BMJ open

Predicting Risk of Hospitalization Using a Population Based Longitudinal Database

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-005223
Article Type:	Research
Date Submitted by the Author:	07-Mar-2014
Complete List of Authors:	Louis, Daniel; Jefferson Medical College, Center for Research in Medical Education and Health Care Robeson, Mary; Jefferson Medical College, Center for Research in Medical Education and Health Care McAna, John; Thomas Jefferson University, Jefferson School of Population Health Maio, Vittorio; Thomas Jefferson University, Jefferson School of Population Health Keith, Scott; Thomas Jefferson University, Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics Liu, Mengdan; Jefferson Medical College, Center for Research in Medical Education and Health Care Gonnella, Joseph; Jefferson Medical College, Center for Research in Medical Education and Health Care Grilli, Roberto; Agenzia Sanitaria e Sociale Regione Emilia-Romagna,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Patient-centred medicine
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

Predicting Risk of Hospitalization Using a Population Based Longitudinal Database

Daniel Z. Louis[1], Mary Robeson[1], John McAna[2], Vittorio Maio[2], Scott W. Keith[3], Mengdan Liu[1], Joseph S. Gonnella[1], Roberto Grilli[4]

[1]Center for Research in Medical Education and Health Care, Jefferson Medical College of Thomas Jefferson University. 1015 Walnut Street, Suite 319. Philadelphia, PA 19107, USA

[2]Jefferson School of Population Health, Thomas Jefferson University. 901 Walnut Street, 10th Floor. Philadelphia, PA 19107, USA

[3]Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics, Thomas Jefferson University. 1015 Chestnut Street, Suite M-100. Philadelphia, PA 19107, USA

[4]Agenzia Sanitaria e Sociale Regione Emilia-Romagna, Viale Aldo Moro, 21. Bologna, 40128 Italy.

Corresponding author: Daniel Z. Louis, Jefferson Medical College, 1015 Walnut Street, Suite 319, Philadelphia, PA 19107. USA; Tel: 215-955-6959, daniel.louis@jefferson.edu

Copyright: The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (non-exclusive for Dr. Grilli as a government employee) on a worldwide basis to the BMJ publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any 0ther BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license.

For peer review only

Abstract

Objectives: Develop predictive models using an administrative health care database that provide information for Patient Centered Medical Homes to proactively identify patients at risk of hospitalization for conditions that may be impacted through improved patient care.

Design: Retrospective health care utilization analysis with multivariate logistic regression models.

Data: A population-based longitudinal database of residents served by the Emilia-Romagna, Italy health service in the years 2004-2012 including demographic information and utilization of health services by 3,726,380 people age ≥ 18 years.

Outcome measures: Models designed to predict risk of hospitalization or death for problems that are potentially avoidable were developed and evaluated using the area under the receiver operating curve C-statistic, in terms of their sensitivity, specificity, and positive predictive value, and for calibration to assess performance across levels of predicted risk.

Results: Among the 3,726,380 adult residents of Emilia-Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions or died in 2012 (3.6% hospitalized, 1.3% died). The C-statistic for the model predicting 2012 outcomes was 0.856. The model was well calibrated across categories of predicted risk. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%.

Conclusions: We have developed a population-based model using a longitudinal administrative database that identifies the risk of hospitalization for residents of the Emilia-Romagna region

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

with a level of performance as high as, or higher than, similar models. The results of this model, along with profiles of patients identified as high risk are being provided to the physicians and other health care professionals associated with the Patient Centered Medical Homes to aid in planning for care management and interventions that may reduce their patients' likelihood of a preventable, high-cost hospitalization.

Strengths and Limitations of this study:

- This study included the entire adult population of the Emilia-Romagna regon of Italy, over 3.7 million people.
- The study used an existing longitudinal administrative health care database with both the advantage of much lower cost than new data collection and the disadvantage of potential errors in administrative data.
- The results of the study are being used to assist in the development of newly formed Patient Centered Medical Homes.

Keywords: hospitalization, risk, medical home, patient-centerd care

Word Count: 3,547 (excluding title page, abstract, references, figure, and tables.)

Contributions of authors: DZL, RG, VM, and JSG were responsible for the conceptualiztion of this project. MR, JM, and ML were responsible for creation of the datasets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.

Funding: This reseach was funded by the Agenzia Sanitaria e Sociale of the Emilia-Romagna Region, Italy. The resulting models are being used by the Local Health Authority of Parma, Italy and the Patient Centered Medical Homes located in the Parm Local Health Authority.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

The predominant healthcare delivery system, which has been a passive model, reacting to patients' problems, is shifting to a more proactive model designed to take the initiative in providing care for an increasingly older population that has a greater prevalence of chronic conditions, often with multiple medical and social needs. These changes are driving the reorganization of the primary care system, emphasizing coordination and cooperation among healthcare professionals.[1] Among the approaches to addressing this need has been the establishment of primary care organizations incorporating integrated teams of physicians and other healthcare professionals that "seek to increase the influence of primary care professionals, and in particular general practitioners (GPs), in health planning and resource allocation."[2] Prominent among these new models of primary care is the Patient Centered Medical Home, an organization in which a team of healthcare providers is engaged in delivering comprehensive, coordinated, patient-centered care to patient defined populations.[3]

Primary care has a central role in the Italian National Health Service (NHS). Twenty one regional governments are responsible for ensuring the delivery of a health benefits package through a network of geographically defined, population-based Local Health Authorities. Primary care physicians work for these authorities as independent contractors and act as "gatekeepers" for specialty and other referral services for their patients.[4]

With the belief that a strong primary care system is conducive to improving population health, the NHS initiated reforms that encouraged primary care physicians to organize into collaborative arrangements. To this end, the Regione Emilia-Romagna (RER), a large northern region with a population of about 4.5 million, has recently launched a plan in its 11 Local Health Authorities to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

establish Patient-Centered Medical Homes to better coordinate patient care and help patients avoid unnecessary hospitalizations.

The identification of those patients who would benefit most from outreach efforts is fundamental to achieving these goals of promoting and practicing population health in Patient-Centered Medical Homes. To accomplish this, predictive models and risk stratification tools are needed to identify patients at risk of a worsening of their health status. According to Knutson and Bella,[5] "Predictive models are data-driven, decision-support tools that estimate an individual's future potential health care costs and/or opportunities for care management." A good model will identify as many of these patients as possible while excluding those for whom these interventions would be unnecessary or ineffective.

The RER has established three objectives for this project: 1) develop predictive models to identify patients at high risk of a progression of their medical problems or who are at risk of developing new medical problems, 2) create "risk of hospitalization" patient profiles that provide information about their high-risk patients to the general practitioners in the newly formed Patient-Centered Medical Homes, and 3) assess the extent to which these models and reports provide additional information useful in the identification of patients who may benefit from case management or disease management.

This paper will address the first of the three goals. We describe the development of a predictive model using the RER's regional longitudinal administrative health care database to help identify patients who are most at risk of hospitalization for conditions that may be impacted through improved patient care. This model will then be used to inform the providers associated with the Patient-Centered Medical Homes and aid in their planning for care management and

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

interventions that can reduce their patients' likelihood of a preventable, high-cost hospitalization.

Methods

Study Data and Study Population

The model was developed using the population-based longitudinal health care database of the residents served by the RER Health Service in the years 2004 through 2012.[6] This administrative database includes demographic information for all residents (gender, birth and death dates, location of current residence, and primary care physician), hospital discharge abstract data (ICD-9-CM diagnosis and procedure codes, and admission and discharge dates), emergency room utilization information, outpatient pharmacy data at the individual prescription level, specialty care (laboratory, diagnostics, therapeutic procedures, rehabilitation, and specialist visits), home health data, and information on each primary care physician in the region. Each patient has an anonymous identifier assigned by the RER so that an individual's utilization can be tracked over time without jeopardizing patient privacy.

The study population consisted of all residents of the RER who were at least 18 years of age and still alive as of 31st December 2011.

Dependent Variable

The dependent variable was defined as the occurrence of a hospitalization for problems that are potentially avoidable, or whose progression may have been avoided or delayed through appropriate patient care, or the death of the individual for any reason in 2012. We decided to not

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

limit the hospitalization to emergency admissions, since a planned admission may also be an indicator of a worsening medical problem. In order to operationally define the dependent variable, we reviewed the Disease Staging[7, 8] primary diagnostic category and severity stage of all day and inpatient hospital admissions (for adults age 18+) in RER for one year, to select those admissions that should be included in the dependent variable.

Deliveries, admission for dental diseases or admissions for vague signs or symptoms with no identified etiology were excluded. Admissions for problems that are not predictable/preventable were excluded while those where screening may identify problems that can potentially be treated to avoid progression were included. For example, admissions for stage 1, chronic cholecystitis or cholelithiasis were excluded, but admissions for advanced stage 2 or 3 complications such as ascending cholangitis or pancreatitis were included.

We felt that inclusion of hospitalization for cancer in the dependent variable should depend on the ability to either prevent or avoid progression of the disease. We therefore included colon cancer and cervical cancer in the definition because they are potentially preventable but excluded all other cancers where prevention/prediction is not currently possible.

Inclusion of injuries, burns, or toxic reaction to prescription or non-prescription drugs would ideally be based on the cause of these problems. Since the etiology of these problems is typically not available in the administrative data being used in this project, we made the decision to include or exclude based on our subjective judgment of the likelihood of preventability. For example, adverse drug reactions were included but burns were excluded from the definition of the dependent variable.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

There is no obvious medical reason for a hospital admission for patients with stage 1 diabetes mellitus or stage 1 essential hypertension without complications. These problems are typically treatable in the outpatient setting. A hospitalization implies a potential problem in the care of these patients, so we decided to include these admissions as a part of the dependent variable.[9]

Independent Variables

A broad range of candidate predictor variables was developed taking advantage of the RER administrative data. The independent variables used for modeling were defined from the RER administrative data for the years 2004 through 2011. Demographic data included patient age, sex, and geographic location of residence. We developed a mapping to broad disease categories defined primarily in terms of the affected body system from home health care data, pharmacy data, and hospital discharge abstract data. (See Appendix.)

For those patients who had been hospitalized, more specific diagnostic data were available. We reviewed the classification of patients hospitalized historically using the Disease Staging diagnostic category and disease severity stages.[7,8] Based on the frequencies specific diagnostic category/stage predictor variables were defined for either specific stages of frequent diseases, or by combinations across similar categories. Predictor variables were defined based on the number of emergency room visits using the RER classifications system for the urgency of the visit.

Pharmacy data were used to identify polypharmacy[10] (defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days), potential drug-drug interactions

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(DDI)[11] and potentially inappropriate medication use in patients[12] 65 years and older. Since cardiovascular disease is highly prevalent, we reviewed the use of cardiovascular drugs and created a variable for each of the following 11 classes of drugs (oral anti-coagulants, beta-blockers, angiotensin converting enzyme / angiotensin II receptor blockers, anti-platelets, calcium channel blockers, anti-arrhythmics, digitalis glycosides, nitrates, diuretics, alfa-blockers, statins) to account for the complexity of therapeutic regimen at the patient level.

To take advantage of the fact that the RER database includes multiple years of data, we created history variables using the utilization for each year of data available. Since we were working with the 2011 data to predict hospitalization or death in 2012, we created history variables based on 2004-2010 data. This set included 83 of the diagnostic category/stage variables as well as 11 variables based on pharmacy utilization such as exposure to polypharmacy and use of cardiovascular drugs. If the individual had a history of a disease in any of the years from 2004 to 2010 they were flagged as having a history of that disease and this was used as a potential predictor variable.

Modeling

Logistic regression models were used to estimate predicted probabilities for the occurrence of an inpatient hospital stay for the selected conditions or death for individual patients. Risk of hospitalization or death, and the variables that relate to those risks are highly dependent on age and gender. Regression models were fit in each of 14 gender and age strata using SAS Version 9.2 (SAS Institute, Cary NC). A stepwise process with relaxed entry and retention criteria

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(inclusion p-value <=0.8, retention <=0.5) was used for selecting a reduced, but robust set of independent variables for the model in each age/gender stratum.

Evaluation of the Models

The predictive accuracy of the modelling was evaluated using C-statistics (the area under the receiver operating characteristics curve), along with three measures traditionally used with clinical screening tools: sensitivity, specificity and positive predictive value.

C-statistics were used to evaluate the models in two ways. The first evaluation consisted of fitting the model developed using utilization and demographic data from 2011, along with historical variables based on 2004-2010 data, and outcomes (hospitalization or death) from 2012 and then computing a C-statistic to evaluate how the models performed at predicting those outcomes on which the models were conditioned. However, this evaluation is not consistent with evaluating how the data are used in practice. In practice, we have current predictor information, but the outcomes have not been realized. To better estimate how the models are likely to perform in this setting, we fit models to outcomes data up to a year prior to the most current available (e.g., 2011 outcomes modelled with predictors from 2010, along with historical variables based on 2004-2009 data). We then computed a C-statistic for projections made on the risk of hospitalization or death outcomes (in 2012) using the next year's predictor information (in 2011). This way, the models are forced to make projections into the future, but we have the actual observed outcomes data to evaluate the modelling process as it would be used in practice. The resulting C-statistics obtained from these two model runs were compared.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

In order to evaluate the performance of the model across different risk thresholds we classified predicted risk scores. "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%. These risk thresholds were selected after consultation with physicians practicing in the medical homes to yield a total of about 10% of the 1,500 patients enrolled with a typical primary care physician.

Results

Among the 3,726,380 adult residents of Emilia Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions defined earlier or died in 2012 (3.6% hospitalized, 1.3% died).

Table 1 shows the distribution of the demographics (age and gender), number of chronic conditions, body systems impacted by the selected chronic conditions, polypharmacy and inappropriate prescribing among the eligible RER residents, as of December 31, 2011. The table also compares these characteristics of the total adult population of the region to the subgroups of the population classified in the "very high risk" and "high risk" categories. Based on the model results, 114,255 individuals were identified as having a predicted risk of hospitalization or death in 2012 of ≥ 25% and classified as "very high risk." An additional 134,610 individuals had a predicted risk of hospitalization or death in 2012 of 15-24% and were classified as "high risk."

Table 1. Demographic and clinical characteristics of the Regione Emilia-Romagna population, overall and by risk category

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	Total population*		Very High Risk**		High risk**	
	3,726,380		114,255		134,610	
Gender	N.	%	N.	%	N.	%
Male	1,788,048	48.0%	54,357	47.6%	61,803	45.9%
Female	1,938,332	52.0%	59,898	52.4%	72,807	54.1%
Age groups						
18-24	258,338	6.9%	76	0.1%	105	0.1%
25-34	499,786	13.4%	302	0.3%	391	0.3%
35-44	732,626	19.7%	1,137	1.0%	1,198	0.9%
45-54	676,047	18.1%	2,612	2.3%	2,485	1.8%
55-64	550,689	14.8%	5,391	4.7%	5,287	3.9%
65-74	482,346	12.9%	13,154	11.5%	14,471	10.8%
74-85	364,369	9.8%	33,430	29.3%	44,857	33.3%
85+	162,179	4.4%	58,153	50.9%	65,816	48.9%
Number of Chronic Conditions						
0-1	2,775,888	74.5%	8,176	7.2%	24,618	18.3%
2 or more	950,492	25.5%	106,079	92.8%	109,992	81.7%
5 or more	99,337	2.7%	45,445	39.8%	20,576	15.3%
Selected Conditions/Body Systems						
Cancer	99,328	2.7%	23,872	20.9%	14,305	10.6%
Cardiovascular	967,796	26.0%	96,157	84.2%	103,749	77.1%
Male Genitourinary#	130,609	7.3%	14,616	26.9%	16,776	27.1%
Ear, Nose, Throat	5,364	0.1%	240	0.2%	242	0.2%
Endocrine	429,528	11.5%	40,653	35.6%	37,471	27.8%
Eye	114,117	3.1%	9,558	8.4%	13,478	10.0%
Gastrointestinal	580,946	15.6%	74,718	65.4%	66,305	49.3%
Gynecologic##	21,806	1.1%	333	0.6%	405	0.6%
Hematologic	45,022	1.2%	15,353	13.4%	6,591	4.9%
Hepatobiliary	24,785	0.7%	6,477	5.7%	3,306	2.5%
Immunologic	3,281	0.1%	464	0.4%	273	0.2%
Infectious Disease	4,723	0.1%	2,207	1.9%	727	0.5%
Musculoskeletal	419,184	11.2%	43,436	38.0%	41,000	30.5%
Neurologic	173,751	4.7%	34,494	30.2%	24,838	18.5%
Psychological	291,308	7.8%	43,387	38.0%	33,715	25.0%
Respiratory	176,830	4.7%	39,082	34.2%	21,763	16.2%
Skin	28,339	0.8%	7,645	6.7%	3,008	2.2%
Urogenital	37,728	1.0%	16,501	14.4%	5,740	4.3%
Polypharmacy^	609,278	16.4%	92,153	80.7%	92,156	68.5%

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any potentially inappropriate medications (age 65 years or older)^^	257,033	25.5%	51,055	48.7%	49,003	39.2%

* Adults (age 18 or older) and alive at 31 December 2011.

** "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of > 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%.

Men only.

Women only.

^ Polypharmacy is defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days.

^^ The list of potentially inappropriate medications can be found in: Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *Journal of Clinical Pharmacy and Therapeutics* 2010;35:219-229.

There was little difference across the risk categories by gender. Age distributions for the "very high risk" and "high risk" groups were shifted more towards the older age groups than those in the overall study population. Residents age 85 or older represented about 4.5% of the RER population, but about 50% of the "very high" and "high" predicted risk groups. More than 75% of the residents over age 85 were classified as "very high " or "high" risk. However, age alone was not sufficient to their predict risk. For example, residents between 75 and 84 years of age made up 23% of the "very high" risk group and 41% of the "high" risk group, but over 85% of the residents in this age category had neither "very high" nor "high" predicted risk.

Across age and gender strata, demographics and heath care utilization experience in 2011 were the most commonly used independent variables for predicting hospitalization or death in 2012. Selected history variables flagging chronic problems such as cardiovascular disease, diabetes mellitus and chronic renal failure and a history of prescriptions for cardiovascular medications and polypharmacy were also significant predictors.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The residents in the two higher risk groups were more likely than others to have multiple chronic diseases and to experience polypharmacy and inappropriate medication use. The residents identified as "very high risk" or "high risk" by the model also showed a number of striking differences from others in terms of the occurrence of some of the most prevalent health conditions by type and body system. Although cardiovascular conditions were not uncommon in the total adult population (26.0%), they were far more common among those classified as "very high risk" and "high risk" (84.2% and 77.1%, respectively). Similarly, gastrointestinal conditions affected 15.6% of the total population, but were diagnosed in 65.4% of the "very high risk" and 49.3% of the "high risk" patients. Cancer occurred in 2.7% of the total population, but 20.9% of the "very high risk" and 10.6% of the "high risk" patients had a cancer diagnosis. Mental health problems were identified in 7.8% of the adult population, but in 34.2% of the "very high risk" and 25.0% of the "high risk" patients.

The C-statistic for the model of 2012 outcomes developed using 2011 predictors and the C-statistic based on the parameters from the model of 2011 outcomes regressed on 2010 predictors applied to the 2011 predictors and 2012 outcomes, were very similar (0.856 and 0.853, respectively). These results suggest that the relationship between predictors and risk of hospitalization changed little in one year and that model parameters developed in a prior year can be used reliably with the most current year's data to predict unknown outcomes in the next year with only a minimal loss in performance in this population.

Table 2 shows the sensitivity, specificity, positive predictive value and number of true positives for the model at the two selected cut-off points. The sensitivity (percentage of patients actually hospitalized who had been identified by the model as having a predicted risk higher than the cut-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

off point) was 29.8% for those with the "very high" risk scores. This percentage represents 46,950 of the 157,550 residents of the region who were hospitalized for a selected condition or died in 2012. If we modify the risk score threshold to include individuals with a predicted risk of hospitalization for selected conditions or death of $\geq$ 15% (i.e., both the "very high risk" and the "high risk" patients) the sensitivity is .471. The true negative rate (specificity) is very high for both risk thresholds (.981 and .951, respectively).

Table 2. Performance of the "Risk of Hospitalization" model for residents identified as "Very High Risk" and "High or Very High Risk"

	Cut-off points for comparison	
Measure	**"Very high risk"***	**"Very high risk"* + "High risk"****
Sensitivity[3]	0.298	0.471
Specificity[4]	0.981	0.951
Positive Predictive Value[5]	0.411	0.298
True positives[6]	46,950	74,196

* "Very high risk" is defined as patients with a predicted risk of hospitalization of > =25%.
** "Very high risk" + "High risk", is defined as patients with a predicted risk of hospitalization of >=15%.
# Sensitivity is defined as the proportion of those hospitalized who were predicted to be hospitalized (true positive rate).
Specificity is the proportion of those not hospitalized who were not predicted to be hospitalized (true negative rate).
^ Positive Predictive Value is the proportion of those predicted to be hospitalized who were actually hospitalized.
^^ True positives are the number of residents who were predicted to be at risk for hospitalization at the predicted risk threshold and were actually hospitalized.

The model appears to be well calibrated across levels of risk. The Figure depicts the RER population divided into groups by deciles of predicted risk of hospitalization or death from the models. The observed prevalence of hospitalization or death is compared to the average predicted risk among individuals in each of the ten predicted risk groups. For example, the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

overall rate of hospitalization for the selected conditions or death in 2012 was 4.2%. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%.

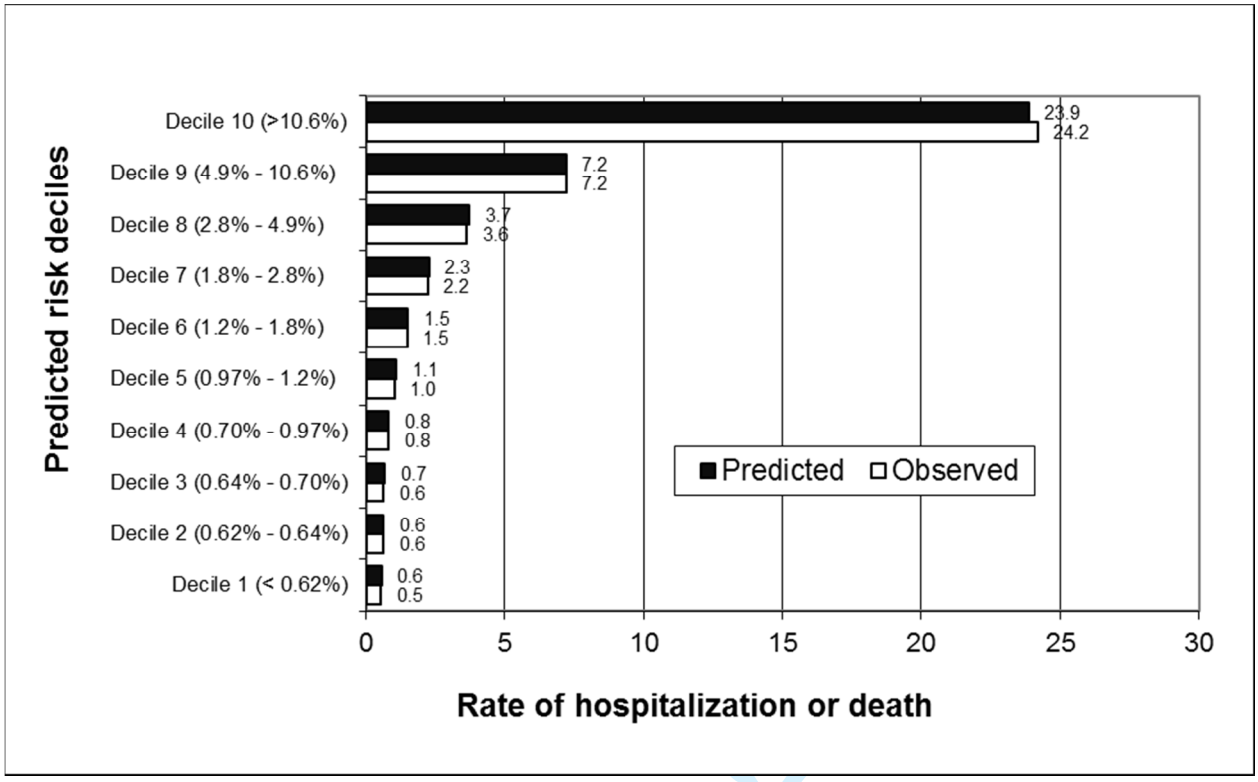


Figure. Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups.

Discussion

We have developed a population-based model that identifies the risk of hospitalization for all adult RER residents and does so with a level of performance (c=0.85) as high as, or higher than, similar models. In addition, we believe that the definition of the dependent variable chosen for our models increases the probability that they are identifying patients who risk can potentially be improved by appropriate care. A systematic review by Kansagara[13] of models designed to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

predict readmissions, showed C-statistic results in the range of 0.55 to 0.83. Recent work by Billings et al[14] to develop models predictive of emergent admissions in the UK had results ranging from 0.73 to 0.78. Li Wang, et al. (2013),[15] using information available through the US Veteran's Administration that also included lab data, demonstrated c-statistics of 0.81 and 0.79, respectively, for their models of 90-day or 12-month hospitalization or death outcomes.

Although previous studies have developed models predictive of hospital care, these models fall short of the needs of the Patient Centered Medical Homes being implemented in RER. Typically, these models have focused on specific age groups,[16] conditions, or types of admissions, such as emergent[14] or unplanned admissions or rehospitalizations, or health insurance plans in the United States, including private insurance plans, Medicare and Medicaid plans.[17,18] The models we have developed are applied to the entire adult population of RER. They use existing administrative data, which makes them cost effective to apply.

