Payne, R. A., Mendonca, S. C., Elliott, M. N., Saunders, C. L., Edwards, D. A., Marshall, M., & Roland, M. (2020). Development and validation of the Cambridge Multimorbidity Score. *CMAJ: Canadian Medical Association Journal*, *192*(5), E107-E114. https://doi.org/10.1503/cmaj.190757

Development and validation of the Cambridge Multimorbidity Score

Rupert A Payne[1], MB ChB PhD

Silvia C Mendonca[2], BSc MSc

Marc N Elliott[3], BA BA MA PhD

Catherine L Saunders[2], PhD

Duncan A Edwards[4], BSc MB BS MPH

Martin Marshall[4], BSc MB BS MSc MD

Martin Roland[2], MA BM BCh DM

1. Centre for Academic Primary Care, Population Health Sciences, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS

2. Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, Institute of Public Health, Forvie Site, Robinson Way, Cambridge, CB2 0SR

3. RAND Corporation, 1776 Main Street, Santa Monica, CA 90401-3208

4. University College London, UCL Research Department of Primary Care & Population Health, UCL Medical School, Upper 3rd Floor, Royal Free Campus, Rowland Hill Street, London, NW3 2PF

Short title: Cambridge Multimorbidity Score

Corresponding author: Dr Rupert Payne, r.payne@bristol.ac.uk

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

Competing interests: The authors have no competing interests.

Contributors' statement: RP, MM and MR were responsible for obtaining study funding. The study was conceived by RP, DE, MM and MR. RP, SM, ME, CS, DE and MR were involved in the study design, approvals and data acquisition. CS and SM conducted the analyses, supported by ME. All authors contributed to interpretation of the findings. RP, CS and SM drafted the manuscript. All authors critically reviewed and revised the manuscript for intellectual content. All authors gave final approval of the final version and agree to be guarantors of the work.

Abstract

Background

Health services have failed to respond to the pressures of multimorbidity. There is a need for improved measures of multimorbidity for research, planning services and resource allocation.

Methods

Development (N=300,000) and validation (N=150,000) data samples were extracted from the UK Clinical Practice Research Database. We modelled the association between 37 morbidities and key outcomes (primary care consultations, unplanned hospitalization, mortality) at one and five years using Cox and zero-inflated negative binomial regression. A general-outcome multimorbidity score was constructed by averaging the standardised weights of the separate outcome scores. Performance was compared with the Charlson co-morbidity index.

Results

Models including all 37 conditions were acceptable predictors of GP consultations, hospitalisation and mortality at one-year (C-indices 0.732 [95% confidence interval 0.731-0.734], 0.742 [0.737-0.747] and 0.912 [0.905-0.918] respectively, adjusted for age/gender). Reducing the models to the 20 conditions which had the greatest combined prevalence/weight made little difference to the predictive value of the models (C-indices 0.727 [0.725-0.728], 0.738 [0.732-0.743] and 0.910 [0.904-0.917] respectively). Prediction of outcomes at five years for the 20-condition model remained similar for consultations and mortality (C-indices 0.735 [0.734-0.736], 0.889 [0.885-0.892]) but performed less well for admissions (C-index 0.708 [0.705-0.712]). The general-outcome score performed similarly to the outcome-specific models. Models performed significantly better than those based on Charlson.

Conclusions

This analysis provides several robust, simple-to-use multimorbidity scores, both tailored and not tailored to specific health outcomes. The scores will be valuable to those planning clinical services, policymakers allocating resources, and researchers seeking to account for the effect of multimorbidity.

Introduction

Patients with multiple long-term health conditions are familiar to clinicians in generalist and specialist settings.[1,2] Services and policies have failed to respond to the pressures multimorbidity places on primary and secondary care. These pressures are driven by the ageing population, by policies which promote rapid access over longer consultations and continuity of care, and by single disease guidelines and performance targets which lead to over-prescribing and fail to address the priorities of multimorbid individuals.[3,4]

Several approaches have been used to quantify multimorbidity. Simple counts of conditions show a clear association with various outcomes, including primary care utilisation, unplanned hospitalisation, and death.[5,6] Weighted approaches allow for differences in the strength of association between specific morbidities and a given outcome, such as the Charlson index which provides a composite morbidity score with condition weightings based on mortality.[7] Although its performance has been shown to exceed a number of other metrics[4], clinical practice has advanced considerably since its' development in the 1980s and the high weightings of particular conditions has been questioned. A further problem with such indices is that weightings are generally based on a specific outcome such as mortality, and may not predict other outcomes. The lists of conditions are also problematic. A minimum list of 12 conditions has been proposed.[8] 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 UK general practitioner records and weighted on different clinical outcomes, that could be used in future studies of multimorbidity and for resource allocation.

