]	Conditions (25–28)	Weighted according to the degree in which each condition interferes with daily activities	Patient involvement in the questionnaire is required
C-10	EI (original and modified)[132]	Condition (21–31)	Based on impact on in-hospital mortality; summing of beta coefficients	Insurance claims' or medical services database
C-11	ERA[56]	Condition (6–9)	Weighted (based on impact on future hospitalisation); sum of weighted regression coefficients	EMR and administrative database

Continued

Table 2 Continued

Category	Instrument	System/Condition based	Weightage; Scoring method	Data sources and resources required
C-12	HCC[133]	Condition (70)	Based on Medicare capitation payments for health expenditure; more severe manifestations of a condition dominating (and zeroing out the effect of) less serious ones. Other diseases are summed additively.	EMR and HCC software licensing and fees
C-13	M3 Index[109]	Condition (55)	Weighted based on 1-year mortality; summing of beta coefficients	Linked patients' records
C-14	MDMS[24]	Condition (7 chronic conditions, 2 health behaviours for first dimension and 5 symptoms for second dimension)	Weighted but not based on any specific outcome; sum of the value for the weighted absolute contributions of each of the dimensions.	Standardised medical evaluation (interviewer-administered); participant involvement is required
C-15	MM weighted by ADL Scale[134]	Condition (19)	Weighted based on OARS functional status scale measuring ADL; sum of weighted conditions	Face-to-face or telephone interviews where participant involvement is required
C-16	MM weighted by HUI[135]	Condition (19)	Weighted based on correlation with health utility index; sum of weighted conditions	Face-to-face or telephone interviews where participant involvement is required
C-17	MM weighted by HUI betas[135]	Condition (19)	Weighted based on correlation with health utility index and adjusted for age and sex; summing of beta coefficients	Face-to-face or telephone interviews where participant involvement is required
C-18	MWI[136]	Condition (81)	Weighted based on impact on SF-36 physical functioning scale; sum of weights	Interviewer-administered or mail questionnaire where participant involvement is required
C-19	QOF standard (weighted)[50]	Condition (14)	0–6, based on impact on 1-year mortality (RR); sum of weighted conditions	EMR
C-20	QOF extended (weighted)[50]	Condition (9)	1–3, based on impact on 1-year mortality (RR); sum of weighted conditions	EMR
C-21	Severity Burden Score[130]	Condition (21)	Sum of weights of diseases by the level of interference for each condition	Interviewer-administered structured questionnaire by nurses where participant involvement is required
D: Other approaches (D1=Case mix, D2=Pharmaceutical-based)				
D1-22	ACG[137 138]	Condition (93 mutually exclusive ACGs. Some are modified to 68 ACGs)	Incorporated into ACGs based on impact on resource use (proprietary); variable	EMR and ACG software licensing and fees
D1-23	ADG[137 138]	Condition (32 groups)	Based on duration, severity, diagnostic certainty, aetiology and need for specialty care; variable	EMR and ACG software licensing and fees

Continued

Table 2 Continued

Category	Instrument	System/Condition based	Weightage; Scoring method	Data sources and resources required
D1-24	CRG[26]	NA; diagnostic categories derived from organ systems or clinical category (37)	Pre-formulated based on the 3M clinical risk groups and consists of 9 core health ranks	EMR—inpatient and outpatient and 3M Clinical Risk Grouping software V.1.6 and service fees
D1-25	Adjusted Morbidity Groups (GMA)[139]	NA; mutually exclusive categories (31)	Based on multimorbidity and levels of patient complexity	Registry data
D1-26	HM[21]	Condition (NS)	Only MN tier 4+MN tier 3 with ERA>10; variable	EMR, HCC software licensing, fees and administrative data
D1-27	HSMI[78]	Condition (73 chronic and acute conditions)	Based on yearly healthcare costs directly derived from primary care setting; sum of regression coefficients (range from –0.06 to 1.04)	EMR
D1-28	Minnesota Tiering[137 138]	Condition (NS)	Grouping patients into 'complexity tiers' based on the number of major condition categories; condition count	EMR or administrative data and MN Tiering software licensing and fees
D1-29	Resource Utilisation Band[137 138]	Condition (six mutually exclusive bands)	Based on ACG algorithm on impact on resource use (proprietary); variable	EMR and ACG software licensing and fees
D2-30	CDS[33]	Condition (17)	Weighted 1–5; sum of weights based on pharmacological database	Prescription drug database
D2-31	Drug Count	NA; variable. Some may be based on pharmacologic-therapeutic classification system	Weighted; medication count	Self-reported questionnaire where participant involvement is required
D2-32	Modified Pra tool using RxRisk-V[34 119 140]	NA; Pra tool+RxRisk V	Weighted due to RxRisk-V; 4 categories	GP medical record+linked pharmacy claims database
D2-33	RxRisk-V[34]	NA; WHO-ATC classification system	Weighted according to the diagnostic group of drugs to predict future healthcare costs; sum of weights	GP medical record+linked pharmacy claims database

ACE, Adult Comorbidity Evaluation; ACG, Adjusted Clinical Groups; ADG, Aggregated Diagnosis Groups; ADL, Activities of Daily Living; ATC, Anatomical Therapeutic Chemical; CCC, Chronic Condition Count; CCI, Charlson Comorbidity Index; CDC, Chronic Disease Count; CDS, Chronic Disease Score; CGI-S, Clinical Global Impression-Severity Scale; CIRS, Cumulative Illness Rating Scale; CLS, Comorbidity Linked Score; CRG, Clinical Risk Groups; DBIS, Disease Burden Impact Scale; DC, Disease Count; EDC, Expanded Diagnosis Clusters; EI, Elixhauser Index; EMR, Electronic Medical Records; ERA, Elder Risk Assessment; GP, General Practitioner; HCC, Hierarchical Condition Categories; HM, hybrid model (MN Tier+ERA); HSMI, Health Search Morbidity Index; HUI, Health Utility Index; mCCI, modified Charlson Comorbidity Index; MDMS, Multidimensional Multimorbidity Score; M3 Index, Multimorbidity Measure Index; MM, Multimorbidity; MWI, Multimorbidity-Weighted Index; OARS, Older Americans Resources and Services; Pra tool, Probability of repeated admission risk prediction tool; QOF, Quality and Outcomes Framework; RxRisk-V, A Veterans Association adapted pharmacy-based case-mix instrument; SF-36, 36-item Short Form Survey.

Table 3 Summary of multimorbidity instruments and their associations with outcomes measured from all the included studies

Multimorbidity measures	Association between outcomes and multimorbidity: Evidence of an association	Association between outcomes and multimorbidity: No evidence of an association
A=Count of individual conditions		
DC (many different groupings ranging from 7[64 76 113] to 147[66] conditions and some are further categorised[21])	ADL limitations,[43] activity limitations,[52] affective well-being,[126] cognitive function,[82] continuity of care (3 years),[105] deficits of perceived general health,[104] depression,[92] depressive symptoms,[67 104] disability,[17 65 84 85] emergency hospital admission (2 years),[120] fall-related injuries,[71] fall risk,[76] frequency of contacts with physicians (1 year),[116] functional capacity,[93] functional decline (2 years),[120] functional Impairment,[111] functional limitations,[74] future physical functioning,[28] general health,[52] healthcare costs,[23 91] health-related quality of life,[68 75 81 90 115] hospitalisation (3 years),[18] hospital admissions (1 year),[95 96 99] hospital outpatient visits (1 year),[114] hospitalisation/emergency department visits,[114] life satisfaction,[22] mental distress,[52] mortality (1 year),[99] (3 years),[18 48] (5 years),[110] (10 years),[28] number of contacts with general practice (1 year),[117] number of medications prescribed (1 year),[117] number of mentally unhealthy days,[43 44] number of physically unhealthy days,[43 44] number of different ambulatory physicians contacted (1 year),[116] number of primary care consultations (1 year),[48] (3 years),[48] number of referrals (1 year),[117] outpatient/Inpatient service use,[53] physical distress,[52] physical function,[55] prescription costs,[70] perceived health status,[22] presence of mental health disorder,[66] primary care consultations (1 year period),[105] (3 years),[105] primary care experience—self-reported,[83] primary healthcare cost,[47] primary healthcare use,[89] potentially preventable unplanned admission (1-year period),[95] quality-adjusted life years,[72 73] quality of life,[93 98 101 102] self-rated health,[118] self-reported falls (12 months),[108] symptom burden,[101] self-rated Health,[64 85] self-perceived health,[93] total number of consultation,[42] total health care costs[42]	Functional decline,[18] quality of cardiovascular preventive care,[27] quality of preventive care[27]
CCC	Healthcare costs,[21] hospital admissions (1 year),[21] number of emergency department visits (1 year),[21] readmission within 30 days (1 year)[21]	
B=Organ or system-based approaches		
Organ systems with CDC	Presence of depressive or anxiety disorder[19]	
CIRS	Disability,[17] frailty,[17] healthcare utilisation,[54] hospitalisation (3 years),[18] mortality[18]	Functional decline[18]
C=Weighted indices		
ACE	Healthcare expenditure[23]	
Cambridge MM Score	Mortality,[97] primary care consultation,[97] unplanned admission[97]	
CCI	Ambulatory care-sensitive hospitalisations (acute and chronic),[106] disability,[17] emergency department visits (1 year),[21] emergency hospital admission (2 years),[119] frailty,[17] functional decline (2 years),[119] future physical functioning,[28] healthcare expenditure,[21] hip fractures,[103] hospitalisation (1 year),[21 64 80 99 109] hospitalisation (3 years),[18] mortality (1 year),[50 63 99 109] mortality (3 years),[18] (5 years),[51 63] (10 years),[51 63] (15, 20, 25 years),[51] number of primary care consultations (3 years),[48] number of primary care physician visits (1 year),[107] number of specialist visits (1 year),[107] potentially preventable unplanned admission (1 year),[96] presence of critical illness,[46] primary healthcare cost,[47] mortality (1 year),[57 58] (3 years),[48 59] (5 years),[58 60 62] (7 years),[86] (10 years),[28] readmission within 30 days (1 year),[21] successful ageing[61]	Functional decline,[18] primary care visits,[69] quality of cardiovascular preventive care,[27] quality of preventive care[27]
CLS	Mortality (1 year)[57]	