Patient Centered Medical Homes, including those instituted in RER, are responsible for addressing the needs of their population and making the best use of their finite resources to accomplish this. Preventing unnecessary admissions could improve both the quality of care and health status of the enrolled population, and result in a substantial savings. To accomplish this, predictive models and risk stratification tools such as those developed for this project are needed to identify patients at risk of preventable admissions and provide information that can be used by the medical homes to help manage care.

There are some limitations to the model. The model is developed from administrative data. Administrative data are collected for reimbursement and tracking utilization and not for medical research. They lack the clinical specificity that would desirable in assessing an individual's

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

medical problems. While the hospital discharge abstract data do include diagnostic information coded using ICD-9-CM, no similar data are available for outpatient encounters in the RER database. In addition, our models use prior utilization among the predictor variables. With the administrative data we cannot distinguish between appropriate and inappropriate treatment which may bias our results.

Despite the limitations of administrative data, they have many advantages for this project: they are readily available, relatively inexpensive to analyze and cover large populations over many years. They are ideal for uncovering patterns of care. If information from the medical records is needed, the results of these analyses can then be supplemented by focused clinical reviews at the local level. Also, The RER has a system in place to monitor the quality of diagnosis and procedure coding in their hospital discharge abstract data. Controls at both the hospital and regional level assess the validity of coding and the consistency of codes assigned such as congruity between sex, age and diagnosis and between diagnosis and procedure. The existence of the RER administrative database made it feasible to develop the models described in this article at relatively low cost and to update the models over time without additional data collection that others have found necessary.[14]

Currently, these risk scores are being integrated with other information in profiles of high-risk patients furnished to providers in 12 newly formed medical homes, including 83 primary care physicians serving a total of about 100,000 patients, in the Parma Local Health Authority located in RER. Along with the risk scores, this information includes data about previous hospitalizations, use of referrals, medications, long-term care and home care services, and a number of process-like quality indicators for diabetic and cardiovascular patients, and for

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

appropriate medication use in older patients. An evaluation of the use and usefulness of these profiles is under way.

In summary, these models provide a means of identifying patients at high risk for hospitalization. The risk predictions, in conjunction with the risk profile, show promise as a useful organizational tool for the regional Patient Centered Medical Homes to develop and implement proactive case management and disease management programs. The RER is reviewing the results of the Parma Local Health Authority pilot project of the profiles. Once their usefulness has been further evaluated, their use will be expanded to other Medical Homes in development in the other Local Health Authorities in the Emilia-Romagna region. If similar data are available, these models can be applied in other Italian regions and other countries investing in organization similar to the Patient Centered Medical Home.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Ethical approval: This study was approved by the Institutional Review Board of Thomas Jefferson University as an expedited retrospective database/record review. The IRB granted a waiver of informed consent.

Contributorship Statement: DZL, RG, VM, and JSG were responsible for the conceptualiztion of this project. MR, JM, and ML were responsible for creation of the data sets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form and declare grants from the Regione Emilia-Romagna (Louis, Robeson, McAna, Maio, Keith, Liu, Gonnella) during the conduct of the study. Louis and Gonnella declare personal fees from Truven Health Analytics, The National School of Public Health, Portugal, INSIEL Mercato.

Data Sharing Statement: No additional data

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1 Saltman RB, Rico A, Boerma W, eds. Primary care in the driver's seat? Organizational reform in European primary care. New York, New York: World Health Organization European Observatory on Health Systems and Policies Series 2006.

2 Smith J, Goodwin N. Towards managed primary care: The role and experience of primary care organizations. Burlington, VT: Ashgate Publishing Company 2006.

3 Defining the PCMH. Agency for Healthcare Research and Quality Web site. http://www.pcmh.ahrq.gov/portal/server.pt/community/pcmh__home/1483/pcmh_defining_the_pcmh_v2 (accessed June 30, 2013).

4 Lo Scalzo A, Donatini A, Orzella l, et al. Italy: Health system review. *Health Syst Transit*. 2009;**11**(6):1-216.

5 Knutson D, Bella M, Llanos K. Predictive Modeling: A Guide for State Medicaid Purchasers. Center for Health Care Strategies, Inc. 2009.

6 The Emilia-Romagna Regional Health Service and the new welfare system: Facilities, expenditure, and activities as of 31.12.2010. Programs, agreements and organisational models. Regione Emilia-Romagna Assessorato Politiche per la Salute. Bologna, Italy 2012.

7 Gonnella JS, Hornbrook MC, Louis DZ. Staging of Disease: A Case-Mix Measurement. *JAMA* 1984;**251**:637–44.

8 Gonnella JS, Louis DZ, Gozum ME, et al, eds. Clinical Criteria for Disease Staging. 6th edition. Ann Arbor, MI: Truven Health Analytics 2012.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

9 Louis DZ, F Taroni, R Melotti, et al. Increasing appropriateness of hospital admissions in the Emilia-Romagna region of Italy. *J Health Serv Res Policy* 2008;**13**(4):202-8.

10 Slabaugh L, Maio V, Templin M, et al. Prevalence and risk of polypharmacy amongst elderly primary care patients in Regione Emilia-Romagna, Italy. *Drug Aging* 2010;**27**:1019-28.

11 Gagne JJ, Maio V, Rabinowitz C. Prevalence and Predictors of Potential Drug-Drug Interactions among Ambulatory Patients in Regione Emilia-Romagna, Italy. *J Clin Pharm Ther* 2008;**33**:141-51.

12 Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *J Clin Pharm Ther* 2010;**35**:219-29.

13 Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: A systematic review. Washington, DC: Department of Veterans Affairs US 2011.

14 Billings J, Georghiou T, Blunt I, et al. Choosing a model to predict hospital admission: An observational study of new variants of predictive models for case finding. *BMJ Open* 2013;**3**(8):e003352. doi: 10.1136/bmjopen-2013-003352 [published Online First: 26 August 2013].

15 Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care* 2013;**51**(4):368-73.

16 Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. *Med Care* 2008;**46**(7):726-31.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

17 Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care* 2012;**50**(2):131-9.

18 McAna JF, Crawfrod AG, Novinger BW, et al. A predictive model of hospitalization risk among diabled Medicaid enrollees. *Am J Manag Care* 2013;**19**(5):166-74.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Appendix to: **Predicting Risk of Hospitalization Using a Population Based Longitudinal Database**

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Cancer	Neoplasm, Malignant: Cardiovascular, Hypopharynx, Oral Cavity, Oropharynx, Salivary Glands and Mandible, Other Endocrine System, Larynx, Glottis, Larynx, Subglottic, Larynx, Supraglottic, Nasopharyngeal, Sinuses, Ocular Melanoma, Other Eye and Periocular, Colon and Rectum, Esophagus, Small Bowel, Stomach, Other Gastrointestinal System, Bladder, Urinary, Kidneys, Other Genitourinary System, Breast (Female), Cervix Uteri, Endometrium, Ovaries, Vagina, Vulva, Other Female Genitalia, Hodgkin's Lymphoma, Multiple Myeloma, Mastocytosis, Pancreas, Other Hepatobiliary Tract, Breast (Male), Penile, Prostate, Testicular, Primary Bone, Waldenstrom's Macroglobulinemia, Nonspecific Sites, Unspecified Primary Site, Lungs, Bronchi, or Mediastinum, Hodgkin's Disease Lymphocytic Depletion, Hodgkin's Disease Lymphocytic Predominance, Hodgkin's Disease Mixed Cellularity, Hodgkin's Disease Nodular Sclerosis, Lymphatic and Hematopoietic (Other Types), Lymphoma, Cutaneous T Cell (Mycosis Fungoides), Lymphoma (Diffuse Mixed Small and Large Cell), Lymphoma (Diffuse Large Cell), Lymphoma (Follicular Predominantly Large Cell), Lymphoma (Histiocytic Cell), Lymphoma (Lymphoblastic), Other Respiratory System, Carcinoma (Basal Cell stage 2/3), Carcinoma (Squamous Cell), Melanoma, Other Skin and Soft	Antineoplastics 5HT3 Antagonists	2005-2009: Visits prescribed due to the presence of cancer. Beginning in 2010, the following ICD-9-CM codes were in the record: 140-208, 235-239, V10, V16	Visits prescribed for radiation therapy, or for Injection or infusion of chemotherapeutic Substances for cancer treatment

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Tissue Neoplasm: Pheochromocytoma, Eyelid, Central Nervous System, Lymphatic or Hematopoietic Leukemia: Acute Lymphocytic, Acute Nonlymphocytic, Chronic Lymphocytic, Chronic Myelogenous, Other Types Encounter for: Chemotherapy, Radiation Therapy ICD-9-CM Procedure codes: 99.25, 99.28, 00.10,00.15,92.2x			
Cardiovascular	Aneurysm: Abdominal, Thoracic Anomaly: Patent Ductus Arteriosus, Atrial Septal Defect, Atrioventricular Defects, Coarctation of the Aorta, Other Congenital Heart Disease, Pulmonary Valve Stenosis, Tetralogy of Fallot (stage 3), Transposition of the Great Arteries, Ventricular Septal Defects, Other Circulatory System Aortic: Regurgitation, Stenosis Mitral: Regurgitation, Stenosis Neoplasm: Benign of the Cardiovascular System Arrhythmias, Cardiomyopathies, Conduction Disorders, Congestive Heart Failure, Coronary Artery Disease Prior Coronary Revascularization, Coronary Artery Disease w/o Prior, Coronary Revascularization, Essential Hypertension, Infective, Endocarditis, Pericarditis: Chronic (stage 2/3), Viral or Traumatic (stage 2/3) Periarteritis Nodosa, Raynaud's Disease, Thromboangiitis,	Oral anti-coagulants beta-blockers ACE/ARB anti-platelets calcium channel blockers anti-arrhythmics digitalis glycosides nitrates diuretics alfa-blockers statins	2005-2009: Visits prescribed due to the presence of Congestive Heart Failure or not-well defined cardiopathy, and other diseases of cardiovascular system. Beginning in 2010, the following ICD-9-CM-CM codes were in the record: 390-454,456-459	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Obliterans, Thrombophlebitis, Tibial, lliac, Femoral, or Popliteal Artery Disease, Varicose Veins of Lower Extremities, Secondary Hypertension, Budd Chiari Syndrome, Rheumatic Fever (stage 2/3) Vasculitis Other: Atherosclerosis, Cardiac Conditions, Cardiovascular Symptoms, Circulatory Disorders, Diseases of Arteries, Diseases of Veins, Disorders of Pulmonary Circulation, Lymphatic Disorders			
Endocrine	Adrenal Insufficiency, Cushing's Syndrome, Diabetes insipidus, Diabetes Mellitus Type 1, Diabetes Mellitus Type 2 and Hyperglycemic States,Hyperthyroidism, Hypoglycemia, Hypothyroidism, Monotropic Hormone Deficiency, Primary Amyloidosis, Thyroiditis, Klinefelter's Syndrome, Turner's or Noonan's Syndrome, Obesity Goiter: Nontoxic or Euthyroid (stage 2/3) Neoplasm, Benign: Acromegaly, Adenoma, Parathyroid, Hyperparathyroidism, Primary Hyperaldosteronism, Other Endocrine System Neoplasm, Malignant:Thyroid Other: Endocrine Disorders, Electrolyte Disorders, Nutritional and Metabolic Disorders Anomaly: Adrenal Hyperplasia	Insulins biguanides sulfonylureas vasopressin thyroid replacement antithyroid agents	2005-2009: Visits prescribed due to the presence of diabetes mellitus Beginning in 2010, the following ICD-9-CM codes were in the record: 240-278	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Ear, Nose, Throat	Diseases of Salivary Gland, Incl. Parotitis or Benign Tumors, Other Disorders of Oral Cavity (stage 2), Cholesteatoma, Meniere's Disease, Otitis Media, Sinusitis Hearing Loss due to: Acoustic Trauma, Otosclerosis Neoplasm, Benign: Larynx, Sinuses, Oral Cavity and Pharyngeal Structures Pharyngitis: Non-Streptococcal (stage 2)			
Eye	Cataract, Conjunctivitis: Bacterial, Contusion or Ruptured Globe, Dacryostenosis or Dacryocystitis, Detachment of the Retina, Ectropion or Entropion (Abnormal Lower Lid Position), Endophthalmitis, Foreign Body: Orbit, Fracture: Orbit, Blow-Out, Fungal Infection of the Eye, Glaucoma, Hypovitaminosis A, Laceration: Cornea, Macular Degeneration, Orbital Infection, Prematurity: Retinopathy, Ptosis of Upper Lid, Retrobulbar Orbital Hemorrhage, Trachoma, Other Eye Disorders Injury or Laceration: Eyelid, Periocular, Cornea, Conjunctiv Injury: Eyes, Nonionizing Radiation Keratitis: Acanthamoeba, Bacterial Neoplasm, Benign: Eye	Sympaticomimetic agents parasympaticomimetic agents anhydrase inhibitors ophthalmic beta blockers		
Gastrointestinal	Anorectal Suppuration, Celiac Disease, Clostridium difficile Colitis, Crohn's Disease, Diverticular Disease, Food Poisoning: Other Organisms (stage 3), Functional Digestive Disorders, Gastritis, Hemorrhoids, Hernia (External), Hernia (Hiatal or Reflux Esophagitis), Intussusception (stage 2), Irritable Bowel Syndrome, Gastroenteritis	Intestinal corticosteroids agents H2 antagonists prostaglandins proton pump inhibitors	2005-2009: Visits prescribed due to the presence of Gastrointestinal Diseases Beginning in 2010,	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Neoplasm, Benign: Adenomatous Polyps, Colon, Small Bowel, Other Gastrointestinal System Peptic Ulcer Disease, Salmonellosis (stage 3), Ulcerative Colitis, Vascular Insufficiency of the Bowels, Complications of Gastrointestinal Treatment, Gastroenteritis (stage 2/3) Other Diseases of Esophagus, Stomach, and Duodenum Other Gastrointestinal Disorders, Other Gastrointestinal Infections (stage 2), Other Gastrointestinal or Abdominal Symptoms Anomaly: Congenital Megacolon, Other Digestive or Hepatobiliary System Burns, Chemical: Esophagus, Stomach, or Small Intestine, Laceration: Esophagus		the following ICD-9-CM codes were in the record: 520-539,550-579	
Genitourinary	Bladder Disorders, Calculus of the Urinary Tract, Glomerulonephritis, Acute, Injury: Urinary Tract, Nephrotic Syndrome (stage 2/3), Renal Failure (stage 2/3), Urethritis, Urinary Tract Infections, Neoplasm, Benign: Urinary Tract, Other Disorders of Kidney or Ureter, Other Urinary Symptoms, Encounter for Dialysis, Anomaly: Defects of Kidney, Defects of Lower Genitourinary Tract, Syphilis: Congenital	Agents for hyperkalemia and hyperphosphatemia	2005-2009: Visits prescribed due to the presence of renal failure and Other diseases of the genito-urinary system Beginning in 2010, the following ICD-9-CM codes were in the record: 580-629	Visits prescribed for dialysis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Gynecological and Obstetrics	Anomaly: External Female Genitalia, Anomaly: Uterus, Dysfunctional Uterine Bleeding, Endometriosis, Neoplasm, Benign: Ovary (stage 2), Pelvic Inflammatory Disease, Uterine Infection, Uterovaginal Prolapse, Vulvovaginitis, Other Disorders of Female Genital System			
Hematological	Agranulocytosis, Anemia: Aplastic, Acquired (stage 2/3), Folic Acid Deficiency, Hemolytic (stage 2/3), Iron Deficiency, Sickle Cell, Thalassemia, Vitamin B-12 Deficiency, Other Graft versus Host reaction, Hemophilia A or B, Polycythemia Vera, Other Disorders of Blood and Blood-Forming Organs	Iron vitamin B12 folic acids	Beginning in 2010, the following ICD-9-CM codes were in the record: 280-289	
Hepatobiliary	Cholecystitis and Cholelithiasis, Cirrhosis of the Liver (stage 2/3), Disorders of Bilirubin Excretion, Hepatitis A, Hepatitis B, Hepatitis C, Hepatitis D, Hepatitis E, Hepatitis G, Hepatitis (Chemical), Pancreatitis, Wilson's Disease, Neoplasm, Benign: Hepatobiliary System, Other Hepatobiliary and Pancreatic Disorders, Other Hepatobiliary Infections, Other Pancreatic Disorders	Interferons blood substitutives and plasmatic protein fractions		
Immunologic Diseases	Human Immunodeficiency Virus Type I (HIV) Infection, Other Immunodeficient Disorders, Pneumonia: Pneumocystis carinii	Nucleosides and nucleotides reverse transcriptase inhibitors	2005-2009: Visits prescribed due to HIV Infections Beginning in 2010, the following ICD-9-CM codes were in the record: 279	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Infectious Diseases	Aspergillosis, Chlamydial Infection Except Trachoma or Pneumonia, Cryptococcosis, Cytomegalovirus Disease (Acquired), Infectious Mononucleosis (stage 2), Mucormycosis, Reye's Syndrome (stage 3), Rubella: Acquired (stage 3), Schistosomiasis, Other Bacterial Infections, Other Fungal Infections, Other Infectious and Parasitic Infections, Other Viral Infections, Cytomegalovirus Disease (Congenital), Parainfluenza Virus Infection, Pneumonia: Chlamydial, Sarcoidosis, Other Respiratory Infections, Scabies		Beginning in 2010, the following ICD-9-CM codes were in the record: 001-139	
Male Genital	Benign Prostatic Hypertrophy, Gonorrhea: Male, Prostatitis	Alfa-adrenoreceptor antagonists testosterone 5-alfa reductase inhibitors		
Musculoskeletal	Vitamin D Deficiency, Dislocation: Knee, Eosinophilia Myalgia Syndrome, Fracture: Acetabulum, Fracture: Calcaneus (stage 2), Fracture: Femur, Except Head or Neck, Fracture: Femur, Head or Neck, Fracture: Fibula (stage 2), Fracture: Humerus (Shaft), Fracture: Humerus (Supracondylar) (stage 2), Fracture: Radial Shaft, Ulna or Olecranon (stage 2), Fracture: Radius, Lower End (stage 2), Fracture: Tibia (stage 2/3), Fracture or Dislocation: Patella (stage 2), Fracture or Sprain: Ankle (stage 2), Fracture, Dislocation, or Sprain: Facial Bones (stage 2/3), Fracture, Dislocation, or Sprain: Foot (stage 2), Fracture, Dislocation, or Sprain: Hip or Pelvis (stage 2/3), Fracture, Dislocation, or	Colchicine uric acid inhibitors antiinflammatory non-steroids gold salts aminoquinolines bisphosphonates calcitonin	2005-2009: Visits prescribed due to the presence of Arthrosis, Arthritis and other osteo-muscular and connective diseases, and Fractures of femurs and other consequences of fractures.	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Sprain: Humerus (Head) or Shoulder (stage 2), Fracture, Dislocation, or Sprain: Wrist or Hand or Fingers (stage 2), Gout, Herniated Intervertebral Disc, Infectious Arthritis (stage 2/3), Injury, Chest Wall, Injury, Knee, Semilunar Cartilages (stage 2), Injury, Open Wound, or Blunt Trauma: Lower Extremity (stage 2), Injury, Open Wound, or Blunt Trauma: Upper Extremity (stage 2/3), Muscular Dystrophy, Osteoarthritis, Osteochondrodysplasia, Osteomalacia, Osteomyelitis (stage 2/3), Osteoporosis, Progressive Systemic Sclerosis, Rheumatoid Arthritis, Scoliosis of the Thoracic Spine, Spondylitis, Ankylosing, Systemic Lupus Erythematosus, Anomaly: Musculoskeletal System, Injury: Other and Ill-Defined Musculoskeletal Sites, Neoplasm, Benign: Musculoskeletal Syst. or Connective Tissue, Other Arthropathies, Bone and Joint Disorders, Other Disorders of Connective Tissue, Other Spinal and Back Disorders, Myasthenia Gravis, Complications of Surgical and Medical Care (stage 1), Injury, Open Wound, or Blunt Trauma: Abdomen or Trunk (stage 2/3), Injury: Other (stage 3)		Beginning in 2010, the following ICD-9-CM codes were in the record: 710-739	
Neurologic Diseases	Down's Syndrome, Herpes zoster, Poliomyelitis, Post-Polio Syndrome, Syphilis: Acquired, Tetanus (stage 1), Toxoplasmosis: Acquired (stage 3), Amyotrophic Lateral Sclerosis, Cerebral Palsy, Cerebrovascular Disease, Disease of Nervous System Secondary to Implants or Grafts, Epilepsy, Guillain-Barre Syndrome (stage 2), Headache (stage 2), Huntington's Chorea, Injury: Craniocerebral,	Anticholinesterase agents anticonvulsivant barbiturates and congeners alprostadil ergot alkaloids 5HT1 agonists dopamine	2005-2009: Visits prescribed due to the presence of Dementia and Alzheimer's syndrome, Parkinson's and	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Injury: Spine and spinal cord, Meningitis, Encephalitis, and Myelitis: Viral, Meningitis: Bacterial, Mental Retardation, Multiple Sclerosis, Neurofibromatosis Type I [Von Recklinghausen's Disease], Parkinson's Disease, Other CNS Inflammation, Infection, or Disorder, Other Cranial Nerve Disorders, Other Neurological Conditions, Other Peripheral Nerve Disorders, Other Spinal Lesions, Anomaly: Neural Tube Defects, Rubella: Congenital (stage 2), Anomaly: Other Nervous System, Injury: Other	MAO b inhibitors	other CNS degenerative disease, hemiplegia, monoplegia, and other associated syndroms, and acute and chronic cerebrovascular diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 320-389,797	
Psychological	Dementia: Primary Degenerative (Alzheimer's or Pick's), Antisocial Personality Disorder, Bipolar Disorder - Major Depressive Episode, Bipolar Disorder - Manic Episode, Depression, Generalized Anxiety Disorder, Obsessive-Compulsive Neurosis, Schizophrenia, Autism, Other Neuroses, Other Psychoses Drug Abuse, Dependence, Intoxication: Alcohol, Amphetamine, Barbiturate, Cannabis, Cocaine, Hallucinogen, Opioid, Other Eating disorders: Anorexia Nervosa, Bulimia Nervosa	Antidepressants antipsychotics agents	2005-2009: Visits prescribed due to the presence of psychoses, neuroses, and mental retardation Beginning in 2010, the following ICD-9-CM codes were in the record: 290-319	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Respiratory	Coxsackie and ECHO Infections (stage 2/3), Anomaly: Tracheoesophageal Malformations, Asbestosis, Asthma, Berylliosis, Byssinosis, Chronic Obstructive Pulmonary Disease, Coal Miner's Pneumoconiosis, Croup, Cystic Fibrosis, Emphysema, Hypersensitivity Pneumonitis, Influenza, Mycoplasma pneumoniae Infection, Parainfluenza Virus Infection (stage 2), Pneumonia: Bacterial, Pneumonia: Legionella, Pulmonary Alveolar Proteinosis, Pulmonary Embolism (stage 3), Radiation Pneumonitis, Silicosis, Tuberculosis, Complications of Tracheostomy, Other Disorders of Respiratory System, Other Respiratory Disease Due to External Agents, Other Respiratory Symptoms, Pneumonia: Aspiration, Neoplasm, Benign: Respiratory System	Inhaled corticosteroids beta-2-adrenoreceptor agonists xanthines leucotrienies antagonists cromolyn pancreatic enzymes mucolytics antituberculosis antibiotics isoniazid	2005-2009: Visits prescribed due to the presence of respiratory diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 460-519	
Skin	Herpes Virus Ocular Infection (stage 1), Urticaria, Candida (Monilial) Infections, Clostridial Wound Infection (stage 2), Herpes Simplex Infections, Complications of Surgical and Medical Care (stage 2/3), Anomaly: Integument (Genodermatoses), Decubitus Ulcers, Erythema Multiforme, Erythroderma, Immunologically Mediated Blistering Skin Diseases, Infections of Skin and Subcutaneous Tissue, Neoplasm, Malignant: Carcinoma, Basal Cell (stage 1), Neoplasm: Atypical Nevus (stage 1), Psoriasis Vulgaris, Other Inflammations & Infections of Skin & SubQ Tissue, Burns, Neoplasm, Benign: Skin or Subcutaneous Tissue (stage 1)	Oral and topical antipsoriasis agents	2005-2009: Visits prescribed due to the presence of decubitus ulcers and othere skin diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 680-709	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Predicting Risk Using a Population Based Longitudinal Database
Louis et al
STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	See abstract section: Design.
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	The outcome measures used and what was found are summarized in the sections of the abstract: Main outcome measures and Results
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	See the Introduction section pages 4 and 5
Objectives	3	State specific objectives, including any prespecified hypotheses	Our objectives are described in the last two paragraphs of the Introduction
Methods			
Study design	4	Present key elements of study design early in the paper	See last paragraph of the Introduction and the Methods section
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	See Study Data and Study Population at beginning of Methods section
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	Our study includes 100% of the adult population See Study Data and Study Population at beginning of Methods section
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	See the Dependent variable and Independent variable sections in the Methods section
Data sources/ measurement	8*	For each variable of interest, give sources of data	See the Dependent variable

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	and Independent variable sections in the Methods section. In addition we have included an Appendix with detailed mapping to independent variable.
Bias	9	Describe any efforts to address potential sources of bias	See Evaluation of the Models (page 10) in the Methods section
Study size	10	Explain how the study size was arrived at	Our study includes 100% of the adult population
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	See the Dependent variable and Independent variable sections in the Methods section. In addition we have included an Appendix with detailed mapping to independent variable.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	See the Modelling section and Evaluation of the Models section (pages 9 and 10) of the Methods
		(*b*) Describe any methods used to examine subgroups and interactions	See the Modelling section (pages 9 and 10) of the Methods
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	Our study includes 100% of the adult population
		(*e*) Describe any sensitivity analyses	See Evaluation of the Models last paragraph of Methods.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	See Page 11 Results section
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	See Table 1 and Results section (page 11-14)
		(b) Indicate number of participants with missing data for each variable of interest	

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	See Table 2 and figure and pages 14 and 15 of Results section
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	See pages 13
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	See Table 2 and page 13 and 14.
Discussion			
Key results	18	Summarise key results with reference to study objectives	See Discussion section page 16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	See Discussion section pages 17 and 18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	See Discussion section page 18
Generalisability	21	Discuss the generalisability (external validity) of the study results	See Discussion section pages 18 and 19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	The study was funded by the Emilia-Romagna, region of Italy. See page 3.

