An exploration of the relationship between multimorbidity and use of primary health care resources in England[¶]

Sam Brilleman[*] Hugh Gravelle[**] Sandra Hollinghurst[*] Sarah Purdy[*]
Chris Salisbury[*] Frank Windmeijer[***]

Abstract

We investigate the effect of multimorbidity on patients' primary care costs (consultations, drugs, tests) using a sample of 86,100 individuals in 174 English practices. Deprivation and practice style explain little of the variation in practice costs. Age and gender are important but their effect is greatly reduced when measures of multimorbidity are included in the regression model. We consider four measures of multimorbidity. A count of QOF conditions performs better than a set of 68 mutually exclusive Adjusted Clinical Group (ACG) morbidity types and better than inclusion of 17 separate indicators for QOF diseases, but slightly worse than a count of Extended Diagnostic Cluster (EDC) morbidity types. A version of the Charlson Index score has the worst performance but still improves markedly on models containing only age, gender, deprivation and practice effects.

1 Introduction

Almost all major chronic diseases[1] are more common with increasing age and also show a socio-economic gradient in their incidence and prevalence. As the population ages, an increasing proportion of health care provision is likely to relate to these conditions and the number of people with multiple problems - multimorbidity - is likely to increase (Starfield 2006). Descriptive studies have identified trends in incidence and prevalence among different populations (Uijen et al. 2008), others have related multimorbidity to quality of life (Fortin et al. 2004) and to healthcare utilisation (Westert et al. 2001). It has been suggested that caring for patients with multiple problems may require resources over and above those needed to treat each condition individually. This has implications for the financing of health care providers.

A number of indices of multimorbidity have been derived with the main aims of describing prevalence and predicting outcomes. Most of these have been developed in the United

---

[¶] PRELIMINARY. PLEASE CHECK WITH AUTHORS BEFORE CITING

[*] Academic Unit of Primary Care, University of Bristol.
[**] Centre for Health Economics, University of York.
[***] Department of Economics, University of Bristol.

[1] Chronic disease has been defined as a condition which displays "…one or more of the following characteristics: are permanent; leave residual disability; are caused by nonreversable pathological alteration; require special training of the patient for rehabilitation; may be expected to require a long period of supervision, observation, or care" (Uijen et al. 2008).

States (US) where they have been used in risk adjustment schemes for predicting costs and setting insurance premiums. Two measures that are widely used are the Charlson index (Charlson et al. 1987) and the John Hopkins Adjusted Clinical Groups (ACG) Case-Mix System (Johns Hopkins Bloomberg School of Public Health 2008). The Charlson Index is a diagnosis-based measure that weights diseases on the basis of strength of association with mortality. It has been studied extensively and used for a variety of purposes in its original and adapted forms. The John Hopkins ACG Case-Mix System is also diagnosis-based and was developed using administrative claims data in the US (Starfield et al. 1991; Weiner et al. 1991). It uses categories that are clinically based but operates on the principle that all patients in any category use similar amounts of resources. The system has been validated in a number of studies in the US and has been studied in primary care in, for example, Sweden (Halling, Fridh, & Ovhed 2006) and Canada (Reid et al. 2002). Further studies have used the ACG system to predict hospital referrals and prescribing rates in the United Kingdom (UK) (Sullivan et al. 2005; Omar et al. 2008). Other diagnosis-based measures of varying complexity, scope and predictive ability include the Index of Coexisiting Disease (ICED) (Greenfield et al. 1993), useful for measuring mortality and disability, the Kaplan Index (Kaplan et al. 1974) developed for use in diabetes and the Cumulative Illness Rating Scale (CIRS) (Linn, Linn, & Gurel 1968) which focuses more on body systems than specific diagnoses. Medication-based measures have also been used to identify multimorbidity.

None of the established measures of multimorbidity has been fully tested as a tool for predicting healthcare costs in a UK primary care setting. Recent changes in the way primary care is managed and financed has already led to more emphasis on budgetary management and the planned changes in commissioning will stimulate an even greater interest in being able to predict accurately expenditure in primary care. There is, therefore, value in having a model of patient level cost of primary care that will predict expenditure of individual patients within practices in order to align funding more closely to patient costs. Such models are likely to include measures of multimorbidity, but measures of multimorbidity developed in countries with very different institutional and financial arrangements and in non-primary care settings may not be useful in UK primary care.

In this study we address three questions:

1. How useful are measures of multimorbidity in predicting costs in UK primary care when used in the absence of other measures of morbidity for particular conditions?
2. How do different measures compare; are sophisticated measures more predictive than simple measures based on disease counts?
3. Are better predictions obtained using multimorbidity measures which reflect the interactions between diseases than by taking account of the diseases separately?

We attempt to answer these questions by using information on morbidity up to 31/3/2007 to predict primary care expenditure in 2007/8 for patients from a sample of patients in English practices.

2 Methods

2.1 Data

2.1.1 Sample

An initial random patient sample aged 18 years and over was drawn from the 182 practices included in the General Practice Research Database (GPRD) which had 'research standard' data continuously from 1st April 2005 to 31 March 2008, and which had given consent to link their data to measures of area deprivation. We dropped 8 practices with entirely missing deprivation data. In order to use the most up-to-date resource use data and the largest possible observation period for diagnosing multimorbidity, we included those 86,100 individuals alive and registered at one of the remaining 174 practices on 1st April 2007. The original sample was stratified by age, gender and practice. The GPRD is considered broadly representative of the general population in the UK (Lawrenson, Williams, & Farmer 1999).

2.1.2 Choice of indices of multimorbidity

The health of individuals can be characterised by a vector of indicators for different diagnoses and the diagnoses can be used to explain their healthcare expenditure. The most complete (saturated) model would investigate the effects of each morbidity type which is defined by all possible combinations of diagnoses. But with anything other than a very coarse set of diagnoses, including all such morbidity types is infeasible. Thus we require some means of aggregating the information in the diagnosis vector into more manageable descriptions of morbidity. One possibility is to use individual diagnostic indicators. With the patient datasets now available it is possible to estimate models where the number of diagnoses is large (PBRA Team, 2010). But such models will not allow for the possibility that diagnoses interact: the cost of two patients one with diabetes and the other with depression will differ from the cost of one patient with both diabetes and depression. Such interactions may arise because of the characteristics of the particular diagnoses or they arise generally if say consultation costs do not vary linearly with the number of diagnoses with which a patient presents. We therefore investigated a set of multimorbidity measures which combine information on underlying diagnoses in different ways and address different aspects of the interactions between diagnoses.

We reviewed the previous use of established indices of multimorbidity to determine which might be appropriate to a UK primary care setting. The Charlson Index and the John Hopkins ACG system were included because they are widely used and potentially straightforward to

operationalise with routine data. The ICED, the Kaplan Index and the CIRS, however, are clinically driven and less likely to be relevant in predicting primary care cost. Medication count was considered for inclusion as a measure as it has proved to be a useful predictor of resource use in some settings (Perkins et al. 2004) but we felt it would not be a suitable management tool because of its discretionary nature and the potential for manipulation. We also included a simple count of chronic diseases, based on the Quality and Outcomes Framework (QOF).