Continued

Table 3 Continued

Multimorbidity measures	Association between outcomes and multimorbidity: Evidence of an association	No evidence of an association
DBIS	Quality of life[98]	
EI (original and modified)	Hospitalisation (1 year),[99 109] mortality (1 year)[57 99 109]	
ERA	Healthcare expenditure,[21] mortality (2 years),[113] number of days hospitalised (1 year),[56] number of emergency department visits (1 year),[21 56] number of hospital admissions (1 year),[21 56] number of hospital visits (1 year),[21 56] nursing home placement (2 years),[113] presence of critical illness,[46] readmission within 30 days (1 year)[21]	
HCC	Hospitalisation (1 year),[21] ED visits (1 year),[21] readmission within 30 days (1 year),[21] healthcare expenditure (1 year)[21]	
M3 Index	Hospitalisation (1 year),[109] mortality (1 year)[109]	
MDMS	Sickness absence episodes taken in 2 years (male)[24]	Sickness absence episodes taken in 2 years (female)[24]
MM weighted by ADL scale	Life satisfaction,[22] perceived health status[22]	
MM weighted by HUI	Life satisfaction,[22] perceived health status[22]	
MM weighted by HUI betas	Life satisfaction,[22] perceived health status[22]	
MWI	ADL limitations,[121] IADL limitations,[121] mortality (10 years),[28] cognitive performance,[121] future physical functioning,[28 125] grip strength,[121] health-related quality of life,[123] mortality,[125] subjective physical functioning,[121] suicide mortality[123]	Gait speed[28]
QOF (standard)	Mortality (1 year)[50]	
QOF (extended)	Mortality (1 year)[50]	
Severity Burden Score	Mental component score (SF-12)[94]	
D=Other approaches (D-1=Case Mix, D2=Pharmaceutical-based)		
ACG	Hospitalisation (1 year),[80] mortality (3 years),[48] number of primary care consultations (3 years),[48] primary healthcare cost,[47] readmission within 30 days (1 year)[21]	
ADG	Hospitalisation (1 year),[107] number of primary care physician visits (1 year),[107] number of specialist visits (1 year)[107]	
CRG	Healthcare expenditure,[49] HRQoL using EQ-5D-3L[87]	
Adjusted Morbidity Groups (GMA)	Use of healthcare resources[88]	
HM	Emergency department visits (1 year),[21] healthcare expenditure,[21] hospitalisation (1 year),[21] readmission within 30 days (1 year)[21]	
HSMI	Healthcare cost (primary care)[78]	
Minnesota Tiering	Emergency department visits (1 year),[21 114] healthcare expenditure,[21] hospitalisation (1 year),[21 114] readmission within 30 days (1 year)[21]	
Resource Utilisation Band	Fee-for-service expenditures,[77] primary healthcare cost,[47] mortality (3 years),[48] number of primary care consultations (3 years),[48] registered active listing in primary care,[100] registered active listing in all healthcare[100]	
CDS	Hospitalisation (1 year),[99] mortality (1 year)[99]	
Drug Count	Emergency hospital admission (2 years),[120] functional decline (2 years),[120] hospitalisation (1 year),[99] mortality (1 year),[99] (3 years),[48] number of primary care consultations (3 years)[48]	
Pra tool Modified using RxRisk-V	Emergency hospital admission (1 year)[119]	
RxRisk-V	Emergency hospital admission (2 years),[120] functional decline (2 years)[120]	

Continued

Table 3 Continued

Multimorbidity measures	Association between outcomes and multimorbidity: Evidence of an association	No evidence of an association

ACE-27, Adult Comorbidity Evaluation; ACG, Adjusted Clinical Groups; ADG, Aggregated Diagnosis Groups; ADL, Activities of Daily Living; CCC, Chronic Condition Count; CCI, Charlson Comorbidity Index; CDC, Chronic Disease Count; CDS, Chronic Disease Score; CIRS, Cumulative Illness Rating Scale; CLS, Comorbidity Linked Score; CRG, Clinical Risk Groups; DBIS, Disease Burden Impact Scale; DC, Disease Count; EI, Elixhauser Index; ERA, Elder Risk Assessment; HCC, Hierarchical Condition Categories; HM, Hybrid Model (MN Tier+ERA); HRQoL, health-related quality of life; HSMI, Health Search Morbidity Index; HUI, Health Utility Index; MDMS, Multidimensional Multimorbidity Score; M3 Index, Multimorbidity Measure (M3) Index; MM, Multimorbidity; MWI, Multimorbidity-Weighted Index; Pra tool, Probability of repeated admission risk prediction tool; QOF, Quality and Outcomes Framework; RxRisk-V, A Veterans Association adapted pharmacy-based case-mix instrument; SF-12, Short Form-12.

Organ or system-based approaches

There were two instruments in this category. They were Cumulative Illness Rating Scale (CIRS)[17 18] and Organ Systems with Chronic Disease Count (Organ-CDC).[19]

Weighted indices

There were 17 unique weighted instruments found in the included studies. The original CCI with its different modifications was the most frequently used instrument and was used in 29 studies. The CCI was based on Disease Count, but the 17 conditions were weighted originally based on their impact on 1-year mortality.[20] The final score was derived by the summation of all the weighted conditions. There were many variations and modifications of the score including the addition of psychosocial factors. The CCI instrument was found to be associated with multiple outcomes other than 1-year mortality.

Most of the other weighted index instruments were novel, like the Multimorbidity-Weighted Index (MWI), in which the investigators built multivariable prognostic models from a set of potential predictor conditions and weighted the conditions based on an outcome of clinical interest. The most common outcomes chosen were mortality and physical function. Other outcomes included health expenditure,[21] health utility index[22] and severity of the most severe condition.[23] The Multidimensional Multimorbidity Score (MDMS)[24] was unique as it was weighted based on health behaviours and patient symptoms and not based on any specific outcome.

Other approaches to measuring multimorbidity

Other approaches included case-mix and pharmaceutical-based instruments. For case-mix approach, the ACG and Resource Utilisation Band were the most commonly used instruments.[25] Most of the case-mix instruments required proprietary software licenses from the USA and obtained data from electronic medical records or administrative data. The Clinical Risk Groups instrument was similar but took into account the severity of individual conditions.[26]

The second group of instruments in this category was related to pharmaceutical data. The most frequent type was the unweighted Drug Count. The other three (Chronic Disease Score, A Veterans Association adapted pharmacy-based case-mix instrument like RxRisk-V and modified Probability of repeated admission risk prediction tool using RxRisk-V) were all weighted indices. Except for the Drug Count that was based on a self-report questionnaire, the rest required a prescription drug database to obtain the data.

Outcomes

We classified the 150 outcomes into 17 categories as reported in the core outcomes set of multimorbidity research (COSmm).[10] The most commonly reported outcomes were healthcare use (n=45), mortality (n=18), health-related quality of life (n=18) and physical function (n=13). The different studies unanimously showed that higher levels of multimorbidity was associated with higher healthcare use and mortality, lower health-related quality of life and poorer physical function. Seven outcomes in the COSmm were not found in all the 96 studies. These were treatment burden, self-management behaviour, self-efficacy, adherence, communications, shared decision-making and prioritisation. There were 19 outcomes that were not described in the COSmm. These included cognitive function, risk of suicide, frailty and falls. The outcomes not found to have any association with the instruments for measuring the level of multimorbidity were preventive care,[27] sickness absence episodes (female)[24] and gait speed.[28]

DISCUSSION

Summary of findings

Thirty-three unique instruments for measuring the level of multimorbidity were identified and categorised according to the classification by Sarfati.[8] The most commonly used instrument was 'Disease Count'. It was also the only instrument that was associated with the three essential outcomes from the core outcomes set of multimorbidity research (COSmm),[10] that is, quality of life, mental health and mortality.

Comparison with previous research

Although the most common instrument identified in this systematic review was similar to that of Huntley *et al*,[12] several instruments including Duke Severity of Illness Checklist (DUSOI) and Functional Comorbidity Index identified in their article were not found in this systematic review. The possible reasons for not finding these instruments in this review could be due to the lack of interest in the instrument by the research community in recent years (to our knowledge, the last publication using DUSOI was

in 2004),[29] or the exclusion of studies specifying an index condition.

