Development and validation of the Cambridge Multimorbidity Score

Rupert A. Payne MBChB PhD, Silvia C. Mendonca MSc, Marc N. Elliott PhD, Catherine L. Saunders MBBS PhD, Duncan A. Edwards MBBS MPH, Martin Marshall MBBS MD, Martin Roland BMBCh DM

Cite as: *CMAJ* 2020 February 3;191:E107-14. doi: 10.1503/cmaj.190757

ABSTRACT

BACKGROUND: Health services have failed to respond to the pressures of multimorbidity. Improved measures of multimorbidity are needed for conducting research, planning services and allocating resources.

METHODS: We modelled the association between 37 morbidities and 3 key outcomes (primary care consultations, unplanned hospital admission, death) at 1 and 5 years. We extracted development (n = 300 000) and validation (n = 150 000) samples from the UK Clinical Practice Research Datalink. We constructed a general-outcome multimorbidity score by averaging the standardized weights of the separate outcome scores. We compared performance with the Charlson Comorbidity Index.

RESULTS: Models that included all 37 conditions were acceptable predictors of general practitioner consultations (C-index 0.732, 95% confidence interval [CI] 0.731–0.734), unplanned hospital admission (C-index 0.742, 95% CI 0.737–0.747) and death at 1 year (C-index 0.912, 95% CI 0.905–0.918). Models reduced to the 20 conditions with the greatest combined prevalence/weight showed similar predictive ability (C-indices 0.727, 95% CI 0.725–0.728; 0.738, 95% CI 0.732–0.743; and 0.910, 95% CI 0.904–0.917, respectively). They also predicted 5-year outcomes similarly for consultations and death (C-indices 0.735, 95% CI 0.734–0.736, and 0.889, 95% CI 0.885–0.892, respectively) but performed less well for admissions (C-index 0.708, 95% CI 0.705–0.712). The performance of the general-outcome score was similar to that of the outcome-specific models. These models performed significantly better than those based on the Charlson Comorbidity Index for consultations (C-index 0.691, 95% CI 0.690–0.693) and admissions (C-index 0.703, 95% CI 0.697–0.709) and similarly for mortality (C-index 0.907, 95% CI 0.900–0.914).

INTERPRETATION: The Cambridge Multimorbidity Score is robust and can be either tailored or not tailored to specific health outcomes. It will be valuable to those planning clinical services, policymakers allocating resources and researchers seeking to account for the effect of multimorbidity.

Patients with multiple long-term health conditions are commonly seen by clinicians in generalist and specialist settings.[1,2] Services and policies have failed to respond to the pressures that multimorbidity places on primary and secondary care. These pressures are driven by the aging population, by policies that promote rapid access over longer consultations and continuity of care, and by single-disease guidelines and performance targets, which lead to overprescribing without addressing the priorities of the patients themselves.[3,4]

Several approaches have been used to quantify multimorbidity. Simple counts of conditions show a clear association with various outcomes, including primary care utilization, unplanned hospital admission and death.[5,6] Weighted approaches allow for differences in the strength of association between specific morbidities and a given outcome, as is the case for the Charlson Comorbidity Index, a composite morbidity score with condition weightings based on mortality.[7] Although its performance has exceeded that of several other metrics,[4] clinical practice has advanced considerably since its development in the 1980s, and the high weightings of particular conditions have been questioned.[8] A further problem with such indices is that weightings are generally based on a specific outcome such as death, and the indices may not predict other outcomes. The lists of conditions are also problematic. A minimum list of 12 conditions has been proposed.[9] However, a limited list may fail to capture important health problems, and comprehensive lists such as the Adjusted Clinical Groups (ACG) system may be challenging to implement.

The aim of the current study was to develop and validate a transparent, simple measure of multimorbidity based on data from United Kingdom general practitioner (GP) records and weighted on different clinical outcomes, for use in future studies of multimorbidity and for resource allocation.