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open

Predicting Risk of Hospitalization or Death: A Retrospective Population Based Analysis

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-005223.R1
Article Type:	Research
Date Submitted by the Author:	02-Jul-2014
Complete List of Authors:	Louis, Daniel; Jefferson Medical College, Center for Research in Medical Education and Health Care Robeson, Mary; Jefferson Medical College, Center for Research in Medical Education and Health Care McAna, John; Thomas Jefferson University, Jefferson School of Population Health Maio, Vittorio; Thomas Jefferson University, Jefferson School of Population Health Keith, Scott; Thomas Jefferson University, Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics Liu, Mengdan; Jefferson Medical College, Center for Research in Medical Education and Health Care Gonnella, Joseph; Jefferson Medical College, Center for Research in Medical Education and Health Care Grilli, Roberto; Agenzia Sanitaria e Sociale Regione Emilia-Romagna,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Patient-centred medicine
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

Predicting Risk of Hospitalization or Death: A Retrospective Population Based Analysis

Daniel Z. Louis[1], Mary Robeson[1], John McAna[2], Vittorio Maio[2], Scott W. Keith[3], Mengdan Liu[1], Joseph S. Gonnella[1], Roberto Grilli[4]

[1]Center for Research in Medical Education and Health Care, Jefferson Medical College of Thomas Jefferson University. 1015 Walnut Street, Suite 319. Philadelphia, PA 19107, USA

[2]Jefferson School of Population Health, Thomas Jefferson University. 901 Walnut Street, 10th Floor. Philadelphia, PA 19107, USA

[3]Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics, Thomas Jefferson University. 1015 Chestnut Street, Suite M-100. Philadelphia, PA 19107, USA

[4]Agenzia Sanitaria e Sociale Regione Emilia-Romagna, Viale Aldo Moro, 21. Bologna, 40128 Italy.

Corresponding author: Daniel Z. Louis, Jefferson Medical College, 1015 Walnut Street, Suite 319, Philadelphia, PA 19107. USA; Tel: 215-955-6959, daniel.louis@jefferson.edu

Copyright: The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (non-exclusive for Dr. Grilli as a government employee) on a worldwide basis to the BMJ publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license.

Ethical approval: This study was approved by the Institutional Review Board of Thomas Jefferson University as an expedited retrospective database/record review. The IRB granted a waiver of informed consent.

For peer review only

Abstract

Objectives: Develop predictive models using an administrative health care database that provide information for Patient Centered Medical Homes to proactively identify patients at risk of hospitalization for conditions that may be impacted through improved patient care.

Design: Retrospective health care utilization analysis with multivariate logistic regression models.

Data: A population-based longitudinal database of residents served by the Emilia-Romagna, Italy health service in the years 2004-2012 including demographic information and utilization of health services by 3,726,380 people age ≥ 18 years.

Outcome measures: Models designed to predict risk of hospitalization or death for problems that are potentially avoidable were developed and evaluated using the area under the receiver operating curve C-statistic, in terms of their sensitivity, specificity, and positive predictive value, and for calibration to assess performance across levels of predicted risk.

Results: Among the 3,726,380 adult residents of Emilia-Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions or died in 2012 (3.6% hospitalized, 1.3% died). The C-statistic for the model predicting 2012 outcomes was 0.856. The model was well calibrated across categories of predicted risk. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%.

Conclusions: We have developed a population-based model using a longitudinal administrative database that identifies the risk of hospitalization for residents of the Emilia-Romagna region

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

with a level of performance as high as, or higher than, similar models. The results of this model, along with profiles of patients identified as high risk are being provided to the physicians and other health care professionals associated with the Patient Centered Medical Homes to aid in planning for care management and interventions that may reduce their patients' likelihood of a preventable, high-cost hospitalization.

Strengths and Limitations of this study:

- This study included the entire adult population of the Emilia-Romagna Region of Italy, over 3.7 million people.
- The study used an existing longitudinal administrative health care database with both the advantage of much lower cost than new data collection and the disadvantage of potential errors in administrative data.
- The results of the study are being used to assist in the development of newly formed Patient Centered Medical Homes.

Keywords: hospitalization, risk, medical home, patient-centered care

Word Count: 3,861 (excluding title page, abstract, references, figure, and tables.)

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

The predominant healthcare delivery system, which has been a passive model, reacting to patients' problems, is shifting to a more proactive model designed to take the initiative in providing care for an increasingly older population that has a greater prevalence of chronic conditions, often with multiple medical and social needs. These changes are driving the reorganization of the primary care system, emphasizing coordination and cooperation among healthcare professionals.[1] Among the approaches to addressing this need has been the establishment of primary care organizations incorporating integrated teams of physicians and other healthcare professionals that "seek to increase the influence of primary care professionals, and in particular general practitioners (GPs), in health planning and resource allocation."[2] Prominent among these new models of primary care is the Patient Centered Medical Home, an organization in which a team of healthcare providers is engaged in delivering comprehensive, coordinated, patient-centered care to patient defined populations.[3]

Primary care has a central role in the Italian National Health Service (NHS). Twenty one regional governments are responsible for ensuring the delivery of a health benefits package through a network of geographically defined, population-based Local Health Authorities. Primary care physicians work for these authorities as independent contractors and act as "gatekeepers" for specialty and other referral services for their patients.[4]

With the belief that a strong primary care system is conducive to improving population health, the NHS initiated reforms that encouraged primary care physicians to organize into collaborative arrangements. To this end, the Regione Emilia-Romagna (RER), a large northern region with a population of about 4.5 million, has recently launched a plan in its 11 Local Health Authorities to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

establish Patient-Centered Medical Homes to better coordinate patient care and help patients avoid unnecessary hospitalizations.

The identification of those patients who would benefit most from outreach efforts is fundamental to achieving these goals of promoting and practicing population health in Patient-Centered Medical Homes. To accomplish this, predictive models and risk stratification tools are needed to identify patients at risk of a worsening of their health status. According to Knutson and Bella,[5] "Predictive models are data-driven, decision-support tools that estimate an individual's future potential health care costs and/or opportunities for care management." A good model will identify as many of these patients as possible while excluding those for whom these interventions would be unnecessary or ineffective.

The RER has established three objectives for this project: 1) develop predictive models to identify patients at high risk of hospitalization or death, 2) create "risk of hospitalization" patient profiles that provide information about their high-risk patients to the general practitioners in the newly formed Patient-Centered Medical Homes, and 3) assess the extent to which these models and reports provide additional information useful in the identification of patients who may benefit from case management or disease management.

This paper will address the first of the three goals. We describe the development of a predictive model using the RER's regional longitudinal administrative health care database to help identify patients who are most at risk of hospitalization for conditions that may be impacted through improved patient care. This model will then be used to inform the providers associated with the Patient-Centered Medical Homes and aid in their planning for care management and interventions that can reduce their patients' likelihood of a preventable, high-cost hospitalization.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Methods

Study Data and Study Population

The model was developed using the population-based longitudinal health care database of the residents served by the RER Health Service in the years 2004 through 2012. [6] This administrative database includes demographic information for all residents (gender, birth and death dates, location of current residence, and primary care physician), hospital discharge abstract data (ICD-9-CM diagnosis and procedure codes, and admission and discharge dates), emergency room utilization information, outpatient pharmacy data at the individual prescription level, specialty care (laboratory, diagnostics, therapeutic procedures, rehabilitation, and specialist visits), home health data, and information on each primary care physician in the region. Each patient has an anonymous identifier assigned by the RER so that an individual's utilization can be tracked over time without jeopardizing patient privacy.

The study population consisted of all residents of the RER who were at least 18 years of age and still alive as of 31st December 2011.

Dependent Variable

The dependent variable was defined as the occurrence of a hospitalization for problems that are potentially avoidable, or whose progression may have been avoided or delayed through appropriate patient care, or the death of the individual, either in or out of the hospital, for any reason in 2012. We decided to not limit the hospitalization to emergency admissions, since a planned admission may also be an indicator of a worsening medical problem. In order to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

operationally define the dependent variable, we (authors JSG and DZL) reviewed the Disease Staging[7, 8] primary diagnostic category and severity stage of all day and inpatient hospital admissions (for adults age 18+) in RER for one year, to select those admissions that should be included in the dependent variable.

Deliveries, admission for dental diseases or admissions for vague signs or symptoms with no identified etiology were excluded. Admissions for problems that are not predictable/preventable were excluded while those where screening may identify problems that can potentially be treated to avoid progression were included. For example, admissions for stage 1, chronic cholecystitis or cholelithiasis were excluded, but admissions for advanced stage 2 or 3 complications such as ascending cholangitis or pancreatitis were included.

We felt that inclusion of hospitalization for cancer in the dependent variable should depend on the ability to either prevent or avoid progression of the disease. We therefore included colon cancer and cervical cancer in the definition because they are potentially preventable but excluded all other cancers where prevention/prediction is not currently possible.

Inclusion of injuries, burns, or toxic reaction to prescription or non-prescription drugs would ideally be based on the cause of these problems. Since the etiology of these problems is typically not available in the administrative data being used in this project, we made the decision to include or exclude based on our subjective judgment of the likelihood of preventability. For example, adverse drug reactions were included but burns were excluded from the definition of the dependent variable.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

There is no obvious medical reason for a hospital admission for patients with stage 1 diabetes mellitus or stage 1 essential hypertension without complications. These problems are typically treatable in the outpatient setting. A hospitalization implies a potential problem in the care of these patients, so we decided to include these admissions as a part of the dependent variable.[9]

Independent Variables

A broad range of candidate predictor variables was developed taking advantage of the RER administrative data. The independent variables used for modeling were defined from the RER administrative data for the years 2004 through 2011. Demographic data included patient age, sex, and geographic location of residence. We developed a mapping to broad disease categories defined primarily in terms of the affected body system from home health care data, pharmacy data, and hospital discharge abstract data. (See Appendix 1.)

For those patients who had been hospitalized, more specific diagnostic data were available. We reviewed the classification of patients hospitalized historically using the Disease Staging diagnostic category and disease severity stages.[7,8] Based on the frequencies specific diagnostic category/stage predictor variables were defined for either specific stages of frequent diseases, or by combinations across similar categories. Predictor variables were defined based on the number of emergency room visits using the RER classifications system for the urgency of the visit.

Pharmacy data were used to identify polypharmacy[10] (defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days), potential drug-drug interactions

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(DDI)[11] and potentially inappropriate medication use in patients[12] 65 years and older. Since cardiovascular disease is highly prevalent, we reviewed the use of cardiovascular drugs and created a variable for each of the following 11 classes of drugs (oral anti-coagulants, beta-blockers, angiotensin converting enzyme / angiotensin II receptor blockers, anti-platelets, calcium channel blockers, anti-arrhythmics, digitalis glycosides, nitrates, diuretics, alfa-blockers, statins) to account for the complexity of therapeutic regimen at the patient level.

To take advantage of the fact that the RER database includes multiple years of data, we created history variables using the utilization for each year of data available. Since we were working with the 2011 data to predict hospitalization or death in 2012, we created history variables based on 2004-2010 data. This set included 83 of the diagnostic category/stage variables as well as 11 variables based on pharmacy utilization such as exposure to polypharmacy and use of cardiovascular drugs. If the individual had a history of a disease in any of the years from 2004 to 2010 they were flagged as having a history of that disease and this was used as a potential predictor variable.

Modeling

Logistic regression models were used to estimate predicted probabilities for the occurrence of an inpatient hospital stay for the selected conditions or death for individual patients. Risk of hospitalization or death, and the variables that relate to those risks are highly dependent on age and gender. Regression models were fit in each of 14 gender and age strata using SAS Version 9.2 (SAS Institute, Cary NC). A stepwise process with relaxed covariate entry and retention criteria (inclusion p-value <=0.8, retention <=0.5) was used. At each step in this process, an

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

attempt is made to remove any unimportant variables from the model before adding a potentially important variable. Each addition or deletion of a variable to or from a potential model is a separate step and, at each step, a new model is fitted. This process results in a reduced, but robust set of independent variables that predict outcome or that might have importance as adjustment terms for the model in each age/gender stratum.

Evaluation of the Models

The predictive accuracy of the modelling was evaluated using C-statistics (the area under the receiver operating characteristics curve), along with three measures traditionally used with clinical screening tools: sensitivity, specificity and positive predictive value.

C-statistics were used to evaluate the models in two ways. The first evaluation consisted of fitting the model developed using utilization and demographic data from 2011, along with historical variables based on 2004-2010 data, and outcomes (hospitalization or death) from 2012 and then computing a C-statistic to evaluate how the models performed at predicting those outcomes on which the models were conditioned. However, this evaluation is not consistent with evaluating how the data are used in practice. In practice, we have current predictor information, but the outcomes have not been realized. To better estimate how the models are likely to perform in this setting, we fit models to outcomes data up to a year prior to the most current available (e.g., 2011 outcomes modelled with predictors from 2010, along with historical variables based on 2004-2009 data). We then computed a C-statistic for projections made on the risk of hospitalization or death outcomes (in 2012) using the next year's predictor information (in 2011). This way, the models are forced to make projections into the future, but we have the actual

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

observed outcomes data to evaluate the modelling process as it would be used in practice. The resulting C-statistics obtained from these two model runs were compared.

In order to evaluate the performance of the model across different risk thresholds we classified predicted risk scores. "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%. These risk thresholds were selected after consultation with physicians practicing in the medical homes to yield a total of about 10% of the 1,500 patients enrolled with a typical primary care physician.

Results

Among the 3,726,380 adult residents of Emilia Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions defined earlier or died in 2012 (3.6% hospitalized, 1.3% died).

Table 1 shows the distribution of the demographics (age and gender), number of chronic conditions, body systems impacted by the selected chronic conditions, polypharmacy and inappropriate prescribing among the eligible RER residents, as of December 31, 2011. The table also compares these characteristics of the total adult population of the region to the subgroups of the population classified in the "very high risk" and "high risk" categories. Based on the model results, 114,255 individuals were identified as having a predicted risk of hospitalization or death in 2012 of ≥ 25% and classified as "very high risk." An additional 134,610 individuals had a predicted risk of hospitalization or death in 2012 of 15-24% and were classified as "high risk."

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1. Demographic and clinical characteristics of the Regione Emilia-Romagna population, overall and by risk category

	Total Population*		Very High Risk**		High Risk**	
	3,726,380		114,255		134,610	
Gender	N.	%	N.	%	N.	%
Male	1,788,048	48.0%	54,357	47.6%	61,803	45.9%
Female	1,938,332	52.0%	59,898	52.4%	72,807	54.1%
Age groups						
18-24	258,338	6.9%	76	0.1%	105	0.1%
25-34	499,786	13.4%	302	0.3%	391	0.3%
35-44	732,626	19.7%	1,137	1.0%	1,198	0.9%
45-54	676,047	18.1%	2,612	2.3%	2,485	1.8%
55-64	550,689	14.8%	5,391	4.7%	5,287	3.9%
65-74	482,346	12.9%	13,154	11.5%	14,471	10.8%
74-85	364,369	9.8%	33,430	29.3%	44,857	33.3%
85+	162,179	4.4%	58,153	50.9%	65,816	48.9%
Number of Chronic Conditions						
0-1	2,775,888	74.5%	8,176	7.2%	24,618	18.3%
2 or more	950,492	25.5%	106,079	92.8%	109,992	81.7%
5 or more	99,337	2.7%	45,445	39.8%	20,576	15.3%
Selected Conditions/Body Systems						
Cancer	99,328	2.7%	23,872	20.9%	14,305	10.6%
Cardiovascular	967,796	26.0%	96,157	84.2%	103,749	77.1%
Male Genitourinary#	130,609	7.3%	14,616	26.9%	16,776	27.1%
Ear, Nose, Throat	5,364	0.1%	240	0.2%	242	0.2%
Endocrine	429,528	11.5%	40,653	35.6%	37,471	27.8%
Eye	114,117	3.1%	9,558	8.4%	13,478	10.0%
Gastrointestinal	580,946	15.6%	74,718	65.4%	66,305	49.3%
Gynecologic##	21,806	1.1%	333	0.6%	405	0.6%
Hematologic	45,022	1.2%	15,353	13.4%	6,591	4.9%
Hepatobiliary	24,785	0.7%	6,477	5.7%	3,306	2.5%
Immunologic	3,281	0.1%	464	0.4%	273	0.2%
Infectious Disease	4,723	0.1%	2,207	1.9%	727	0.5%
Musculoskeletal	419,184	11.2%	43,436	38.0%	41,000	30.5%
Neurologic	173,751	4.7%	34,494	30.2%	24,838	18.5%
Psychological	291,308	7.8%	43,387	38.0%	33,715	25.0%
Respiratory	176,830	4.7%	39,082	34.2%	21,763	16.2%
Skin	28,339	0.8%	7,645	6.7%	3,008	2.2%

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Urogenital	37,728	1.0%	16,501	14.4%	5,740	4.3%
Polypharmacy^	609,278	16.4%	92,153	80.7%	92,156	68.5%
Any potentially inappropriate medications (age 65 years or older)^^	257,033	25.5%	51,055	48.7%	49,003	39.2%

* Adults (age 18 or older) and alive at 31 December 2011.

** "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%.

Men only.

Women only.

^ Polypharmacy is defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days.

^^ The list of potentially inappropriate medications can be found in: Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *Journal of Clinical Pharmacy and Therapeutics* 2010;35:219-229.

There was little difference across the risk categories by gender. Age distributions for the "very high risk" and "high risk" groups were shifted more towards the older age groups than those in the overall study population. Residents age 85 or older represented about 4.5% of the RER population, but about 50% of the "very high" and "high" predicted risk groups. More than 75% of the residents over age 85 were classified as "very high" or "high" risk. However, age alone was not sufficient to their predict risk. For example, residents between 75 and 84 years of age made up 23% of the "very high" risk group and 41% of the "high" risk group, but over 85% of the residents in this age category had neither "very high" nor "high" predicted risk.

Across age and gender strata, demographics and heath care utilization experience in 2011 were the most commonly used independent variables for predicting hospitalization or death in 2012. Selected history variables flagging chronic problems such as cardiovascular disease, diabetes mellitus and chronic renal failure and a history of prescriptions for cardiovascular medications and polypharmacy were also significant predictors.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The residents in the two higher risk groups were more likely than others to have multiple chronic diseases and to experience polypharmacy and inappropriate medication use. The residents identified as "very high risk" or "high risk" by the model also showed a number of striking differences from others in terms of the occurrence of some of the most prevalent health conditions by type and body system. Although cardiovascular conditions were not uncommon in the total adult population (26.0%), they were far more common among those classified as "very high risk" and "high risk" (84.2% and 77.1%, respectively). Similarly, gastrointestinal conditions affected 15.6% of the total population, but were diagnosed in 65.4% of the "very high risk" and 49.3% of the "high risk" patients. Cancer occurred in 2.7% of the total population, but 20.9% of the "very high risk" and 10.6% of the "high risk" patients had a cancer diagnosis. Mental health problems were identified in 7.8% of the adult population, but in 34.2% of the "very high risk" and 25.0% of the "high risk" patients.

The C-statistic for the model of 2012 outcomes developed using 2011 predictors and the C-statistic based on the parameters from the model of 2011 outcomes regressed on 2010 predictors applied to the 2011 predictors and 2012 outcomes were very similar (0.856 and 0.853, respectively). These results suggest that the relationship between predictors and risk of hospitalization changed little in one year and that model parameters developed in a prior year can be used reliably with the most current year's data to predict unknown outcomes in the next year with only a minimal loss in performance in this population.

Table 2 shows the sensitivity, specificity, positive predictive value and number of true positives for the model at the two selected cut-off points. The sensitivity (percentage of patients actually hospitalized who had been identified by the model as having a predicted risk higher than the cut-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

off point) was 29.8% for those with the "very high" risk scores. This percentage represents 46,950 of the 157,550 residents of the region who were hospitalized for a selected condition or died in 2012. If we modify the risk score threshold to include individuals with a predicted risk of hospitalization for selected conditions or death of $\geq$ 15% (i.e., both the "very high risk" and the "high risk" patients) the sensitivity is .471. The true negative rate (specificity) is very high for both risk thresholds (.981 and .951, respectively).

Table 2. Performance of the "Risk of Hospitalization" model for residents identified as "Very High Risk" and "High or Very High Risk"

	Cut-off points for comparison	
Measure	**"Very high risk"***	**"Very high risk"* + "High risk"****
Sensitivity#	0.298	0.471
Specificity##	0.981	0.951
Positive Predictive Value^	0.411	0.298
True positives^^	46,950	74,196

* "Very high risk" is defined as patients with a predicted risk of hospitalization of $\geq$ 25%.
** "Very high risk" + "High risk", is defined as patients with a predicted risk of hospitalization of $\geq$15%.
# Sensitivity is defined as the proportion of those hospitalized who were predicted to be hospitalized (true positive rate).
Specificity is the proportion of those not hospitalized who were not predicted to be hospitalized (true negative rate).
^ Positive Predictive Value is the proportion of those predicted to be hospitalized who were actually hospitalized.
^^ True positives are the number of residents who were predicted to be at risk for hospitalization at the predicted risk threshold and were actually hospitalized.

The model appears to be well calibrated across levels of risk. The Figure depicts the RER population divided into groups by deciles of predicted risk of hospitalization or death from the models. The observed prevalence of hospitalization or death is compared to the average predicted risk among individuals in each of the ten predicted risk groups. For example, the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

overall rate of hospitalization for the selected conditions or death in 2012 was 4.2%. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%. (Regression coefficients and significance levels of independent variables for models for each ot 14 age and gender strata are displayed in Appendix 2.)

[Insert Figure about here]

Discussion

We have developed a population-based model that identifies the risk of hospitalization for all adult RER residents and does so with a level of performance (c=0.85) as high as, or higher than, similar models. In addition, we believe that the definition of the dependent variable chosen for our models increases the probability that they are identifying patients who risk can potentially be improved by appropriate care. A systematic review by Kansagara[13] of models designed to predict readmissions, showed C-statistic results in the range of 0.55 to 0.83. Recent work by Billings et al[14] to develop models predictive of emergent admissions in the UK had results ranging from 0.73 to 0.78. Li Wang, et al. (2013),[15] using information available through the US Veteran's Administration that also included lab data, demonstrated c-statistics of 0.81 and 0.79, respectively, for their models of 90-day or 12-month hospitalization or death outcomes. At a predicted risk of ≥25% our model had a Positive Predictive Value (PPV) of .411. Billings et al[14] reported a PPV of .417 at a risk threshold of 30. There is a trade-off in using our model, or any predictive model, between the threshold for follow-up and predictive accuracy. A lower risk threshold would identify more patients but with a lower prevalence of hospitalization or death.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Although previous studies have developed models predictive of hospital care, these models fall short of the needs of the Patient Centered Medical Homes being implemented in RER. Typically, these models have focused on specific age groups,[16] conditions, or types of admissions, such as emergent[14] or unplanned admissions or rehospitalizations, or health insurance plans in the United States, including private insurance plans, Medicare and Medicaid plans.[17,18] The models we have developed are applied to the entire adult population of RER. They use existing administrative data, which makes them cost effective to apply.

Patient Centered Medical Homes, including those instituted in RER, are responsible for addressing the needs of their population and making the best use of their finite resources to accomplish this. Preventing unnecessary admissions could improve both the quality of care and health status of the enrolled population, and result in a substantial savings. To accomplish this, predictive models and risk stratification tools such as those developed for this project are needed to identify patients at risk of preventable admissions and provide information that can be used by the medical homes to help manage care.

There are some limitations to the model. The model is developed from administrative data. Administrative data are collected for reimbursement and tracking utilization and not for medical research. They lack the clinical specificity that would desirable in assessing an individual's medical problems. While the hospital discharge abstract data do include diagnostic information coded using ICD-9-CM, no similar data are available for outpatient encounters in the RER database. The mortality data available for this project did not include information about cause of death. Therefore, some proportion of patients whose death was not predictable were included, limiting model performance. In addition, our models use prior utilization among the predictor

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

variables. With the administrative data we cannot distinguish between appropriate and inappropriate treatment which may bias our results.

Despite the limitations of administrative data, they have many advantages for this project: they are readily available, relatively inexpensive to analyze and cover large populations over many years. They are ideal for uncovering patterns of care. If information from the medical records is needed, the results of these analyses can then be supplemented by focused clinical reviews at the local level. Also, The RER has a system in place to monitor the quality of diagnosis and procedure coding in their hospital discharge abstract data. Controls at both the hospital and regional level assess the validity of coding and the consistency of codes assigned such as congruity between sex, age and diagnosis and between diagnosis and procedure. The existence of the RER administrative database made it feasible to develop the models described in this article at relatively low cost and to update the models over time without additional data collection that others have found necessary.[14]

Currently, these risk scores are being integrated with other information in profiles of high-risk patients furnished to providers in 12 newly formed medical homes, including 83 primary care physicians serving a total of about 100,000 patients, in the Parma Local Health Authority located in RER. Along with the risk scores, this information includes data about previous hospitalizations, use of referrals, medications, long-term care and home care services, and a number of process-like quality indicators for diabetic and cardiovascular patients, and for appropriate medication use in older patients.