Advantages and disadvantages of selected instruments

Disease Count

The advantage of using 'Disease Count' is its simplicity and the ease of data ascertainment with minimal resources required. However, using 'Disease Count' does not consider the severity of each condition where the complexity of multimorbidity may not be properly addressed.[30] The other disadvantage noted was the lack of transparency in the operational definition of multimorbidity, especially regarding the list of conditions considered for multimorbidity and the cut-points used. Despite its simplicity, the level of multimorbidity measured using 'Disease Count' was the only instrument that was found to be associated with the three essential core outcomes (quality of life, mental health and mortality).

Weighted indices

The common weighted indices identified in this systematic review were CCI, Elders Risk Assessment (ERA), Elixhauser Index (EI) and MWI. These weighted indices were often used in prognostic models to build complex multivariable regression models in which the weights were calculated from hazard ratios, odds ratios or regression coefficients.[31]

The advantage of these weighted indices is that the weights allow the adaptation of an index to a specific outcome. An investigator could recalibrate the correct weight by creating a prognostic model to produce a contextualised instrument for a different setting. Prognostic models can provide clinically relevant risk stratification and help to allocate resources.[32] The disadvantage of such indices is that calculated weights are greatly influenced by the population, outcomes used, and the instrument's original conception and purpose, hampering the ability to compare across studies.

Case-mix

The ACG system has a good track record in the USA and several other countries, especially for measuring the outcomes of healthcare utilisation. However, the instrument is proprietary, and the exact algorithm of the instrument is not open to the public and may not be suitable in certain settings. The Clinical Risk Group (CRG) system has a good track record in Spain. It measures the severity of each condition and its algorithm is fully transparent. The common disadvantage of both systems is the financial costs involved in obtaining the license.

Pharmaceutical-based instruments

Medication-based indices include versions of the Chronic Disease Score,[33] which later became known as the RxRisk,[34] and its adaptation for use in the veteran population, the RxRisk-V.[35] Like the Disease Count, its main advantage is the ease of use with minimal resources required. However, many studies were not transparent regarding which type of drugs were included.

Data sources

Data sources used by these instruments included medical record information, patient self-report, clinical judgement and large administrative databases. Each data source has its inherent advantages and disadvantages. For patient self-report, patients with cognitive impairment may under-report symptoms and may be seen less frequently by their physicians, resulting in an under-recognition or undertreatment of conditions.[36] It has also been shown that health administrative data based on billing system underestimated the prevalence of many chronic conditions.[37]

The available data in a particular setting may strongly influence the ultimate instrument chosen for multimorbidity research. As there is currently no consensus on the gold standard for sources of data, it is difficult to assess which data source was superior from this review.

Outcomes

There were 17 multimorbidity outcomes identified by a Delphi process involving a panel of international experts in multimorbidity intervention studies.[10] However, only 10 out of the 17 outcomes were reported in the 96 studies identified in this systematic review. The most common outcome that was investigated was healthcare use. The seven missing outcomes belong to 'patient-reported impact and behaviours' and 'consultation-related' outcome groups, most likely indicating that there is a dearth of multimorbidity studies looking at these two groups of outcomes measures.

Clinical implications

Ideally, a single instrument measuring the level of multimorbidity should be able to predict a variety of relevant outcomes. However, Byles *et al*[38] reported that a single instrument could not be used to predict different outcomes, in different patient groups and settings, unless different weights were assigned to these factors in calculating a score. Such multiple-scoring instruments may be the way forward for validation of prognostic models for different outcomes and different populations with established multimorbidity instruments. For example, depending on the outcome, study population and setting, the choice of conditions included in the multiple-scoring instrument should include those with a high prevalence in that study population and the weights should be determined by their significant impact (ie, outcome) on the affected population.

For pragmatic reasons, the final selection of the conditions to be included in such a multiple-scoring instrument may still have to take into account the availability of relevant and reliable data. A certain degree of reductionism will also have to be accepted because a single instrument will not be able to encompass all the nuances of the different interactions of chronic conditions on an individual living in his/her unique milieu. We recommend that researchers perform validation studies using the instruments listed in this systematic review to adjust

the weights according to the specific outcome of interest for the study population relevant to their setting.

Strengths and limitations of the study

The main strengths of this systematic review were the involvement of a health science librarian in our search strategy, a published protocol, adherence to the protocol without major changes during the systematic review process,[39] and the critical appraisal of all the primary studies with a risk of bias assessment tool.

The systematic review had several limitations. We excluded grey literature and included only studies that were published in the English language. We also did not contact authors directly for a suggestion of studies, nor identified a list of instruments from the preliminary search and then performed an additional search using the same databases.[40] Additionally, this systematic review did not review the validity and reliability of all the instruments as it was beyond the scope of the intended work. We have, however, included the references of the original articles or validation studies in table 2 for each of the instrument where available. Finally, this review specifically aimed to look at the association of the level of multimorbidity as the main independent variable and excluded the level of multimorbidity as a mediating, confounding or effect-modifying variable. This strict criterion excluded 17 studies (figure 1) as a result. Excluding these 17 studies did not alter the findings as the instruments used in all the 17 studies were Disease Count (n=9), CIRS (n=3), CCI (n=3), EI (n=1) and Aggregated Diagnosis Groups (n=1) where no new instruments were identified.

CONCLUSIONS

In this systematic review, we found 33 instruments for measuring the level of multimorbidity in community-dwelling individuals that predict or explore the association of multimorbidity with at least one specified outcome. Disease Count and weighted indices like the CCI, the ERA and EI were commonly used for measuring the level of multimorbidity. Other approaches to measuring the level of multimorbidity included case-mix or pharmaceutical-based instruments.

We found continuing interest in measuring the level of multimorbidity with Disease Count and Drug Count. There has also been a rise in the development of novel weighted indices using prognostic models or validation of existing well-established instruments like the CCI over the last few years. There is currently an absence of a gold standard for where to obtain chronic disease information. The most suitable instrument will depend on the specified outcome of interest, the study population and the type of data and resources available.

Finally, there is still much work to improve on the body of knowledge of multimorbidity when most investigators in the last decade measured multimorbidity without including some of the important outcome measures of multimorbidity. We also suggest that a clear description of the instruments is required in the publication of multimorbidity studies to counter the frequent lack of information currently seen so as to contribute to robust multimorbidity research in future.

Author affiliations
[1]Clinical Research Unit, National Healthcare Group Polyclinics, Singapore
[2]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
[3]Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[4]Department of Epidemiology and Biostatistics, Western University Schulich School of Medicine and Dentistry, London, Ontario, Canada
[5]Centre for Studies in Family Medicine, Department of Family Medicine, Western University Schulich School of Medicine and Dentistry, London, Ontario, Canada
[6]Department of Family Medicine and Emergency Medicine, Université de Sherbrooke, Sherbrooke, Quebec, Canada

Acknowledgements We would like to thank Mr John Costella, Collections and Content Strategies Librarian, Western University for his assistance and guidance in the search strategy. We would also like to thank Dr Mohammed Zulhakim, Family Medicine Resident Doctor, Ministry of Health Holdings, for his help in testing out the various risk assessment tools and his suggestions on the modified Newcastle-Ottawa Scale.

Contributors All authors made substantial contribution to the design of the work. BLR, MF and MS mentored and provided guidance to the review. ESL, together with a health science librarian, created the search strategy and conducted the search. ESL, EH, TSH, HLK and FYW screened articles, extracted and analysed data. ESL drafted the manuscript with valuable inputs from BLR, MF and MS and all authors reviewed drafts and approved the final version.

Funding This study was supported by the Singapore Ministry of Health's National Medical Research Council under the Centre Grant Programme (Ref No: NMRC/CG/C019/2017).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. All data relevant to the study are included in the article or uploaded as supplemental information. There are no additional, unpublished data arising from this research.