Methods

Population and data sources

We undertook a retrospective cohort study using anonymous coded GP electronic health record data obtained from the UK Clinical Practice Research Datalink (CPRD).[10]

We restricted our analysis to the 148 practices contributing data classified by the CPRD as "up to standard" from 2010 to 2015, with primary care records linked to national data on mortality (Office for National Statistics),[11] hospital admission (Hospital Episode Statistics)[12] and socioeconomic deprivation (Index of Multiple Deprivation).[13] To ensure independence of samples, we randomly sampled practices into 1 of 3 data sets (at a ratio of 2:1:1). The development data set consisted of 300 000 randomly sampled adults at least 20 years of age registered on Jan. 1, 2014 (study start), with data classified by the CPRD as acceptable for use in research. We determined the presence of morbidities at an index date 12 months after the study start (Jan. 1, 2015), to ensure at least 1 year of registration and to maximize recording of prevalent cases. We followed patient records for 1 year after the index date (study end Dec. 31, 2015). The first validation data set consisted of 150 000 patients with the same specification as the development data set. A second, similar validation data set of 150 000 patients provided up to 5 years of follow-up, as well as a 1-year asynchronous follow-up (study start Jan. 1, 2010; index date Jan. 1, 2011; data available until Dec. 31, 2015). The sample size was selected to limit the width of the 95% confidence interval (CI) for a condition with 2% prevalence to about 0.5 on the log-odds scale for a dichotomous outcome such as death. Flow charts for the selection of practices and patients are presented in Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1), with details of the time periods and dates of each data set given in Appendix 2 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1).

Exposures

We defined comorbidities by relevant Read codes and/or prescribing before the index date, according to a list of 37 long-term conditions as described by Cassell and colleagues[2] and adapted from work by Barnett and colleagues,[14] the latter of which is considered one of the definitive epidemiologic studies of multimorbidity (Appendix 3, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). The conditions were chosen and defined (on the basis of clinical expert consensus) as those having a substantial impact on patients. The current study and that of Cassell and colleagues[2] align closely, both aiming to develop better means of quantifying multimorbidity. An additional condition list was included for the Charlson Comorbidity Index.[15] The code lists that we used (which have been published online[16]) were subject to considerable clinical attention, and we thus consider all of the comorbidities to have face validity; previous studies of the CPRD data set have shown that most long-term conditions have positive predictive values in excess of 80%.[17] Sex and age were included as covariables.

Outcomes

We used Office for National Statistics data and Hospital Episode Statistics care data for admitted patients to determine the occurrence of death and unplanned (emergency) inpatient hospital admission, respectively, during the follow-up period. We established the number of primary care consultations from GP records of face-to-face (including telephone) clinical encounters; multiple encounters in a single day were counted as 1 consultation.

Statistical analysis

We developed morbidity scores using 3 separate models, 1 model for each outcome, in the 2015 development data set. We modelled consultations using zero-inflated negative binomial regression, and unplanned hospital admission and death using Cox regression. In addition to the extended scores containing all 37 conditions, we constructed a set of simplified primary scores with the most important 20 conditions. In addition, we constructed a general-outcome multimorbidity score by averaging the standardized weights of the 3 simple scores. Details of the statistical modelling, including data cleaning, are provided in Appendix 3.

We independently evaluated the performance of each of the three 37-condition and 20-condition outcome-specific scores, as well as the 20-condition general-outcome score, at 1-year follow-up in the 2015 (synchronous) data set, as well as at 1-year and 5-year follow-up in the 2011 (asynchronous) data set. We examined the performance of each score for predicting each of the 3 outcomes, and also compared performance against the Charlson Comorbidity Index. We assessed model discrimination using Harrell's C-index.[18]

Ethics approval

Permission for the CPRD to receive and supply anonymous patient data for generic public health research is granted directly to the CPRD by the national Research Ethics Service of the UK Health Research Authority. Regulatory approvals to use CPRD data for the current project were granted by the CPRD Independent Scientific Advisory Committee (ISAC protocol 17_051).

Results

The characteristics of the 3 cohorts are shown in Table 1, and descriptive statistics for the multimorbidity scores are presented in Appendix 4 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). The development cohort had a mean age of 50.7 years, with 22.7% over age 65 years and 5.8% over age 80 years; 51.1% of the patients were women. The most socioeconomically deprived group was underrepresented. The mean number of morbidities was 1.3 (standard deviation 1.7), with 31.7% of individuals having 2 or more recorded conditions. In general, similar patterns of age, sex, socioeconomic deprivation and multimorbidity were observed across all cohorts (Table 1).