Of course, model results need to lead to an effective intervention to have a positive impact on patient care. To this end, we are working with the physicians, nurses, and other health care

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

professionals as well as the administration of the newly formed Medical Homes in Parma to assist them in understanding how to use the results of these models and in developing potentially effective interventions. The individual profiles of high risk patients provided to the health care team in the Medical Homes allow them to trigger specific actions such as inviting patients to enroll in disease management programs for chronic problems such as heart failure, chronic obstructive pulmonary disease, or diabetes mellitus, activating home health assistance, initiating a medication review, or recommending that the patient come in for an office visit. An evaluation of the use and usefulness of the profiles and intervention is under way.

In summary, these models provide a means of identifying patients at high risk for hospitalization. The risk predictions, in conjunction with the risk profile, show promise as a useful organizational tool for the regional Patient Centered Medical Homes to develop and implement proactive case management and disease management programs. The RER is reviewing the results of the Parma Local Health Authority pilot project of the profiles. Once their usefulness has been further evaluated, their use will be expanded to other Medical Homes in development in the other Local Health Authorities in the Emilia-Romagna region. If similar data are available, these models can be applied in other Italian regions and other countries investing in organization similar to the Patient Centered Medical Home.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Contributorship statement: DZL, RG, VM, and JSG were responsible for the conceptualization of this project. MR, JM, and ML were responsible for creation of the datasets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form and declare grants from the Regione Emilia-Romagna (Louis, Robeson, McAna, Maio, Keith, Liu, Gonnella) during the conduct of the study. Louis and Gonnella declare personal fees from Truven Health Analytics, The National School of Public Health, Portugal, INSIEL Mercato.

Funding: This research was funded by the Agenzia Sanitaria e Sociale of the Emilia-Romagna Region, Italy. The resulting models are being used by the Local Health Authority of Parma, Italy and the Patient Centered Medical Homes located in the Parma Local Health Authority.

Data Sharing: No additional data available.

For peer review only

1 Saltman RB, Rico A, Boerma W, eds. Primary care in the driver's seat? Organizational reform in European primary care. New York, New York: World Health Organization European Observatory on Health Systems and Policies Series 2006.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

2 Smith J, Goodwin N. Towards managed primary care: The role and experience of primary care organizations. Burlington, VT: Ashgate Publishing Company 2006.

3 Defining the PCMH. Agency for Healthcare Research and Quality Web site. http://www.pcmh.ahrq.gov/portal/server.pt/community/pcmh__home/1483/pcmh_defining_the_pcmh_v2 (accessed June 30, 2013).

4 Lo Scalzo A, Donatini A, Orzella l, et al. Italy: Health system review. *Health Syst Transit*. 2009;**11**(6):1-216.

5 Knutson D, Bella M, Llanos K. Predictive Modeling: A Guide for State Medicaid Purchasers. Center for Health Care Strategies, Inc. 2009.

6 The Emilia-Romagna Regional Health Service and the new welfare system: Facilities, expenditure, and activities as of 31.12.2010. Programs, agreements and organisational models. Regione Emilia-Romagna Assessorato Politiche per la Salute. Bologna, Italy 2012.

7 Gonnella JS, Hornbrook MC, Louis DZ. Staging of Disease: A Case-Mix Measurement. *JAMA* 1984;**251**:637–44.

8 Gonnella JS, Louis DZ, Gozum ME, et al, eds. Clinical Criteria for Disease Staging. 6th edition. Ann Arbor, MI: Truven Health Analytics 2012.

9 Louis DZ, F Taroni, R Melotti, et al. Increasing appropriateness of hospital admissions in the Emilia-Romagna region of Italy. *J Health Serv Res Policy* 2008;**13**(4):202-8.

10 Slabaugh L, Maio V, Templin M, et al. Prevalence and risk of polypharmacy amongst elderly primary care patients in Regione Emilia-Romagna, Italy. *Drug Aging* 2010;**27**:1019-28.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

11 Gagne JJ, Maio V, Rabinowitz C. Prevalence and Predictors of Potential Drug-Drug Interactions among Ambulatory Patients in Regione Emilia-Romagna, Italy. *J Clin Pharm Ther* 2008;**33**:141-51.

12 Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *J Clin Pharm Ther* 2010;**35**:219-29.

13 Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: A systematic review. Washington, DC: Department of Veterans Affairs US 2011.

14 Billings J, Georghiou T, Blunt I, et al. Choosing a model to predict hospital admission: An observational study of new variants of predictive models for case finding. *BMJ Open* 2013;**3**(8):e003352. doi: 10.1136/bmjopen-2013-003352 [published Online First: 26 August 2013].

15 Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care* 2013;**51**(4):368-73.

16 Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. *Med Care* 2008;**46**(7):726-31.

17 Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care* 2012;**50**(2):131-9.

18 McAna JF, Crawfrod AG, Novinger BW, et al. A predictive model of hospitalization risk among diabled Medicaid enrollees. *Am J Manag Care* 2013;**19**(5):166-74.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure Legend:

Figure. Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Predicting Risk of Hospitalization or Death: A Retrospective ~~Using a~~ Population Based ~~Longitudinal Database~~Analysis

Daniel Z. Louis[1], Mary Robeson[1], John McAna[2], Vittorio Maio[2], Scott W. Keith[3], Mengdan Liu[1], Joseph S. Gonnella[1], Roberto Grilli[4]

[1]Center for Research in Medical Education and Health Care, Jefferson Medical College of Thomas Jefferson University. 1015 Walnut Street, Suite 319. Philadelphia, PA 19107, USA

[2]Jefferson School of Population Health, Thomas Jefferson University. 901 Walnut Street, 10th Floor. Philadelphia, PA 19107, USA

[3]Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics, Thomas Jefferson University. 1015 Chestnut Street, Suite M-100. Philadelphia, PA 19107, USA

[4]Agenzia Sanitaria e Sociale Regione Emilia-Romagna, Viale Aldo Moro, 21. Bologna, 40128 Italy.

Corresponding author: Daniel Z. Louis, Jefferson Medical College, 1015 Walnut Street, Suite 319, Philadelphia, PA 19107. USA; Tel: 215-955-6959, daniel.louis@jefferson.edu

Copyright: The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (non-exclusive for Dr. Grilli as a government employee) on a worldwide basis to the BMJ publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any ~~0ther~~ other BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license.

Ethical approval: This study was approved by the Institutional Review Board of Thomas Jefferson University as an expedited retrospective database/record review. The IRB granted a waiver of informed consent.

~~**Competing interests:** All authors have completed the Unified Competing Interest form and declare grants from the Regione Emilia-Romagna (Louis, Robeson, McAna, Maio, Keith, Liu, Gonnella) during the conduct of the study. Louis and Gonnella declare personal fees from Truven Health Analytics, The National School of Public Health, Portugal, INSIEL Mercato.~~

For peer review only

Abstract

Objectives: Develop predictive models using an administrative health care database that provide information for Patient Centered Medical Homes to proactively identify patients at risk of hospitalization for conditions that may be impacted through improved patient care.

Design: Retrospective health care utilization analysis with multivariate logistic regression models.

Data: A population-based longitudinal database of residents served by the Emilia-Romagna, Italy health service in the years 2004-2012 including demographic information and utilization of health services by 3,726,380 people age ≥ 18 years.

Outcome measures: Models designed to predict risk of hospitalization or death for problems that are potentially avoidable were developed and evaluated using the area under the receiver operating curve C-statistic, in terms of their sensitivity, specificity, and positive predictive value, and for calibration to assess performance across levels of predicted risk.

Results: Among the 3,726,380 adult residents of Emilia-Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions or died in 2012 (3.6% hospitalized, 1.3% died). The C-statistic for the model predicting 2012 outcomes was 0.856. The model was well calibrated across categories of predicted risk. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%.

Conclusions: We have developed a population-based model using a longitudinal administrative database that identifies the risk of hospitalization for residents of the Emilia-Romagna region

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

with a level of performance as high as, or higher than, similar models. The results of this model, along with profiles of patients identified as high risk are being provided to the physicians and other health care professionals associated with the Patient Centered Medical Homes to aid in planning for care management and interventions that may reduce their patients' likelihood of a preventable, high-cost hospitalization.

Strengths and Limitations of this study:

- This study included the entire adult population of the Emilia-Romagna ~~regon~~ Region of Italy, over 3.7 million people.
- The study used an existing longitudinal administrative health care database with both the advantage of much lower cost than new data collection and the disadvantage of potential errors in administrative data.
- The results of the study are being used to assist in the development of newly formed Patient Centered Medical Homes.

Keywords: hospitalization, risk, medical home, patient-~~centerd~~centered care

Word Count: 3,~~547~~ 861 (excluding title page, abstract, references, figure, and tables.)

~~**Contributions of authors:** DZL, RG, VM, and JSG were responsible for~~responsible for the ~~conceptualiztion~~conceptualization of this project. MR, JM, and ML were responsible for creation of the datasets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.~~

~~**Funding:** This reseach~~research ~~was funded by the Agenzia Sanitaria e Sociale of the Emilia-Romagna Region, Italy. The resulting models are being used by the Local Health Authority of Parma, Italy and the Patient Centered Medical Homes located in the Parm~~Parma ~~Local Health Authority.~~

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

The predominant healthcare delivery system, which has been a passive model, reacting to patients' problems, is shifting to a more proactive model designed to take the initiative in providing care for an increasingly older population that has a greater prevalence of chronic conditions, often with multiple medical and social needs. These changes are driving the reorganization of the primary care system, emphasizing coordination and cooperation among healthcare professionals.[1] Among the approaches to addressing this need has been the establishment of primary care organizations incorporating integrated teams of physicians and other healthcare professionals that "seek to increase the influence of primary care professionals, and in particular general practitioners (GPs), in health planning and resource allocation."[2] Prominent among these new models of primary care is the Patient Centered Medical Home, an organization in which a team of healthcare providers is engaged in delivering comprehensive, coordinated, patient-centered care to patient defined populations.[3]

Primary care has a central role in the Italian National Health Service (NHS). Twenty one regional governments are responsible for ensuring the delivery of a health benefits package through a network of geographically defined, population-based Local Health Authorities. Primary care physicians work for these authorities as independent contractors and act as "gatekeepers" for specialty and other referral services for their patients.[4]

With the belief that a strong primary care system is conducive to improving population health, the NHS initiated reforms that encouraged primary care physicians to organize into collaborative arrangements. To this end, the Regione Emilia-Romagna (RER), a large northern region with a population of about 4.5 million, has recently launched a plan in its 11 Local Health Authorities to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

establish Patient-Centered Medical Homes to better coordinate patient care and help patients avoid unnecessary hospitalizations.

The identification of those patients who would benefit most from outreach efforts is fundamental to achieving these goals of promoting and practicing population health in Patient-Centered Medical Homes. To accomplish this, predictive models and risk stratification tools are needed to identify patients at risk of a worsening of their health status. According to Knutson and Bella,[5] "Predictive models are data-driven, decision-support tools that estimate an individual's future potential health care costs and/or opportunities for care management." A good model will identify as many of these patients as possible while excluding those for whom these interventions would be unnecessary or ineffective.

The RER has established three objectives for this project: 1) develop predictive models to identify patients at high risk of ~~a progression of their medical problems or who are at risk of developing new medical problems~~hospitalization or death, 2) create "risk of hospitalization" patient profiles that provide information about their high-risk patients to the general practitioners in the newly formed Patient-Centered Medical Homes, and 3) assess the extent to which these models and reports provide additional information useful in the identification of patients who may benefit from case management or disease management.

This paper will address the first of the three goals. We describe the development of a predictive model using the RER's regional longitudinal administrative health care database to help identify patients who are most at risk of hospitalization for conditions that may be impacted through improved patient care. This model will then be used to inform the providers associated with the Patient-Centered Medical Homes and aid in their planning for care management and

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

interventions that can reduce their patients' likelihood of a preventable, high-cost hospitalization.

Methods

Study Data and Study Population

The model was developed using the population-based longitudinal health care database of the residents served by the RER Health Service in the years 2004 through 2012.[6] This administrative database includes demographic information for all residents (gender, birth and death dates, location of current residence, and primary care physician), hospital discharge abstract data (ICD-9-CM diagnosis and procedure codes, and admission and discharge dates), emergency room utilization information, outpatient pharmacy data at the individual prescription level, specialty care (laboratory, diagnostics, therapeutic procedures, rehabilitation, and specialist visits), home health data, and information on each primary care physician in the region. Each patient has an anonymous identifier assigned by the RER so that an individual's utilization can be tracked over time without jeopardizing patient privacy.

The study population consisted of all residents of the RER who were at least 18 years of age and still alive as of 31st December 2011.

Dependent Variable

The dependent variable was defined as the occurrence of a hospitalization for problems that are potentially avoidable, or whose progression may have been avoided or delayed through appropriate patient care, or the death of the individual, either in or out of the hospital, for any

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

reason in 2012. We decided to not limit the hospitalization to emergency admissions, since a planned admission may also be an indicator of a worsening medical problem. In order to operationally define the dependent variable, we (authors JSG and DZL) reviewed the Disease Staging[7, 8] primary diagnostic category and severity stage of all day and inpatient hospital admissions (for adults age 18+) in RER for one year, to select those admissions that should be included in the dependent variable.

Deliveries, admission for dental diseases or admissions for vague signs or symptoms with no identified etiology were excluded. Admissions for problems that are not predictable/preventable were excluded while those where screening may identify problems that can potentially be treated to avoid progression were included. For example, admissions for stage 1, chronic cholecystitis or cholelithiasis were excluded, but admissions for advanced stage 2 or 3 complications such as ascending cholangitis or pancreatitis were included.

We felt that inclusion of hospitalization for cancer in the dependent variable should depend on the ability to either prevent or avoid progression of the disease. We therefore included colon cancer and cervical cancer in the definition because they are potentially preventable but excluded all other cancers where prevention/prediction is not currently possible.

Inclusion of injuries, burns, or toxic reaction to prescription or non-prescription drugs would ideally be based on the cause of these problems. Since the etiology of these problems is typically not available in the administrative data being used in this project, we made the decision to include or exclude based on our subjective judgment of the likelihood of preventability. For example, adverse drug reactions were included but burns were excluded from the definition of the dependent variable.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

There is no obvious medical reason for a hospital admission for patients with stage 1 diabetes mellitus or stage 1 essential hypertension without complications. These problems are typically treatable in the outpatient setting. A hospitalization implies a potential problem in the care of these patients, so we decided to include these admissions as a part of the dependent variable.[9]

Independent Variables

A broad range of candidate predictor variables was developed taking advantage of the RER administrative data. The independent variables used for modeling were defined from the RER administrative data for the years 2004 through 2011. Demographic data included patient age, sex, and geographic location of residence. We developed a mapping to broad disease categories defined primarily in terms of the affected body system from home health care data, pharmacy data, and hospital discharge abstract data. (See Appendix 1.)

For those patients who had been hospitalized, more specific diagnostic data were available. We reviewed the classification of patients hospitalized historically using the Disease Staging diagnostic category and disease severity stages.[7,8] Based on the frequencies specific diagnostic category/stage predictor variables were defined for either specific stages of frequent diseases, or by combinations across similar categories. Predictor variables were defined based on the number of emergency room visits using the RER classifications system for the urgency of the visit.

Pharmacy data were used to identify polypharmacy[10] (defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days), potential drug-drug interactions

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(DDI)[11] and potentially inappropriate medication use in patients[12] 65 years and older. Since cardiovascular disease is highly prevalent, we reviewed the use of cardiovascular drugs and created a variable for each of the following 11 classes of drugs (oral anti-coagulants, beta-blockers, angiotensin converting enzyme / angiotensin II receptor blockers, anti-platelets, calcium channel blockers, anti-arrhythmics, digitalis glycosides, nitrates, diuretics, alfa-blockers, statins) to account for the complexity of therapeutic regimen at the patient level.

To take advantage of the fact that the RER database includes multiple years of data, we created history variables using the utilization for each year of data available. Since we were working with the 2011 data to predict hospitalization or death in 2012, we created history variables based on 2004-2010 data. This set included 83 of the diagnostic category/stage variables as well as 11 variables based on pharmacy utilization such as exposure to polypharmacy and use of cardiovascular drugs. If the individual had a history of a disease in any of the years from 2004 to 2010 they were flagged as having a history of that disease and this was used as a potential predictor variable.

Modeling

Logistic regression models were used to estimate predicted probabilities for the occurrence of an inpatient hospital stay for the selected conditions or death for individual patients. Risk of hospitalization or death, and the variables that relate to those risks are highly dependent on age and gender. Regression models were fit in each of 14 gender and age strata using SAS Version 9.2 (SAS Institute, Cary NC). A stepwise process with relaxed covariate entry and retention criteria (inclusion p-value <=0.8, retention <=0.5) was used. At each step in this process, an

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

attempt is made to remove any unimportant variables from the model before adding a potentially important variable. Each addition or deletion of a variable to or from a potential model is a separate step and, at each step, a new model is fitted. This process results in a ~~for selecting a~~ reduced, but robust set of independent variables that predict outcome or that might have importance as adjustment terms for the model in each age/gender stratum.

Evaluation of the Models

The predictive accuracy of the modelling was evaluated using C-statistics (the area under the receiver operating characteristics curve), along with three measures traditionally used with clinical screening tools: sensitivity, specificity and positive predictive value.

C-statistics were used to evaluate the models in two ways. The first evaluation consisted of fitting the model developed using utilization and demographic data from 2011, along with historical variables based on 2004-2010 data, and outcomes (hospitalization or death) from 2012 and then computing a C-statistic to evaluate how the models performed at predicting those outcomes on which the models were conditioned. However, this evaluation is not consistent with evaluating how the data are used in practice. In practice, we have current predictor information, but the outcomes have not been realized. To better estimate how the models are likely to perform in this setting, we fit models to outcomes data up to a year prior to the most current available (e.g., 2011 outcomes modelled with predictors from 2010, along with historical variables based on 2004-2009 data). We then computed a C-statistic for projections made on the risk of hospitalization or death outcomes (in 2012) using the next year's predictor information (in 2011). This way, the models are forced to make projections into the future, but we have the actual

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

observed outcomes data to evaluate the modelling process as it would be used in practice. The resulting C-statistics obtained from these two model runs were compared.

In order to evaluate the performance of the model across different risk thresholds we classified predicted risk scores. "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%. These risk thresholds were selected after consultation with physicians practicing in the medical homes to yield a total of about 10% of the 1,500 patients enrolled with a typical primary care physician.

Results

Among the 3,726,380 adult residents of Emilia Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions defined earlier or died in 2012 (3.6% hospitalized, 1.3% died).

Table 1 shows the distribution of the demographics (age and gender), number of chronic conditions, body systems impacted by the selected chronic conditions, polypharmacy and inappropriate prescribing among the eligible RER residents, as of December 31, 2011. The table also compares these characteristics of the total adult population of the region to the subgroups of the population classified in the "very high risk" and "high risk" categories. Based on the model results, 114,255 individuals were identified as having a predicted risk of hospitalization or death in 2012 of ≥ 25% and classified as "very high risk." An additional 134,610 individuals had a predicted risk of hospitalization or death in 2012 of 15-24% and were classified as "high risk."

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1. Demographic and clinical characteristics of the Regione Emilia-Romagna population, overall and by risk category

	Total ~~population~~Population*		Very High Risk**		High ~~risk~~Risk**	
	3,726,380		114,255		134,610	
Gender	N.	%	N.	%	N.	%
Male	1,788,048	48.0%	54,357	47.6%	61,803	45.9%
Female	1,938,332	52.0%	59,898	52.4%	72,807	54.1%
Age groups						
18-24	258,338	6.9%	76	0.1%	105	0.1%
25-34	499,786	13.4%	302	0.3%	391	0.3%
35-44	732,626	19.7%	1,137	1.0%	1,198	0.9%
45-54	676,047	18.1%	2,612	2.3%	2,485	1.8%
55-64	550,689	14.8%	5,391	4.7%	5,287	3.9%
65-74	482,346	12.9%	13,154	11.5%	14,471	10.8%
74-85	364,369	9.8%	33,430	29.3%	44,857	33.3%
85+	162,179	4.4%	58,153	50.9%	65,816	48.9%
Number of Chronic Conditions						
0-1	2,775,888	74.5%	8,176	7.2%	24,618	18.3%
2 or more	950,492	25.5%	106,079	92.8%	109,992	81.7%
5 or more	99,337	2.7%	45,445	39.8%	20,576	15.3%
Selected Conditions/Body Systems						
Cancer	99,328	2.7%	23,872	20.9%	14,305	10.6%
Cardiovascular	967,796	26.0%	96,157	84.2%	103,749	77.1%
Male Genitourinary#	130,609	7.3%	14,616	26.9%	16,776	27.1%
Ear, Nose, Throat	5,364	0.1%	240	0.2%	242	0.2%
Endocrine	429,528	11.5%	40,653	35.6%	37,471	27.8%
Eye	114,117	3.1%	9,558	8.4%	13,478	10.0%
Gastrointestinal	580,946	15.6%	74,718	65.4%	66,305	49.3%
Gynecologic##	21,806	1.1%	333	0.6%	405	0.6%
Hematologic	45,022	1.2%	15,353	13.4%	6,591	4.9%
Hepatobiliary	24,785	0.7%	6,477	5.7%	3,306	2.5%
Immunologic	3,281	0.1%	464	0.4%	273	0.2%
Infectious Disease	4,723	0.1%	2,207	1.9%	727	0.5%
Musculoskeletal	419,184	11.2%	43,436	38.0%	41,000	30.5%
Neurologic	173,751	4.7%	34,494	30.2%	24,838	18.5%
Psychological	291,308	7.8%	43,387	38.0%	33,715	25.0%

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Respiratory	176,830	4.7%	39,082	34.2%	21,763	16.2%
Skin	28,339	0.8%	7,645	6.7%	3,008	2.2%
Urogenital	37,728	1.0%	16,501	14.4%	5,740	4.3%
Polypharmacy^	609,278	16.4%	92,153	80.7%	92,156	68.5%
Any potentially inappropriate medications (age 65 years or older)^^	257,033	25.5%	51,055	48.7%	49,003	39.2%

* Adults (age 18 or older) and alive at 31 December 2011.

** "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%.

Men only.

Women only.

^ Polypharmacy is defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days.

^^ The list of potentially inappropriate medications can be found in: Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *Journal of Clinical Pharmacy and Therapeutics* 2010;35:219-229.

There was little difference across the risk categories by gender. Age distributions for the "very high risk" and "high risk" groups were shifted more towards the older age groups than those in the overall study population. Residents age 85 or older represented about 4.5% of the RER population, but about 50% of the "very high" and "high" predicted risk groups. More than 75% of the residents over age 85 were classified as "very ~~high "~~high" or "high" risk. However, age alone was not sufficient to their predict risk. For example, residents between 75 and 84 years of age made up 23% of the "very high" risk group and 41% of the "high" risk group, but over 85% of the residents in this age category had neither "very high" nor "high" predicted risk.

Across age and gender strata, demographics and heath care utilization experience in 2011 were the most commonly used independent variables for predicting hospitalization or death in 2012. Selected history variables flagging chronic problems such as cardiovascular disease, diabetes

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

mellitus and chronic renal failure and a history of prescriptions for cardiovascular medications and polypharmacy were also significant predictors.

The residents in the two higher risk groups were more likely than others to have multiple chronic diseases and to experience polypharmacy and inappropriate medication use. The residents identified as "very high risk" or "high risk" by the model also showed a number of striking differences from others in terms of the occurrence of some of the most prevalent health conditions by type and body system. Although cardiovascular conditions were not uncommon in the total adult population (26.0%), they were far more common among those classified as "very high risk" and "high risk" (84.2% and 77.1%, respectively). Similarly, gastrointestinal conditions affected 15.6% of the total population, but were diagnosed in 65.4% of the "very high risk" and 49.3% of the "high risk" patients. Cancer occurred in 2.7% of the total population, but 20.9% of the "very high risk" and 10.6% of the "high risk" patients had a cancer diagnosis. Mental health problems were identified in 7.8% of the adult population, but in 34.2% of the "very high risk" and 25.0% of the "high risk" patients.

The C-statistic for the model of 2012 outcomes developed using 2011 predictors and the C-statistic based on the parameters from the model of 2011 outcomes regressed on 2010 predictors applied to the 2011 predictors and 2012 ~~outcomes,~~outcomes were very similar (0.856 and 0.853, respectively). These results suggest that the relationship between predictors and risk of hospitalization changed little in one year and that model parameters developed in a prior year can be used reliably with the most current year's data to predict unknown outcomes in the next year with only a minimal loss in performance in this population.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2 shows the sensitivity, specificity, positive predictive value and number of true positives for the model at the two selected cut-off points. The sensitivity (percentage of patients actually hospitalized who had been identified by the model as having a predicted risk higher than the cut-off point) was 29.8% for those with the "very high" risk scores. This percentage represents 46,950 of the 157,550 residents of the region who were hospitalized for a selected condition or died in 2012. If we modify the risk score threshold to include individuals with a predicted risk of hospitalization for selected conditions or death of ≥ 15% (i.e., both the "very high risk" and the "high risk" patients) the sensitivity is .471. The true negative rate (specificity) is very high for both risk thresholds (.981 and .951, respectively).

Table 2. Performance of the "Risk of Hospitalization" model for residents identified as "Very High Risk" and "High or Very High Risk"

	Cut-off points for comparison	
Measure	**"Very high risk"***	**"Very high risk"* + "High risk"****
Sensitivity[3#]	0.298	0.471
Specificity[4##]	0.981	0.951
Positive Predictive Value[5^]	0.411	0.298
True positives[6^^]	46,950	74,196

* "Very high risk" is defined as patients with a predicted risk of hospitalization of ≥ =25%.