Methods

Population and data sources

A retrospective cohort study was undertaken using anonymous coded UK general practice (GP) electronic health record data from the Clinical Practice Research Datalink (CPRD).[9]

We restricted our analysis to the 148 practices contributing data classified by CPRD as "up to standard" from 2010 to 2015, with primary care records linked to national data on mortality (Office for National Statistics, ONS)[10], hospitalisation (Hospital Episode Statistics, HES)[11] and socioeconomic deprivation (Index of Multiple Deprivation, IMD).[12] To ensure independence of samples, practices were randomly sampled into one of three datasets (2:1:1). The development dataset included 300,000 randomly-sampled adults aged ≥20 years registered on 1/1/2014 (study start), with data classified by CPRD as acceptable for use in research. The presence of morbidities was determined at an index date 12-months after the study start (1/1/2015), to ensure at least one year of registration and maximise recording of prevalent cases. Patient records were followed for one year after the index date (study end 31/12/2015). The first validation dataset included 150,000 patients with the same specification as the development dataset. A similar, second validation dataset of 150,000 patients provided up to 5-year follow-up, as well as a 1-year asynchronous follow-up (study start 1/1/2010; index date 1/1/2011; data available until 31/12/2015). The sample size was selected to limit the width of a 95% confidence interval for a condition with 2% prevalence to approximately 0.5 on the log-odds scale for a dichotomous outcome such as mortality. Flow charts for the selection of practices and patients are presented in Appendix S1, with details of the time periods and dates of each dataset given in Appendix S2.

Exposures

Co-morbidities were defined by relevant Read codes and/or prescribing prior to the index date, based on a list of 37 long-term conditions as described by Cassell and colleagues, and adapted from work by Barnett and colleagues which is considered one of the definitive epidemiological studies of

multimorbidity (Appendix S3). [2, 13] The conditions were chosen and defined based on clinical expert consensus as those having a significant impact on patients. The current study and that of Cassell align closely, both aiming to develop better means of quantifying multimorbidity. An additional condition list was included for the Charlson index.[14] The code lists used were subject to considerable clinical attention, and we thus consider all the comorbidities to have face-validity; previous studies of the CPRD dataset have shown the majority of long-term conditions to have positive predictive value in excess of 80%.[15] Gender and age were included as covariates.

Outcomes

ONS data and HES admitted patient care data were used to determine the occurrence of death and unplanned (emergency) inpatient hospitalization respectively during the follow-up period. Number of primary care consultations was established from GP records of face-to-face (including telephone) clinical encounters; multiple encounters per day were counted as one consultation.

Statistical analysis

Morbidity scores were developed using three separate models, one for each outcome, in the 2015 development dataset. Consultations were modelled using zero-inflated negative binomial regression, and mortality and unplanned hospitalisation using Cox regression. In addition to the extended scores containing all 37 conditions, we constructed a set of simplified primary scores including the most important 20 conditions. In addition, a general-outcome multimorbidity score was constructed by averaging the standardised weights of the three simple scores. Details of the statistical modelling (including data cleaning) are provided in Appendix S3.

Performance of each of the three 37-condition and 20-condition outcome-specific scores, as well as the 20-condition general-outcome score, was independently evaluated at 1-year follow-up in the 2015 (synchronous) dataset, as well as at 1-year and 5-years follow-up in the 2011 (asynchronous) dataset. We examined the performance of each score for predicting each of the three outcomes,

and additionally compared performance against the Charlson index. Model fit was assessed using Harrell's C-index.

Results

The characteristics of the different cohorts is shown in Table 1, and descriptive statistics for the multimorbidity scores are presented in Appendix S4. The development cohort had a mean age of 50.7 years, with 23% and 6% aged over 65 years and 80 years respectively. 51% of patients were female. The most socioeconomically deprived were under-represented. The mean (standard deviation) number of morbidities was 1.3 (1.7), with 31.7% of individuals having 2 or more recorded conditions. The commonest conditions (Table 2) were hypertension (19.2%), anxiety or depression (12.9%), painful condition (11.6%) and hearing loss (11.3%); the full list of disease prevalence and score weightings is provided in Appendix S5. In general, similar patterns of age, gender, socioeconomic deprivation and multimorbidity were observed across all cohorts (Table 1).