ORCID iD
Eng Sing Lee http://orcid.org/0000-0003-4963-535X

REFERENCES

1 WHO. *The World Health Report 2008. Primary Care - Now more than ever*, 2008.
2 Huber M, Knottnerus JA, Green L, *et al*. How should we define health? *BMJ* 2011;343:d4163.
3 Coventry PA, Small N, Panagioti M, *et al*. Living with complexity; marshalling resources: a systematic review and qualitative meta-

synthesis of lived experience of mental and physical multimorbidity. *BMC Fam Pract* 2015;16:171.
4 Liddy C, Blazkho V, Mill K. Challenges of self-management when living with multiple chronic conditions: systematic review of the qualitative literature. *Can Fam Physician* 2014;60:1123–33.
5 Lugtenberg M, Burgers JS, Clancy C, *et al*. Current guidelines have limited applicability to patients with comorbid conditions: a systematic analysis of evidence-based guidelines. *PLoS One* 2011;6:e25987.
6 Tinetti ME, Fried T. The end of the disease era. *Am J Med* 2004;116:179–85.
7 Fortin M, Stewart M, Poitras M-E, *et al*. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012;10:142–51.
8 Sarfati D. Review of methods used to measure comorbidity in cancer populations: no gold standard exists. *J Clin Epidemiol* 2012;65:924–33.
9 Sarfati D. How Do We Measure Comorbidity? In: Koczwara B, ed. *Cancer and Chronic Conditions - Addressing the problem of multimorbidity in cancer patients and survivors*. Springer, 2016: 35–70.
10 Smith SM, Wallace E, Salisbury C, *et al*. A core outcome set for multimorbidity research (COSmm). *Ann Fam Med* 2018;16:132–8.
11 Kirkham JJ, Gorst S, Altman DG, *et al*. Core outcome Set-STAndards for reporting: the COS-STAR statement. *PLoS Med* 2016;13:e1002148.
12 Huntley AL, Johnson R, Purdy S, *et al*. 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.
13 Lee ES, EQ H, Koh HL. A systematic review on the instruments used for measuring the severity of multimorbidity, 2018. Available: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=105297
14 van den Akker M, Buntinx F, Metsemakers JF, *et al*. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J Clin Epidemiol* 1998;51:367–75.
15 Alshabanat A, Zafari Z, Albanyan O, *et al*. Asthma and COPD overlap syndrome (ACOs): a systematic review and meta analysis. *PLoS One* 2015;10:e0136065.
16 Wells GA, Shhea B, O'Connell D. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [Accessed 6 Apr 2018].
17 Boeckxstaens P, Vaes B, Legrand D, *et al*. The relationship of multimorbidity with disability and frailty in the oldest patients: a cross-sectional analysis of three measures of multimorbidity in the BELFRAIL cohort. *Eur J Gen Pract* 2015;21:39–44.
18 Boeckxstaens P, Vaes B, Van Pottelbergh G, *et al*. Multimorbidity measures were poor predictors of adverse events in patients aged ≥80 years: a prospective cohort study. *J Clin Epidemiol* 2015;68:220–7.
19 DiNapoli EA, Bramoweth AD, Whiteman KL, *et al*. Mood disorders in middle-aged and older veterans with multimorbidity. *J Aging Health* 2017;29:657–68.
20 Charlson ME, Pompei P, Ales KL, *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
21 Haas LR, Takahashi PY, Shah ND, *et al*. Risk-Stratification methods for identifying patients for care coordination. *Am J Manag Care* 2013;19:725–32.
22 Wister AV, Levasseur M, Griffith LE, *et al*. 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.
23 Hwang W, LaClair M, Camacho F, *et al*. Persistent high utilization in a privately insured population. *Am J Manag Care* 2015;21:309–16.
24 Ubalde-Lopez M, Delclos GL, Benavides FG, *et al*. Measuring multimorbidity in a working population: the effect on incident sickness absence. *Int Arch Occup Environ Health* 2016;89:667–78.
25 The Johns Hopkins ACG® system excerpt from version 11.0 technical reference guide 2014. Available: https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/cntrb_035024.pdf [Accessed 10 Sep 2019].
26 Hughes JS, Averill RF, Eisenhandler J, *et al*. Clinical risk groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management. *Med Care* 2004;42:81–90.
27 Streit S, da Costa BR, Bauer DC, *et al*. Multimorbidity and quality of preventive care in Swiss university primary care cohorts. *PLoS One* 2014;9:e96142.
28 Wei MY, Mukamal KJ. Multimorbidity, mortality, and long-term physical functioning in 3 prospective cohorts of community-dwelling adults. *Am J Epidemiol* 2018;187:103–12.
29 Moll van Charante E, Hartman E, Yzermans J, *et al*. The first general practitioner hospital in the Netherlands: towards a new form of integrated care? *Scand J Prim Health Care* 2004;22:38–43.
30 Salive ME. Multimorbidity in older adults. *Epidemiol Rev* 2013;35:75–83.
31 Moons KGM, Royston P, Vergouwe Y, *et al*. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.
32 Alonso-Morán E, Nuño-Solinis R, Onder G, *et al*. Multimorbidity in risk stratification tools to predict negative outcomes in adult population. *Eur J Intern Med* 2015;26:182–9.
33 Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 1992;45:197–203.
34 Fishman PA, Goodman MJ, Hornbrook MC, *et al*. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. *Med Care* 2003;41:84–99.
35 Sloan KL, Sales AE, Liu C-F, *et al*. Construction and characteristics of the RxRisk-V: a VA-adapted pharmacy-based case-mix instrument. *Med Care* 2003;41:761–74.
36 Lash TL, Mor V, Wieland D, *et al*. Methodology, design, and analytic techniques to address measurement of comorbid disease. *J Gerontol A Biol Sci Med Sci* 2007;62:281–5.
37 Fortin M, Haggerty J, Sanche S, *et al*. Self-reported versus health administrative data: implications for assessing chronic illness burden in populations. A cross-sectional study. *CMAJ Open* 2017;5:E729–33.
38 Byles JE, D'Este C, Parkinson L, *et al*. Single index of multimorbidity did not predict multiple outcomes. *J Clin Epidemiol* 2005;58:997–1005.
39 Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. *PLoS One* 2010;5:e9810.
40 Yurkovich M, Avina-Zubieta JA, Thomas J, *et al*. A systematic review identifies valid comorbidity indices derived from administrative health data. *J Clin Epidemiol* 2015;68:3–14.
41 Agborsangaya CB, Lau D, Lahtinen M, *et al*. Health-Related quality of life and healthcare utilization in multimorbidity: results of a cross-sectional survey. *Qual Life Res* 2013;22:791–9.
42 Bähler C, Huber CA, Brüngger B, *et al*. Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. *BMC Health Serv Res* 2015;15:23.
43 Barile JP, Thompson WW, Zack MM, *et al*. Multiple chronic medical conditions and health-related quality of life in older adults, 2004–2006. *Prev Chronic Dis* 2013;10:E162.
44 Barile JP, Thompson WW, Zack MM, *et al*. Activities of daily living, chronic medical conditions, and health-related quality of life in older adults. *J Ambul Care Manage* 2012;35:292–303.
45 Barnett K, Mercer SW, Norbury M, *et al*. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37–43.
46 Biehl M, Takahashi PY, Cha SS, *et al*. Prediction of critical illness in elderly outpatients using elder risk assessment: a population-based study. *Clin Interv Aging* 2016;11:829–34.
47 Brilleman SL, Gravelle H, Hollinghurst S, *et al*. Keep it simple? Predicting primary health care costs with clinical morbidity measures. *J Health Econ* 2014;35:109–22.
48 Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. *Fam Pract* 2013;30:172–8.
49 Caballer-Tarazona V, Guadalajara-Olmeda N, Vivas-Consuelo D. Predicting healthcare expenditure by multimorbidity groups. *Health Policy* 2019;123:427–34.
50 Carey IM, Shah SM, Harris T, *et al*. 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.
51 Chapman BP, Weiss A, Fiscella K, *et al*. Mortality risk prediction: can comorbidity indices be improved with psychosocial data? *Med Care* 2015;53:909–15.
52 Chen H-Y, Baumgardner DJ, Rice JP. Health-Related quality of life among adults with multiple chronic conditions in the United States, behavioral risk factor surveillance system, 2007. *Prev Chronic Dis* 2011;8:A09.
53 Chen H, Cheng M, Zhuang Y, *et al*. Multimorbidity among middle-aged and older persons in urban China: prevalence, characteristics and health service utilization. *Geriatr Gerontol Int* 2018;18:1447–52.
54 Chu T, Lau P, Cheng RSY. A study on the prevalence of multi - morbidities of diseases and utilisation of public healthcare services