** "Very high risk" + "High risk", is defined as patients with a predicted risk of hospitalization of ≥=15%.

Sensitivity is defined as the proportion of those hospitalized who were predicted to be hospitalized (true positive rate).

Specificity is the proportion of those not hospitalized who were not predicted to be hospitalized (true negative rate).

^ Positive Predictive Value is the proportion of those predicted to be hospitalized who were actually hospitalized.

^^ True positives are the number of residents who were predicted to be at risk for hospitalization at the predicted risk threshold and were actually hospitalized.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The model appears to be well calibrated across levels of risk. The Figure depicts the RER population divided into groups by deciles of predicted risk of hospitalization or death from the models. The observed prevalence of hospitalization or death is compared to the average predicted risk among individuals in each of the ten predicted risk groups. For example, the overall rate of hospitalization for the selected conditions or death in 2012 was 4.2%. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%. (Regression coefficients and significance levels of independent variables for models for each ot 14 age and gender strata are displayed in Appendix 2.)

~~Figure. Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups.~~

[Insert Figure about here]

Discussion

We have developed a population-based model that identifies the risk of hospitalization for all adult RER residents and does so with a level of performance (c=0.85) as high as, or higher than, similar models. In addition, we believe that the definition of the dependent variable chosen for our models increases the probability that they are identifying patients who risk can potentially be improved by appropriate care. A systematic review by Kansagara[13] of models designed to predict readmissions, showed C-statistic results in the range of 0.55 to 0.83. Recent work by Billings et al[14] to develop models predictive of emergent admissions in the UK had results ranging from 0.73 to 0.78. Li Wang, et al. (2013),[15] using information available through the US

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Veteran's Administration that also included lab data, demonstrated c-statistics of 0.81 and 0.79, respectively, for their models of 90-day or 12-month hospitalization or death outcomes. At a predicted risk of >25% our model had a Positive Predictive Value (PPV) of .411. Billings et al[14] reported a PPV of .417 at a risk threshold of 30. There is a trade-off in using our model, or any predictive model, between the threshold for follow-up and predictive accuracy. A lower risk threshold would identify more patients but with a lower prevalence of hospitalization or death.

Although previous studies have developed models predictive of hospital care, these models fall short of the needs of the Patient Centered Medical Homes being implemented in RER. Typically, these models have focused on specific age groups,[16] conditions, or types of admissions, such as emergent[14] or unplanned admissions or rehospitalizations, or health insurance plans in the United States, including private insurance plans, Medicare and Medicaid plans.[17,18] The models we have developed are applied to the entire adult population of RER. They use existing administrative data, which makes them cost effective to apply.

Patient Centered Medical Homes, including those instituted in RER, are responsible for addressing the needs of their population and making the best use of their finite resources to accomplish this. Preventing unnecessary admissions could improve both the quality of care and health status of the enrolled population, and result in a substantial savings. To accomplish this, predictive models and risk stratification tools such as those developed for this project are needed to identify patients at risk of preventable admissions and provide information that can be used by the medical homes to help manage care.

There are some limitations to the model. The model is developed from administrative data. Administrative data are collected for reimbursement and tracking utilization and not for medical

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

research. They lack the clinical specificity that would desirable in assessing an individual's medical problems. While the hospital discharge abstract data do include diagnostic information coded using ICD-9-CM, no similar data are available for outpatient encounters in the RER database. The mortality data available for this project did not include information about cause of death. Therefore, some proportion of patients whose death was not predictable were included, limiting model performance. In addition, our models use prior utilization among the predictor variables. With the administrative data we cannot distinguish between appropriate and inappropriate treatment which may bias our results.

Despite the limitations of administrative data, they have many advantages for this project: they are readily available, relatively inexpensive to analyze and cover large populations over many years. They are ideal for uncovering patterns of care. If information from the medical records is needed, the results of these analyses can then be supplemented by focused clinical reviews at the local level. Also, The RER has a system in place to monitor the quality of diagnosis and procedure coding in their hospital discharge abstract data. Controls at both the hospital and regional level assess the validity of coding and the consistency of codes assigned such as congruity between sex, age and diagnosis and between diagnosis and procedure. The existence of the RER administrative database made it feasible to develop the models described in this article at relatively low cost and to update the models over time without additional data collection that others have found necessary.[14]

Currently, these risk scores are being integrated with other information in profiles of high-risk patients furnished to providers in 12 newly formed medical homes, including 83 primary care physicians serving a total of about 100,000 patients, in the Parma Local Health Authority located

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

in RER. Along with the risk scores, this information includes data about previous hospitalizations, use of referrals, medications, long-term care and home care services, and a number of process-like quality indicators for diabetic and cardiovascular patients, and for appropriate medication use in older patients.

Of course, model results need to lead to an effective intervention to have a positive impact on patient care. To this end, we are working with the physicians, nurses, and other health care professionals as well as the administration of the newly formed Medical Homes in Parma to assist them in understanding how to use the results of these models and in developing potentially effective interventions. The individual profiles of high risk patients provided to the health care team in the Medical Homes allow them to trigger specific actions such as inviting patients to enroll in disease management programs for chronic problems such as heart failure, chronic obstructive pulmonary disease, or diabetes mellitus, activating home health assistance, initiating a medication review, or recommending that the patient come in for an office visit. An evaluation of the use and usefulness of the~~se~~ profiles and intervention is under way.

In summary, these models provide a means of identifying patients at high risk for hospitalization. The risk predictions, in conjunction with the risk profile, show promise as a useful organizational tool for the regional Patient Centered Medical Homes to develop and implement proactive case management and disease management programs. The RER is reviewing the results of the Parma Local Health Authority pilot project of the profiles. Once their usefulness has been further evaluated, their use will be expanded to other Medical Homes in development in the other Local Health Authorities in the Emilia-Romagna region. If similar data are available, these models can

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

be applied in other Italian regions and other countries investing in organization similar to the Patient Centered Medical Home.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure Legend:

Figure. Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups.

Contributorship statement: DZL, RG, VM, and JSG were responsible for the conceptualization of this project. MR, JM, and ML were responsible for creation of the datasets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form and declare grants from the Regione Emilia-Romagna (Louis, Robeson, McAna, Maio, Keith, Liu, Gonnella) during the conduct of the study. Louis and Gonnella declare personal fees from Truven Health Analytics, The National School of Public Health, Portugal, INSIEL Mercato.

Funding: This research was funded by the Agenzia Sanitaria e Sociale of the Emilia-Romagna Region, Italy. The resulting models are being used by the Local Health Authority of Parma, Italy and the Patient Centered Medical Homes located in the Parma Local Health Authority.

Data Sharing: No additional data available.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1 Saltman RB, Rico A, Boerma W, eds. Primary care in the driver's seat? Organizational reform in European primary care. New York, New York: World Health Organization European Observatory on Health Systems and Policies Series 2006.

2 Smith J, Goodwin N. Towards managed primary care: The role and experience of primary care organizations. Burlington, VT: Ashgate Publishing Company 2006.

3 Defining the PCMH. Agency for Healthcare Research and Quality Web site. http://www.pcmh.ahrq.gov/portal/server.pt/community/pcmh__home/1483/pcmh_defining_the_pcmh_v2 (accessed June 30, 2013).

4 Lo Scalzo A, Donatini A, Orzella l, et al. Italy: Health system review. *Health Syst Transit*. 2009;**11**(6):1-216.

5 Knutson D, Bella M, Llanos K. Predictive Modeling: A Guide for State Medicaid Purchasers. Center for Health Care Strategies, Inc. 2009.

6 The Emilia-Romagna Regional Health Service and the new welfare system: Facilities, expenditure, and activities as of 31.12.2010. Programs, agreements and organisational models. Regione Emilia-Romagna Assessorato Politiche per la Salute. Bologna, Italy 2012.

7 Gonnella JS, Hornbrook MC, Louis DZ. Staging of Disease: A Case-Mix Measurement. *JAMA* 1984;**251**:637–44.

8 Gonnella JS, Louis DZ, Gozum ME, et al, eds. Clinical Criteria for Disease Staging. 6th edition. Ann Arbor, MI: Truven Health Analytics 2012.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

9 Louis DZ, F Taroni, R Melotti, et al. Increasing appropriateness of hospital admissions in the Emilia-Romagna region of Italy. *J Health Serv Res Policy* 2008;**13**(4):202-8.

10 Slabaugh L, Maio V, Templin M, et al. Prevalence and risk of polypharmacy amongst elderly primary care patients in Regione Emilia-Romagna, Italy. *Drug Aging* 2010;**27**:1019-28.

11 Gagne JJ, Maio V, Rabinowitz C. Prevalence and Predictors of Potential Drug-Drug Interactions among Ambulatory Patients in Regione Emilia-Romagna, Italy. *J Clin Pharm Ther* 2008;**33**:141-51.

12 Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *J Clin Pharm Ther* 2010;**35**:219-29.

13 Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: A systematic review. Washington, DC: Department of Veterans Affairs US 2011.

14 Billings J, Georghiou T, Blunt I, et al. Choosing a model to predict hospital admission: An observational study of new variants of predictive models for case finding. *BMJ Open* 2013;**3**(8):e003352. doi: 10.1136/bmjopen-2013-003352 [published Online First: 26 August 2013].

15 Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care* 2013;**51**(4):368-73.

16 Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. *Med Care* 2008;**46**(7):726-31.

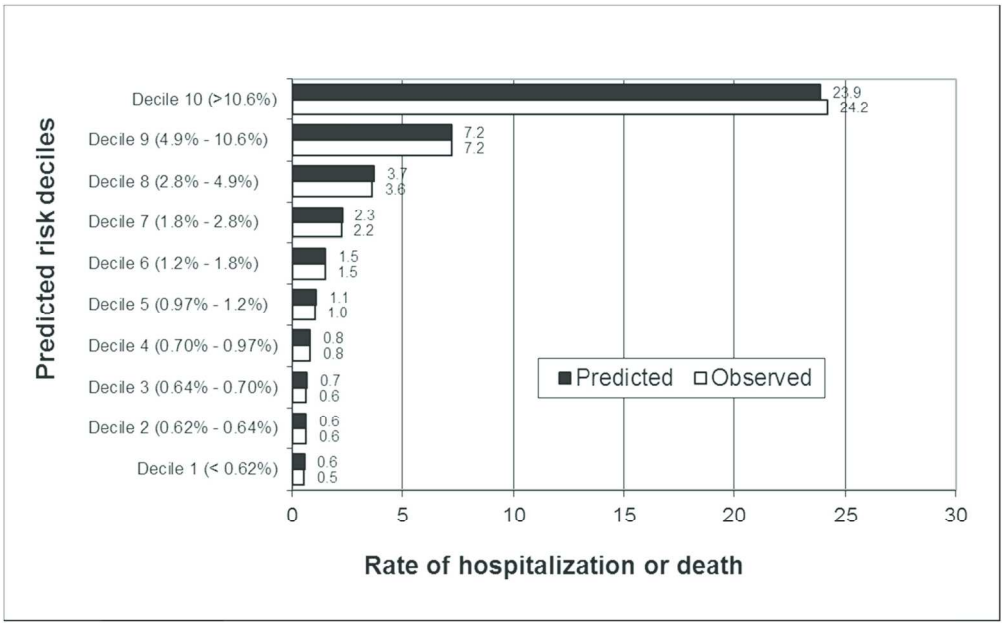
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

17 Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care* 2012;**50**(2):131-9.

18 McAna JF, Crawfrod AG, Novinger BW, et al. A predictive model of hospitalization risk among diabled Medicaid enrollees. *Am J Manag Care* 2013;**19**(5):166-74.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups
165x102mm (300 x 300 DPI)

eview only

Appendix 1: Mapping to Body System or Etiology Groups

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Cancer	Neoplasm, Malignant: Cardiovascular, Hypopharynx, Oral Cavity, Oropharynx, Salivary Glands and Mandible, Other Endocrine System, Larynx, Glottis, Larynx, Subglottic, Larynx, Supraglottic, Nasopharyngeal, Sinuses, Ocular Melanoma, Other Eye and Periocular, Colon and Rectum, Esophagus, Small Bowel, Stomach, Other Gastrointestinal System, Bladder, Urinary, Kidneys, Other Genitourinary System, Breast (Female), Cervix Uteri, Endometrium, Ovaries, Vagina, Vulva, Other Female Genitalia, Hodgkin's Lymphoma, Multiple Myeloma, Mastocytosis, Pancreas, Other Hepatobiliary Tract, Breast (Male), Penile, Prostate, Testicular, Primary Bone, Waldenstrom's Macroglobulinemia, Nonspecific Sites, Unspecified Primary Site, Lungs, Bronchi, or Mediastinum, Hodgkin's Disease Lymphocytic Depletion, Hodgkin's Disease Lymphocytic Predominance, Hodgkin's Disease Mixed Cellularity, Hodgkin's Disease Nodular Sclerosis, Lymphatic and Hematopoietic (Other Types), Lymphoma, Cutaneous T Cell (Mycosis Fungoides), Lymphoma (Diffuse Mixed Small and Large Cell), Lymphoma (Diffuse Large Cell), Lymphoma (Follicular Predominantly Large Cell), Lymphoma (Histiocytic Cell), Lymphoma (Lymphoblastic), Other Respiratory System, Carcinoma (Basal Cell stage 2/3), Carcinoma (Squamous Cell), Melanoma, Other Skin and Soft	Antineoplastics 5HT3 Antagonists	2005-2009: Visits prescribed due to the presence of cancer. Beginning in 2010, the following ICD-9-CM codes were in the record: 140-208, 235-239, V10, V16	Visits prescribed for radiation therapy, or for Injection or infusion of chemotherapeutic Substances for cancer treatment

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Tissue Neoplasm: Pheochromocytoma, Eyelid, Central Nervous System, Lymphatic or Hematopoietic Leukemia: Acute Lymphocytic, Acute Nonlymphocytic, Chronic Lymphocytic, Chronic Myelogenous, Other Types Encounter for: Chemotherapy, Radiation Therapy ICD-9-CM Procedure codes: 99.25, 99.28, 00.10,00.15,92.2x			
Cardiovascular	Aneurysm: Abdominal, Thoracic Anomaly: Patent Ductus Arteriosus, Atrial Septal Defect, Atrioventricular Defects, Coarctation of the Aorta, Other Congenital Heart Disease, Pulmonary Valve Stenosis, Tetralogy of Fallot (stage 3), Transposition of the Great Arteries, Ventricular Septal Defects, Other Circulatory System Aortic: Regurgitation, Stenosis Mitral: Regurgitation, Stenosis Neoplasm: Benign of the Cardiovascular System Arrhythmias, Cardiomyopathies, Conduction Disorders, Congestive Heart Failure, Coronary Artery Disease Prior Coronary Revascularization, Coronary Artery Disease w/o Prior, Coronary Revascularization, Essential Hypertension, Infective, Endocarditis, Pericarditis: Chronic (stage 2/3), Viral or Traumatic (stage 2/3) Periarteritis Nodosa, Raynaud's Disease, Thromboangiitis,	Oral anti-coagulants beta-blockers ACE/ARB anti-platelets calcium channel blockers anti-arrhythmics digitalis glycosides nitrates diuretics alfa-blockers statins	2005-2009: Visits prescribed due to the presence of Congestive Heart Failure or not-well defined cardiopathy, and other diseases of cardiovascular system. Beginning in 2010, the following ICD-9-CM-CM codes were in the record: 390-454,456-459	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Obliterans, Thrombophlebitis, Tibial, lliac, Femoral, or Popliteal Artery Disease, Varicose Veins of Lower Extremities, Secondary Hypertension, Budd Chiari Syndrome, Rheumatic Fever (stage 2/3) Vasculitis Other: Atherosclerosis, Cardiac Conditions, Cardiovascular Symptoms, Circulatory Disorders, Diseases of Arteries, Diseases of Veins, Disorders of Pulmonary Circulation, Lymphatic Disorders			
Endocrine	Adrenal Insufficiency, Cushing's Syndrome, Diabetes insipidus, Diabetes Mellitus Type 1, Diabetes Mellitus Type 2 and Hyperglycemic States,Hyperthyroidism, Hypoglycemia, Hypothyroidism, Monotropic Hormone Deficiency, Primary Amyloidosis, Thyroiditis, Klinefelter's Syndrome, Turner's or Noonan's Syndrome, Obesity Goiter: Nontoxic or Euthyroid (stage 2/3) Neoplasm, Benign: Acromegaly, Adenoma, Parathyroid, Hyperparathyroidism, Primary Hyperaldosteronism, Other Endocrine System Neoplasm, Malignant:Thyroid Other: Endocrine Disorders, Electrolyte Disorders, Nutritional and Metabolic Disorders Anomaly: Adrenal Hyperplasia	Insulins biguanides sulfonylureas vasopressin thyroid replacement antithyroid agents	2005-2009: Visits prescribed due to the presence of diabetes mellitus Beginning in 2010, the following ICD-9-CM codes were in the record: 240-278	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Ear, Nose, Throat	Diseases of Salivary Gland, Incl. Parotitis or Benign Tumors, Other Disorders of Oral Cavity (stage 2), Cholesteatoma, Meniere's Disease, Otitis Media, Sinusitis Hearing Loss due to: Acoustic Trauma, Otosclerosis Neoplasm, Benign: Larynx, Sinuses, Oral Cavity and Pharyngeal Structures Pharyngitis: Non-Streptococcal (stage 2)			
Eye	Cataract, Conjunctivitis: Bacterial, Contusion or Ruptured Globe, Dacryostenosis or Dacryocystitis, Detachment of the Retina, Ectropion or Entropion (Abnormal Lower Lid Position), Endophthalmitis, Foreign Body: Orbit, Fracture: Orbit, Blow-Out, Fungal Infection of the Eye, Glaucoma, Hypovitaminosis A, Laceration: Cornea, Macular Degeneration, Orbital Infection, Prematurity: Retinopathy, Ptosis of Upper Lid, Retrobulbar Orbital Hemorrhage, Trachoma, Other Eye Disorders Injury or Laceration: Eyelid, Periocular, Cornea, Conjunctiv Injury: Eyes, Nonionizing Radiation Keratitis: Acanthamoeba, Bacterial Neoplasm, Benign: Eye	Sympaticomimetic agents parasympaticomimetic agents anhydrase inhibitors ophthalmic beta blockers		
Gastrointestinal	Anorectal Suppuration, Celiac Disease, Clostridium difficile Colitis, Crohn's Disease, Diverticular Disease, Food Poisoning: Other Organisms (stage 3), Functional Digestive Disorders, Gastritis, Hemorrhoids, Hernia (External), Hernia (Hiatal or Reflux Esophagitis), Intussusception (stage 2), Irritable Bowel Syndrome, Gastroenteritis	Intestinal corticosteroids agents H2 antagonists prostaglandins proton pump inhibitors	2005-2009: Visits prescribed due to the presence of Gastrointestinal Diseases Beginning in 2010,	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Neoplasm, Benign: Adenomatous Polyps, Colon, Small Bowel, Other Gastrointestinal System Peptic Ulcer Disease, Salmonellosis (stage 3), Ulcerative Colitis, Vascular Insufficiency of the Bowels, Complications of Gastrointestinal Treatment, Gastroenteritis (stage 2/3) Other Diseases of Esophagus, Stomach, and Duodenum Other Gastrointestinal Disorders, Other Gastrointestinal Infections (stage 2), Other Gastrointestinal or Abdominal Symptoms Anomaly: Congenital Megacolon, Other Digestive or Hepatobiliary System Burns, Chemical: Esophagus, Stomach, or Small Intestine, Laceration: Esophagus		the following ICD-9-CM codes were in the record: 520-539,550-579	
Genitourinary	Bladder Disorders, Calculus of the Urinary Tract, Glomerulonephritis, Acute, Injury: Urinary Tract, Nephrotic Syndrome (stage 2/3), Renal Failure (stage 2/3), Urethritis, Urinary Tract Infections, Neoplasm, Benign: Urinary Tract, Other Disorders of Kidney or Ureter, Other Urinary Symptoms, Encounter for Dialysis, Anomaly: Defects of Kidney, Defects of Lower Genitourinary Tract, Syphilis: Congenital	Agents for hyperkalemia and hyperphosphatemia	2005-2009: Visits prescribed due to the presence of renal failure and Other diseases of the genito-urinary system Beginning in 2010, the following ICD-9-CM codes were in the record: 580-629	Visits prescribed for dialysis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Gynecological and Obstetrics	Anomaly: External Female Genitalia, Anomaly: Uterus, Dysfunctional Uterine Bleeding, Endometriosis, Neoplasm, Benign: Ovary (stage 2), Pelvic Inflammatory Disease, Uterine Infection, Uterovaginal Prolapse, Vulvovaginitis, Other Disorders of Female Genital System			
Hematological	Agranulocytosis, Anemia: Aplastic, Acquired (stage 2/3), Folic Acid Deficiency, Hemolytic (stage 2/3), Iron Deficiency, Sickle Cell, Thalassemia, Vitamin B-12 Deficiency, Other Graft versus Host reaction, Hemophilia A or B, Polycythemia Vera, Other Disorders of Blood and Blood-Forming Organs	Iron vitamin B12 folic acids	Beginning in 2010, the following ICD-9-CM codes were in the record: 280-289	
Hepatobiliary	Cholecystitis and Cholelithiasis, Cirrhosis of the Liver (stage 2/3), Disorders of Bilirubin Excretion, Hepatitis A, Hepatitis B, Hepatitis C, Hepatitis D, Hepatitis E, Hepatitis G, Hepatitis (Chemical), Pancreatitis, Wilson's Disease, Neoplasm, Benign: Hepatobiliary System, Other Hepatobiliary and Pancreatic Disorders, Other Hepatobiliary Infections, Other Pancreatic Disorders	Interferons blood substitutives and plasmatic protein fractions		
Immunologic Diseases	Human Immunodeficiency Virus Type I (HIV) Infection, Other Immunodeficient Disorders, Pneumonia: Pneumocystis carinii	Nucleosides and nucleotides reverse transcriptase inhibitors	2005-2009: Visits prescribed due to HIV Infections Beginning in 2010, the following ICD-9-CM codes were in the record: 279	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Infectious Diseases	Aspergillosis, Chlamydial Infection Except Trachoma or Pneumonia, Cryptococcosis, Cytomegalovirus Disease (Acquired), Infectious Mononucleosis (stage 2), Mucormycosis, Reye's Syndrome (stage 3), Rubella: Acquired (stage 3), Schistosomiasis, Other Bacterial Infections, Other Fungal Infections, Other Infectious and Parasitic Infections, Other Viral Infections, Cytomegalovirus Disease (Congenital), Parainfluenza Virus Infection, Pneumonia: Chlamydial, Sarcoidosis, Other Respiratory Infections, Scabies		Beginning in 2010, the following ICD-9-CM codes were in the record: 001-139	
Male Genital	Benign Prostatic Hypertrophy, Gonorrhea: Male, Prostatitis	Alfa-adrenoreceptor antagonists testosterone 5-alfa reductase inhibitors		
Musculoskeletal	Vitamin D Deficiency, Dislocation: Knee, Eosinophilia Myalgia Syndrome, Fracture: Acetabulum, Fracture: Calcaneus (stage 2), Fracture: Femur, Except Head or Neck, Fracture: Femur, Head or Neck, Fracture: Fibula (stage 2), Fracture: Humerus (Shaft), Fracture: Humerus (Supracondylar) (stage 2), Fracture: Radial Shaft, Ulna or Olecranon (stage 2), Fracture: Radius, Lower End (stage 2), Fracture: Tibia (stage 2/3), Fracture or Dislocation: Patella (stage 2), Fracture or Sprain: Ankle (stage 2), Fracture, Dislocation, or Sprain: Facial Bones (stage 2/3), Fracture, Dislocation, or Sprain: Foot (stage 2), Fracture, Dislocation, or Sprain: Hip or Pelvis (stage 2/3), Fracture, Dislocation, or	Colchicine uric acid inhibitors antiinflammatory non-steroids gold salts aminoquinolines bisphosphonates calcitonin	2005-2009: Visits prescribed due to the presence of Arthrosis, Arthritis and other osteo-muscular and connective diseases, and Fractures of femurs and other consequences of fractures.	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Sprain: Humerus (Head) or Shoulder (stage 2), Fracture, Dislocation, or Sprain: Wrist or Hand or Fingers (stage 2), Gout, Herniated Intervertebral Disc, Infectious Arthritis (stage 2/3), Injury, Chest Wall, Injury, Knee, Semilunar Cartilages (stage 2), Injury, Open Wound, or Blunt Trauma: Lower Extremity (stage 2), Injury, Open Wound, or Blunt Trauma: Upper Extremity (stage 2/3), Muscular Dystrophy, Osteoarthritis, Osteochondrodysplasia, Osteomalacia, Osteomyelitis (stage 2/3), Osteoporosis, Progressive Systemic Sclerosis, Rheumatoid Arthritis, Scoliosis of the Thoracic Spine, Spondylitis, Ankylosing, Systemic Lupus Erythematosus, Anomaly: Musculoskeletal System, Injury: Other and Ill-Defined Musculoskeletal Sites, Neoplasm, Benign: Musculoskeletal Syst. or Connective Tissue, Other Arthropathies, Bone and Joint Disorders, Other Disorders of Connective Tissue, Other Spinal and Back Disorders, Myasthenia Gravis, Complications of Surgical and Medical Care (stage 1), Injury, Open Wound, or Blunt Trauma: Abdomen or Trunk (stage 2/3), Injury: Other (stage 3)		Beginning in 2010, the following ICD-9-CM codes were in the record: 710-739	
Neurologic Diseases	Down's Syndrome, Herpes zoster, Poliomyelitis, Post-Polio Syndrome, Syphilis: Acquired, Tetanus (stage 1), Toxoplasmosis: Acquired (stage 3), Amyotrophic Lateral Sclerosis, Cerebral Palsy, Cerebrovascular Disease, Disease of Nervous System Secondary to Implants or Grafts, Epilepsy, Guillain-Barre Syndrome (stage 2), Headache (stage 2), Huntington's Chorea, Injury: Craniocerebral,	Anticholinesterase agents anticonvulsivant barbiturates and congeners alprostadil ergot alkaloids 5HT1 agonists dopamine	2005-2009: Visits prescribed due to the presence of Dementia and Alzheimer's syndrome, Parkinson's and	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Injury: Spine and spinal cord, Meningitis, Encephalitis, and Myelitis: Viral, Meningitis: Bacterial, Mental Retardation, Multiple Sclerosis, Neurofibromatosis Type I [Von Recklinghausen's Disease], Parkinson's Disease, Other CNS Inflammation, Infection, or Disorder, Other Cranial Nerve Disorders, Other Neurological Conditions, Other Peripheral Nerve Disorders, Other Spinal Lesions, Anomaly: Neural Tube Defects, Rubella: Congenital (stage 2), Anomaly: Other Nervous System, Injury: Other	MAO b inhibitors	other CNS degenerative disease, hemiplegia, monoplegia, and other associated syndroms, and acute and chronic cerebrovascular diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 320-389,797	
Psychological	Dementia: Primary Degenerative (Alzheimer's or Pick's), Antisocial Personality Disorder, Bipolar Disorder - Major Depressive Episode, Bipolar Disorder - Manic Episode, Depression, Generalized Anxiety Disorder, Obsessive-Compulsive Neurosis, Schizophrenia, Autism, Other Neuroses, Other Psychoses Drug Abuse, Dependence, Intoxication: Alcohol, Amphetamine, Barbiturate, Cannabis, Cocaine, Hallucinogen, Opioid, Other Eating disorders: Anorexia Nervosa, Bulimia Nervosa	Antidepressants antipsychotics agents	2005-2009: Visits prescribed due to the presence of psychoses, neuroses, and mental retardation Beginning in 2010, the following ICD-9-CM codes were in the record: 290-319	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Respiratory	Coxsackie and ECHO Infections (stage 2/3), Anomaly: Tracheoesophageal Malformations, Asbestosis, Asthma, Berylliosis, Byssinosis, Chronic Obstructive Pulmonary Disease, Coal Miner's Pneumoconiosis, Croup, Cystic Fibrosis, Emphysema, Hypersensitivity Pneumonitis, Influenza, Mycoplasma pneumoniae Infection, Parainfluenza Virus Infection (stage 2), Pneumonia: Bacterial, Pneumonia: Legionella, Pulmonary Alveolar Proteinosis, Pulmonary Embolism (stage 3), Radiation Pneumonitis, Silicosis, Tuberculosis, Complications of Tracheostomy, Other Disorders of Respiratory System, Other Respiratory Disease Due to External Agents, Other Respiratory Symptoms, Pneumonia: Aspiration, Neoplasm, Benign: Respiratory System	Inhaled corticosteroids beta-2-adrenoreceptor agonists xanthines leucotrienies antagonists cromolyn pancreatic enzymes mucolytics antituberculosis antibiotics isoniazid	2005-2009: Visits prescribed due to the presence of respiratory diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 460-519	
Skin	Herpes Virus Ocular Infection (stage 1), Urticaria, Candida (Monilial) Infections, Clostridial Wound Infection (stage 2), Herpes Simplex Infections, Complications of Surgical and Medical Care (stage 2/3), Anomaly: Integument (Genodermatoses), Decubitus Ulcers, Erythema Multiforme, Erythroderma, Immunologically Mediated Blistering Skin Diseases, Infections of Skin and Subcutaneous Tissue, Neoplasm, Malignant: Carcinoma, Basal Cell (stage 1), Neoplasm: Atypical Nevus (stage 1), Psoriasis Vulgaris, Other Inflammations & Infections of Skin & SubQ Tissue, Burns, Neoplasm, Benign: Skin or Subcutaneous Tissue (stage 1)	Oral and topical antipsoriasis agents	2005-2009: Visits prescribed due to the presence of decubitus ulcers and othere skin diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 680-709	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Appendix: Regression coefficients and significance levels