The most common conditions were hypertension (19.24%), anxiety or depression (12.85%), painful condition (11.63%) and hearing loss (11.27%) (Table 2). The full list of disease prevalence and score weightings is provided in Appendix 5 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1).

In the development cohort, the mean primary care consultation rate was 5.92 per person-year, the unplanned admission rate was 69.5 per 1000 person-years, and the mortality rate was

10.7 per 1000 person-years (Table 3). Similar values were observed for the 2015 validation cohort. The mortality rates were similar for the 2011 validation cohort, whereas admission rates were considerably lower in the 2011 cohort, especially when based on 5-year follow-up. Nearly all patients (93.7%) in the development data set had complete follow-up, with similar proportions for the other 1-year follow-up validation groups; for the 2011 cohort, follow-up was 75.1% complete at 5 years.

Primary care consultation models

The C-index for prediction of primary care consultations in the 2015 validation data set, using a model incorporating the 37-condition weighted multimorbidity score with adjustment for age and sex, was 0.732 (95% CI 0.731–0.734). Comparisons of this model against other models are presented in Appendix 6 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1), and model output is presented in Appendix 7 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). Using the score directly, without additional adjustment for age and sex, resulted in poorer performance (C-index 0.702, 95% CI 0.701–0.704). An adjusted model incorporating each condition as a binary variable performed only slightly better (C-index 0.737, 95% CI 0.736–0.739) than the single weighted score. Performance was only very slightly worse for predicting consultations over 1 year from 2011 (C-index 0.724, 95% CI 0.722–0.725) and was a little better for 5-year prediction (C-index 0.739, 95% CI 0.738–0.740).

Table 1: Characteristics of the 3 data sets used

Characteristic	Data set; % of patients*		
	Development (2015)	Validation (2015)	Validation (2011)
Sex			
Male	48.9	49.4	49.2
Female	51.1	50.6	50.8
Age at index date, yr			
Mean ± SD	50.7 ± 17.6	51.0 ± 17.8	50.1 ± 17.3
Range	21–95	21–95	21–95
> 65 yr	22.7	23.8	20.6
> 80 yr	5.8	6.0	5.1
Deprivation quintiles (IMD 2010)			
1 (least deprived)	25.9	29.8	20.3
2	24.5	19.3	20.3
3	19.1	17.5	20.8
4	16.6	18.0	22.6
5 (most deprived)	13.9	15.4	16.0
Multimorbidity			
No. of conditions, mean ± SD	1.3 ± 1.7	1.3 ± 1.8	1.2 ± 1.7
No. of conditions, range	0–15	0–15	0–15
With 0 conditions	45.0	43.3	46.9
With 1 condition	23.3	23.5	23.4
With ≥ 2 conditions	31.7	33.2	29.7

Note: IMD = Index of Multiple Deprivation, SD = standard deviation.
*Except where indicated otherwise.

Unplanned admission models

The C-index for prediction of unplanned admissions in the 2015 validation data set, based on an adjusted 37-condition weighted score, was 0.742 (95% CI 0.737–0.747; Appendix 8, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). Model output is presented in Appendix 9 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). Performance was only marginally worse when age and sex were excluded (C-index 0.738, 95% CI 0.733–0.744) and was almost identical with an adjusted model incorporating separate conditions (C-index 0.743, 95% CI 0.738–0.748). One-year and 5-year performance using the 2011 data set was similar (C-index 0.739, 95% CI 0.733–0.744) and substantially worse (C-index 0.712, 0.709–0.715), respectively.

Mortality models

Prediction of death in the 2015 validation data set, based on an adjusted score for all 37 conditions weighted by age and sex, was excellent (C-index 0.912, 95% CI 0.905–0.918; Appendix 10, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). Model output is presented in Appendix 11 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). Performance was worse (albeit still very good) with exclusion of age and sex (C-index 0.868, 95% CI 0.857–0.878). An adjusted model incorporating all 37 conditions with age and sex separately performed slightly better (C-index 0.920, 95% CI 0.914–0.926). Performance at both 1 year and 5 years in the 2011 data set was only marginally worse (C-index 0.901, 95% CI 0.894–0.908, and C-index 0.890, 95% CI 0.886–0.894, respectively).