in the New Territories West area of Hong Kong. *The Hong Kong Practitioner* 2018;40.
55 Clynes MA, Bevilacqua G, Jameson KA, *et al*. Does self-report of multimorbidity in later life predict impaired physical functioning, and might this be useful in clinical practice? *Aging Clin Exp Res* 2020;32:1443–50.
56 Crane SJ, Tung EE, Hanson GJ, *et al*. Use of an electronic administrative database to identify older community dwelling adults at high-risk for hospitalization or emergency department visits: the elders risk assessment index. *BMC Health Serv Res* 2010;10:338.
57 Crooks CJ, Card TR, West J. The use of a Bayesian hierarchy to develop and validate a co-morbidity score to predict mortality for linked primary and secondary care data from the NHS in England. *PLoS One* 2016;11:e0165507.
58 Crooks CJ, West J, Card TR. A comparison of the recording of comorbidity in primary and secondary care by using the Charlson index to predict short-term and long-term survival in a routine linked data cohort. *BMJ Open* 2015;5:e007974.
59 Formiga F, Ferrer A, Chivite D, *et al*. Utility of geriatric assessment to predict mortality in the oldest old: the Octabaix study 3-year follow-up. *Rejuvenation Res* 2013;16:279–84.
60 Formiga F, Ferrer A, Chivite D, *et al*. Predictors of long-term survival in nonagenarians: the NonaSantfeliu study. *Age Ageing* 2011;40:111–6.
61 Formiga F, Ferrer A, Megido MJ, *et al*. Low co-morbidity, low levels of malnutrition, and low risk of falls in a community-dwelling sample of 85-year-olds are associated with successful aging: the Octabaix study. *Rejuvenation Res* 2011;14:309–14.
62 Formiga F, Ferrer A, Padros G, *et al*. Evidence of functional declining and global comorbidity measured at baseline proved to be the strongest predictors for long-term death in elderly community residents aged 85 years: a 5-year follow-up evaluation, the OCTABAIX study. *Clin Interv Aging* 2016;11:437–44.
63 Fraccaro P, Kontopantelis E, Sperrin M, *et al*. Predicting mortality from change-over-time in the Charlson comorbidity index: a retrospective cohort study in a data-intensive UK health system. *Medicine* 2016;95:e4973.
64 Galenkamp H, Braam AW, Huisman M, *et al*. Somatic multimorbidity and self-rated health in the older population. *J Gerontol B Psychol Sci Soc Sci* 2011;66:380–6.
65 Garin N, Olaya B, Moneta MV, *et al*. Impact of multimorbidity on disability and quality of life in the Spanish older population. *PLoS One* 2014;9:e111498.
66 Glynn LG, Valderas JM, Healy P, *et al*. The prevalence of multimorbidity in primary care and its effect on health care utilization and cost. *Fam Pract* 2011;28:516–23.
67 Gunn JM, Ayton DR, Densley K, *et al*. The association between chronic illness, multimorbidity and depressive symptoms in an Australian primary care cohort. *Soc Psychiatry Psychiatr Epidemiol* 2012;47:175–84.
68 Hanmer J, Vanness D, Gangnon R, *et al*. Three methods tested to model SF-6D health utilities for health states involving comorbidity/co-occurring conditions. *J Clin Epidemiol* 2010;63:331–41.
69 Hu T, Dattani ND, Cox KA, *et al*. Effect of comorbidities and medications on frequency of primary care visits among older patients. *Can Fam Physician* 2017;63:45–50.
70 Isaacs AA, Manga N, Le Grange C, *et al*. A snapshot of noncommunicable disease profiles and their prescription costs at ten primary healthcare facilities in the in the Western half of the Cape town Metropole. *South African Family Practice* 2014;56:43–9.
71 Jennings LA, Reuben DB, Kim S-B, *et al*. Targeting a high-risk group for fall prevention: strategies for health plans. *Am J Manag Care* 2015;21:e519–26.
72 Jia H, Lubetkin EI, Barile JP, *et al*. Quality-Adjusted life years (QALY) for 15 chronic conditions and combinations of conditions among US adults aged 65 and older. *Med Care* 2018;56:740–6.
73 Jia H, Lubetkin EI. Impact of nine chronic conditions for US adults aged 65 years and older: an application of a hybrid estimator of quality-adjusted life years throughout remainder of lifetime. *Qual Life Res* 2016;25:1921–9.
74 Jindai K, Nielson CM, Vorderstrasse BA, *et al*. Multimorbidity and functional limitations among adults 65 or older, NHANES 2005-2012. *Prev Chronic Dis* 2016;13:E151.
75 Kim K-I, Lee JH, Kim C-H. Impaired health-related quality of life in elderly women is associated with multimorbidity: results from the Korean National health and nutrition examination survey. *Gend Med* 2012;9:309–18.
76 Kojima T, Akishita M, Nakamura T, *et al*. Association of polypharmacy with fall risk among geriatric outpatients. *Geriatr Gerontol Int* 2011;11:438–44.
77 Kristensen T, Olsen KR, Schroll H, *et al*. Association between fee-for-service expenditures and morbidity burden in primary care. *Eur J Health Econ* 2014;15:599–610.
78 Lapi F, Bianchini E, Cricelli I, *et al*. Development and validation of a score for adjusting health care costs in general practice. *Value Health* 2015;18:884–95.
79 Lawson KD, Mercer SW, Wyke S, *et al*. Double trouble: the impact of multimorbidity and deprivation on preference-weighted health related quality of life a cross sectional analysis of the Scottish health survey. *Int J Equity Health* 2013;12:67.
80 Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care* 2012;50:131–9.
81 Li J, Green M, Kearns B, *et al*. Patterns of multimorbidity and their association with health outcomes within Yorkshire, England: baseline results from the Yorkshire health study. *BMC Public Health* 2016;16:649–49.
82 Loprinzi PD. Multimorbidity, cognitive function, and physical activity. *Age* 2016;38:8..
83 Macinko J, Andrade FCD, Nunes BP, *et al*. Primary care and multimorbidity in six Latin American and Caribbean countries. *Rev Panam Salud Publica* 2019;43:e8:1–9.
84 Marengoni A, Angleman S, Fratiglioni L. Prevalence of disability according to multimorbidity and disease clustering: a population-based study. *J Comorb* 2011;1:11–18.
85 McDaid O, Hanly MJ, Richardson K, *et al*. The effect of multiple chronic conditions on self-rated health, disability and quality of life among the older populations of Northern Ireland and the Republic of ireland: a comparison of two nationally representative cross-sectional surveys. *BMJ Open* 2013;3:e002571.
86 Md Yusof MY, Horan MA, Jones M, *et al*. Developing a self-reported comorbidity index to predict mortality of community-dwelling older adults. *Arch Gerontol Geriatr* 2010;50:e63–7.
87 Millá-Perseguer M, Guadalajara-Olmeda N, Vivas-Consuelo D, *et al*. Measurement of health-related quality by multimorbidity groups in primary health care. *Health Qual Life Outcomes* 2019;17:8.
88 Monterde D, Vela E, Clèries M, *et al*. Multimorbidity as a predictor of health service utilization in primary care: a registry-based study of the Catalan population. *BMC Fam Pract* 2020;21:39.
89 Muggah E, Graves E, Bennett C, *et al*. The impact of multiple chronic diseases on ambulatory care use; a population based study in Ontario, Canada. *BMC Health Serv Res* 2012;12:452–52.
90 Mujica-Mota RE, Roberts M, Abel G, *et al*. Common patterns of morbidity and multi-morbidity and their impact on health-related quality of life: evidence from a national survey. *Qual Life Res* 2015;24:909–18.
91 Naessens JM, Stroebel RJ, Finnie DM, *et al*. Effect of multiple chronic conditions among working-age adults. *Am J Manag Care* 2011;17:118–22.
92 Østergaard SD, Foldager L. The association between physical illness and major depressive episode in general practice. *Acta Psychiatr Scand* 2011;123:290–6.
93 Palladino R, Pennino F, Finbarr M, *et al*. Multimorbidity and health outcomes in older adults in ten European health systems, 2006–15. *Health Aff* 2019;38:613–23.
94 Pati S, Swain S, Knottnerus JA, *et al*. Health related quality of life in multimorbidity: a primary-care based study from Odisha, India. *Health Qual Life Outcomes* 2019;17:116.
95 Payne RA, Abel GA, Guthrie B, *et al*. The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study. *CMAJ* 2013;185:E221–8.
96 Payne RA, Abel GA, Avery AJ, *et al*. Is polypharmacy always hazardous? A retrospective cohort analysis using linked electronic health records from primary and secondary care. *Br J Clin Pharmacol* 2014;77:1073–82.
97 Payne RA, Mendonca SC, Elliott MN, *et al*. Development and validation of the Cambridge multimorbidity score. *CMAJ* 2020;192:E107–14.
98 Peters M, Kelly L, Potter CM, *et al*. Quality of life and burden of morbidity in primary care users with multimorbidity. *Patient Relat Outcome Meas* 2018;9:103–13.
99 Quail JM, Lix LM, Osman BA, *et al*. Comparing comorbidity measures for predicting mortality and hospitalization in three population-based cohorts. *BMC Health Serv Res* 2011;11:146.
100 Ranstad K, Midlöv P, Halling A. Importance of healthcare utilization and multimorbidity level in choosing a primary care provider in Sweden. *Scand J Prim Health Care* 2014;32:99–105.
101 Reinke LF, Vig EK, Tartaglione EV, *et al*. Symptom burden and palliative care needs among high-risk veterans with multimorbidity. *J Pain Symptom Manage* 2019;57:880–9.