2.1.3 Measuring multimorbidity

Using all historic diagnoses on patient general practice records up until 1st April 2007 we constructed four distinct measures of multimorbidity as shown in **Table 1**.

Number of QOF diseases. We counted the number of diseases an individual had been previously diagnosed with from a list of 17 chronic diseases included in the 2006/7 version of the Quality and Outcomes Framework (QOF). These were asthma, atrial fibrillation, cancer, coronary heart disease (CHD), chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD), dementia, depression, diabetes, epilepsy, heart failure, hypertension, learning difficulties, mental health, obesity, stroke, and hyperthyroidism. We excluded smoking and palliative care. We used the QOF Business Rules Version 11 which outlines the clinical READ Version 2 codes and any additional criteria required to include a patient on the relevant QOF disease register.

Charlson Index score. A Charlson Index score was calculated for each individual using the adaption contained in Khan et al. 2010. The Charlson Index is a weighted score which includes 17 chronic diseases associated with inpatient mortality. A higher disease weighting reflects a stronger association with the risk of death. The diseases included in the Charlson Index score, with associated weight in parentheses, were cerebrovascular disease (1), chronic pulmonary disease (1), congestive heart disease (1), dementia (1), diabetes (1), mild liver disease (1), myocardial infarction (1), peptic ulcer disease (1), peripheral vascular disease (1), rheumatological disease (1), cancer (2), diabetes with complications (2), hemiplegia and paraplegia (2), renal disease (2), moderate or severe liver disease (3), AIDS (6), and metastatic tumour (6). Khan et al. provides the clinical READ Version 2 codes for diagnosing each index disease, and is based on a translation from the widely used Deyo adaption of the Charlson Index for ICD-9 codes (Deyo, Cherkin, & Ciol 1992).

EDC count. The John Hopkins University ACG Case-Mix System is software designed in the US using secondary care data and originally intended for use in the US healthcare insurance market. It has recently been expanded to allow the input of UK-based READ codes. The software provides a large range of output; we use two measures in our analysis which most closely relate to the concept of multimorbidity. Expanded Diagnostic Clusters (EDCs) are groupings of diagnostic codes which are clinically similar. An individual was assigned to an

EDC if any of the READ Version 2 diagnostic codes relating to the specific EDC appeared in their clinical data. We counted the number of chronic EDCs an individual was included in (Salisbury et al. 2011).

Adjusted Clinical Groups (ACGs) are mutually exclusive categories, defined by patterns of morbidity, age and gender. Individuals in an ACG are assumed to have similar types of co-morbidity and expected healthcare resource demands. We used the ACG software to place every patient into one of 68 mutually exclusive ACGs.

2.1.4 Costing of primary care resources

We estimated the total cost of primary care resources used by each patient during the year 1st April 2007 to 31st March 2008. This included consultations, prescription drugs, and tests initiated within the primary care setting. All costs were valued in £ sterling at 2007/08 prices, adjusted for inflation where necessary.

Consultations included all face-to-face (including surgery consultation, follow-up/routine visit, clinic, out of hours) and telephone consultations. The unit cost of each consultation was based on a combination of both consultation type and primary staff role (type of general practitioner, nurse or other specialist leading the consultation). The unit costs for consultations are shown in **Table 2**. An additional cost was attached to cover administrative activities such as the recording of results or sending mail to a patient when this was performed by a receptionist, administrator, or secretary. Unit costs were taken from Curtis (Curtis 2008) and GP Earnings and Expenses (Technical Steering Committee 2010).

Unit costs for prescription drugs were based on standardised unit cost information provided by the GPRD. This combines data from several sources, including the National Drug Tariff for information on generic products, and branded product information sourced from the manufacturer. Each prescription drug observed in the patient level data was matched to cost using drug name, strength and formulation. In situations where there were multiple unit costs for a single prescription drug the median unit cost was assumed.

With advice from a general practitioner member of the research team (SP) we determined which tests were performed within a standard surgery consultation and applied a zero unit cost, save for the cost associated with any consumables such as a pregnancy test or urine dipstick. Unit costs for the remaining tests were based on the National Health Service (NHS) Reference Costs (Department of Health 2009). For laboratory tests the unit cost was based on pathology discipline. For hospital-based tests and investigations requested through the primary care setting we matched directly to entries in the NHS Reference Costs.

2.1.5 Covariates

Age, at 1st April 2007, was categorised into ten-year age bands with the exception of 90+ years as the upper category. Deprivation was measured by deciles of the Index of Multiple Deprivation (IMD) 2007 (Department for Communities and Local Government 2008) which is based on seven dimensions of deprivation and was attributed to the individuals by their Lower Layer Super Output Area (LSOA).

2.1.6 Modelling

We used Generalised Linear Models (GLMs) assuming a log link ($\ln Ec = \beta x$) and a gamma distribution (errors increase with the square of expected cost Ec) to relate total patient cost to age, gender, deprivation, and multimorbidity. The log form allows for the right skewness of the patient cost data and because we use a GLM specification, rather than an OLS with log dependent variable, we do not have to correct for retransformation bias (Manning 1998) or the fact that a proportion of patients have zero cost.

We included an indicator for the general practice the individual was registered with on 1 April 2007 to reflect idiosyncratic unobserved practice factors including GP 'practice styles'. We also included an exposure adjustment which corrected for any individuals who deregistered from the general practice during the resource use year.

We estimated separate models using the four multimorbidity measures and the covariates. We used 68 ACG categories and the three numerical multimorbidity measures (QOF count, Charlson Index, EDC count) were included in their models as categories to give a fully flexible non-linear relationship between multimorbidity and cost. In our dataset, the maximum QOF count was 10, the maximum Charlson score was 13, and the maximum EDC count was 27. As there were few patients with large numerical scores we used 6 categories for the QOF count (1,2,..,6 or more), 7 for the Charlson (1,2,…,7 or more) and 18 for the EDC count (1,2,…,18 or more).

We compared the models using the Akaike Information Criterion (AIC) (Akaike 1974) and Bayesian Information Criterion (BIC) (Schwarz 1978), for which a smaller value provides indication of a better model fit, the mean absolute prediction error (MAPE), and a deviance-based R-squared measure which may be interpreted as the fraction of empirical uncertainty in total patient cost which has been explained by the model (Cameron & Windmeijer 1997).

For analysing marginal effects, we predicted mean patient cost at each level or category of multimorbidity, evaluated at the sample means of all other covariates (age, gender, and deprivation) and using the average practice fixed effect.

2.1.7 Sensitivity analyses

We examined the sensitivity of the regression model estimates to trimming of the upper 1% of total patient costs.