Primary (20-condition) outcome-specific multimorbidity scores

We constructed simplified primary versions of the scores based on the 20 most important conditions. The conditions selected were those with highest average rankings of both prevalence and effect size. This selection of 20 conditions was considered clinically most relevant and was associated with better model performance, relative to selection based on prevalence or effect size alone. Compared with the 37-condition score, model performance was only marginally worse for each outcome (Table 4: for consultations, C-index 0.727, 95% CI 0.725–0.728; for unplanned admissions, C-index 0.738, 95% CI 0.732–0.743; for death, C-index 0.910, 95% CI 0.904–0.917).

Comparison of performance for different outcomes

A multimorbidity score may also be used to predict outcomes for which it was not originally designed. For example, a score weighted on the basis of one particular outcome (e.g., death) may be used to predict a different outcome (e.g., admissions).

Table 2: Prevalence and weights for the 20 conditions in the multimorbidity scores

Condition	Prevalence, %*	Outcome of interest; weight†			
		Primary care consultations‡	Unplanned admissions§	Mortality§	General outcome¶
Hypertension	19.24	0.66	10.76	−2.09	0.08
Anxiety/depression	12.85	2.12	46.61	7.04	0.50
Painful condition	11.63	3.43	84.93	16.46	0.92
Hearing loss	11.27	1.04	8.93	−3.94	0.09
Irritable bowel syndrome	7.61	1.82	8.55	−1.33	0.21
Asthma	7.20	1.32	22.78	−2.73	0.19
Diabetes mellitus	6.58	3.77	55.33	10.23	0.75
Coronary heart disease	4.79	1.49	70.87	4.22	0.49
Chronic kidney disease	4.50	0.98	52.13	16.61	0.53
Atrial fibrillation	2.72	5.94	105.21	22.14	1.34
Constipation	2.67	3.42	72.73	35.42	1.12
Stroke and TIA	2.55	1.54	90.84	20.63	0.80
COPD	2.46	3.43	134.51	42.50	1.46
Connective tissue disorder	2.33	3.10	28.87	−0.39	0.43
Cancer	2.15	2.58	104.80	62.00	1.53
Alcohol problems	1.60	0.97	93.59	12.72	0.65
Heart failure	1.04	2.90	73.20	43.47	1.18
Dementia	1.02	1.81	156.90	124.42	2.50
Psychosis/bipolar disorder	0.98	2.24	77.28	7.20	0.64
Epilepsy	0.97	2.13	113.42	18.26	0.92

Note: COPD = chronic obstructive pulmonary disease, TIA = transient ischemic attack.
*Based on development data set.
†Negative weights can be interpreted as reflecting a negative association with the outcome of interest after controlling for other conditions.
‡Per person-year.
§Per 1000 person-years.
¶Unit change associated with a change of 1 standard deviation in each of the 3 outcomes.

Therefore, we also examined performance for each of the different scores (i.e., consultations, admissions, death) not just against the corresponding outcome but for the alternative outcomes as well (Table 5). In general, all adjusted models predicted death well, with the admissions model performing best (C-index 0.913, 95% CI 0.906–0.919). The consultation and admission models each performed similarly in predicting the alternative outcome. However, the mortality model was notably worse at predicting either consultations (C-index 0.694, 95% CI 0.692–0.696) or admissions (C-index 0.712, 95% CI 0.706–0.717). We also explored the correlation between multimorbidity scores at the person level (Table 6). In particular, this analysis showed the weakest correlation between the consultation- and mortality-based scores (0.777, 95% CI 0.775–0.779) and the strongest correlation between the consultation- and admission-based scores (0.947, 95% CI 0.946–0.947).

Primary general-outcome multimorbidity score

A general (i.e., not outcome-specific) 20-condition score, based on the combined weights, had performance for each of the 3 outcomes similar to that of the outcome-specific models (Tables 4 and 5: consultations, C-index 0.723, 95% CI 0.722–0.725; admissions, C-index 0.735, 95% CI 0.729–0.740; death, C-index 0.913 95% CI 0.907–0.920), with a strong correlation between general-outcome and outcome-specific scores (Table 6).