102 Renne I, Gobbens RJ. Effects of frailty and chronic diseases on quality of life in Dutch community-dwelling older adults: a cross-sectional study. *Clin Interv Aging* 2018;13:325–34.
103 Reyes C, Estrada P, Nogués X, *et al*. The impact of common co-morbidities (as measured using the Charlson index) on hip fracture risk in elderly men: a population-based cohort study. *Osteoporos Int* 2014;25:1751–8. [Erratum appears in Osteoporos Int. 2014 Sep;25(9):2333 Note: Macias, J G [corrected to Gonzalez-Macias, J]].
104 Ryu E, Takahashi PY, Olson JE, *et al*. Quantifying the importance of disease burden on perceived general health and depressive symptoms in patients within the Mayo clinic Biobank. *Health Qual Life Outcomes* 2015;13:95.
105 Salisbury C, Johnson L, Purdy S, *et al*. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011;61:e12–21.
106 Saver BG, Wang C-Y, Dobie SA, *et al*. The central role of comorbidity in predicting ambulatory care sensitive hospitalizations*. *Eur J Public Health* 2014;24:66–72.
107 Shadmi E, Balicer RD, Kinder K, *et al*. Assessing socioeconomic health care utilization inequity in Israel: impact of alternative approaches to morbidity adjustment. *BMC Public Health* 2011;11:609.
108 Sibley KM, Voth J, Munce SE, *et al*. Chronic disease and falls in community-dwelling Canadians over 65 years old: a population-based study exploring associations with number and pattern of chronic conditions. *BMC Geriatr* 2014;14:22.
109 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.
110 St John PD, Tyas SL, Menec V, *et al*. Multimorbidity, disability, and mortality in community-dwelling older adults. *Can Fam Physician* 2014;60:e272–80.
111 St John PD, Tyas SL, Menec V, *et al*. Multimorbidity predicts functional decline in community-dwelling older adults: prospective cohort study. *Can Fam Physician* 2019;65:e56–63.
112 Sullivan PW, Ghushchyan VH, Bayliss EA. The impact of co-morbidity burden on preference-based health-related quality of life in the United States. *Pharmacoeconomics* 2012;30:431–42.
113 Takahashi PY, Tung EE, Crane SJ, *et al*. Use of the elderly risk assessment (era) index to predict 2-year mortality and nursing home placement among community dwelling older adults. *Arch Gerontol Geriatr* 2012;54:34–8.
114 Takahashi PY, Heien HC, Sangaralingham LR, *et al*. Enhanced risk prediction model for emergency department use and hospitalizations in patients in a primary care medical home. *Am J Manag Care* 2016;22:475–83.
115 Tyack Z, Frakes K-A, Barnett A, *et al*. Predictors of health-related quality of life in people with a complex chronic disease including multimorbidity: a longitudinal cohort study. *Qual Life Res* 2016;25:2579–92.
116 van den Bussche H, Schön G, Kolonko T, *et al*. Patterns of ambulatory medical care utilization in elderly patients with special reference to chronic diseases and multimorbidity--results from a claims data based observational study in Germany. *BMC Geriatr* 2011;11:54.
117 van Oostrom SH, Picavet HSJ, de Bruin SR, *et al*. Multimorbidity of chronic diseases and health care utilization in general practice. *BMC Fam Pract* 2014;15:61.
118 Vos HMM, Bor HH, Rangelrooij-Minkels MJAvan, *et al*. Multimorbidity in older women: the negative impact of specific combinations of chronic conditions on self-rated health. *Eur J Gen Pract* 2013;19:117–22.
119 Wallace E, McDowell R, Bennett K, *et al*. External validation of the probability of repeated admission (PRA) risk prediction tool in older community-dwelling people attending general practice: a prospective cohort study. *BMJ Open* 2016;6:e012336.
120 Wallace E, McDowell R, Bennett K, *et al*. 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.
121 Wei MY, Kabeto MU, Langa KM, *et al*. Multimorbidity and physical and cognitive function: performance of a new Multimorbidity-Weighted index. *J Gerontol A Biol Sci Med Sci* 2018;73:225–32.
122 Wei MY, Kabeto MU, Galecki AT, *et al*. Physical functioning decline and mortality in older adults with multimorbidity: joint modeling of longitudinal and survival data. *J Gerontol A Biol Sci Med Sci* 2019;74:226–32.
123 Wei MY, Mukamal KJ. Multimorbidity and mental health-related quality of life and risk of completed suicide. *J Am Geriatr Soc* 2019;67:511–9.
124 Wei MY, Levine DA, Zahodne LB, *et al*. Multimorbidity and cognitive decline over 14 years in older Americans. *J Gerontol A Biol Sci Med Sci* 2020;75:1206–13.
125 Wei MY, Ratz D, Mukamal KJ. Multimorbidity in Medicare beneficiaries: performance of an ICD-Coded Multimorbidity-Weighted index. *J Am Geriatr Soc* 2020;68:999–1006.
126 Wikman A, Wardle J, Steptoe A. Quality of life and affective well-being in middle-aged and older people with chronic medical illnesses: a cross-sectional population based study. *PLoS One* 2011;6:e18952.
127 Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc* 1968;16:622–6.
128 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.
129 Fleming ST, Sabatino SA, Kimmick G, *et al*. Developing a claim-based version of the ACE-27 comorbidity index: a comparison with medical record review. *Med Care* 2011;49:752–60.
130130 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.
131 Bayliss EA, Ellis JL, Steiner JF. Seniors' self-reported multimorbidity captured biopsychosocial factors not incorporated into two other data-based morbidity measures. *J Clin Epidemiol* 2009;62:550–7.
132 Elixhauser A, Steiner C, Harris DR, *et al*. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8–27.
133 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.
134 Fillenbaum GG, Smyer MA. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. *J Gerontol* 1981;36:428–34.
135 Feeny D, Furlong W, Torrance GW, *et al*. Multiattribute and single-attribute utility functions for the health Utilities index mark 3 system. *Med Care* 2002;40:113–28.
136 Wei MY, Kawachi I, Okereke OI, *et al*. 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.
137 Starfield B, Weiner J, Mumford L, *et al*. Ambulatory care groups: a categorization of diagnoses for research and management. *Health Serv Res* 1991;26:53–74.
138 Weiner JP, Starfield BH, Steinwachs DM, *et al*. Development and application of a population-oriented measure of ambulatory care case-mix. *Med Care* 1991;29:452–72.
139 Monterde D, Vela E, Clèries M, *et al*. [Adjusted morbidity groups: A new multiple morbidity measurement of use in Primary Care]. *Aten Primaria* 2016;48:674.
140 Boult C, Dowd B, McCaffrey D, *et al*. Screening elders for risk of hospital admission. *J Am Geriatr Soc* 1993;41:811–7.

Appendix 1. Search Strategy

Table 1. Medline Search 1946 to October 23, 2020

#	Searches	Results
1	exp comorbidity/ or multiple chronic conditions/	111191
2	(multimorbid* or multi-morbid* or comorbid* or co-morbid*).ab,ti,kw.	180089
3	((multiple or coexist* or co-exist* or concurrent* or simultaneous*) adj2 (disease* or illness* or diagnos* or condition* or morbid* or disorder*)).ab,ti,kw.	52139
4	1 or 2 or 3	292316
5	primary health care/ or "continuity of patient care"/ or exp general practice/ or ambulatory care/ or physicians, family/ or physicians, primary care/ or community health services/ or general practitioners/	249484
6	((ambulatory or community or general or family or primary) adj2 (care or health* or practi* or physician* or medicine or center* or centre* or facilit* or clinic*)).ab,ti,kw.	391013
7	5 or 6	508037
8	epidemiologic measurements/ or risk assessment/ or "Outcome Assessment (Health Care)"/ or patient reported outcome measures/ or health status indicators/ or "severity of illness index"/ or sickness impact profile/ or diagnosis-related groups/ or case mix/	594279
9	((health status or risk or outcome* or sickness impact) adj2 (appraisal* or assessment* or evaluation* or measur* or index* or indic* or profile*)).ab,ti,kw.	398319
10	((severity or burden) adj2 (illness* or diseas* or disorder* or condition* or diagnos*)).ab,ti,kw.	85595
11	(charlson comorbidity index or charlson index or charlson score or elixhauser* or cumulative illness rating scale* or CIRS or adjusted clinical group* or ACG* or disease count or Duke severity* or DUSOI or casemix or case-mix).ab,ti,kw.	21547
12	8 or 9 or 10 or 11	939031
13	4 and 7 and 12	4862
14	limit 13 to (english language and yr="2010 - Current")	2827

Table 2. Embase Classic+Embase 1947 to October 23, 2020

#	Searches	Results
1	comorbidity/ or multiple chronic conditions/	265602
2	(multimorbid* or multi-morbid* or comorbid* or co-morbid*).ab,ti,kw.	329397
3	((multiple or coexist* or co-exist* or concurrent* or simultaneous*) adj2 (disease* or illness* or diagnos* or condition* or morbid* or disorder*)).ab,ti,kw.	81506
4	1 or 2 or 3	493137
5	primary health care/ or patient care/ or primary medical care/ or general practice/ or general practitioner/ or ambulatory care/ or family medicine/ or community care/ or community health services/	664384
6	((ambulatory or community or general or family or primary) adj2 (care or health* or practi* or physician* or medicine or center* or centre* or facilit* or clinic*)).ab,ti,kw.	539525
7	5 or 6	943743
8	risk assessment/ or outcome assessment/ or patient reported outcome/ or health status indicator/ or disease activity score/ or global disease burden/ or organ dysfunction score/ or "severity of illness index"/ or sickness impact profile/ or general health status assessment/ or disease severity/ or diagnosis related group/ or charlson comorbidity index/ or comorbidity assessment/ or elixhauser comorbidity index/ or case mix/	1590115
9	((health status or risk or outcome* or sickness impact) adj2 (appraisal* or assessment* or evaluation* or measur* or index* or indic* or profile*)).ab,ti,kw.	547083
10	((severity or burden) adj2 (illness* or diseas* or disorder* or condition* or diagnos*)).ab,ti,kw.	138554
11	(charlson comorbidity index or charlson index or charlson score or elixhauser* or cumulative illness rating scale* or CIRS or adjusted clinical group* or ACG* or disease count or Duke severity* or DUSOI or casemix or case-mix).ab,ti,kw.	42616
12	8 or 9 or 10 or 11	2022565
13	4 and 7 and 12	12245
14	limit 13 to (english language and yr="2010 - Current")	8896