Females 18-34

Variable	Coefficient	p-value
Intercept	-5.0771	<.0001
Number of Chronic Conditions (from any data source = 1)	0.4347	<.0001
Number of Chronic Conditions (from any data source = 2)	0.8614	<.0001
Number of Chronic Conditions (from any data source = 3)	1.1305	<.0001
Number of Chronic Conditions (from any data source = 4 or more)	1.7194	<.0001
Number of Chronic Conditions (from hospital data = 1)	0.9074	<.0001
Number of Chronic Conditions (from hospital data = 2 or more)	0.8834	<.0001
Total number of ER visits	0.2634	<.0001
History of Obesity-Stage 2 or 3 *	1.6342	<.0001
History of polypharmacy *	0.5968	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	2.1087	<.0001

Males 18-34

Variable	Coefficient	p-value
Intercept	-5.2835	<.0001
Number of Chronic Conditions (from any data source) = 1	0.6534	<.0001
Number of Chronic Conditions (from any data source) = 2	1.2390	<.0001
Number of Chronic Conditions (from any data source) = 3	1.5240	<.0001
Number of Chronic Conditions (from any data source) = 4 or more	2.0556	<.0001
Neurologic Diseases (from home health prescription)	1.6802	<.0001
Renal Failure-Stage 2 or 3	1.3802	<.0001
Any Gastrointestinal Disease - Stage 2 (from hospital data)	1.1512	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	0.8614	<.0001
Any Psychologic Disease - Stage 2 (from hospital data)	1.0198	<.0001
Any Respiratory Disease - Stage 2 (from hospital data)	0.9451	<.0001
Anti-arrhythmics	1.5415	<.0001
Total number of ER visits	0.2371	<.0001
History of Neurologic Diseases (from drug prescriptions) *	0.4880	<.0001
History of Crohns Disease-Stage 2 or 3 *	1.4684	0.0004
History of Neoplasm, Malignant: Colon and Rectum-Stage 2 *	2.9037	0.0063
History of Calculus of the Urinary Tract-Stage 1 *	1.0806	<.0001
History of Cirrhosis of the Liver-Stage 2 or 3 *	1.2212	<.0001
History of Pancreatitis-any stage *	1.7777	<.0001
History of Cerebrovascular Disease-Stage 2 *	2.0588	0.0004
History of Obesity-Stage 2 or 3 *	1.6569	<.0001
History of polypharmacy *	0.4747	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	1.6103	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Hospitalization	0.6338	<.0001

Females 35-44

Variable	Coefficient	p-value
Intercept	-4.9905	<.0001
Number of Chronic Conditions (from any data source) = 1	0.5265	<.0001
Number of Chronic Conditions (from any data source)= 2	0.8446	<.0001
Number of Chronic Conditions (from any data source) = 3	0.8519	<.0001
Number of Chronic Conditions (from any data source)= 4 or more	0.6525	0.0155
Number of Chronic Conditions (from any data source) = 5 or more	0.5419	0.0934
Number of Chronic Conditions (from hospital data) = 1	0.8319	<.0001
Number of Chronic Conditions (from hospital data) = 2	1.0387	<.0001
Number of Chronic Conditions (from hospital data) = 3 or more	1.3840	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.5128	0.0390
Number of Chronic Conditions (from drug prescriptions)=1	-0.0948	0.3171
Number of Chronic Conditions (from drug prescriptions)=2	0.0670	0.6542
Number of Chronic Conditions (from drug prescriptions)=3 or more	0.2362	0.2746
Reside in Mountain area on 12/ 31/ 2012	0.1865	0.0196
Reside in Hill area on 12/ 31/ 2012	-0.0128	0.7341
Cardiovascular Disease (from home health prescription)	1.7641	0.0006
Endocrine Disease (from home health prescription)	1.5904	0.0930
Infectious Disease (from home health prescription)	1.6836	0.0468
Genitourinary (dialysis)	0.7208	0.0081
Aortic Stenosis-Stage 1	2.2652	0.0004
Arrhythmias-Stage 2	1.0050	0.0016
Neoplasm, Malignant: Stomach-Stage 3	1.8592	0.0204
Neoplasm, Malignant: Breast, Female-Stage 3	0.7628	0.0204
Progressive Systemic Sclerosis-Stage 1	0.9852	0.0321
Progressive Systemic Sclerosis-Stage 2 or 3	1.6206	0.0087
Obesity-Stage 2 or 3	0.4604	0.0039
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	0.6965	0.0223
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.7457	0.0325
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	1.6501	0.0187
Any Cancer - Stage 3 (from hospital data)	1.2094	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.5051	0.0014
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.4724	0.0065
Any Genitourinary Disease - Stage 2 (from hospital data)	0.3895	0.0581
Any Gynecologic Disease - Stage 1 (from hospital data)	-0.5549	<.0001
Any Hepatobiliary Disease - Stage 1 (from hospital data)	-0.5393	0.0029
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.4858	0.0002

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any Neurologic Disease - Stage 3 (from hospital data)	0.8922	<.0001
Any Psychologic Disease - Stage 2 (from hospital data)	0.5166	0.0038
Any Respiratory Disease - Stage 2 (from hospital data)	0.5162	0.0436
Endocrine Disease (from drug prescriptions)	-0.4482	<.0001
Genitourinary Disease (from drug prescriptions)	0.9552	0.0001
Respiratory Disease (from drug prescriptions)	-0.2258	0.0107
Cardiovascular Disease (from any data source)	-0.2700	0.0118
Day hospitalization	-0.2627	0.0004
ACE/ARB	0.2465	0.0266
Digitalis glycosides	0.9678	0.0285
Number of ER visits labeled 'Yellow'	-0.2515	0.0378
Total number of ER visits	0.4548	0.0001
Eye Disease (from any data source)	-0.5158	0.0174
History of Cancer (from drug prescriptions) *	0.2289	0.0247
History of Endocrine Disease (from drug prescriptions) *	0.1529	0.0624
History of Psychological Disease (from drug prescriptions) *	0.2272	<.0001
History of Arrhythmias-Stage 2 *	0.4871	0.0794
History of Cardiomyopathies-Stage 3 *	1.1771	0.0262
History of Thrombophlebitis-Stage 2 or 3 *	0.9344	0.0030
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	1.1944	<.0001
History of Crohns Disease-Stage 2 or 3 *	0.7513	0.0377
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	2.0723	<.0001
History of Calculus of the Urinary Tract-Stage 1 *	0.4914	0.0146
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.8776	0.0003
History of Neoplasm, Malignant: Kidneys-Stage 1 *	1.3977	0.0102
History of Neoplasm, Malignant: Kidneys-Stage 3 *	3.2491	0.0100
History of Pancreatitis-any stage *	0.8084	0.0199
History of Cerebrovascular Disease-Stage 1 *	0.8172	0.0057
History of Cerebrovascular Disease-Stage 2 *	1.1315	0.0013
History of Obesity-Stage 2 or 3 *	1.2321	<.0001
History of polypharmacy *	0.3345	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.6431	0.0246
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.8255	<.0001
History of Depression-Stage 1 or 2 *	0.2207	0.0716
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.5358	0.0330
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.9824	0.0003
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	1.8886	0.0964
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	1.6543	0.0122
History of Pneumonia: Bacterial-Stage 3 *	1.3040	0.0015
Immunologic Disease (from any data source)	0.8590	0.0001
Polypharmacy	0.2838	<.0001
Number of the other 9 Cardiovascular drugs	0.1330	0.0036

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Males 35-44

Variable	Coefficient	p-value
Intercept	-4.8083	<.0001
Number of Chronic Conditions (from any data source)=1	0.5439	<.0001
Number of Chronic Conditions (from any data source)=2	0.7994	<.0001
Number of Chronic Conditions (from any data source)=3	0.7949	<.0001
Number of Chronic Conditions (from any data source)=4	1.1832	<.0001
Number of Chronic Conditions (from any data source)=5 or more	0.9615	<.0001
Age on 12/ 31/ 2012	-0.00005	0.0031
Cancer (from home health prescription)	1.6698	0.0023
Blood Diseases (from home health prescription)	2.7427	0.0385
Infectious Disease (from home health prescription)	3.1524	0.0129
Neurologic Diseases (from home health prescription)	2.0454	<.0001
Genitourinary (dialysis)	1.0696	<.0001
Cardiomyopathies-Stage 3	1.7220	0.0016
Infective Endocarditis-Stage 3	3.7783	0.0015
Mitral Stenosis-Stage 3	3.0377	0.0345
Pericarditis: Chronic-Stage 2 or 3	1.2938	0.0240
Crohns Disease-Stage 1	1.6408	<.0001
Neoplasm, Malignant: Colon and Rectum-Stage 2	1.4659	0.0030
Cirrhosis of the Liver-Stage 2 or 3	0.6646	0.0066
Neoplasm, Malignant: Pancreas-Stage 1	2.4864	0.0143
Pancreatitis-all stages	0.8241	0.0053
Cerebrovascular Disease-Stage 3	0.9540	<.0001
Epilepsy-all stages	0.5515	0.0247
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	0.6065	0.0050
Any Cancer - Stage 3 (from hospital data)	1.7602	<.0001
Any Endocrine Disease - Stage 1 (from hospital data)	-0.5267	0.0028
Any Endocrine - Stage 2 (from hospital data)	0.5688	0.0028
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-1.5927	<.0001
Any Immunologic Disease - All stages (from hospital data)	0.6498	0.0068
Any Psychologic Disease - Stage 2 (from hospital data)	0.7963	<.0001
Any Psychologic Disease - Stage 3 (from hospital data)	1.4603	0.0014
Endocrine Disease (from drug prescriptions)	0.2204	0.0058
Number of day hospitalizations	0.2078	0.0003
Oral anti-coagulants	0.4822	0.0028
Anti-arrhythmics	0.7069	0.0027
Total number of ER visits	0.2455	<.0001
Eye Disease (from any data source)	-0.4757	0.0200
History of Neurological Disease (from drug prescriptions) *	0.3728	<.0001
History of Psychological Disease (from drug prescriptions) *	0.2678	<.0001
History of Arrhythmias-Stage 2 *	0.7259	0.0008
History of Cardiomyopathies-Stage 2 *	0.8655	0.0009

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Essential Hypertension-Stage 3 *	0.8004	0.0062
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	1.0253	0.0008
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.7425	0.0034
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.7380	0.0071
History of Crohns Disease-Stage 2 or 3 *	1.3269	<.0001
History of Calculus of the Urinary Tract-Stage 1 *	0.7919	<.0001
History of Renal Failure-Stage 2 or 3 *	0.5494	0.0107
History of Cholecystitis and Cholelithiasis-Stage 3 *	1.4633	0.0012
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.5857	0.0009
History of Pancreatitis-any stage *	1.2530	<.0001
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	3.8605	0.0026
History of Obesity-Stage 2 or 3 *	0.8764	<.0001
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.8467	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.5825	0.0005
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.7745	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.7316	0.0018
History of Pneumonia: Bacterial-Stage 3 *	1.0518	0.0006
History of Other Cardiovascular drugs *	0.2342	0.0043
Hospitalization	0.7414	<.0001
Gastrointestinal Disease (from hospital data)	0.9745	<.0001
Respiratory Disease (from hospital data)	0.4067	0.0025
Any of the other 9 Cardiovascular drugs	-0.5914	<.0001
Number of the other 9 Cardiovascular drugs	0.2767	<.0001

Females 45-54

Variable	Coefficient	p-value
Intercept	-4.9051	<.0001
Number of Chronic Conditions (from any data source)=1	0.3066	<.0001
Number of Chronic Conditions (from any data source)=2	0.4393	<.0001
Number of Chronic Conditions (from any data source)=3	0.4533	<.0001
Number of Chronic Conditions (from any data source)=4	0.3924	0.0002
Number of Chronic Conditions (from any data source)=5 or more	0.3544	0.0079
Number of Chronic Conditions (from hospital data)=1	0.4819	<.0001
Number of Chronic Conditions (from hospital data)=2	0.6828	<.0001
Number of Chronic Conditions (from hospital data)=3 or more	0.8788	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.9174	<.0001
Reside in Mountain area on 12/ 31/ 2012	0.2243	0.0008
Reside in Hill area on 12/ 31/ 2012	-0.0104	0.7596
Essential Hypertension-Stage 1	-0.3166	0.0082
Pericarditis: Chronic-Stage 2 or 3	2.0018	0.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Neoplasm, Malignant: Colon and Rectum-Stage 2	0.8110	0.0074
Cholecystitis and Cholelithiasis-Stage 2	0.8570	0.0043
Cirrhosis of the Liver-Stage 2 or 3	0.4736	0.0176
Pancreatitis-all stages	0.8873	0.0217
Progressive Systemic Sclerosis-Stage 2 or 3	0.9823	0.0173
Cerebrovascular Disease-Stage 1	0.5623	0.0340
Cerebrovascular Disease-Stage 3	-0.6910	0.0224
Obesity-Stage 2 or 3	0.4054	0.0063
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	1.3912	0.0132
Any Cancer - Stage 3 (from hospital data)	1.7066	<.0001
Any Ear,Nose,Throat Disease - Stage 1 (from hospital data)	-0.6186	0.0401
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.4730	<.0001
Any Gynecologic Disease - Stage 1 (from hospital data)	-0.5578	<.0001
Any Infectious Disease - Stage 3 (from hospital data)	-1.8556	0.0017
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.4338	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	1.0267	<.0001
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2591	0.0256
Any Psychologic Disease - Stage 2 (from hospital data)	0.4123	0.0034
Any Psychologic Disease - Stage 3 (from hospital data)	1.3041	0.0042
Any Respiratory Disease - Stage 2 (from hospital data)	0.4618	0.0149
Any Respiratory Disease - Stage 3 (from hospital data)	0.7372	0.0009
Any Skin Disease - Stage 1 (from hospital data)	-0.3787	0.0245
Gastrointestinal Disease (from drug prescriptions)	0.2535	<.0001
Genitourinary Disease (from drug prescriptions)	1.1262	<.0001
Oral anti-coagulants	0.3753	0.0063
Anti-arrhythmics	0.7321	<.0001
Digitalis glycosides	0.9270	0.0003
Total number of ER visits	0.2068	<.0001
History of Cancer (from drug prescriptions) *	0.1500	0.0286
History of Psychological Disease (from drug prescriptions) *	0.2010	<.0001
History of Aortic Stenosis-Stage 3 *	1.6666	0.0170
History of Coronary Artery Disease-Stage 3 *	0.5560	0.0019
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	1.0444	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.6990	0.0012
History of Crohns Disease-Stage 2 or 3 *	1.3534	<.0001
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.9831	0.0009
History of Neoplasm, Malignant: Stomach-Stage 3 *	2.1435	0.0007
History of Ulcerative Colitis-any stage *	0.6736	0.0177
History of Calculus of the Urinary Tract-Stage 1 *	0.7453	<.0001
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.5209	0.0218
History of Renal Failure-Stage 2 or 3 *	0.5702	0.0006
History of Neoplasm, Malignant: Ovaries-Stage 2 or 3 *	0.7030	0.0286
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.7484	0.0026
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.4895	0.0014

*History variables are calculated on previous 5 years of exposure data

For peer review only

History of Pancreatitis-any stage *	0.7645	0.0043
History of Cerebrovascular Disease-Stage 3 *	0.3947	0.0087
History of Obesity-Stage 2 or 3 *	0.8330	<.0001
History of polypharmacy *	0.2928	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.6454	0.0038
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.6630	0.0001
History of Depression-Stage 1 or 2 *	0.4957	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.6357	0.0005
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	1.1324	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.4127	0.0104
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	1.3247	0.0050
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	1.6041	<.0001
History of Neoplasm, Malignant: Melanoma-Stage 3 *	1.6975	0.0026
History of Other Cardiovascular drugs *	0.1255	0.0236
Immunologic Disease (from any data source)	0.7145	0.0003
Infectious Disease (from any data source)	0.5052	0.0335
Neurologic Disease (from any data source)	0.2075	0.0003
Hospitalization	0.3975	<.0001
Polypharmacy	0.2655	<.0001
Any of the other 9 Cardiovascular drugs	-0.2999	0.0008
Number of the other 9 Cardiovascular drugs	0.1482	<.0001

Males 45-54

Variable	Coefficient	p-value
Intercept	-4.4469	<.0001
Number of Chronic Conditions (from any data source)=1	0.3859	<.0001
Number of Chronic Conditions (from any data source)=2	0.6634	<.0001
Number of Chronic Conditions (from any data source)=3	0.7465	<.0001
Number of Chronic Conditions (from any data source)=4	0.7901	<.0001
Number of Chronic Conditions (from any data source)=5 or more	0.5246	<.0001
Number of Chronic Conditions (from hospital data)=1	0.2577	0.0009
Number of Chronic Conditions (from hospital data)=2	0.3237	0.0015
Number of Chronic Conditions (from hospital data)=3 or more	0.4067	0.0021
Number of Chronic Conditions (from home health prescription)=1 or more	0.8811	<.0001
Age on 12 /31/ 2012	-0.00008	<.0001
Cancer (chemo or radiation)	0.5498	0.0011

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Genitourinary (dialysis)	0.9242	<.0001
Aortic Stenosis-Stage 3	2.0591	0.0014
Arrhythmias-Stage 2	0.5607	0.0005
Essential Hypertension-Stage 2	-0.6186	0.0030
Neoplasm, Malignant Hematologic-Stage 3	-1.5129	0.0096
Cirrhosis of the Liver-Stage 2 or 3	0.8760	<.0001
Pancreatitis-all stages	0.9702	<.0001
Chronic Obstructive Pulmonary Disease-Stage 3	1.2772	0.0051
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	1.5413	<.0001
Any Cancer - Stage 3 (from hospital data)	1.5337	<.0001
Any Cardiovascular - Stage 3 (from hospital data)	0.4528	<.0001
Any Endocrine Disease - Stage 1 (from hospital data)	-0.2681	0.0085
Any Endocrine - Stage 2 (from hospital data)	0.3403	0.0055
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.7672	<.0001
Any Immunologic Disease - All stages (from hospital data)	0.7049	<.0001
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.3976	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	0.4660	0.0008
Any Psychologic Disease - Stage 2 (from hospital data)	0.8976	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.1612	0.0030
Eye Disease (from drug prescriptions)	-0.5308	<.0001
Genitourinary Disease (from drug prescriptions)	0.4405	0.0099
Hematologic Disease (from drug prescriptions)	0.5097	0.0070
Hepatobiliary Disease (from drug prescriptions)	0.3691	0.0179
Number of day hospitalizations	0.1160	0.0093
Statins	-0.1389	0.0112
Anti-platelets	0.2288	<.0001
Anti-arrhythmics	0.3517	0.0063
Nitrates	0.4390	<.0001
Total number of ER visits	0.1627	<.0001
History of Cancer (from drug prescriptions) *	0.3251	0.0004
History of Aortic Stenosis-Stage 1 *	-0.9794	0.0150
History of Arrhythmias-Stage 2 *	0.3919	0.0019
History of Cardiomyopathies-Stage 3 *	0.7836	<.0001
History of Coronary Artery Disease-Stage 1 *	0.3743	<.0001
History of Thrombophlebitis-Stage 2 or 3 *	0.7954	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.3220	0.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.8677	<.0001
History of Calculus of the Urinary Tract-Stage 1 *	0.3374	0.0033
History of Renal Failure-Stage 2 or 3 *	0.4042	0.0033
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.5888	0.0001
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.4954	<.0001
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	1.9346	0.0029
History of Pancreatitis-any stage *	0.5981	0.0009
History of Obesity-Stage 2 or 3 *	0.5126	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only

History of polypharmacy *	0.2306	<.0001
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.8192	<.0001
History of Depression-Stage 1 or 2 *	0.2814	0.0098
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.4287	0.0020
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.9474	<.0001
History of Pneumonia: Bacterial-Stage 3 *	1.1780	<.0001
History of Other Cardiovascular drugs *	0.2404	<.0001
Male Genital System (from any data source)	-0.3177	0.0021
Neurologic Disease (from any data source)	0.2173	0.0002
Hospitalization	0.4249	<.0001
Number of hospitalizations	0.0445	0.1777
Polypharmacy	0.2976	<.0001
Gastrointestinal Disease (from hospital data)	-0.2445	0.0085

Females 55-64

Variable	Coefficient	p-value
Intercept	0.9467	0.6785
Number of Chronic Conditions (from any data source)=1	0.5017	<.0001
Number of Chronic Conditions (from any data source)=2	0.6666	<.0001
Number of Chronic Conditions (from any data source)=3	0.7010	<.0001
Number of Chronic Conditions (from any data source)=4	0.7868	0.0001
Number of Chronic Conditions (from any data source)=5	0.7545	0.0024
Number of Chronic Conditions (from any data source)=6 or more	0.5597	0.0587
Number of Chronic Conditions (from hospital data)=1	0.5017	<.0001
Number of Chronic Conditions (from hospital data)=2	0.6365	<.0001
Number of Chronic Conditions (from hospital data)=3	0.7653	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.7953	0.0007
Number of Chronic Conditions (from home health prescription)=1 or more	0.4889	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.3395	0.0003
Number of Chronic Conditions (from drug prescriptions)=2	-0.3996	0.0020
Number of Chronic Conditions (from drug prescriptions)=3	-0.4436	0.0073
Number of Chronic Conditions (from drug prescriptions)=4	-0.5404	0.0083
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.4198	0.0955
Age on 12/ 31/ 2012	-0.1094	0.0134
Endocrine Disease (from home health prescription)	0.9016	0.0026
Gastrointestinal Disease (from home health prescription)	-0.6807	0.1725
Genitourinary Disease (from home health prescription)	1.1277	0.1537
Blood Diseases (from home health prescription)	-1.8804	0.1597
Infectious Disease (from home health prescription)	1.2458	0.0624
Musculoskeletal Disease (from home health prescription)	0.7627	0.0235