Comparison against Charlson Comorbidity Index

The Charlson Comorbidity Index, adjusted for age and sex, performed less well than the primary (20-condition) outcome-specific and general-outcome models, for all 3 outcomes, although the performance difference for mortality was minimal (Table 4 and Appendix 12, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). Of note, without adjustment for age or sex, performance dropped relatively more with the Charlson Comorbidity Index across all 3 outcomes, particularly for death.

Model calibration

Calibration plots are presented in Appendix 13 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190757/-/DC1). These show reasonable calibration for mortality and unplanned

Table 3: Descriptive statistics for number of primary care consultations, unplanned hospital admissions, mortality and time of follow-up for each data set

Variable	Development data set* *n* = 300 000	Validation data sets		
		2015, 1-yr follow-up *n* = 150 000	2011, 1-yr follow-up *n* = 150 000	2011, 5-yr follow-up *n* = 150 000
Primary care consultations				
Consultation rate, per person-year	5.92	5.84	5.52	5.68
No. of events in follow-up, mean ± SD	5.7 ± 7.5	5.6 ± 7.1	5.3 ± 6.4	24.5 ± 26.5
Range	0–162	0–237	0–110	0–493
Zero consultations, % of patients	21.9	21.3	20.3	8.7
Unplanned admissions				
No. of first events in follow-up	19 509	10 045	8878	30 075
Rate, per 1000 person-yr	69.5	71.6	63.1	51.8
Mortality				
No. of deaths in follow-up	3106	1558	1505	7087
Mortality rate, per 1000 person-yr	10.7	10.7	10.4	11.0
Follow-up period				
Mean, yr/patient	0.968	0.967	0.967	4.3
No. (%) of patients with complete follow-up	281 150 (93.7)	140 353 (93.6)	140 325 (93.6)	112 586 (75.1)

Note: SD = standard deviation.
*The development data set was based on data from 2015, with 1-year follow-up.

Table 4: Performance of the 3 outcome-specific and general-outcome primary (20-condition) scores in 2 validation data sets*

Outcome	Adjusted score†; C-index value (95% CI)			Unadjusted score; C-index value (95% CI)		
	Cambridge outcome-specific score	Cambridge general-outcome score‡	Charlson Comorbidity Index	Cambridge outcome-specific score	Cambridge general-outcome score‡	Charlson Comorbidity Index
Primary care consultations						
2015, 1 yr	0.727 (0.725–0.728)	0.723 (0.722–0.725)	0.691 (0.690–0.693)	0.692 (0.691–0.694)	0.690 (0.689–0.691)	0.605 (0.603–0.606)
2011, 5 yr	0.735 (0.734–0.736)	0.729 (0.728–0.730)	0.709 (0.708–0.711)	0.669 (0.668–0.671)	0.667 (0.665–0.668)	0.585 (0.583–0.586)
Unplanned admissions						
2015, 1 yr	0.738 (0.732–0.743)	0.735 (0.729–0.740)	0.703 (0.697–0.709)	0.733 (0.728–0.739)	0.731 (0.726–0.737)	0.660 (0.656–0.664)
2011, 5 yr	0.708 (0.705–0.712)	0.706 (0.703–0.709)	0.683 (0.680–0.686)	0.694 (0.691–0.698)	0.692 (0.689–0.695)	0.623 (0.621–0.625)
Mortality						
2015, 1 yr	0.910 (0.904–0.917)	0.913 (0.907–0.920)	0.907 (0.900–0.914)	0.868 (0.857–0.879)	0.880 (0.872–0.889)	0.804 (0.792–0.815)
2011, 5 yr	0.889 (0.885–0.892)	0.891 (0.887–0.894)	0.887 (0.883–0.890)	0.795 (0.788–0.801)	0.824 (0.819–0.830)	0.742 (0.736–0.748)

Note: CI = confidence interval.
*Validation data sets: 2015 data set with 1 year of follow-up and 2011 data set with up to 5 years of follow-up, with the Charlson Comorbidity Index for comparison.
†Scores adjusted by age and sex.
‡The general-outcome score was constructed by averaging the standardized weights of the 3 simple scores.