We also investigated the reliability of our methods for calculating prescribing costs. There is some uncertainty about the consistency of general practitioners' recording of total quantity of drugs prescribed. As a primary method, we attached costs to the prescription drug data using a per unit basis. For example, per tablet for a tablet formulation, and per millilitre for a liquid formulation. In order to validate the results from the analysis, we also attached costs to the prescription drug data using a second method and compared the results. The prescription cost analysis (PCA) (NHS Information Centre 2008) provides information on the net ingredient cost of all prescriptions dispensed in the community in England, and information is listed using a British National Formulary (BNF) code (British Medical Association and Royal Pharmaceutical Society 2010). It was therefore possible to attach the average cost of an average prescription item dispensed in the community under the related BNF code. When using this method, the specific quantity actually prescribed on a given prescription is ignored and the average prescribed quantity is assumed.

Finally we considered the impact of dropping the exposure term for patients who deregistered during the resource use year.

2.1.8 Multimorbidity and specific disease

We undertook a preliminary investigation of whether, in the context of English general practice, the multimorbidity measures add anything to the ability to predict primary care costs once account is taken of separate specific chronic disease indicators. Multimorbidty measures may improve on the predictive ability of separate chronic disease indicators because there may be interactions in the effects of different diseases on primary care costs. A patient with say diabetes and depression may require more GP time than two separate patients with diabetes and depression because it is more difficult to control blood sugar levels in patients with mental health problems. On the other hand there may be cost savings if the same consultation can be used to deal with two problems in the same patient.

We therefore considered a model using both the separate QOF disease indicators and the QOF count multimorbidity indicator. We compared the predictive ability of models with and without the 17 QOF indicator variables and with and without the QOF count multimorbidity indicator.

2.1.9 Software

STATA Version 11.2 (StataCorp. 2009) was used for the statistical analysis. John Hopkins ACG System Version 8.2 was used to obtain the EDC and ACG classifications.

3 Results

3.1 Summary statistics

There were 86,100 individuals alive and registered at the beginning of the resource use year. **Table 3** gives the frequency distribution by age and gender alongside the mean, standard deviation and median number of consultations, prescriptions, and tests per individual for the resource use year. Also included is the percentage of individuals who had at least one consultation, prescription, or test. Across all areas of resource use, rates were higher among females than males, and there were large increases in resource use with increasing age. The mean number of consultations and prescriptions is similar to those in other UK datasets for the same period (NHS Information Centre 2008).

Table 4 gives the mean total cost per patient, stratified by age and gender, categorised by consultations, prescription drugs, and tests and investigations. The mean total patient cost for males ranged from £92 in the 20-29 years age category to £871 in the 90+ years age category. For females the mean total patient cost ranged from £183 in the 20-29 years age category to £745 in the 80-89 years age category. The proportional contribution of prescription drugs to total patient cost increased with age and was higher for males than for females. Prescription drug costs accounted for around one-half of mean total cost for the male 20-29 age category, compared with around three-quarters of mean total cost for the male 90+ age category. For females, prescription drugs costs accounted for around one-third of mean total cost for the 20-29 age category, compared with around two-thirds for the 90+ age category. The proportional contribution of consultations decreased as age increased, whereas the relative cost of tests stayed fairly stable.

Figure 1 shows the mean total patient cost, separately, by age category and deprivation decile. Mean total patient cost increases from £136 (95% confidence interval (CI): £130 to £141) in the 20-29 age category to £784 (95% CI: £747 to £821) in the 80-89 age category. The difference in mean total patient cost between the least and most deprived deciles is relatively smaller than across age categories. Mean total patient cost for the least deprived is £312 (95% CI: £286 to £338) compared with £399 (95% CI: £384 to £415) for the most deprived.

Figure 2 shows the frequency distribution for the QOF chronic disease count, Charlson Index score, EDC count, and ACG categories. The distribution of the EDC count has a larger range than the QOF chronic disease count and Charlson Index score. This is due to the greater number of relatively more minor diseases that the EDC count includes. There are 20%, 12%, and 64% of individuals with a score of two or more for the QOF chronic disease count, Charlson Index score, and EDC count, respectively. Females had slightly higher scores on all

three numerical multimorbidity measures. There were 68 mutually exclusive ACG categories, although the eight largest categories accounted for over 58% of the sample.

There were highly significant positive Spearman rank correlations amongst the three count measures (for the collapsed values used in the models) - QOF and Charlson: 0.59; QOF and EDC: 0.65; Charlson and EDC: 0.51.

3.2 Multimorbidity and cost

For the regression modelling we dropped 154 individuals with missing deprivation data to leave an estimation sample of 85,946. **Table 5** shows AIC, BIC and deviance-based R-squared values for the models predicting total patient cost of primary care resources used over one year beginning 1st April 2007.

Both the AIC and BIC values indicate that the inclusion of deprivation and/or practice significantly improves the fit of the model. This is reinforced by the likelihood ratio test (LRT) between the model including just age and gender (Model 1), and the model including age, gender, and deprivation (Model 2) which showed a highly significant result (p<.0001) suggesting that the inclusion of deprivation is justified. Similarly the LRT between Model 1 and the model including age, gender, and practice (Model 3) also showed a highly significant result (p<.0001). Even so, there appears very little improvement in the deviance-based R-squared values with the inclusion of either of these variables.

The model including age, gender, deprivation, and practice (Model 4) explained 14% of the uncertainty in total patient cost. Inclusion of the QOF chronic disease count, Charlson Index score, count of EDCs, or mutually exclusive ACG categories increased this to 27%, 21%, 29%, and 25% respectively. The rankings from smallest to largest of both AIC and BIC suggest that the model including a count of EDCs provided the best fit. This is followed in order by the model including the ACG categories, QOF chronic disease count, and Charlson Index score.

Table 6 gives estimated risk ratios[2] (with 95% confidence intervals) for the best fitting model which uses EDC count and age, gender, deprivation, and practice effects. (The practice effects are omitted to save space.) The effects of age, gender and deprivation are very similar in all four models. Cost increases with age and are greater for females than males but, compared with the unconditional effects of age in Table 4 and Figure 1 which do not allow for differences in multimorbidity, the effect of age is much reduced. Allowing for multimorbidity makes little difference to the negligible unconditional relationship between deprivation and cost.

[2] The model estimates $\ln Ec = x'\beta$ or $Ec = \exp(x'\beta)$. Since all explanatory variables are categorical the risk ratio for a variable x_k is the ratio of expected cost when $x_k = 1$ to the expected cost when $x_k = 0$ and is $\exp\beta_k$ since $\exp(\beta_k * 1 + \sum_{j \neq k} \beta_j x_j) / \exp(\beta_k * 0 + \sum_{j \neq k} \beta_j x_j) = \exp\beta_k$

The risk ratios in Table 6 show that total cost increases almost proportionately with the number of diseases.

Table 7 reports the risk ratios for the categorical QOF disease count variables and the categorical Charlson Index score variables. As in Table 6 the higher the count the greater the risk ratio and all categories are highly significant. The estimated risk ratios again show a reasonably proportional relationship between total patient cost and the number of chronic diseases or Charlson Index score.