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Neurologic Diseases (from home health prescription)	0.7672	0.0003
Respiratory Diseases (from home health prescription)	1.0350	0.0211
Skin Disease (from home health prescription)	0.4575	0.2643
Cancer (chemo or radiation)	0.2011	0.0741
Genitourinary (dialysis)	-0.2335	0.2676
Aneurysm, Thoracic-all stages	0.7181	0.1737
Arrhythmias-Stage 1	-0.8531	0.1853
Arrhythmias-Stage 3	-0.4355	0.3839
Congestive Heart Failure-Stage3	0.5326	0.0248
Essential Hypertension-Stage 2	-0.4672	0.0156
Mitral Stenosis-Stage 1	0.3593	0.1871
Mitral Stenosis-Stage 2	0.4834	0.2609
Pericarditis: Viral or Traumatic-Stage 2 or 3	-1.2494	0.1428
Thrombophlebitis-Stage 2 or 3	0.3434	0.2209
Tibial/lliac/Femoral/Popliteal Artery Disease-Stage 1	0.2126	0.4831
Diabetes Mellitus Type 1 or Type 2-Stage 3	-1.2846	0.0001
Hyperthyroidism-Stage 1	-0.6448	0.1376
Hypothyroidism-Stage 1	-0.1514	0.2913
Hypothyroidism-Stage 2 or 3	-0.4319	0.2665
Crohns Disease-Stage 2 or 3	-1.3093	0.2190
Diverticular Disease-Stage 1	-0.3203	0.2326
Diverticular Disease-Stage 2 or 3	-1.7244	0.0142
Gastritis-Stage 2 or 3	-0.4282	0.3525
Hernia, Hiatal or Reflux Esophagitis-Stage 1	-0.4908	0.0465
Hernia, Hiatal or Reflux Esophagitis-Stage 2 or 3	-0.6378	0.2767
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.6795	0.0391
Neoplasm, Malignant: Colon and Rectum-Stage 3	0.9264	<.0001
Neoplasm, Malignant: Stomach-Stage 1	0.5593	0.2519
Neoplasm, Malignant: Stomach-Stage 3	0.6181	0.1650
Ulcerative Colitis-all stages	0.5059	0.1961
Neoplasm, Malignant: Bladder, Urinary-Stage 1	-0.3433	0.3353
Neoplasm, Malignant: Kidneys-Stage 1	0.3389	0.3869
Renal Failure-Stage 2 or 3	0.3465	0.0464
Neoplasm, Malignant: Breast, Female-Stage 1	-0.9369	<.0001
Neoplasm, Malignant: Ovaries-Stage 1	-0.7428	0.0657
Anemia: Aplastic, Acquired-Stage 2 or 3	0.3219	0.2666
Neoplasm, Malignant Hematologic-Stage 1	-0.4013	0.0971
Neoplasm, Malignant Hematologic-Stage 2	-0.6307	0.0974
Neoplasm, Malignant Hematologic-Stage 3	-0.5380	0.2971
Cholecystitis and Cholelithiasis-Stage 1	0.7769	0.0033
Cholecystitis and Cholelithiasis-Stage 2	1.3013	0.0057
Cirrhosis of the Liver-Stage 2 or 3	0.7679	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.8360	0.0386
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.3469	0.0051

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Pancreatitis-all stages	1.0560	0.0079
Progressive Systemic Sclerosis-Stage 2 or 3	1.0431	0.0155
Cerebrovascular Disease-Stage 1	0.2701	0.2083
Cerebrovascular Disease-Stage 2	-0.2745	0.2425
Cerebrovascular Disease-Stage 3	-0.2845	0.3180
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.6421	0.0743
Epilepsy-all stages	-0.4298	0.1025
Bipolar Disorder - Major Depressive Episode-Stage 2 or 3	-0.5924	0.2896
Depression-Stage 1 or 2	0.1495	0.4002
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	-0.6312	0.2187
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.4370	0.1568
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.1433	0.4617
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.4562	0.0554
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.4995	0.0489
Pneumonia: Bacterial-Stage 1	0.3737	0.0753
Pneumonia: Bacterial-Stage 3	-0.4891	0.1195
Pulmonary Embolism-Stage 3	0.8005	0.0250
Any Cancer - Stage 1 (from hospital data)	0.2989	0.0931
Any Cancer - Stage 2 (from hospital data)	0.2308	0.2732
Any Cancer - Stage 3 (from hospital data)	0.8552	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.2364	0.0023
Any Cardiovascular Disease - Stage 2 (from hospital data)	0.0953	0.3950
Any Cardiovascular - Stage 3 (from hospital data)	0.1844	0.1691
Any Endocrine - Stage 2 (from hospital data)	0.4515	<.0001
Any Endocrine Disease - Stage 3 (from hospital data)	0.7544	0.0013
Any Ear,Nose,Throat Disease - Stage 2 (from hospital data)	-0.8908	0.3948
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.3647	0.0445
Any Hepatobiliary Disease - Stage 1 (from hospital data)	-0.6278	0.0950
Any Hepatobiliary Disease - Stage 2 (from hospital data)	-0.4898	0.3504
Any Infectious Disease - Stage 3 (from hospital data)	-0.3543	0.3858
Any Neurologic Disease - Stage 2 (from hospital data)	0.5032	0.0016
Any Neurologic Disease - Stage 3 (from hospital data)	0.4457	0.0740
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2030	0.1733
Any Psychologic Disease - Stage 2 (from hospital data)	0.1866	0.3379
Any Psychologic Disease - Stage 3 (from hospital data)	0.9858	0.0302
Any Respiratory Disease – Stage 1 (from hospital data)	-0.3063	0.0617
Any Respiratory Disease - Stage 3 (from hospital data)	0.9969	<.0001
Any Skin Disease - Stage 2 (from hospital data)	-0.5257	0.1185
Neoplasm, Malignant: Melanoma-Stage 2	-1.2688	0.1880
Neoplasm, Malignant: Melanoma-Stage 3	1.5149	0.0037
Cancer (from any data source)	-0.3957	0.0036
Cancer (from drug prescription)	0.4367	0.0002
Eye Disease (from drug prescriptions)	0.4615	0.0460
Gastrointestinal Disease (from drug prescriptions)	0.0544	0.1584

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Genitourinary Disease (from drug prescriptions)	0.2762	0.1949
Immunologic Disease (from drug prescriptions)	-1.0597	0.0780
Psychological Disease (from drug prescriptions)	-0.1619	0.3523
Respiratory Disease (from drug prescriptions)	0.2058	0.0005
Skin Disease (from drug prescriptions)	0.1296	0.4503
Day hospitalization	-0.0504	0.4576
Drug-Drug interactions	0.2803	0.0229
Statins	-0.2104	<.0001
Beta-blockers	0.0461	0.1974
Anti-platelets	0.1061	0.0107
Calcium channel blockers	0.0933	0.0329
Anti-arrhythmics	0.2835	0.0060
Digitalis glycosides	-0.2094	0.2138
Nitrates	0.4632	<.0001
Diuretics	0.2687	<.0001
Ear,Nose,Throat Disease (from any data source)	-0.3916	0.1154
Number of ER visits labeled 'Yellow'	0.0881	0.2672
Total number of ER visits	0.1621	0.0341
Eye Disease (from any data source)	-0.6320	0.0043
Genitourinary Disease (from any data source)	0.7024	0.0014
Gynecologic Disease (from any data source)	-0.6085	<.0001
Hematologic Disease (from any data source)	0.2896	0.0019
Hepatobiliary Disease (from any data source)	0.4386	0.0713
History of Cancer (from drug prescriptions) *	0.2544	<.0001
History of Neurological Disease (from drug prescriptions) *	0.0571	0.2898
History of Respiratory Disease (from drug prescriptions) *	0.0775	0.0855
History of Aortic Stenosis-Stage 1 *	0.1809	0.4275
History of Arrhythmias-Stage 2 *	0.3253	0.0041
History of Arrhythmias-Stage 3 *	0.3799	0.2085
History of Cardiomyopathies-Stage 2 *	0.4667	0.0155
History of Congestive Heart Failure-Stage 3 *	0.1210	0.4597
History of Coronary Artery Disease-Stage 1 *	0.3581	0.0002
History of Coronary Artery Disease-Stage 2 *	0.1121	0.4469
History of Coronary Artery Disease-Stage 3 *	0.3913	0.0023
History of Essential Hypertension-Stage 1 *	0.0703	0.1748
History of Infective Endocarditis-Stage 3 *	-0.5804	0.3469
History of Mitral Stenosis-Stage 2 *	0.4323	0.1047
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	-0.5354	0.2771
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.5543	0.0659
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.3017	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.3810	0.0005
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.4846	0.0020
History of Hypothyroidism-Stage 2 or 3 *	0.2600	0.3833

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Crohns Disease-Stage 2 or 3 *	0.7668	0.0458
History of Diverticular Disease-Stage 1 *	0.3579	0.0123
History of Diverticular Disease-Stage 2 or 3 *	0.7500	0.0026
History of Neoplasm, Malignant: Stomach-Stage 3 *	1.3562	0.0009
History of Calculus of the Urinary Tract-Stage 1 *	0.1328	0.4254
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.1600	0.4523
History of Neoplasm, Malignant: Kidneys-Stage 1 *	-0.2475	0.4638
History of Neoplasm, Malignant: Kidneys-Stage 3 *	1.1246	0.0296
History of Renal Failure-Stage 2 or 3 *	0.5356	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 1 *	-0.2892	0.0025
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.5799	0.0001
History of Neoplasm, Malignant: Ovaries-Stage 2 or 3 *	0.4641	0.0675
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.2917	0.1730
History of Neoplasm, Malignant Hematologic-Stage 2 *	0.2692	0.3031
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.6343	0.2195
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.4606	0.0057
History of Cholecystitis and Cholelithiasis-Stage 3 *	-0.3862	0.3756
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.1037	0.4226
History of Neoplasm, Malignant: Pancreas-Stage 1 *	0.9151	0.0222
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	-1.0822	0.1459
History of Pancreatitis-any stage *	0.2106	0.4490
History of Progressive Systemic Sclerosis-Stage 1 *	0.2697	0.2504
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	-0.4137	0.2802
History of Cerebrovascular Disease-Stage 1 *	0.2919	0.0479
History of Cerebrovascular Disease-Stage 2 *	0.1387	0.2983
History of Cerebrovascular Disease-Stage 3 *	0.3321	0.0030
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.2370	0.3913
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	1.0816	0.1258
History of Obesity-Stage 2 or 3 *	0.2556	0.0066
History of polypharmacy *	0.1792	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.5506	0.0209
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.6414	0.0002
History of Depression-Stage 1 or 2 *	0.3655	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.6588	0.0024
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.6840	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.4496	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.9085	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2 *	1.1063	0.0136

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	0.4570	0.1957
History of Pneumonia: Bacterial-Stage 2 *	0.7303	0.1047
History of Oral Anti-coagulants *	0.4605	<.0001
History of Other Cardiovascular drugs *	0.0431	0.3044
History of Statins *	0.0710	0.1377
Immunologic Disease (from any data source)	0.7304	0.0366
Neurologic Disease (from any data source)	0.2104	0.0008
Hospitalization	0.2413	0.0005
Polypharmacy	0.3614	<.0001
Psychological Disease (from any data source)	0.2990	0.0875
Cancer (from hospital data)	0.2380	0.2926
Gastrointestinal Disease (from hospital data)	-0.2031	0.0659
Genitourinary Disease (from hospital data)	-0.4981	0.0234
Hepatobiliary (from hospital data)	-0.5335	0.2235
Musculoskeletal Disease (from hospital data)	-0.4348	<.0001
Any of the other 9 Cardiovascular drugs	-0.1434	0.0142

Males 55-64

Variable	Coefficient	p-value
Intercept	-4.2367	<.0001
Number of Chronic Conditions (from any data source)=1	0.3641	<.0001
Number of Chronic Conditions (from any data source)=2	0.7530	<.0001
Number of Chronic Conditions (from any data source)=3	1.0181	<.0001
Number of Chronic Conditions (from any data source)=4	1.2055	<.0001
Number of Chronic Conditions (from any data source)=5	1.4339	<.0001
Number of Chronic Conditions (from any data source)=6 or more	1.4674	<.0001
Number of Chronic Conditions (from hospital data)=1	0.4141	<.0001
Number of Chronic Conditions (from hospital data)=2	0.5725	<.0001
Number of Chronic Conditions (from hospital data)=3	0.7463	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.6436	0.0066
Number of Chronic Conditions (from home health prescription)=1 or more	-0.5216	0.0939
Number of Chronic Conditions (from drug prescriptions)=1	-0.0825	0.3624
Number of Chronic Conditions (from drug prescriptions)=2	-0.3715	0.0074
Number of Chronic Conditions (from drug prescriptions)=3	-0.5199	0.0060
Number of Chronic Conditions (from drug prescriptions)=4	-0.7343	0.0025
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.8378	0.0069
Age on 12 /31/ 2012	-0.00009	<.0001
Cancer (from home health prescription)	1.5149	<.0001
Cardiovascular Disease (from home health prescription)	0.6241	0.0555
Endocrine Disease (from home health prescription)	1.2243	0.0016
Genitourinary Disease (from home health prescription)	0.5007	0.4113

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Blood Diseases (from home health prescription)	1.2676	0.2946
Musculoskeletal Disease (from home health prescription)	1.2147	0.0136
Neurologic Diseases (from home health prescription)	1.5200	<.0001
Mental Disorders (from home health prescription)	0.9790	0.0086
Respiratory Diseases (from home health prescription)	0.5207	0.3572
Skin Disease (from home health prescription)	1.2618	0.0224
Cancer (chemo or radiation)	0.4095	0.0006
Genitourinary (dialysis)	0.4977	0.0006
Aneurysm, Abdominal-all stages	-0.4983	0.0674
Aneurysm, Thoracic-all stages	-1.0860	0.0027
Aortic Stenosis-Stage 3	0.3053	0.4855
Arrhythmias-Stage 1	0.3575	0.2891
Arrhythmias-Stage 3	0.4253	0.1234
Cardiomyopathies-Stage 2	0.1958	0.2233
Cardiomyopathies-Stage 3	0.3422	0.0864
Congestive Heart Failure-Stage3	0.2219	0.2411
Coronary Artery Disease-Stage 1	-0.0683	0.3883
Coronary Artery Disease-Stage 2	0.0967	0.3847
Essential Hypertension-Stage 2	-0.1509	0.2134
Mitral Stenosis-Stage 2	0.4867	0.0815
Thrombophlebitis-Stage 1	0.3111	0.1622
Thrombophlebitis-Stage 2 or 3	0.2554	0.2361
Tibial/lliac/Femoral/Popliteal Artery Disease-Stage 1	0.1621	0.2869
Crohns Disease-Stage 1	1.2423	0.0002
Diverticular Disease-Stage 1	0.1565	0.4435
Diverticular Disease-Stage 2 or 3	-0.6513	0.1725
Functional Digestive Disorders-Stage 1	0.2423	0.3863
Hernia, Hiatal or Reflux Esophagitis-Stage 1	0.4054	0.0335
Hernia, Hiatal or Reflux Esophagitis-Stage 2 or 3	0.3879	0.3142
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.3750	0.1667
Neoplasm, Malignant: Colon and Rectum-Stage 3	0.1798	0.4389
Neoplasm, Malignant: Stomach-Stage 1	0.3938	0.2408
Neoplasm, Malignant: Stomach-Stage 3	0.5800	0.1449
Calculus of the Urinary Tract-Stage 2 or 3	0.2830	0.3339
Neoplasm, Malignant: Bladder, Urinary-Stage 3	0.8356	0.0617
Neoplasm, Malignant: Kidneys-Stage 3	0.9161	0.0054
Renal Failure-Stage 2 or 3	0.2518	0.0215
Anemia: Aplastic, Acquired-Stage 2 or 3	0.3357	0.2123
Neoplasm, Malignant Hematologic-Stage 2	-0.4371	0.1590
Neoplasm, Malignant Hematologic-Stage 3	-0.8421	0.0510
Cholecystitis and Cholelithiasis-Stage 1	0.1946	0.3670
Cholecystitis and Cholelithiasis-Stage 2	0.7374	0.0003
Cirrhosis of the Liver-Stage 2 or 3	0.6437	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	1.1672	0.0009

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Neoplasm, Malignant: Pancreas-Stage 2 or 3	0.4115	0.3686
Pancreatitis-all stages	0.6998	0.0093
Rheumatic Fever- Stage 2	0.2881	0.3795
Rheumatic Fever- Stage 3	0.7440	0.2101
Neoplasm, Malignant: Prostate-Stage 2	-0.9665	0.0008
Progressive Systemic Sclerosis-Stage 1	1.5792	0.0739
Cerebrovascular Disease-Stage 3	-0.3096	0.1750
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.3701	0.2504
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	-0.8789	0.3960
Bipolar Disorder - Major Depressive Episode-Stage 2 or 3	0.4031	0.3584
Bipolar Disorder - Manic Episode-Stage 2	0.3873	0.2024
Depression-Stage 1 or 2	-0.1959	0.3188
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.5678	0.0022
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.4929	0.0002
Chronic Obstructive Pulmonary Disease-Stage 3	0.7526	0.0053
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.4482	0.0104
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.7027	0.0037
Pneumonia: Bacterial-Stage 1	0.4147	0.0065
Pneumonia: Bacterial-Stage 2	-1.0459	0.0841
Any Cancer - Stage 1 (from hospital data)	0.1447	0.3765
Any Cancer - Stage 2 (from hospital data)	0.4529	0.0235
Any Cancer - Stage 3 (from hospital data)	0.8561	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.1085	0.2067
Any Cardiovascular - Stage 3 (from hospital data)	0.1646	0.0649
Any Endocrine - Stage 2 (from hospital data)	0.1840	0.0457
Any Endocrine Disease - Stage 3 (from hospital data)	0.1092	0.4222
Any Eye Disease - All stages (from hospital data)	0.1685	0.4757
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.3016	0.0049
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.2142	0.0818
Any Genitourinary Disease - Stage 2 (from hospital data)	0.1929	0.4584
Any Genitourinary Disease - Stage 3 (from hospital data)	-0.3144	0.2654
Any Male Genital System - All stages (from hospital data)	-0.4482	0.0010
Any Neurologic Disease - Stage 1 (from hospital data)	0.2517	0.1493
Any Neurologic Disease - Stage 2 (from hospital data)	0.3684	0.0263
Any Neurologic Disease - Stage 3 (from hospital data)	0.6947	0.0025
Any Respiratory Disease - Stage 2 (from hospital data)	0.3146	0.0525
Any Respiratory Disease - Stage 3 (from hospital data)	0.5118	0.0036
Neoplasm, Malignant: Melanoma-Stage 2	-1.0279	0.1274
Neoplasm, Malignant: Melanoma-Stage 3	0.8045	0.0919
Cancer (from any data source)	-0.5680	<.0001
Cancer (from drug prescription)	0.7548	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.2032	0.0068
Eye Disease (from drug prescriptions)	0.4930	0.0727
Gastrointestinal Disease (from drug prescriptions)	0.2771	0.0301

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Genitourinary Disease (from drug prescriptions)	0.5082	0.0008
Hematologic Disease (from drug prescriptions)	0.7320	<.0001
Hepatobiliary Disease (from drug prescriptions)	0.7078	<.0001
Musculoskeletal Disease (from drug prescriptions)	0.2658	0.0704
Neurologic Diseases (from drug prescriptions)	-0.1106	0.4344
Respiratory Disease (from drug prescriptions)	0.3397	0.0140
Skin Disease (from drug prescriptions)	-0.8534	0.0737
Day hospitalization	-0.3167	0.0005
Number of day hospitalizations	0.1936	0.0009
Oral anti-coagulants	0.2097	0.0073
Alpha-blockers	-0.0815	0.1273
Statins	-0.1195	0.0067
Beta-blockers	0.0758	0.0200
ACE/ARB	-0.1067	0.0015
Anti-platelets	0.2416	<.0001
Anti-arrhythmics	0.2796	<.0001
Nitrates	0.3750	<.0001
Diuretics	0.1841	<.0001
Ear,Nose,Throat Disease (from any data source)	-0.4299	0.0218
Total number of ER visits	0.1667	<.0001
Eye Disease (from any data source)	-0.5605	0.0459
Gastrointestinal Disease (from any data source)	-0.2670	0.0340
Genitourinary Disease (from any data source)	-0.2029	0.0436
Hematologic Disease (from any data source)	-0.2489	0.0209
History of Endocrine Disease (from drug prescriptions) *	0.1446	0.0036
History of Neurological Disease (from drug prescriptions) *	0.1136	0.0491
History of Aortic Stenosis-Stage 1 *	0.1778	0.2745
History of Aortic Stenosis-Stage 3 *	-0.3400	0.2752
History of Arrhythmias-Stage 2 *	0.1806	0.0180
History of Cardiomyopathies-Stage 2 *	0.2800	0.0099
History of Cardiomyopathies-Stage 3 *	0.2745	0.0450
History of Congestive Heart Failure-Stage 3 *	0.3791	0.0032
History of Coronary Artery Disease-Stage 1 *	0.2123	<.0001
History of Coronary Artery Disease-Stage 2 *	0.2602	0.0002
History of Coronary Artery Disease-Stage 3 *	0.1210	0.0479
History of Essential Hypertension-Stage 3 *	0.1309	0.1997
History of Mitral Stenosis-Stage 2 *	0.1638	0.4910
History of Pericarditis: Chronic-Stage 2 or 3 *	-0.6191	0.0980
History of Thrombophlebitis-Stage 2 or 3 *	0.2574	0.0966
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.3798	0.0282
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.0741	0.1711
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.2216	0.0034
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.2666	0.0234

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Hypothyroidism-Stage 1 *	-0.1031	0.4735
History of Crohns Disease-Stage 2 or 3 *	0.5055	0.1168
History of Diverticular Disease-Stage 1 *	0.1351	0.3093
History of Diverticular Disease-Stage 2 or 3 *	0.3463	0.1463
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.6400	0.0005
History of Neoplasm, Malignant: Stomach-Stage 1 *	0.7035	0.0151
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.7735	0.0405
History of Ulcerative Colitis-any stage *	0.1614	0.4858
History of Calculus of the Urinary Tract-Stage 1 *	0.1264	0.1702
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.3105	0.0119
History of Neoplasm, Malignant: Kidneys-Stage 1 *	-0.3428	0.0948
History of Renal Failure-Stage 2 or 3 *	0.3768	<.0001
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.4850	0.0332
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.2115	0.1691
History of Neoplasm, Malignant Hematologic-Stage 2 *	0.2152	0.3377
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.5625	0.1068
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.3497	0.0041
History of Cholecystitis and Cholelithiasis-Stage 3 *	0.1889	0.4655
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.3937	<.0001
History of Neoplasm, Malignant: Pancreas-Stage 1 *	0.3556	0.3596
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	-0.7080	0.2534
History of Pancreatitis-any stage *	0.5502	0.0003
History of Rheumatic Fever-Stage 3 *	0.6045	0.2213
History of Neoplasm, Malignant: Prostate-Stage 2 *	-0.3042	0.0181
History of Progressive Systemic Sclerosis-Stage 1 *	-1.8861	0.1260
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	1.2042	0.1939
History of Cerebrovascular Disease-Stage 1 *	0.1150	0.3529
History of Cerebrovascular Disease-Stage 2 *	0.1580	0.0735
History of Cerebrovascular Disease-Stage 3 *	0.2798	0.0003
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.5172	0.0355
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.9102	0.1461
History of Obesity-Stage 2 or 3 *	0.2153	0.0147
History of polypharmacy *	0.1251	0.0003
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.4806	0.0162
History of Depression-Stage 1 or 2 *	0.4028	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.3627	0.009
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.6949	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.3625	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.9574	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.1444	0.4290

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2 *	0.3890	0.3220
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	0.7587	0.0080
History of Pneumonia: Bacterial-Stage 2 *	-0.9199	0.0638
History of Pneumonia: Bacterial-Stage 3 *	0.4583	0.0128
History of Neoplasm, Malignant: Melanoma-Stage 2 *	0.2292	0.4962
History of Neoplasm, Malignant: Melanoma-Stage 3 *	0.6301	0.2916
History of Oral Anti-coagulants *	0.0666	0.3850
History of Other Cardiovascular drugs *	0.0303	0.4200
History of Statins *	0.0592	0.1378
Immunologic Disease (from any data source)	0.2942	0.1479
Male Genital System (from any data source)	-0.1980	0.0015
Musculoskeletal Disease (from any data source)	-0.3422	0.0208
Neurologic Disease (from any data source)	0.2313	0.1007
Hospitalization	0.2053	0.0038
Number of hospitalizations	0.0291	0.2995
Polypharmacy	0.2601	<.0001
Psychological Disease (from any data source)	0.1041	0.0984
Respiratory Disease (from any data source)	-0.1691	0.2366
Cancer (from hospital data)	0.2023	0.3581
Cardiovascular Disease (from hospital data)	-0.0893	0.3710
Hepatobiliary (from hospital data)	-0.6082	<.0001
Musculoskeletal Disease (from hospital data)	-0.4016	0.0016
Neurologic Disease (from hospital data)	-0.6060	0.0072
Psychological Disease (from hospital data)	-0.2465	0.1059
Respiratory Disease (from hospital data)	-0.2243	0.1884
Skin Disease (from hospital data)	-0.5825	0.2155
Skin Disease (from any data source)	0.6403	0.1830
Any of the other 9 Cardiovascular drugs	0.0866	0.1859