Table 5: Performance of each primary (20-condition) score on all other outcomes assessed in the 1-year follow-up sample for 2015, with adjustment for age and sex

Outcome	Primary multimorbidity score; adjusted C-index (95% CI)			
	Primary care consultation score	Unplanned admission score	Mortality score	General-outcome score
Primary care consultations	0.727 (0.725–0.728)	0.723 (0.722–0.725)	0.694 (0.692–0.696)	0.723 (0.722–0.725)
Unplanned admissions	0.735 (0.729–0.740)	0.738 (0.732–0.743)	0.712 (0.706–0.717)	0.735 (0.729–0.740)
Mortality	0.906 (0.900–0.913)	0.913 (0.906–0.919)	0.910 (0.904–0.917)	0.913 (0.907–0.920)

Note: CI = confidence interval.

Table 6: Correlation* of the 4 primary (20-condition) scores at the person level

Outcome	Outcome; Pearson correlation coefficient (95% CI)			
	Primary care consultations	Unplanned admissions	Mortality	General outcome
Primary care consultations	1			
Unplanned admissions	0.947 (0.946–0.947)	1		
Mortality	0.777 (0.775–0.779)	0.889 (0.888–0.890)	1	
General outcome	0.950 (0.950–0.951)	0.989 (0.988–0.989)	0.929 (0.929–0.930)	1

Note: CI = confidence interval.
*Weights for each patient were summed, and the within-individual Pearson correlation was calculated between different scores. Results are derived from the 1-year follow-up validation data set obtained in 2015.

admissions, although consultation rates are underestimated (which is to be expected, given that persons with no long-term conditions are still likely to consult their GP from time to time).

Interpretation

In this study, we developed several robust, outcome-specific multimorbidity scores, with acceptable predictive validity for primary care utilization, unplanned hospital admission and death. The primary (simplified) models performed nearly as well as the more complex extended ones, and the general-outcome multimorbidity score performed similarly across all outcomes and over time. The scores outperformed the widely used Charlson Comorbidity Index across all outcomes. Performance was best for the outcome of death, particularly after adjustment for age and sex, and was least good for consultations.

These scores have benefits over commonly used existing measures, including weightings for several outcomes and a pragmatic balance of number and choice of conditions (which in the UK align with those recently proposed for practice multimorbidity registries[19]). A person's score can be calculated by summing the weights of their individual conditions, according to the outcome considered most appropriate for the given context.

Multimorbidity scores offer a means of identifying those patients in the population who are most likely to benefit from a tailored approach to care, helping clinicians to prioritize their efforts accordingly,[20] but they are unlikely to have a direct role in individual patient care. The scores that we have described specifically quantify multimorbidity, as opposed to focusing on the identification of a specific priority problem, such as unplanned admissions (e.g., QAdmissions[21]) or frailty (e.g., the electronic Frailty Index or eFI[22]), and as such may be more relevant to optimizing the delivery of care for those with multimorbidity. Morbidity scores can also inform health policy decision-making, including resource allocation. Patient case-mix, as measured using the comprehensive ACG system, was shown, in a study of Swedish primary care, to explain most of the variance in patient costs.[23] However, in UK primary care, the funding allocation (Carr–Hill) formula does not account for patient morbidity directly.[24] Scores developed through a transparent process, with "real world" contemporary data and weightings that incorporate a range of key outcomes, should help policy-makers and

clinicians to understand and support their use for priority-setting purposes. In addition, multimorbidity scores provide an opportunity to capture clinical complexity and to identify what matters most in general practice, for example, by moving away from the UK's current payment-for-performance Quality and Outcomes Framework (also known as QOF) incentivization system, which is based on individual conditions.[25]

Finally, having a robust method of quantifying multimorbidity facilitates research, including descriptive epidemiologic analysis and matching of individuals on morbidity status. In particular, multimorbidity scores can be added to routine data sets to evaluate how the response to many clinical and health service interventions varies with morbidity. Future work should also be undertaken to explore the utility of the scores in practice, as well as to gain a better understanding of how the scores are associated with other important clinical outcomes such as function, quality of life and experience of care.