Figure 3 shows predicted mean patient cost at each level or category of multimorbidity. For the models that use a numerical multimorbidity measure, a reasonably linear relationship between mean patient cost and multimorbidity can be observed. As the ACGs are mutually exclusive and have no underlying ordinal structure, the horizontal axis of the plot is based on a ranking of predicted mean patient cost under the model.

3.3 Sensitivity analyses

Figure 4 shows predicted mean patient cost at each level or category of multimorbidity, having trimmed the upper 1% of the distribution of costs before fitting the model. The overall trend is a reduction in predicted mean patient cost at each level of multimorbidity. Although the magnitude of the reduction in predicted mean patient cost is somewhat relative across the different levels of multimorbidity, the greatest shifts are seen in the upper categories of multimorbidity, particularly for the QOF chronic disease count and Charlson Index score, creating slight non-linearity between patient cost and multimorbidity. A similar pattern was observed in the estimated risk ratios for the multimorbidity measures in each of the models (not shown here).

Our results were not sensitive to the method for calculating prescribing costs. Using the prescription cost analysis (PCA) data produced similar results to our main analysis. Both the estimated risk ratios and the predictions of mean patient cost were similar and the model rankings were unchanged.

Dropping the exposure term for time to deregistration also made very little difference to the results.

3.4 Multimorbidity and specific disease

Table 8 shows estimated risk ratios (with 95% confidence interval) for three models: with the QOF chronic disease count only (model A); with separate indicators for each of the 17 QOF diseases (model B); and with separate indicators for the 17 QOF diseases *and* the QOF chronic disease count (model C). All models also include age, gender, deprivation, and practice effects.

The AIC and BIC values suggest that model B (with the 17 separate disease indicators) is worse than model A (with the QOF count only), and both are worse than model C (with the 17 separate disease indicators *and* the QOF count). The deviance based R-squared value suggests there is almost no additional uncertainty in the patient costs that can be explained by the addition of either the QOF disease indicators or the QOF disease count over and above what one contributes alone without the other. The difference in MAPE is almost negligible when model C (with the 17 separate disease indicators *and* the QOF count) is compared with model A (with the QOF count only). Although model C (with the 17 separate disease indicators *and* the QOF count) has the best goodness of fit statistics, note that the confidence intervals on the risk ratios on the QOF indicators and the QOF count are much wider than in models with these entered separately. This is possibly because of multi-collinearity: the QOF count will usually increase when an individual has an additional QOF disease.

4 Discussion

Deprivation is often suggested as having a strong relationship with health status and until 2004 the capitation payments received by general practices depended on the deprivation level of the areas of residence of their patients. The results here suggest there is only a weak relationship between deprivation and primary care expenditure, whether or not allowance is made for patient multimorbidity. As we are considering health expenditure here, rather than health status, the surprisingly weak relationship in the data may be evidence of unrecognised need: poorer health status in more deprived groups may not be translating into more diagnosed health conditions and thus into increased health resource use.

Age is one of the factors affecting capitation payments to general practices and our results show that age does have a marked effect on expenditure, with a roughly five fold variation in expenditure between the lowest age band (20-29) and those aged 80 to 90 if no allowance is made for multimorbidity. However, once multimorbidity is allowed for the variation in cost with age is much less marked: those aged 80 to 90 have costs which are about twice as large as those aged 20-90, rather than five times as large.

It is often suggested that unobserved factors lead to general practice idiosyncracies in prescribing, consultation, and referrals for tests. We found that including a practice indicator in the regression models made only a small difference to the ability of the models to explain the primary care costs of patients. One possible explanation is that 'practice style' manifests itself in differences in case finding and thus the number and type of diagnoses that are recorded for patients. Hence models of cost which include measures of morbidity will not find any difference across practices. But we find that practice effects are not important in explaining costs even when no account is taken of morbidity.

Including any of our four measures of multimorbidity in the regression models greatly improves their explanatory power. Surprisingly, the most elaborate multimorbidity measure which uses 68 mutually exclusive ACG morbidity types performs slightly worse than the simplest measure which is a count of the number of QOF diseases, and the QOF count measure performs only slightly worse than the EDC measure which uses more than twice as many count categories.

Three of the measures were developed specifically as multimorbidity measures, albeit in healthcare systems with different institutional financial systems. We included the simple count of QOF conditions because it uses information readily available in all UK general practices. We plan to extend our analysis of the use of QOF diagnoses to construct more sophisticated measures of multimorbidity. Even with only 17 QOF conditions we cannot consider all possible combinations of morbidity. We will therefore use clinical judgements to restrict attention to a limited set of diagnosis combinations (for example depression combined, separately, with diabetes, CHD, and asthma, or asthma combined with COPD). We can also examine whether allowing age to interact with some of the QOF diagnoses improves the predictive power of the multimorbidity measure.

Acknowledgments

This project was funded by the NIHR School for Primary Care Research.

References

Akaike, H. 1974, "New Look at Statistical-Model Identification", *Ieee Transactions on Automatic Control*, vol. AC19, no. 6, pp. 716-723.

British Medical Association and Royal Pharmaceutical Society. 2010. [online] http://bnf.org. Accessed 12 May 2011.

Cameron, A. C. & Windmeijer, F. A. G. 1997, "An R-squared measure of goodness of fit for some common nonlinear regression models", *Journal of Econometrics*, vol. 77, no. 2, pp. 329-342.

Charlson, M. E., Pompei, P., Ales, K. L., & Mackenzie, C. R. 1987, "A New Method of Classifying Prognostic Co-Morbidity in Longitudinal-Studies - Development and Validation", *Journal of Chronic Diseases*, vol. 40, no. 5, pp. 373-383.

Curtis, L. 2008, *Unit costs of health and social care 2007* PSSRU, University of Kent.

Department for Communities and Local Government. 2008. "The English Indices of Deprivation 2007".

Department of Health. 2009. "NHS Reference Costs 2007/08". [online] http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsLibrary/index.htm. Accessed 10 May 2011.

Deyo, R. A., Cherkin, D. C., & Ciol, M. A. 1992, "Adapting A Clinical Comorbidity Index for Use with Icd-9-Cm Administrative Databases", *Journal of Clinical Epidemiology*, vol. 45, no. 6, pp. 613-619.

Fortin, M., Lapointe, L., Hudon, C., Vanasse, A., Ntetu, A. L., Maltais, D., Fortin, M., Lapointe, L., Hudon, C., Vanasse, A., Ntetu, A. L., & Maltais, D. 2004, "Multimorbidity and quality of life in primary care: a systematic review", *Health & Quality of Life Outcomes*, vol. 2, p. 51.

Greenfield, S., Apolone, G., Mcneil, B. J., & Cleary, P. D. 1993, "The Importance of Coexistent Disease in the Occurrence of Postoperative Complications and One-Year Recovery in Patients Undergoing Total Hip-Replacement - Comorbidity and Outcomes After Hip-Replacement", *Medical Care*, vol. 31, no. 2, pp. 141-154.