In the development cohort, mean consultation rate was 5.9 per person-year, mortality 10.7 per 1000 person-years, and unplanned admission rate 69.5 per 1000 person-years (Table 3). Similar figures were observed for the 2015 validation cohort. Mortality rates were similar for the 2011 validation cohort, although admission rates in particular were considerably lower in the 2011 cohorts, especially when based on 5-year follow-up. 93.7% of patients had complete follow-up in the development dataset, with similar numbers for the other 1-year follow-up validation groups; follow-up was 75.1% complete at 5-years for the 2011 cohort.

Consultation models

The C-index for prediction of consultations in the 2015 validation dataset, using a model incorporating the 37-condition weighted multimorbidity score and adjusting for age and gender, was 0.732 (95% confidence interval, 95% CI 0.731-0.734). Comparison of this model against other models is presented in Appendix S6, and model output presented in Appendix S7. Using the score directly, without additionally adjusting for age and gender resulted in poorer performance (C=0.702, 95% CI 0.701-0.704). An adjusted model incorporating each condition as a binary variable performed only slightly better (C=0.737, 95% CI 0.736-0.739) than using the single weighted score. Performance was

only very slightly worse for predicting consultations over 1 year from 2011 (C=0.724, 95% CI 0.722-0.725), and a little better for 5-year prediction (C=0.739, 95% CI 0.738-0.740).

Unplanned admission models

The C-index for prediction of unplanned admissions in the 2015 validation dataset, based on an adjusted 37-condition weighted score, was 0.742 (95% CI 0.737-0.747; Appendix S8). Model output is presented in Appendix S9. Performance was only marginally worse when age and gender were excluded (C=0.738, 95% CI 0.733-0.744), and almost identical to an adjusted model incorporating separate conditions (C=0.743, 95% CI 0.738-0.748). 1-year and 5-year performance using the 2011 dataset was similar (C=0.739, 95% CI 0.733-0.744) and substantially worse (C=0.712, 0.709-0.715) respectively.

Mortality models

Prediction of mortality in the 2015 validation dataset, based on an adjusted score for all 37-conditions weighted by age and gender, was excellent (C=0.912, 95% CI 0.905-0.918; Appendix S10). Model output is presented in Appendix S11. There was worse (albeit still very good) performance upon excluding age and gender (C=0.868, 95% CI 0.857-0.878). An adjusted model incorporating all 37 conditions with age and gender separately performed slightly better (C=0.920, 95% CI 0.914-0.926). Performance at both 1-year and 5-year in the 2011 dataset was only marginally worse (C=0.901, 95% CI 0.894-0.908 and C=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. Selection of conditions was based upon those conditions ranking highest in terms of average ranking of both prevalence and effect size. This selection of 20 conditions was considered clinically most relevant and associated with better model performance, in comparison to selection based on prevalence or effect size alone. Compared to the 37-condition score, model performance

was only marginally worse for each outcome (Table 4: consultations, C=0.727 (95% CI 0.725-0.728); admissions, C=0.738 (95% CI 0.732-0.743); mortality, C=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 based on one particular outcome (e.g. mortality) used to predict a different outcome (e.g. admissions). Therefore we also examined performance for each of the different scores (i.e. mortality, admissions, consultations) not just against the corresponding outcome but for the alternative outcomes as well (Table 5). In general, all adjusted models predicted mortality well, with the admissions model performing best (C=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=0.694, 95% CI 0.692-0.696) or admissions (C=0.712, 95% CI 0.706-0.717). We also explored the correlation in multimorbidity scores at the person level (Table 6). In particular, this 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 admissions and consultation-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 similar performance for each of the three outcomes as the outcome-specific models (Tables 4 and 5: consultations, C=0.723 (95% CI 0.722-0.725); admissions, C=0.735 (95% CI 0.729-0.740); mortality, C=0.913 (95% CI 0.907-0.920)), with a strong correlation between general-outcome and outcome-specific scores (Table 6).