Table 3. CINAHL – since inception to October 23, 2020

#	Searches	Results
1	(MH “comorbidity”)	38970
2	(multimorbid* or multi-morbid* or comorbid* or co-morbid*)	66442
3	(multiple or coexist* or co-exist* or concurrent* or simultaneous*) N2 (disease* or illness* or diagnos* or condition* or morbid* or disorder*)	12822
4	1 or 2 or 3	77266
5	((MH “primary health care”) or (MH “family practice”) or (MH “ambulatory care") or (MH “ambulatory care facilities") or (MM “community health services”) or (MM “community health centers”) or (MH “physicians, family”) or (MH “continuity of patient care"))	100224
6	((ambulatory or community or general or family or primary) N2 (care or health* or practi* or physician* or medicine or center* or centre* or facilit* or clinic*))	332346
7	5 or 6	340014
8	((MH “risk assessment”) or (MH "Outcome Assessment") or (MH “patient-reported outcomes) or (MH “health status indicators”) or (MH "severity of illness indices") or (MH “sickness impact profile”) or (MH “diagnosis-related groups”) or (MH “case mix”))	102350
9	((health status or risk or outcome* or sickness impact) N2 (appraisal* or assessment* or evaluation* or measur* or index* or indic* or profile*))	340802
10	(severity or burden) N2 (illness* or diseas* or disorder* or condition* or diagnos*)	68742
11	(charlson comorbidity index or charlson index or charlson score or elixhauser* or cumulative illness rating scale* or CIRS or adjusted clinical group* or ACG* or disease count or Duke severity* or DUSOI or casemix or case-mix)	6729
12	8 or 9 or 10 or 11	396387
13	4 and 7 and 12	4729
14	limit 13 to (english language and yr="2010 - Current")	3536

Appendix 2. Coding Description for the Modified Newcastle-Ottawa Scale for Cohort Studies

Category	Description
Selection	
1) Representativeness of the sample	This item assesses the representativeness of sample in the community, not from some general population. a) Truly representative* (e.g., everyone from the database) b) Somewhat representative* (with at least 2 criteria but selection method was convincing due to random sampling) c) Selected group (e.g., only certain socio-economic groups or areas) d) No description of sampling strategy
2) Ascertainment of multimorbidity	This item assesses the method by which multimorbidity was confirmed. a) Secure record* (e.g., GP questionnaire) b) Structured interview* (e.g., interviewer-administered questionnaire) c) Written self-report (e.g., mailed survey, if items are unable to be confirmed by objective measure) d) No description / Other
3) Demonstration that outcome of interest was not present at start of study	A statement of no history of disease earns a star. In the case of mortality studies, outcome of interest is still the presence of a disease, rather than death. a) Yes* b) No
Comparability	
1) Study controls for age and sex	Covariates must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the relative risk for the exposure of interest is adjusted for the covariates listed, then the groups will be considered to be comparable on each variable used. a) Yes* b) No
2) Study controls for other factors	a) Yes* b) No
Outcome	
1) Statistical test	Statistical test(s) must be clearly described and appropriate before assessing the other items under the Outcome category. a) Clearly described and appropriate b) Not described, incomplete or inappropriate
2) Assessment of outcome	This item assesses the method by which the outcome of interest was confirmed. For some outcomes, reference to the medical record is sufficient to satisfy the requirement for confirmation. This may not be adequate for other outcomes where reference to specific tests or measures would be required. a) Independent or blind assessment* (e.g., interviewer-administered questionnaire) b) Record linkage* (e.g., identified through ICD codes on database records) c) Self-report (i.e., no reference to original medical records) d) No description / Other

3)	Was follow-up long enough for outcomes to occur?	An acceptable length of time was at least 1 year of follow-up. a) Yes* b) No
4)	Adequacy of follow-up of cohorts	This item assesses the follow-up of the sample to ensure that losses are not related to the outcome. a) Complete follow-up - all subject accounted for* b) Subjects lost to follow-up unlikely to introduce bias* (Number lost ≤20% or description of those lost suggested no different from those followed.) c) Follow-up rate less than 80% and no description of those lost d) No statement

Thresholds for converting the modified Newcastle-Ottawa scales to Good, Fair, and Poor Quality:

- Good Quality - 2-3 stars in Selection category (Representativeness of Sample item must be fulfilled) AND 1-2 stars in Comparability category AND 2-3 stars in Outcome category.
- Fair Quality - 1 star in Selection category AND 1-2 stars in Comparability category AND 2-3 stars in Outcome category.
- Poor Quality - 0 star in Selection category OR 0 star in Comparability category OR 0-2 stars in Outcome category.

Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Modifications from original NOS include:

- 'Statistical test' was added to the Outcome category
- 'Representativeness of sample' item under Selection category and 'Statistical test' item under Outcome category must both be fulfilled for study to be considered Good Quality.
- 'Selection of the non-exposed cohort' item was removed from Selection category as most studies did not describe a non-exposed cohort, and this review sought to compare the different levels of multimorbidity within the group.
- 'Representativeness of the exposed cohort' item under Selection category was renamed 'Representativeness of the sample' since the 'non-exposed cohort' was removed above.
- 'Ascertainment of exposure' under the Selection category was renamed as 'Ascertainment of multimorbidity' to specify that multimorbidity is the exposure in this review.
- 'Study controls for age and sex' and 'Study controls for other factors' were revised to be items under Comparability category
- The thresholds for converting the scales to quality were amended accordingly due to the above modifications.

Consensus:

- The study team decided that a period of one-year was a reasonable period of follow-up under 'Was follow-up long enough for outcomes to occur?' item under Outcome category.

Appendix 3. Coding Description for the Modified Newcastle-Ottawa Scale for Cross-sectional Studies

Category	Description
Selection	
1) Representativeness of the sample	This item assesses the representativeness of sample in the specified population, not from some general population. a) Truly representative* (e.g., everyone from the database, random sampling) b) Somewhat representative* (with at least 2 criteria but selection method was convincing due to random sampling) c) Selected group (e.g., only certain socio-economic groups or areas) d) No description of sampling strategy
2) Ascertainment of multimorbidity	This item assesses the method by which multimorbidity was confirmed. a) Secure record* (e.g., Clinical records, GP questionnaire) b) Structured interview* (e.g., interviewer-administered questionnaire) c) Written self-report (e.g., mailed survey, if items are unable to be confirmed by objective measure) d) No description / Other
3) Sample Size	If there is no description, a reported sample size of 800 and above is satisfactory. a) Justified and satisfactory* b) Not justified
4) Non-respondents	Acceptable response rates for surveys through various methods† • In-person: 57% • Mail: 50% • Average: 33% • Email: 30% • Internet: 29% • Telephone: 18% • In-app: 13% a) Comparability between respondents and non-respondents characteristics is established, and the response rate is satisfactory.* b) The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory. c) No description of the response rate or the characteristics of the responders and the non-responders
Comparability	
1) Study controls for age and sex	Confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. a) Yes* b) No
2) Study controls for other factors	a) Yes* b) No

Outcome	
1) Statistical test	Statistical test(s) must be clearly described and appropriate before assessing the other items under the Outcome category. a) Clearly described and appropriate b) Not described, incomplete or inappropriate
2) Assessment of outcome	This item assesses the method by which the outcome of interest was confirmed. For some outcomes, reference to the medical record is sufficient to satisfy the requirement for confirmation. This may not be adequate for other outcomes where reference to specific tests or measures would be required. a) Independent or blind assessment***** b) Record linkage***** (e.g., identified through ICD codes on database records) c) Self-report (i.e., no reference to original medical records) d) No description / Other

Thresholds for converting the modified Newcastle-Ottawa scales to Good, Fair, and Poor Quality:

- Good Quality – 3-4 stars in Selection category (Representativeness of Sample item must be fulfilled) AND 1-2 stars in Comparability category AND 1 star in Outcome category.
- Fair Quality - 2 stars in Selection category AND 1-2 stars in Comparability category AND 1 star in Outcome category.
- Poor Quality – 0-1 star in Selection category OR 0 star in Comparability category OR 0 star in Outcome category.

†Lindermann N. What's the average survey response rate? [2018 benchmark]. SurveyAnyplace. Available from: https://surveyanyplace.com/average-survey-response-rate/.

Alshabanat A, Zafari Z, Albanyan O, Dairi M, FitzGerald JM. Asthma and COPD Overlap Syndrome (ACOS): a systematic review and meta-analysis. PLoS ONE. 2015;10(9):e0136065. Doi:10.1371/journal.pone.0136065.

Modifications from original NOS include:

- 'Ascertainment of exposure' item under Selection category was renamed as 'Ascertainment of multimorbidity' to specify that multimorbidity is the exposure in this review.
- The ratings for the 'Ascertainment of exposure' item under Selection category were revised to 'secure record, structured interview, written self-report, no description/other' to align with the Modified NOS for cohort studies in this review.
- 'Study controls for age and sex' and 'Study controls for other factors' items were revised to be items under Comparability category.
- 'Was follow-up long enough for outcomes to occur' item under Outcome category was renamed as 'Statistical test'. This item must be fulfilled for the study to be considered as Good Quality.
- The thresholds for converting the scales to quality were amended accordingly due to the above modifications.