Halling, A., Fridh, G., & Ovhed, I. 2006, "Validating the Johns Hopkins ACG case-mix system of the elderly in Swedish primary health care", *Bmc Public Health*, vol. 6.

Johns Hopkins Bloomberg School of Public Health. The Johns Hopkins ACG® Case-mix System Version 8.2. Baltimore, 2008.

Hippisley-Cox, J. & Vinogradova, Y. 2009. "Trends in Consultation Rates in General Practice 1995 to 2008: Analysis of the QResearch database. Final Report to the NHS Information Centre and Department of Health." National Health Service (NHS) Information Centre and QResearch.

Kaplan, M. H., Feinstein, A. R., Kaplan, M. H., & Feinstein, A. R. 1974, "The importance of classifying initial co-morbidity in evaluating the outcome of diabetes mellitus 672", *Journal of Chronic Diseases*, vol. 27, no. 7-8, pp. 387-404.

Khan, N. F., Perera, R., Harper, S., & Rose, P. W. 2010, "Adaptation and validation of the Charlson Index for Read/OXMIS coded databases", *Bmc Family Practice*, vol. 11.

Lawrenson, R., Williams, T., & Farmer, R. 1999, "Clinical information for research; the use of general practice databases", *Journal of Public Health Medicine*, vol. 21, no. 3, pp. 299-304.

Linn, B. S., Linn, M. W., & Gurel, L. 1968, "Cumulative Illness Rating Scale", *Journal of the American Geriatrics Society*, vol. 16, no. 5, p. 622-&.

Manning, W. G. 1998, "The logged dependent variable, heteroscedasticity, and the retransformation problem", *Journal of Health Economics*, vol. 17, no. 3, pp. 283-295.

National Health Service (NHS) Information Centre. 2008. Prescription Cost Analysis 2007.

National Health Service (NHS) Information Centre. 2008. Prescriptions Dispensed in the Community, Statistics for 1997 to 2007: England.

Omar, R. Z., O'Sullivan, C., Petersen, I., Islam, A., & Majeed, A. 2008, "A model based on age, sex, and morbidity to explain variation in UK general practice prescribing: cohort study", *British Medical Journal*, vol. 337, no. 7663.

PBRA Team. Developing a Person-Based Resource Allocation Formula for Allocations to General Practices in England. 2010. [online] http://www.nuffieldtrust.org.uk/uploadedFiles/Projects/Developing_a_person-based_resource_allocation_formula_REPORT.pdf?n=6255. Accessed 12 May 2011.

Perkins, A. J., Kroenke, K., Unutzer, J., Katon, W., Williams, J. W., Hope, C., & Callahan, C. M. 2004, "Common comorbidity scales were similar in their ability to predict health care costs and mortality", *Journal of Clinical Epidemiology*, vol. 57, no. 10, pp. 1040-1048.

Reid, R. J., Roos, N. P., MacWilliam, L., Frohlich, N., & Black, C. 2002, "Assessing population health care need using a claims-based ACG morbidity measure: A validation analysis in the province of Manitoba", *Health Services Research*, vol. 37, no. 5, pp. 1345-1364.

Salisbury, C., Johnson, L., Purdy, S., Valderas, J. M., & Montgomery, A. A. 2011, "Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study", *British Journal of General Practice*, vol. 61, no. 582, pp. 18-24.

Schwarz, G. 1978, "Estimating Dimension of A Model", *Annals of Statistics*, vol. 6, no. 2, pp. 461-464.

Starfield, B. 2006, "Threads and yarns: Weaving the tapestry of comorbidity", *Annals of Family Medicine*, vol. 4, no. 2, pp. 101-103.

Starfield, B., Weiner, J., Mumford, L., & Steinwachs, D. 1991, "Ambulatory Care Groups - A Categorization of Diagnoses for Research and Management", *Health Services Research*, vol. 26, no. 1, pp. 53-74.

StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Sullivan, C. O., Omar, R. Z., Ambler, G., & Majeed, A. 2005, "Case-mix and variation in specialist referrals in general practice", *British Journal of General Practice*, vol. 55, no. 516, pp. 529-533.

Technical Steering Committee 2010, *GP Earnings and expenses 2007/08 673*, NHS Information Centre.

Uijen, A. A., van de Lisdonk, E. H., Uijen, A. A., & van de Lisdonk, E. H. 2008, "Multimorbidity in primary care: prevalence and trend over the last 20 years", *European Journal of General Practice*, vol. 14 Suppl 1, pp. 28-32.

Weiner, J. P., Starfield, B. H., Steinwachs, D. M., & Mumford, L. M. 1991, "Development and Application of A Population-Oriented Measure of Ambulatory Care Case-Mix", *Medical Care*, vol. 29, no. 5, pp. 452-472.

Westert, G. P., Satariano, W. A., Schellevis, F. G., & Van den Bos, G. A. M. 2001, "Patterns of comorbidity and the use of health services in the Dutch population", *European Journal of Public Health*, vol. 11, no. 4, pp. 365-372.

Table 1 Overview of multimorbidity measures included in this study.

Multimorbidity measure	Number diseases	Range of measure	Details
QOF chronic disease count	17	0 to 17	A count which includes 17 of the 19 diseases contained in the clinical domain of the Quality and Outcomes Framework (QOF) in use in the UK. The purpose of the QOF is to incentivise improvements in monitoring and care of patients with specific diseases. The QOF Business Rules outline the READ diagnostic codes that are required to classify a patient as having one of the diseases included in the QOF clinical domain. The 17 diseases included in the QOF disease count have been previously identified as chronic (Salisbury et al. 2011).
Charlson Index score	17	0 to 33	A weighted score, where weights of 1, 2, 3, or 6 are given to each of the 17 diseases depending on the strength of their relationship with inpatient mortality. The widely used Deyo adaptation (Deyo, Cherkin, & Ciol 1992) of the index was recently translated from International Classification of Disease (ICD) codes to UK-based READ diagnostic codes (Khan et al. 2010).
Count of Expanded Diagnostic Clusters (EDCs)	114	0 to 114	EDCs are clinically related groupings of administrative diagnostic codes. They allow identification of a clinical disease area whilst helping to eliminate differences in the coding behaviour of general practitioners and other clinicians. In our count we include 114 of the possible 264 EDCs identified by the ACG System. Those 114 included in the EDC count have been previously identified as chronic (Salisbury et al. 2011).
Adjusted Clinical Group (ACG)	-	Mutually exclusive categories (68 in this study)	ACGs are mutually exclusive categories, for which individuals in a single ACG are expected to have similar resource use demands. Classification into an ACG is based on an individual's age, gender, and combination of morbidities. Morbidities are incorporated into the classification process based on expected resource use demand, rather than organ or disease area. The ACG system currently uses 82 default categories; the age range of our sample meant only 68 ACG categories were populated.

Table 2 Unit costs (£) per encounter for each combination of consultation type and primary staff role.