Comparison against Charlson index

The Charlson index, adjusted for age and gender, performed less well than the primary (20-condition) outcome-specific and general-outcome models, for all three outcomes (Table 4), although the performance difference for mortality was minimal (Appendix S12). Of note, when no adjustment for age or gender was carried out, performance dropped relatively more with Charlson across all three outcomes, particularly for mortality.

Model calibration

Calibration plots are presented in Appendix S13. These show reasonable calibration for mortality and emergency admissions, although consultation rates are underestimated (to be expected as persons with no long-term conditions are still likely to consult their GP at times).

Discussion

This study has developed several robust, outcome-specific multimorbidity scores, with acceptable predictive validity for primary care utilisation, unplanned hospitalisation and mortality. The primary (simplified) models perform nearly as well as the more complex extended ones, and the general-outcome multimorbidity score performs similarly across all outcomes and over time. The scores outperform the widely used Charlson index across all outcomes. Performance is best for mortality, particularly following adjustment for age and gender, and least good for consultations.

Policy and research relevance

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.[16] 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 prioritise their efforts accordingly,[17] but are unlikely to have a direct role in individual patient care. The score we have described specifically quantifies multimorbidity, as opposed to focusing on the identification of a specific priority problem such as unplanned admissions (e.g. QAdmissions[18]) or frailty (e.g. electronic Frailty Index, eFI[19]), and as such may be more relevant to optimising 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, has been shown to explain most of the variance in patient costs in a study of Swedish primary care.[20] However, in UK primary care, the funding allocation (Carr-Hill) formula does not account for patient morbidity directly.[21] Scores developed via a transparent process, with "real world" contemporary data and weightings incorporating a range of key outcomes, should help policy-makers and clinicians

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 single-condition based payment-for-performance QOF incentivisation system.[22]

Finally, having a robust method of quantifying multimorbidity facilitates research, including descriptive epidemiological analysis and matching individuals on morbidity status. In particular, multimorbidity scores can be added to routine datasets 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 better understanding how the scores are associated with other important clinical outcomes such as function, quality of life and experience of care.

Strengths and limitations

The study has several important strengths, utilising contemporary data from a large, representative primary care population, and including a range of pertinent long-term clinical conditions. Performance was evaluated for different years and follow-up periods, increasing confidence in external generalizability and performance over time. There are also important limitations. Interpreting C-index values involves a measure of judgement in terms of what constitutes an important threshold. Nevertheless, our conclusions are based on conventional standards, and in particular there is an order of magnitude improvement (2-11%) over Charlson. Diagnostic coding in medical records is undertaken for clinical rather than research purposes, and is subject to misclassification or missingness.[15] However, this also means the scores' performance reflects that expected in the real world. We used established UK Read coding rather than the newer international SNOMED-CT being introduced currently in UK practice. Nevertheless, these can be readily mapped to one another, and most conditions are captured by a small subset of codes so we believe this is unlikely to have significantly impacted upon findings. Furthermore, although the models are based

on UK data, there is no reason to suspect that the findings do not generalize to other non-UK settings; similar scores such as Charlson have demonstrated international applicability. We have only conducted a comparison of performance against Charlson, so are unable to claim superiority or equivalence to alternative metrics. It is also possible to question the list of conditions. Although based on well-established previous work, two morbidities are particularly noteworthy. Firstly, the use of chronic pain rather than specific musculoskeletal conditions: the former is both common and clinically meaningful, and has the advantage of capturing the latter whilst more-readily distinguishing chronic from self-limiting conditions. Secondly, constipation might also be viewed as anomalous, but has a similar prevalence to other important conditions,[2] is common in older people, and can significantly impact quality of life. A further issue was our omission of several important predictors from the models (e.g. previous healthcare utilization). However, the aim of the study was not to develop 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 minimising the number of conditions requiring recorded in practice, we elected not to simplify the weightings (cf. Charlson index) as these will most likely be implemented using electronic systems. A further advantage that the weightings are easily interpreted in terms of predicting outcomes on a natural scale. Finally, the impact of newly diagnosed co-morbidities on healthcare utilisation and mortality is likely to be much higher than longer-term health conditions; further work is required to examine the impact of timing of diagnoses on outcomes.

Conclusions

In conclusion, we have described the development of several robust, simple-to-use, multimorbidity scores, both tailored and 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 optimising the delivery of healthcare to an ageing and increasingly multimorbid population.