Limitations

Although this study has several important strengths — use of contemporary data from a large, representative primary care population, inclusion of a range of pertinent long-term clinical conditions, and evaluation of performance for different years and follow-up periods (which increased confidence in external generalizability and performance over time) — there were also important limitations. Interpreting C-index values involves a measure of judgment in terms of what constitutes an important threshold. Nevertheless, our conclusions are based on conventional standards; in particular, there was a 2%–11% improvement over the Charlson Comorbidity Index. Diagnostic coding in medical records is undertaken for clinical rather than research purposes and is subject to misclassification or missingness.[16] However, this also means that the scores' performance reflects what can be expected in the real world.

We used established UK Read coding rather than the newer international SNOMED-CT system now being introduced in UK practice. Nevertheless, these coding systems can be readily mapped to one another, and most conditions are captured by a small subset of codes; therefore, we believe this limitation is unlikely to have substantially affected our findings. Furthermore, although the models are based on UK data, there is no reason to suspect that the findings would not be generalizable to other, non-UK settings; similar scores such as the Charlson Comorbidity Index have shown international applicability. We compared the performance of our scores only against the Charlson Comorbidity Index, so are unable to claim superiority over or equivalence to alternative metrics.

It is also possible to question the list of conditions that we analyzed. Although our list was based on well-established previous work, 2 morbidities are particularly noteworthy. The use of chronic pain rather than specific musculoskeletal conditions may be questioned, but the former is both common and clinically meaningful and has the advantage of capturing the latter while more readily distinguishing chronic from self-limiting conditions. Constipation might also be viewed as anomalous, but it has a prevalence similar to that of other important conditions,[2] is common among older people and can substantially affect quality of life. Additionally, certain important conditions were not included because they are relatively rare in UK practice (e.g., HIV) or because they were covered by other, broader categories (e.g., opioid use disorder was captured by substance misuse).

A further issue was our omission of several important predictors from the models (e.g., previous health care utilization). However, the aim of our study was to develop not the best risk prediction tools, but rather an optimal approach to describe or adjust for the general health status of individuals in health services and outcomes research. In addition, although we aimed to create a simpler score by minimizing the number of conditions that need to be recorded in practice, we elected not to simplify the weightings (in contrast to the Charlson Comorbidity Index), as these weightings will most likely be implemented using electronic systems. A further advantage is that the weightings are easily interpreted in terms of predicting outcomes on a natural scale.

Finally, the effect of newly diagnosed comorbidities on health care utilization and mortality is likely to be much higher than the effects of longer-term health conditions; further work is required to examine the effect of timing of diagnosis on outcomes.

Conclusion

We have described the development of several robust, simple-to-use multimorbidity scores, some tailored and others not tailored to specific health and health service outcomes. These scores have the potential to be of considerable value for policy development and clinical priority-setting, providing a clinically relevant, pragmatic, transparent and methodologically easy-to-implement means of optimizing the delivery of health care to an aging and increasingly multimorbid population.

References

1. King DE, Xiang J, Pilkerton CS. Multimorbidity trends in United States Adults, 1988–2014. *J Am Board Fam Med* 2018;31:503-13.
2. Cassell A, Edwards D, Harshfield A, et al. The epidemiology of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2018;68:e245-51.
3. Salisbury C. Multimorbidity: time for action rather than words. *Br J Gen Pract* 2013;63:64-5.
4. Marengoni A, Onder G. Guidelines, polypharmacy, and drug-drug interactions in patients with multimorbidity. *BMJ* 2015;350:h1059.
5. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. *Fam Pract* 2013;30:172-8.
6. 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.
7. 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.
8. Zavascki AP, Fuchs SC. The need for reappraisal of AIDS score weight of Charlson Comorbidity Index. *J Clin Epidemiol* 2007;69:867-8.
9. Fortin M, Stewart M, Poitras ME, et al. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012; 10:142-51.
10. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827-36.
11. User guide to mortality statistics: supporting information for mortality statistics, which present figures on deaths registered in England and Wales in a specific week, month, quarter or year. Office for National Statistics; 2019. Available: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsjuly2017 (accessed 2019 Nov. 28).