Staff Type	Surgery consultation	Follow-up/ routine visit	Clinic	Telephone call to/from a patient	Out of hours
GP: partner	24.47	16.32	35.98	14.85	36.97
GP: registrar/associate	15.92	10.61	23.40	9.66	24.05
GP: sole practitioner	27.55	18.37	40.51	16.72	41.63
Practice Nurse	9.00	6.00	9.00	5.46	13.60
Community Nurse	11.88	7.92	11.88	7.21	-
Midwife / Health visitor	18.08	12.06	18.08	10.97	-
Community Psychiatric Nurse	8.01	5.34	8.01	4.86	-
Other Health Care Professional	15.00	10.00	15.00	9.10	22.66
Counsellor	64.00	42.67	64.00	-	-
Chiropodist	11.00	7.33	6.68	-	-

Table 3 Consultations, prescriptions, tests 1/4/2007 – 31/3/2008: mean (standard deviation) [median]

	Age	N	Number of consultations	% with at least one consultation	Number of prescriptions	% with at least one prescription	Number of tests	% with at least one test
Male								
	20-29	6021	1.7 (3.1) [1]	54%	2.6 (10.0) [0]	44%	0.7 (2.2) [0]	19%
	30-39	7204	2.0 (3.6) [1]	54%	4.1 (13.4) [0]	47%	1.0 (2.9) [0]	22%
	40-49	8902	2.6 (4.5) [1]	60%	6.6 (18.3) [1]	53%	1.8 (4.1) [0]	30%
	50-59	7486	3.8 (5.5) [2]	69%	13.4 (27.1) [2]	64%	3.2 (5.7) [0]	44%
	60-69	6481	5.8 (7.0) [4]	83%	26.5 (37.7) [15]	81%	5.7 (7.9) [3]	64%
	70-79	4112	8.3 (8.2) [6]	92%	42.9 (45.2) [33]	92%	8.0 (8.9) [6]	77%
	80-89	1878	10.3 (10.0) [8]	94%	54.4 (55.9) [44]	94%	9.2 (10.2) [7]	79%
	90+	253	7.8 (6.9) [6]	87%	46.1 (52.0) [32]	88%	6.0 (7.1) [4]	68%
Female								
	20-29	5551	4.7 (5.1) [3]	84%	5.8 (9.8) [3]	81%	2.6 (4.6) [0]	47%
	30-39	6930	4.9 (5.9) [3]	82%	7.9 (20.4) [3]	76%	2.9 (5.3) [1]	51%
	40-49	8447	4.6 (5.9) [3]	81%	10.0 (23.6) [3]	74%	3.1 (5.7) [1]	51%
	50-59	7525	5.1 (6.1) [3]	82%	15.5 (28.2) [6]	78%	4.2 (6.7) [1]	65%
	60-69	6563	6.5 (6.9) [5]	89%	28.0 (40.5) [16]	87%	5.9 (8.2) [3]	73%
	70-79	4954	8.6 (8.4) [6]	94%	44.7 (50.8) [33]	93%	7.8 (8.9) [6]	77%
	80-89	3057	9.5 (9.1) [7]	93%	59.2 (68.0) [44]	95%	7.9 (9.0) [6]	77%
	90+	736	8.5 (8.7) [6]	91%	64.6 (75.5) [47]	94%	6.6 (7.8) [4]	72%
Overall		86100	5.0 (6.6) [3]	77%	18.5 (36.3) [4]	72%	3.9 (6.8) [1]	52%

Table 4 Costs (£) 1/4/2007-31/3/2008: Unadjusted mean, (standard deviation), [median]

	Age category	Consultation cost	Prescription drug cost	Test cost	Total patient cost
Male					
	20-29	42 (70) [20]	44 (351) [0]	6 (32) [0]	92 (379) [24]
	30-39	49 (85) [23]	58 (261) [0]	9 (40) [0]	115 (314) [26]
	40-49	63 (100) [27]	103 (884) [2]	13 (50) [0]	179 (912) [41]
	50-59	88 (118) [49]	185 (865) [12]	22 (67) [0]	295 (915) [93]
	60-69	134 (145) [97]	320 (1996) [112]	34 (87) [8]	488 (2048) [269]
	70-79	186 (169) [148]	474 (1259) [270]	49 (102) [14]	709 (1322) [490]
	80-89	231 (201) [183]	557 (1003) [329]	60 (124) [17]	848 (1088) [612]
	90+	182 (154) [149]	655 (5019) [195]	34 (73) [10]	871 (5049) [422]
Female					
	20-29	108 (116) [74]	58 (163) [17]	17 (50) [0]	183 (246) [110]
	30-39	111 (131) [74]	85 (332) [14]	20 (58) [0]	217 (412) [113]
	40-49	107 (135) [71]	113 (326) [16]	24 (68) [0]	244 (431) [113]
	50-59	122 (137) [83]	179 (475) [34]	40 (84) [11]	341 (567) [170]
	60-69	152 (151) [113]	294 (567) [109]	47 (89) [16]	493 (667) [296]
	70-79	197 (173) [156]	399 (546) [239]	48 (101) [14]	644 (662) [471]
	80-89	222 (194) [178]	478 (1397) [291]	45 (95) [13]	745 (1448) [550]
	90+	209 (194) [162]	415 (601) [253]	32 (70) [10]	656 (714) [515]
Overall		115 (143) [71]	199 (899) [24]	27 (74) [0]	342 (960) [137]

Table 5 Fit statistics comparing models predicting total patient cost of primary care resources used over one year.

Model	AIC	BIC	Deviance-based R-squared	MAPE
Age, gender (Model 1)	1140233	1140318	0.12	299.8
Age, gender, and deprivation (Model 2)	1139001	1139170	0.13	298.9
Age, gender, and practice (Model 3)	1137571	1139275	0.14	298.9
Age, gender, deprivation, and practice (Model 4)	1136862	1138650	0.14	298.6
(Model 4) + QOF chronic disease count	1106973	1108817	0.27	262.5
(Model 4) + Charlson Index score	1121359	1123213	0.21	279.5
(Model 4) + EDC count	1099240	1101196	0.29	253.4
(Model 4) + ACG	1101602	1104018	0.25	265.6

Table 6 Risk ratios (and 95% confidence interval) for the model including age, gender, deprivation, practice, and EDC count.