Acknowledgement. We 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 dataset

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

Data sharing: CPRD does not allow the sharing of patient-level data. The data structure/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 by this study are available at
www.phpc.cam.ac.uk/pcu/cprd_cam/codelists/

Table 1

Characteristics of the 3 samples used.

	Development dataset (2015)	Validation dataset (2015)	Validation dataset (2011)
Gender			
% male	48.9	49.4	49.2
% female	51.1	50.6	50.8
Age at index date			
Mean (SD), years	50.7 (17.6)	51.0 (17.8)	50.1 (17.3)
Range, years	21-95	21-95	21-95
% over 65	22.7	23.8	20.6
% over 80	5.8	6.0	5.1
Deprivation quintiles (IMD2010), %			
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			
Mean (SD), number of conditions	1.3 (1.7)	1.3 (1.8)	1.2 (1.7)
Range, number of conditions	0-15	0-15	0-15
% 0 conditions	45.0	43.3	46.9
% 1 condition	23.3	23.5	23.4
% 2 or more conditions	31.7	33.2	29.7

SD, standard deviation; IMD, index of multiple deprivation

Table 2

Prevalence and weights[1] for the 20 conditions in the multimorbidity scores.

	Prevalence[2]	Weight for consultations[3]	Weight for mortality[4]	Weight for emergency admissions[4]	General-outcome weight[5]
Hypertension	19.24	0.66	-2.09	10.76	0.08
Anxiety/Depression	12.85	2.12	7.04	46.61	0.50
Painful condition	11.63	3.43	16.46	84.93	0.92
Hearing loss	11.27	1.04	-3.94	8.93	0.09
Irritable bowel syndrome	7.61	1.82	-1.33	8.55	0.21
Asthma	7.20	1.32	-2.73	22.78	0.19
Diabetes	6.58	3.77	10.23	55.33	0.75
Coronary heart disease	4.79	1.49	4.22	70.87	0.49
Chronic kidney disease	4.50	0.98	16.61	52.13	0.53
Atrial fibrillation	2.72	5.94	22.14	105.21	1.34
Constipation	2.67	3.42	35.42	72.73	1.12
Stroke & TIA	2.55	1.54	20.63	90.84	0.80
COPD	2.46	3.43	42.50	134.51	1.46
Connective tissue disorder	2.33	3.10	-0.39	28.87	0.43
Cancer	2.15	2.58	62.00	104.80	1.53
Alcohol problems	1.60	0.97	12.72	93.59	0.65
Heart failure	1.04	2.90	43.47	73.20	1.18
Dementia	1.02	1.81	124.42	156.90	2.50
Psychosis/bipolar disorder	0.98	2.24	7.20	77.28	0.64
Epilepsy	0.97	2.13	18.26	113.42	0.92

TIA, transient ischaemic attack; COPD, chronic obstructive pulmonary disease

1. Negative weights can be interpreted as reflecting a negative association with the outcome of interest after controlling for other conditions
2. Based on development dataset
3. Per person-year
4. Per 1,000 person-years
5. Unit change associated a 1 standard deviation change in each of the three outcomes

Table 3

Descriptive statistics for number of consultations, emergency hospital admissions, mortality and time of follow up in each cohort.

	Development dataset	Validation datasets		
	2015, 1-year follow-up	2015, 1-year follow-up	2011, 1-year follow-up	2011, 5-year follow-up
Number of consultations				
Consultation rate (per person-year)	5.92	5.84	5.52	5.68
Mean N events in FU (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	21.9	21.3	20.3	8.7
Mortality				
Number of deaths in FU	3,106	1,558	1,505	7,087
Mortality rate (per 1,000 person-years)	10.7	10.7	10.4	11.0
Emergency admissions				
Number of first events in FU	19,509	10,045	8,878	30,075
Rate (per 1,000 person-years)	69.5	71.6	63.1	51.8
Follow-up time				
Mean (years per patient)	0.968	0.967	0.967	4.3
N (%) patients with complete follow-up	281,150 (93.7)	140,353 (93.6)	140,325 (93.6)	112,586 (75.1)

FU, follow-up; SD, standard deviation

Table 4

Performance (C-index) of the 3 outcome-specific and general-outcome primary (20-condition) scores on two validation datasets: the 2015 dataset with up 1 year of follow-up and the 2011 dataset with up to 5 years of follow-up. Comparison against the Charlson score is provided. 95% confidence intervals shown in parentheses.