12. Hospital Episode Statistics (HES). West Yorkshire (UK): NHS Digital; 2019. Available: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics (accessed 2019 Nov. 28).
13. Abel GA, Barclay ME, Payne RA. Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. *BMJ Open* 2016;6:e012750.
14. 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.
15. Khan NF, Perera R, Harper S, et al. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. *BMC Fam Pract* 2010;11:1.
16. *CPRD @ Cambridge – code lists*. Version 1.1. Cambridge (UK): University of Cambridge; 2018. Available: https://www.phpc.cam.ac.uk/pcu/cprd_cam/codelists (accessed 2020 Jan. 17).
17. Herrett E, Thomas SL, Schoonen WM, et al. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010;69:4-14.
18. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
19. NICE quality and outcomes framework indicator. NM184. London (UK): National Institute for Health and Care Excellence (NICE); 2019. Available: www.nice.org.uk/standards-and-indicators/qofindicators/the-practice-can-produce-a-register-of-people-with-multimorbidity-who-would-benefit-from-a-tailored-approach-to-care (accessed 2019 Nov. 28).
20. Multimorbidity: clinical assessment and management. London (UK): National Institute for Health and Care Excellence (NICE); 2016. Available: www.nice.org.uk/guidance/ng56 (accessed 2019 Nov. 28).
21. Hippisley-Cox J, Coupland C. Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. *BMJ Open* 2013;3:e003482.
22. Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing* 2016;45:353-60.
23. Zielinski A, Kronogård M, Lenhoff H, et al. Validation of ACG Case-mix for equitable resource allocation in Swedish primary health care. *BMC Public Health* 2009;9:347.
24. *Financial allocations 2016/17–2020/21*. NHS England; 2016. Available: www.england.nhs.uk/wp-content/uploads/2016/01/allocations-201617-202021.pdf (accessed 2019 Nov. 28).
25. Marshall M, Roland M. The future of the Quality and Outcomes Framework in England. *BMJ* 2017;359:j4681.

Competing interests: None declared.

This article has been peer reviewed.

Affiliations: Centre for Academic Primary Care (Payne), Population Health Sciences, University of Bristol, Bristol, UK; Primary Care Unit (Mendonca, Saunders, Edwards, Roland), Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, UK; RAND Corporation (Elliott), Santa Monica, Calif.; Research Department of Primary Care and Population Health (Marshall), University College London Medical School, Royal Free Campus, London, UK

Contributors: Rupert Payne, Martin Marshall and Martin Roland were responsible for obtaining study funding. The study was conceived by Rupert Payne, Duncan Edwards, Martin Marshall and Martin Roland. Rupert Payne, Silvia Mendonca, Marc Elliott, Catherine Saunders, Duncan Edwards and Martin Roland were involved in the study design, approvals and data acquisition. Silvia Mendonca and Catherine Saunders conducted the analyses, with support from Marc Elliott. All of the authors contributed to interpretation of the findings. Rupert Payne, Silvia Mendonca and Catherine Saunders drafted the manuscript. All of the authors critically reviewed and revised the manuscript for important intellectual content, gave final approval of the version to be published and agreed to be guarantors of the work.

Funding: This paper presents independent research funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR, reference FR10/283). The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health.

Data sharing: The Clinical Practice Research Datalink (CPRD) does not allow the sharing of patient-level data. The structure and format of the CPRD data set is available at: https://cprd.com/sites/default/files/CPRD%20GOLD%20Full%20Data%20Specification%20v2.0_0.pdf. The morbidity code lists used in this study are available at: www.phpc.cam.ac.uk/pcu/cprd_cam/codelists/.

Acknowledgements: The authors would like to thank James Brimicombe, data manager in the Department of Public Health and Primary Care, University of Cambridge, for assistance in coding and management of the CPRD data set.

Disclaimer: This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare Products Regulatory Agency. The data are provided by patients and collected by the NHS as part of their care and support. Linked mortality data were provided by the UK Office for National Statistics through NHS Digital. The interpretation and conclusions contained in this study are those of the authors alone.

Accepted: Dec. 6, 2019

Correspondence to: Rupert Payne, r.payne@bristol.ac.uk