Age category		Risk Ratio	95% CI	EDC count		Risk Ratio	95% CI
	20-29	1.00			**0**	1.00	
	30-39	1.03	(0.97,1.09)		**1**	2.09	(1.97,2.21)
	40-49	1.09	(1.03,1.16)		**2**	3.11	(2.93,3.30)
	50-59	1.34	(1.26,1.43)		**3**	4.25	(3.99,4.52)
	60-69	1.65	(1.54,1.76)		**4**	5.53	(5.16,5.94)
	70-79	1.84	(1.71,1.98)		**5**	6.79	(6.28,7.33)
	80-89	1.90	(1.73,2.08)		**6**	8.27	(7.59,9.02)
	90+	1.72	(1.46,2.03)		**7**	9.41	(8.53,10.38)
Gender					**8**	10.67	(9.54,11.93)
	Male	1.00			**9**	13.06	(11.51,14.83)
	Female	1.22	(1.18,1.26)		**10**	13.31	(11.51,15.39)
IMD decile					**11**	14.12	(11.82,16.88)
	1	1.00	(0.94,1.11)		**12**	14.91	(12.17,18.27)
	2	1.02	(0.93,1.10)		**13**	17.13	(13.51,21.70)
	3	1.01	(0.90,1.06)		**14**	17.04	(12.78,22.71)
	4	0.98	(0.91,1.08)		**15**	26.52	(18.81,37.40)
	5	0.99	(0.91,1.09)		**16**	21.95	(14.12,34.12)
	6	1.00	(0.95,1.14)		**17**	20.22	(11.91,34.35)
	7	1.04	(0.95,1.14)		**18+**	23.15	(15.36,34.89)
	8	1.04	(1.03,1.23)				
	9	1.13	(1.06,1.29)				
	10	1.17	(0.94,1.11)				

Table 7 Risk ratios (and 95% confidence intervals) for QOF chronic disease count and Charlson Index score

Charlson Index score (plus model 4 covariates)			QOF chronic disease count (plus model 4 covariates)		
Score	**Risk ratio**	**95% CI**	**Count**	**Risk ratio**	**95% CI**
0	1.00		**0**	1.00	
1	2.06	(1.98,2.15)	**1**	2.46	(2.36,2.57)
2	2.85	(2.66,3.06)	**2**	3.90	(3.67,4.14)
3	3.29	(2.97,3.63)	**3**	5.04	(4.64,5.47)
4	4.04	(3.48,4.70)	**4**	5.82	(5.18,6.54)
5	3.83	(3.06,4.79)	**5**	8.09	(6.72,9.75)
6	4.17	(2.95,5.90)	**6+**	7.21	(5.61,9.28)
7+	5.66	(3.87,8.27)			

Table 8 Comparison of models with and without QOF multimorbidity count and QOF diseases.

		QOF disease count only (model A)		QOF disease indicators only (model B)		QOF disease indicators and QOF count (model C)	
QOF disease (main effect)		**Risk ratio**	**95% CI**	**Risk ratio**	**95% CI**	**Risk ratio**	**95% CI**
	asthma			2.26	(2.11,2.42)	1.19	(0.84,1.69)
	atrialfib			1.21	(1.07,1.37)	0.81	(0.56,1.17)
	cancer			2.43	(2.13,2.77)	1.35	(0.94,1.95)
	chd			1.75	(1.62,1.89)	1.10	(0.77,1.56)
	ckd			1.26	(1.14,1.38)	0.86	(0.60,1.22)
	copd			2.16	(1.92,2.44)	1.37	(0.96,1.97)
	dementia			1.42	(1.11,1.81)	0.89	(0.58,1.36)
	depression			1.79	(1.70,1.88)	0.91	(0.64,1.29)
	diabetes			2.26	(2.09,2.45)	1.47	(1.04,2.08)
	epilepsy			3.63	(3.04,4.32)	1.86	(1.27,2.74)
	heartfailure			1.35	(1.14,1.59)	1.09	(0.75,1.59)
	hypertension			1.59	(1.52,1.67)	0.85	(0.60,1.21)
	learning			2.16	(1.64,2.85)	1.13	(0.72,1.76)
	mentalhealth			3.07	(2.57,3.67)	1.77	(1.20,2.61)
	obesity			1.39	(1.31,1.48)	0.81	(0.57,1.15)
	stroke			1.36	(1.21,1.52)	0.94	(0.66,1.34)
	thyroid			1.35	(1.24,1.47)	0.80	(0.56,1.13)
QOF disease count							
	0	1.00				1.00	
	1	2.46	(2.36,2.57)			2.48	(1.75,3.50)
	2	3.90	(3.67,4.14)			4.00	(2.01,7.95)
	3	5.04	(4.64,5.47)			5.22	(1.87,14.60)
	4	5.82	(5.18,6.54)			5.93	(1.51,23.33)
	5	8.09	(6.72,9.75)			7.97	(1.43,44.25)
	6+	7.21	(5.61,9.28)			6.97	(0.76,63.64)
AIC		1106973		1107726		1104557	
BIC		1108817		1109673		1106560	
Deviance-based R-squared		0.27		0.27		0.28	
MAPE		262.5		310.4		260.5	

All models have Model 4 covariates as well. na: not applicable

Figure 1 Unadjusted mean total patient cost (with 95% confidence interval) for 1/4/2007-31/3/2008 by age category and deprivation decile

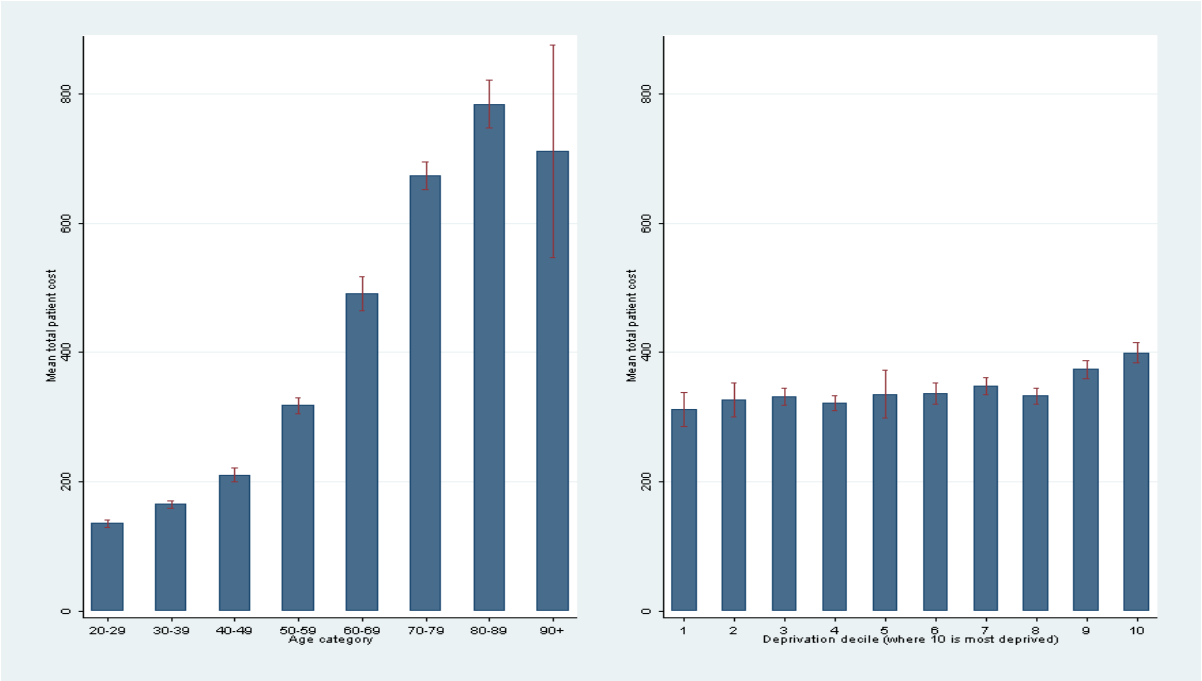

Figure 2 Frequency distribution of QOF chronic disease count, Charlson Index score, EDC count, and ACG categories.