Consensus:

- The study team decided on a list of acceptable response rates for various survey methods for the 'Non-respondents' item under Selection category.
- The study team decided that a sample size of 800 was reasonable under 'Sample size' item under Selection category.

Appendix 4. Summary of risk of bias (RoB) appraisal of included cohort articles using the modified Newcastle-Ottawa Quality Assessment Scale (44 articles)

	Barile et al. (2013)[19]	Biehl et al. (2016)[22]	Boeckxstaens et al. (2015b)[24]	Brilleman et al. (2014)[25]	Brilleman and Salisbury(2013)[26]	Carey et al. (2013)[28]	Chapman et al. (2015)[29]	Charlson et al. (2014)[30]	Crane et al. (2010)[35]	Crooks et al. (2016)[36]	Crooks et al. (2015)[37]	Formiga et al. (2013)[39]	Formiga et al. (2011a)[40]	Formiga et al. (2016)[42]	Fraccaro et al. (2016)[43]	Haas et al. (2013)[48]	Hwang et al. (2015)[51]	Jennings et al. (2015)[53]	Jia et al. (2018)[54]	Jia and Lubetkin (2017)[55]	Lemke et al. (2012)[62]	Md Yusof et al. (2010)[68]	Monterde et al. (2020)[70]	Pavne et al. (2013)[77]	Pavne et al. (2014)[78]	Pavne et al. (2020)[79]	Quail et al. (2011)[81]	Reves et al. (2014)[85]	Salisbury et al. (2011)[87]	Saver et al. (2014)[88]	Stanley and Sarfati (2017)[91]	St John et al. (2014)[92]	St John et al. (2019)[93]	Streit et al. (2014)[94]	Takahashi et al. (2011)[96]	Takahashi et al. (2016)[97]	Tvack et al. (2016)[98]	Wallace et al. (2016a)[103]	Wallace et al. (2016b)[104]	Wei et al. (2019a)[106]	Wei et al. (2019b)[107]	Wei et al. (2020a)[108]	Wei et al. (2020b)[109]	Wei and Mukamal (2018)[110]
Selection Criteria																																												
Representativeness of the sample	*	*	-	*	*	*	*	-	*	*	*	*	*	*	*	*	*	-	*	*	*	-	*	*	*	*	*	*	*	*	*	*	*	*	-	*	*	*	-	*	*	*
Ascertainment of Multimorbidity	-	*	*	*	*	*	-	*	*	*	*	*	*	*	*	*	*	*	-	*	*	-	*	*	*	*	*	*	*	*	*	*	*	*	-	*	*	*	-	*	*	-
Demonstration that outcome of interest was not present at the start of the study	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Comparability Criteria																																												
Study controls for age and sex	*	-	*	*	*	*	*	*	-	*	*	-	*	-	*	*	*	*	-	*	*	-	*	*	*	*	*	-	*	*	*	*	*	*	*	*	*	*	-	*	-	-
Study controls for others	*	*	-	*	*	*	*	*	*	-	*	*	*	*	*	*	*	*	-	-	*	*	*	*	*	-	*	*	*	*	-	*	*	*	*	*	*	*	*	*	-	*
Outcome Criteria																																												
Statistical test†	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Assessment of outcome	-	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Follow-up was long enough for outcomes to occur	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Adequacy of follow-up of cohorts	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	-	*	*	*	*	*	-	*	*	*	-	*	*	*	-	-	*	*
Overall RoB Judgement[a]	G	G	F	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	F	P	P	G	F	G	G	G	G	G	G	G	G	G	G	G	G	G	G	F	G	P	G

Note. G = Good; F = Fair; P = Poor.

†Statistical test must be clearly described and appropriate before assessing the other items under the Outcome category. [a] Good rating is given when there are 2 to 3 stars in Selection category (Representativeness of Sample must be fulfilled) AND 1 to 2 stars in Comparability category AND 2 to 3 stars in Outcome category; Fair rating is given when there are 1 star in Selection category AND 1 to 2 stars in Comparability category AND 2 to 3 stars in Outcome category; Poor rating is given when there are 0 star in Selection category OR 0 star in Comparability category OR 0 to 2 stars in Outcome category.

"*" Study satisfies the criteria; "-" Study did not satisfy the criteria; "✓" Statistical test is clearly described and appropriate; "X" Statistical test is not described, incomplete or inappropriate.

Appendix 5. Summary of risk of bias (RoB) appraisal of included cross-sectional articles using the modified Newcastle-Ottawa Quality Assessment Scale (52 articles)

	Agborsangaya et al. (2013)[17]	Bahler et al. (2015)[18]	Barile et al. (2012)[20]	Barnett et al. (2012)[21]	Boeckxstaens et al. (2015a)[23]	Caballer-Tarazona et al. (2019)[27]	Chen et al. (2011)[31]	Chen et al. (2018)[32]	Chu (2018)[33]	Clynes et al. (2020)[34]	DiNapoli et al. (2017)[38]	Formiga et al. (2011b)[41]	Galenkamp et al. (2011)[44]	Garin et al. (2014)[45]	Givnn et al. (2011)[46]	Gunn et al. (2012)[47]	Hanmer et al. (2010)[49]	Hu et al. (2017)[50]	Isaacs et al. (2014)[52]	Jindai et al. (2016)[56]	Kim et al. (2012)[57]	Koiima et al. (2011)[58]	Kristensen et al. (2014)[59]	Lani et al. (2015)[60]	Lawson et al. (2013)[61]	Li et al. (2016)[63]	Lonrinzi (2016)[64]	Macinko et al. (2019)[65]	Marengoni et al. (2011)[66]	McDaid et al. (2013)[67]	Milla-Perseguer (2019)[69]	Murrah et al. (2012)[71]	Mujica-Mota et al. (2015)[72]	Naessens et al. (2011)[73]	Østergaard and Foldager (2011)[74]	Palladino et al. (2019)[75]	Pati et al. (2019)[76]	Peters et al. (2018)[80]	Ranstad et al. (2014)[82]	Reinke et al. (2019)[83]	Renne and Gobbens (2018)[84]	Ryu et al. (2015)[86]	Shadmi et al. (2011)[89]	Sibley et al. (2014)[90]	Sullivan et al. (2012)[95]	Ubalde-Lopez et al. (2016)[99]	van den Bussche et al. (2014)[100]	van Oostrom et al. (2014)[101]	Vos et al. (2013)[102]	Wei et al. (2018)[105]	Wikman et al. (2011)[111]	Wister et al. (2015)[112]
Selection																																																				
Representativeness of the sample	*	*	*	*	-	*	*	-	*	-	*	*	*	*	*	*	*	-	*	-	*	-	*	*	*	*	*	*	*	*	*	*	*	-	-	*	*	*	-	*	*	*	*	*	*	*	*	*	*	*
Ascertainment of Multimorbidity	*	*	-	*	*	*	*	-	*	*	*	*	*	*	*	*	-	-	*	*	*	*	*	*	*	*	-	*	*	*	*	*	*	-	*	*	-	*	*	*	-	-	*	*	*	*	*	*	*	*
Sample size	*	*	*	*	-	*	*	*	*	-	*	-	*	*	*	*	*	*	-	*	*	*	-	*	*	*	*	*	*	*	*	*	*	-	*	*	*	*	*	*	*	*	*	*	*	-	*	*	*	*
Non-respondents	-	*	*	*	-	*	*	*	*	-	*	*	-	-	*	-	-	*	*	*	*	-	*	*	*	-	*	-	-	-	-	*	-	*	-	*	*	-	*	*	*	*	*	*	*	*	*	*	-
Comparability																																																			
Study controls for age and sex	*	*	*	*	*	-	*	*	*	-	*	-	*	*	*	*	*	*	-	-	-	*	*	*	*	-	*	*	*	*	*	*	-	*	-	-	*	*	*	*	*	*	*	*	-	*	*	*	*	*
Study controls for others	*	*	*	*	*	-	*	*	*	*	*	*	-	*	*	*	*	*	-	*	-	*	*	*	*	-	*	*	*	*	*	*	-	*	-	-	*	*	*	*	*	*	*	*	-	*	*	*	*	-
Outcome																																																			
Statistical test†	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Assessment of outcome	*	*	✓	*	*	*	*	-	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	-	*	*	*	*	*	*	*	*	*
Overall RoB Judgement[a]	G	G	G	G	P	P	G	P	G	P	G	G	G	G	G	F	F	F	P	G	P	P	G	G	G	P	G	G	G	G	G	P	F	P	P	G	G	F	G	G	P	G	G	P	G	G	G	G

Note. G = Good; F = Fair; P = Poor.

†Statistical test must be clearly described and appropriate before assessing the other items under the Outcome category. [a] Good rating is given when there are 3 to 4 stars in Selection category (Representativeness of Sample must be fulfilled) AND 1 to 2 stars in Comparability category AND 1 star in Outcome category; Fair rating is given when there are 2 stars in Selection category AND 1 to 2 stars in Comparability category AND 1 star in Outcome category; Poor rating is given when there are 0 to 1 star in Selection category OR 0 star in Comparability category OR 0 star in Outcome category.

"*" Study satisfies the criteria; "-" Study did not satisfy the criteria; "✓" Statistical test is clearly described and appropriate; "X" Statistical test is not described, incomplete or inappropriate