	Scores adjusted by age and gender			Unadjusted scores		
	Cambridge outcome-specific score	Cambridge general-outcome score[1]	Charlson score	Cambridge outcome-specific score	Cambridge general-outcome score	Charlson score
Consultations						
2015 1 year	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 years	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)
Mortality						
2015 1 year	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 years	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)
Emergency admissions						
2015 1 year	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 years	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)

1. *The general-outcome score was constructed by averaging the standardised weights of the three simple scores.*

Table 5

Performance (C-index) of each primary score on all other outcomes. Performance was assessed in the 2015 1-year follow-up sample including adjustment for age and gender. 95% confidence intervals shown in parentheses.

		How does the score for…			
		N consultations	Mortality	Emergency admissions	General-outcome
…perform on the outcome…	N consultations	0.727 (0.725-0.728)	0.694 (0.692-0.696)	0.723 (0.722-0.725)	0.723 (0.722-0.725)
	Mortality	0.906 (0.900-0.913)	0.910 (0.904-0.917)	0.913 (0.906-0.919)	0.913 (0.907-0.920)
	Emergency admissions	0.735 (0.729-0.740)	0.712 (0.706-0.717)	0.738 (0.732-0.743)	0.735 (0.729-0.740)

Table 6

Correlation of the four primary scores at the person level. 95% confidence intervals shown in parentheses.*

	N consultations	Mortality	Emergency admissions	General-outcome
N consultations	1			
Mortality	0.777 (0.775-0.779)	1		
Emergency admissions	0.947 (0.946-0.947)	0.889 (0.888-0.890)	1	
General-outcome	0.950 (0.950-0.951)	0.929 (0.929-0.930)	0.989 (0.988-0.989)	1

* Weights for each patient summed, and the within-individual Pearson correlation calculated between different scores. Results correspond to first validation sample.

References

1 King DE, Xiang J, Pilkerton CS. Multimorbidity Trends in United States Adults, 1988–2014. Journal of the American Board of Family Medicine 2018;31:503-514

2 Cassell A, Edwards D, Harshfield A, Rhodes K, Brimicombe J, Payne R, Griffin S. The epidemiology of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract. 2018;68:e245-e251

3 Salisbury C. Multimorbidity: time for action rather than words. Br J Gen Pract. 2013;63:64–65

4 Marengoni A, Onder G. Guidelines, polypharmacy, and drug-drug interactions in patients with multimorbidity. A cascade of failure. BMJ 2015;350:h1059

5 Brilleman S, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. Family Practice. 2013;30:172-8

6 Payne RA, Abel GA, Guthrie B, Mercer SW. The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study. Canadian Medical Association Journal. 2013;185:E221-8

7 Charlson M, Pompeiu P, Ales K, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83

8 Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. Annals of Family Medicine 2012;10:142-151

9 Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44:827-36

10 Office for National Statistics. User guide to mortality statistics. 2017. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsjuly2017

11 NHS Digital. Hospital Episode Statistics (HES). 2019. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics

12 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

13 Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet. 2012;380:37-43

14 Khan, N.F., et al., Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. BMC Family Practice 2010;11:1

15 Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. Br J Clin Pharmacol. 2010;69:4-14

16 National Institute for Health and Care Excellence (NICE). Quality and Outcomes Framework indicator NM184. 2019. https://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

17 National Institute for Health and Care Excellence (NICE). Multimorbidity: clinical assessment and management. 2016. https://www.nice.org.uk/guidance/ng56

18 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

19 Clegg A, Bates C, Young J, Ryan R, Nichols L, Teale E, Mohammed MA, Parry J, Marshall T. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age Ageing. 2016;45:353-60

20 Zielinski A, Kronogård M, Lenhoff H, Halling A. Validation of ACG Case-mix for equitable resource allocation in Swedish primary health care. BMC Public Health. 2009;9:347

21 NHS England. Financial Allocations 2016/17-2020/21 https://www.england.nhs.uk/wp-content/uploads/2016/01/allocations-201617-202021.pdf

22 Marshall M, Roland M. The future of the Quality and Outcomes Framework in England. BMJ. 2017;359:j4681