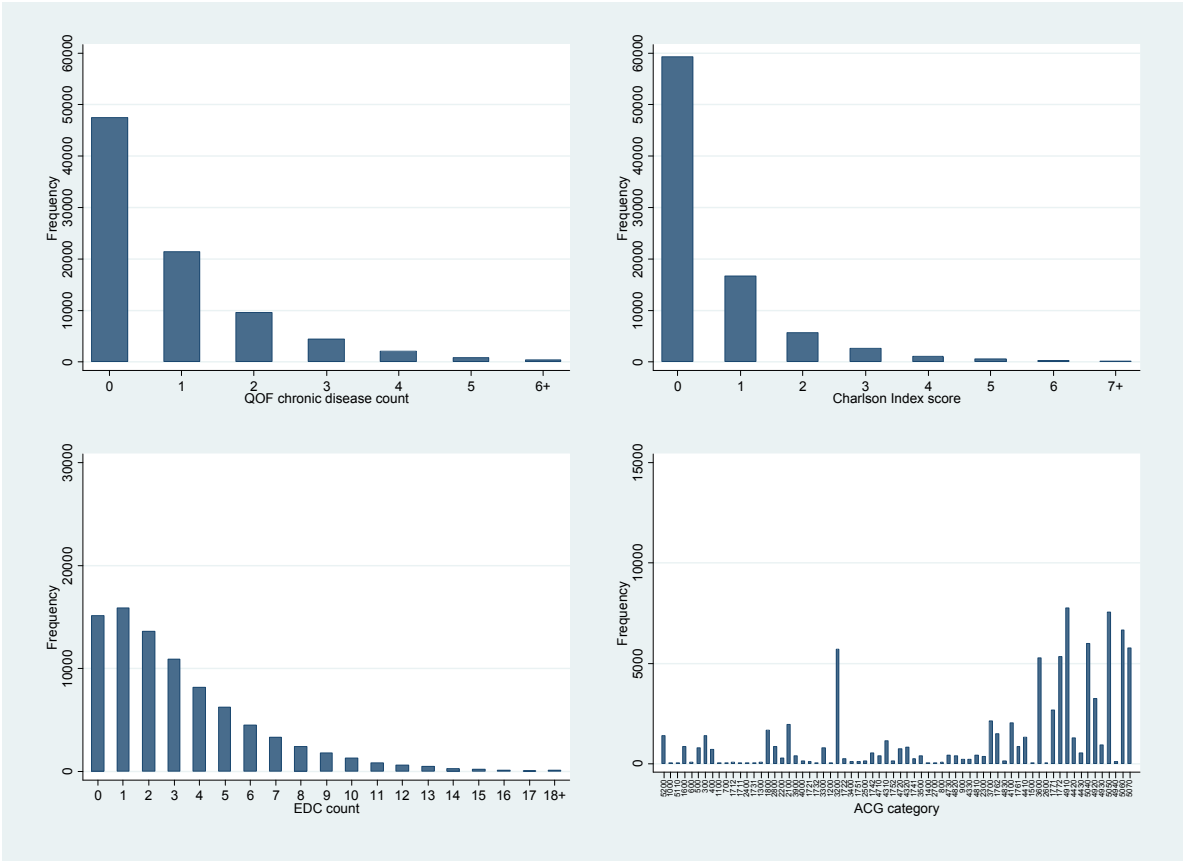

Figure 3 Predicted mean patient cost (with 95% confidence interval) for different levels of multimorbidity, by model.

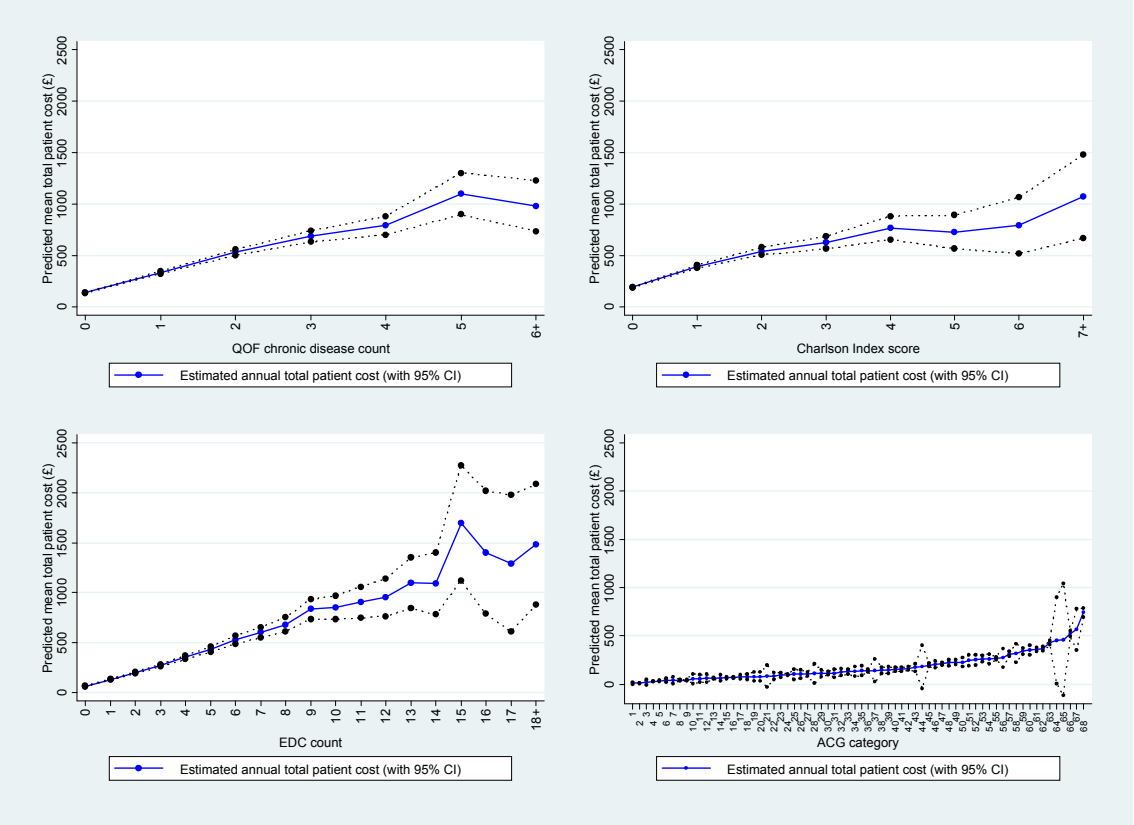

Figure 4 Predicted mean patient cost (with 95% confidence interval) for different levels of multimorbidity, by model, with trimming of the upper 1% of total patient costs before fitting the model.