Females 65-74

Variable	Coefficient	p-value
Intercept	-4.5504	<.0001
Number of Chronic Conditions (from any data source)=1	0.3294	0.0005
Number of Chronic Conditions (from any data source)=2	0.7012	<.0001
Number of Chronic Conditions (from any data source)=3	1.0162	<.0001
Number of Chronic Conditions (from any data source)=4	1.2243	<.0001
Number of Chronic Conditions (from any data source)=5	1.3625	<.0001
Number of Chronic Conditions (from any data source)=6 or more	1.5014	<.0001
Number of Chronic Conditions (from hospital data)=1	0.3904	<.0001
Number of Chronic Conditions (from hospital data)=2	0.4301	<.0001
Number of Chronic Conditions (from hospital data)=3	0.4970	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Number of Chronic Conditions (from hospital data)=4 or more	0.5908	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.4758	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.1393	0.1297
Number of Chronic Conditions (from drug prescriptions)=2	-0.3041	0.0058
Number of Chronic Conditions (from drug prescriptions)=3	-0.3950	0.0024
Number of Chronic Conditions (from drug prescriptions)=4	-0.4051	0.0076
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.4003	0.0256
Age on 12 /31/ 2012	-0.00012	<.0001
Cancer (from home health prescription)	0.6069	<.0001
Cardiovascular Disease (from home health prescription)	-0.2636	0.0301
Blood Diseases (from home health prescription)	0.6193	0.1479
Neurologic Diseases (from home health prescription)	0.4376	0.0016
Mental Disorders (from home health prescription)	-0.2879	0.1224
Cancer (chemo or radiation)	0.1589	0.0714
Aortic Stenosis-Stage 1	0.5120	0.0026
Cardiomyopathies-Stage 3	0.8973	0.0003
Congestive Heart Failure-Stage3	0.3086	0.0103
Coronary Artery Disease-Stage 1	-0.1015	0.2148
Coronary Artery Disease-Stage 2	0.4617	0.0005
Coronary Artery Disease-Stage 3	0.2471	0.0565
Essential Hypertension-Stage 1	-0.1363	0.0191
Essential Hypertension-Stage 3	0.2124	0.0429
Mitral Stenosis-Stage 1	-0.4732	0.0120
Thrombophlebitis-Stage 2 or 3	0.7060	<.0001
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 1	0.3844	0.0228
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3	0.8107	0.0004
Diabetes Mellitus Type 1 or Type 2-Stage 3	-0.6668	0.0022
Crohns Disease-Stage 2 or 3	0.7852	0.1300
Hernia, Hiatal or Reflux Esophagitis-Stage 1	0.1799	0.1780
Neoplasm, Malignant: Stomach-Stage 3	0.4943	0.1142
Neoplasm, Malignant: Bladder, Urinary-Stage 3	0.8348	0.0780
Neoplasm, Malignant: Breast, Female-Stage 1	-0.7289	<.0001
Neoplasm, Malignant: Breast, Female-Stage 3	-0.2375	0.1392
Anemia: Aplastic, Acquired-Stage 2 or 3	0.2640	0.1999
Neoplasm, Malignant Hematologic-Stage 1	0.2313	0.1024
Neoplasm, Malignant Hematologic-Stage 3	-1.0860	0.0017
Cirrhosis of the Liver-Stage 2 or 3	0.6134	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.7125	0.0116
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.7605	<.0001
Rheumatic Fever- Stage 2	0.3447	0.0706
Cerebrovascular Disease-Stage 3	-0.6890	0.0007
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.3887	0.2397
Bipolar Disorder - Major Depressive Episode-Stage 2 or 3	-0.7551	0.1418

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Bipolar Disorder - Manic Episode-Stage 2	0.6386	0.0329
Depression-Stage 1 or 2	-0.1537	0.1428
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	1.4878	0.0036
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.1325	0.0980
Chronic Obstructive Pulmonary Disease-Stage 3	0.5302	0.0063
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.5790	0.0005
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	0.6907	0.0477
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.7408	0.0003
Any Cancer - Stage 1 (from hospital data)	0.3101	<.0001
Any Cancer - Stage 3 (from hospital data)	1.2060	<.0001
Any Endocrine Disease - Stage 3 (from hospital data)	0.4493	0.0062
Any Eye Disease - All stages (from hospital data)	-0.1582	0.1255
Any Genitourinary Disease - Stage 2 (from hospital data)	0.3230	0.0655
Any Genitourinary Disease - Stage 3 (from hospital data)	-0.4009	0.1324
Any Gynecologic Disease - Stage 2 or 3 (from hospital data)	-0.4173	0.0548
Any Hepatobiliary Disease - Stage 1 (from hospital data)	-0.3374	0.0002
Any Hepatobiliary Disease - Stage 3 (from hospital data)	0.3881	0.1255
Any Immunologic Disease - All stages (from hospital data)	-0.5745	0.1396
Any Infectious Disease - Stage 3 (from hospital data)	0.5105	0.0511
Any Neurologic Disease - Stage 3 (from hospital data)	0.8526	<.0001
Any Psychologic Disease - Stage 2 (from hospital data)	0.3067	0.0376
Any Psychologic Disease - Stage 3 (from hospital data)	0.6691	0.0261
Any Respiratory Disease - Stage 3 (from hospital data)	0.1787	0.1601
Any Skin Disease - Stage 1 (from hospital data)	-0.2893	0.0063
Cancer (from any data source)	-0.3239	<.0001
Cancer (from drug prescriptions)	0.3190	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.3086	<.0001
Genitourinary Disease (from drug prescriptions)	0.6302	<.0001
Hematologic Disease (from drug prescriptions)	0.3673	<.0001
Hepatobiliary Disease (from drug prescriptions)	0.6321	0.0005
Musculoskeletal Disease (from drug prescriptions)	0.1877	0.0283
Respiratory Disease (from drug prescriptions)	0.1731	<.0001
Oral anti-coagulants	0.2036	0.0012
Statins	-0.1126	<.0001
ACE/ARB	-0.1348	<.0001
Anti-platelets	0.1104	0.0002
Anti-arrhythmics	0.0912	0.0986
Digitalis glycosides	0.2082	0.0046
Nitrates	0.2180	<.0001
Diuretics	0.0387	0.2255
Endocrine Disease (from any data source)	-0.1114	0.0002
Ear,Nose,Throat Disease (from any data source)	-0.4077	0.0494
Number of ER visits labeled 'Yellow'	-0.3525	<.0001
Total number of ER visits	0.5514	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Eye Disease (from any data source)	-0.2411	<.0001
Gastrointestinal Disease (from any data source)	-0.0917	0.0025
Gynecologic Disease (from any data source)	-0.4752	<.0001
History of Cancer (from drug prescriptions) *	0.1368	0.0038
History of Neurological Disease (from drug prescriptions) *	0.1374	<.0001
History of Aortic Stenosis-Stage 3 *	0.2156	0.1868
History of Arrhythmias-Stage 2 *	0.1567	0.0059
History of Cardiomyopathies-Stage 2 *	0.4598	0.0002
History of Congestive Heart Failure-Stage 3 *	0.2601	0.0033
History of Coronary Artery Disease-Stage 1 *	0.1550	0.0040
History of Coronary Artery Disease-Stage 2 *	0.1778	0.0271
History of Coronary Artery Disease-Stage 3 *	0.2409	0.0019
History of Essential Hypertension-Stage 2 *	0.1060	0.0526
History of Essential Hypertension-Stage 3 *	0.1504	0.0375
History of Mitral Stenosis-Stage 2 *	-0.2678	0.0937
History of Mitral Stenosis-Stage 3 *	0.3696	0.0337
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.2128	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.3147	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.4937	<.0001
History of Crohns Disease-Stage 2 or 3 *	0.5581	0.0954
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.4016	0.0044
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.4353	0.1847
History of Calculus of the Urinary Tract-Stage 1 *	0.2185	0.0607
History of Neoplasm, Malignant: Kidneys-Stage 3 *	-0.7174	0.1125
History of Renal Failure-Stage 2 or 3 *	0.4347	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 1 *	-0.1897	0.0064
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.2713	0.0494
History of Neoplasm, Malignant: Ovaries-Stage 1 *	0.4490	0.0239
History of Neoplasm, Malignant Hematologic-Stage 2 *	0.5429	0.0348
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.1929	0.0880
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.3420	<.0001
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	0.7431	<.0001
History of Cerebrovascular Disease-Stage 1 *	0.2276	0.0043
History of Cerebrovascular Disease-Stage 3 *	0.1967	0.0029
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.4871	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.7255	0.0080
History of Obesity-Stage 2 or 3 *	0.1359	0.0870
History of polypharmacy *	0.1622	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.6969	0.0010
History of Depression-Stage 1 or 2 *	0.1646	0.0034
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.8459	0.0008

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.6589	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.4628	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.8534	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.2386	0.1126
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2 *	0.6212	0.0971
History of Pneumonia: Bacterial-Stage 2 *	0.4624	0.1267
History of Pneumonia: Bacterial-Stage 3 *	0.4924	0.0003
History of Neoplasm, Malignant: Melanoma-Stage 3 *	0.8606	0.1231
History of Oral Anti-coagulants *	0.2508	<.0001
History of Other Cardiovascular drugs *	0.1002	0.0015
Musculoskeletal Disease (from any data source)	-0.3863	<.0001
Hospitalization	0.2562	<.0001
Number of hospitalizations	-0.0651	0.0068
Polypharmacy	0.2370	<.0001
Cardiovascular Disease (from hospital data)	-0.0957	0.1172
Musculoskeletal Disease (from hospital data)	-0.2090	0.0016
Any of the other 9 Cardiovascular drugs	0.0806	0.1274
Number of the other 9 Cardiovascular drugs	0.0425	0.0129

Males 65-74

Variable	Coefficient	p-value
Intercept	-4.0290	<.0001
Number of Chronic Conditions (from any data source)=1	0.3090	<.0001
Number of Chronic Conditions (from any data source)=2	0.4055	<.0001
Number of Chronic Conditions (from any data source)=3	0.6026	<.0001
Number of Chronic Conditions (from any data source)=4	0.7813	<.0001
Number of Chronic Conditions (from any data source)=5	0.7088	<.0001
Number of Chronic Conditions (from any data source)=6 or more	0.8178	<.0001
Number of Chronic Conditions (from hospital data)=1	0.2460	<.0001
Number of Chronic Conditions (from hospital data)=2	0.3676	<.0001
Number of Chronic Conditions (from hospital data)=3	0.3746	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.4616	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.7247	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.0302	0.6751
Number of Chronic Conditions (from drug prescriptions)=2	0.0804	0.3543
Number of Chronic Conditions (from drug prescriptions)=3	0.0770	0.4505
Number of Chronic Conditions (from drug prescriptions)=4	0.1604	0.1774
Number of Chronic Conditions (from drug prescriptions)=5 or more	0.3823	0.0066
Age on 12 /31/ 2012	-0.0001	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Cardiovascular Disease (from home health prescription)	-0.2088	0.0342
Gastrointestinal Disease (from home health prescription)	-0.9964	0.0117
Cancer (chemo or radiation)	0.4296	<.0001
Cardiomyopathies-Stage 3	0.5047	0.0007
Coronary Artery Disease-Stage 2	0.2966	<.0001
Coronary Artery Disease-Stage 3	0.2001	0.0086
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 1	0.2752	0.0019
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3	0.3313	0.0330
Hyperthyroidism-Stage 1	-0.8506	0.0059
Hyperthyroidism-Stage 2 or 3	-0.8443	0.0397
Hypothyroidism-Stage 1	-0.4005	0.0102
Diverticular Disease-Stage 1	0.3805	0.0034
Gastritis-Stage 1	0.2944	0.0367
Hernia, Hiatal or Reflux Esophagitis-Stage 1	0.3404	0.0214
Neoplasm, Malignant: Stomach-Stage 3	0.6357	0.0021
Neoplasm, Malignant: Bladder, Urinary-Stage 1	-0.3858	0.0003
Neoplasm, Malignant: Bladder, Urinary-Stage 3	1.0573	0.0003
Renal Failure-Stage 2 or 3	0.2462	0.0001
Cholecystitis and Cholelithiasis-Stage 2	0.8383	0.0010
Cirrhosis of the Liver-Stage 2 or 3	0.5557	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.8289	0.0004
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.0295	0.0026
Rheumatic Fever- Stage 2	0.4941	0.0138
Rheumatic Fever- Stage 3	0.9219	0.0348
Neoplasm, Malignant: Prostate-Stage 2	-0.5899	<.0001
Cerebrovascular Disease-Stage 1	0.2014	0.0319
Cerebrovascular Disease-Stage 3	-0.3785	0.0234
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.2581	0.1704
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.6623	<.0001
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	1.0595	<.0001
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.9074	<.0001
Pulmonary Embolism-Stage 3	0.5962	0.0059
Any Cancer - Stage 1 (from hospital data)	0.1424	0.1477
Any Cancer - Stage 3 (from hospital data)	0.9400	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.0869	0.0242
Any Ear,Nose,Throat Disease - Stage 1 (from hospital data)	-0.4741	0.0100
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.2854	<.0001
Any Hepatobiliary Disease - Stage 2 (from hospital data)	-0.5419	0.0149
Any Hepatobiliary Disease - Stage 3 (from hospital data)	0.4751	0.0179
Any Immunologic Disease - All stages (from hospital data)	0.6987	0.0029
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.2629	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	0.5546	0.0004
Any Psychologic Disease - Stage 1 (from hospital data)	-0.4107	0.0146
Any Psychologic Disease - Stage 3 (from hospital data)	-0.7275	0.0035

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any Respiratory Disease - Stage 3 (from hospital data)	0.6359	<.0001
Cancer (from any data source)	-0.3827	<.0001
Cancer (from drug prescriptions)	0.2794	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.3439	<.0001
Eye Disease (from drug prescriptions)	-0.3279	<.0001
Genitourinary Disease (from drug prescriptions)	0.2042	0.0754
Hepatobiliary Disease (from drug prescriptions)	0.5858	0.0002
Respiratory Disease (from drug prescriptions)	0.1468	<.0001
Drug-Drug interactions	0.1473	0.0128
Oral anti-coagulants	0.1161	0.0182
Statins	-0.1331	<.0001
Beta-blockers	0.0499	0.0248
ACE/ARB	-0.0997	<.0001
Anti-platelets	0.1992	<.0001
Anti-arrhythmics	0.1064	0.0130
Digitalis glycosides	0.3471	<.0001
Nitrates	0.3349	<.0001
Diuretics	0.1464	<.0001
Total number of ER visits	0.1632	<.0001
Gastrointestinal Disease (from any data source)	-0.1029	<.0001
Genitourinary Disease (from any data source)	0.3139	0.0021
History of Cancer (from drug prescriptions) *	0.1477	0.0020
History of Neurological Disease (from drug prescriptions) *	0.1240	0.0001
History of Cardiomyopathies-Stage 2 *	0.3116	<.0001
History of Cardiomyopathies-Stage 3 *	0.4397	<.0001
History of Congestive Heart Failure-Stage 3 *	0.3374	<.0001
History of Coronary Artery Disease-Stage 1 *	0.2568	<.0001
History of Essential Hypertension-Stage 2 *	0.1079	0.0140
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.1006	0.0031
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.1970	0.0001
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.3887	0.0005
History of Renal Failure-Stage 2 or 3 *	0.3301	<.0001
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.2586	0.0009
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.3021	<.0001
History of Neoplasm, Malignant: Prostate-Stage 3 *	0.3601	0.0252
History of Cerebrovascular Disease-Stage 1 *	0.1414	0.0309
History of Cerebrovascular Disease-Stage 3 *	0.2819	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.6145	<.0001
History of polypharmacy *	0.1051	<.0001
History of Bipolar Disorder - Manic Episode-Stage 1 *	0.6800	0.0265
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.3941	0.0071
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.4235	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.3669	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.7646	<.0001
History of Pneumonia: Bacterial-Stage 3 *	0.3809	0.0002
History of Oral Anti-coagulants *	0.1253	0.0059
History of Other Cardiovascular drugs *	0.1019	0.0003
Male Genital System (from any data source)	-0.3918	<.0001
Musculoskeletal Disease (from any data source)	-0.1502	<.0001
Hospitalization	0.1762	0.0003
Number of hospitalizations	0.0176	0.3845
Polypharmacy	0.2522	<.0001
Cancer (from hospital data)	0.3120	0.0130
Genitourinary Disease (from hospital data)	-0.4713	<.0001
Psychological Disease (from hospital data)	0.5748	0.0003

Females 75-84

Variable	Coefficient	p-value
Intercept	-2.1966	0.0983
Number of Chronic Conditions (from any data source)=1	0.3204	<.0001
Number of Chronic Conditions (from any data source)=2	0.5216	<.0001
Number of Chronic Conditions (from any data source)=3	0.6879	<.0001
Number of Chronic Conditions (from any data source)=4	0.8108	<.0001
Number of Chronic Conditions (from any data source)=5	0.8615	<.0001
Number of Chronic Conditions (from any data source)=6 or more	0.9003	<.0001
Number of Chronic Conditions (from hospital data)=1	0.2734	<.0001
Number of Chronic Conditions (from hospital data)=2	0.4760	<.0001
Number of Chronic Conditions (from hospital data)=3	0.4602	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.5605	<.0001
Number of Chronic Conditions (from home health prescription)=1	0.4539	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.5024	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.2927	<.0001
Number of Chronic Conditions (from drug prescriptions)=2	-0.4346	<.0001
Number of Chronic Conditions (from drug prescriptions)=3	-0.5230	<.0001
Number of Chronic Conditions (from drug prescriptions)=4	-0.5943	<.0001
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.6388	<.0001
Age on 12/ 31/ 2012	-0.0494	0.0537
Cancer (from home health prescription)	0.5074	<.0001
Cardiovascular Disease (from home health prescription)	-0.1174	0.0117
Respiratory Diseases (from home health prescription)	0.4899	<.0001
Cancer (chemo or radiation)	0.3093	0.0003
Genitourinary (dialysis)	0.3668	0.0018
Aneurysm, Thoracic-all stages	0.4180	0.0286

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Aortic Stenosis-Stage 1	0.2859	0.0055
Arrhythmias-Stage 3	-0.4014	0.0049
Coronary Artery Disease-Stage 2	0.3630	<.0001
Essential Hypertension-Stage 3	-0.1678	0.0129
Mitral Stenosis-Stage 2	-0.3215	0.0749
Thrombophlebitis-Stage 1	0.3956	0.0024
Thrombophlebitis-Stage 2 or 3	-0.3711	0.0013
Hyperthyroidism-Stage 1	-0.3290	0.0847
Crohns Disease-Stage 1	0.7565	0.0329
Functional Digestive Disorders-Stage 1	0.2509	0.0124
Gastritis-Stage 1	0.3782	0.0002
Neoplasm, Malignant: Kidneys-Stage 3	-0.7235	0.1071
Renal Failure-Stage 2 or 3	0.2070	0.0002
Neoplasm, Malignant: Breast, Female-Stage 1	-0.3991	0.0001
Neoplasm, Malignant: Breast, Female-Stage 3	-0.4149	0.0077
Neoplasm, Malignant: Ovaries-Stage 1	0.5296	0.0693
Anemia: Aplastic, Acquired-Stage 2 or 3	0.3538	0.0227
Cirrhosis of the Liver-Stage 2 or 3	0.6150	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.8098	0.0011
Neoplasm, Malignant: Pancreas-Stage 2 or 3	0.7321	0.0866
Progressive Systemic Sclerosis-Stage 1	0.5340	0.0203
Cerebrovascular Disease-Stage 1	0.1406	0.0941
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.3594	<.0001
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.2377	0.0716
Bipolar Disorder - Manic Episode-Stage 1	-0.9145	0.1201
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.2567	0.0037
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.5431	0.0015
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	0.9029	0.0394
Pneumonia: Bacterial-Stage 1	0.3367	0.0003
Pneumonia: Bacterial-Stage 3	-0.3613	0.0110
Pulmonary Embolism-Stage 3	-0.3302	0.0333
Any Cancer - Stage 1 (from hospital data)	0.2294	0.0003
Any Cancer - Stage 2 (from hospital data)	0.1477	0.1008
Any Cancer - Stage 3 (from hospital data)	1.3077	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.1001	0.0026
Any Cardiovascular - Stage 3 (from hospital data)	0.3030	<.0001
Any Eye Disease - All stages (from hospital data)	-0.1491	0.0443
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.2093	<.0001
Any Hemotologic Disease - Stage 3 (from hospital data)	0.4744	0.0174
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.1645	<.0001
Any Neurologic Disease - Stage 1 (from hospital data)	-0.1809	0.0076
Any Neurologic Disease - Stage 3 (from hospital data)	0.3131	<.0001
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2321	0.0008
Any Respiratory Disease – Stage 1 (from hospital data)	-0.1657	0.0476

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any Respiratory Disease - Stage 3 (from hospital data)	0.5038	<.0001
Cancer (from any data source)	-0.1228	0.0560
Cancer (from drug prescriptions)	0.2540	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.2097	0.0046
Endocrine Disease (from drug prescriptions)	-0.0828	0.0222
Genitourinary Disease (from drug prescriptions)	0.4347	<.0001
Hematologic Disease (from drug prescriptions)	0.1757	0.0011
Hepatobiliary Disease (from drug prescriptions)	0.6510	0.0006
Neurologic Diseases (from drug prescriptions)	0.1086	0.0014
Respiratory Disease (from drug prescriptions)	0.1865	<.0001
Cardiovascular Disease (from any data source)	-0.0770	0.2400
Day hospitalization	-0.1056	0.0250
Oral anti-coagulants	0.1976	<.0001
Alpha-blockers	-0.1014	0.0038
Statins	-0.1832	<.0001
Beta-blockers	-0.0628	0.0031
ACE/ARB	-0.1956	<.0001
Calcium channel blockers	-0.0715	0.0007
Nitrates	0.1676	<.0001
Number of ER visits labeled 'Red'	0.2174	<.0001
Total number of ER visits	0.2167	<.0001
Eye Disease (from any data source)	-0.1462	<.0001
Gynecologic Disease (from any data source)	-0.4041	<.0001
History of Cancer (from drug prescriptions) *	0.1108	0.0048
History of Endocrine Disease (from drug prescriptions) *	0.1042	0.0016
History of Neurological Disease (from drug prescriptions) *	0.0915	0.0011
History of Psychological Disease (from drug prescriptions) *	0.0506	0.0295
History of Respiratory Disease (from drug prescriptions) *	0.0905	0.0002
History of Aortic Stenosis-Stage 3 *	0.1616	0.0515
History of Arrhythmias-Stage 2 *	0.1000	0.0015
History of Cardiomyopathies-Stage 2 *	0.1681	0.0697
History of Cardiomyopathies-Stage 3 *	0.3151	0.0108
History of Congestive Heart Failure-Stage 3 *	0.3407	<.0001
History of Coronary Artery Disease-Stage 1 *	0.1753	<.0001
History of Coronary Artery Disease-Stage 2 *	0.1517	0.0017
History of Coronary Artery Disease-Stage 3 *	0.1935	<.0001
History of Essential Hypertension-Stage 1 *	0.0413	0.0434
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.1886	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.1483	0.0027
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.1335	0.0715
History of Hypothyroidism-Stage 2 or 3 *	0.1892	0.0534
History of Diverticular Disease-Stage 2 or 3 *	0.1638	0.0572
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.7712	0.0010
History of Neoplasm, Malignant: Bladder, Urinary-Stage 3 *	1.3084	0.0182

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Neoplasm, Malignant: Kidneys-Stage 3 *	0.7219	0.0444
History of Renal Failure-Stage 2 or 3 *	0.2398	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 1 *	-0.2279	0.0002
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.3479	0.0056
History of Neoplasm, Malignant: Ovaries-Stage 1 *	0.3945	0.0551
History of Neoplasm, Malignant: Ovaries-Stage 2 or 3 *	0.4745	0.0314
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.2900	0.0103
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.4900	0.0545
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.1010	0.0831
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.1793	0.0089
History of Cholecystitis and Cholelithiasis-Stage 3 *	0.2803	0.0315
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.2739	<.0001
History of Rheumatic Fever-Stage 3 *	0.4281	0.0005
History of Progressive Systemic Sclerosis-Stage 1 *	0.4962	0.0004
History of Cerebrovascular Disease-Stage 1 *	0.1710	<.0001
History of Cerebrovascular Disease-Stage 2 *	0.1318	0.0003
History of Cerebrovascular Disease-Stage 3 *	0.2019	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.3755	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.4162	<.0001
History of Obesity-Stage 2 or 3 *	0.1187	0.1089
History of polypharmacy *	0.1086	<.0001
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.3557	0.0356
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.2343	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.6055	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.3311	0.0150
History of Pneumonia: Bacterial-Stage 3 *	0.2981	0.0013
History of Oral Anti-coagulants *	0.1008	0.0057
History of Other Cardiovascular drugs *	0.1223	<.0001
Musculoskeletal Disease (from any data source)	-0.0624	0.0065
Hospitalization	0.1280	0.0248
Number of hospitalizations	-0.0643	0.0011
Polypharmacy	0.1500	<.0001
Psychological Disease (from any data source)	0.1240	<.0001
Any of the other 9 Cardiovascular drugs	0.0679	0.1160
Number of the other 9 Cardiovascular drugs	0.1028	<.0001

Males 75-84

Variable	Coefficient	p-value
Intercept	-1.0190	0.4621
Number of Chronic Conditions (from any data source)=1	0.2984	0.0003

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Number of Chronic Conditions (from any data source)=2	0.5009	<.0001
Number of Chronic Conditions (from any data source)=3	0.5987	<.0001
Number of Chronic Conditions (from any data source)=4	0.7284	<.0001
Number of Chronic Conditions (from any data source)=5	0.7507	<.0001
Number of Chronic Conditions (from any data source)=6 or more	0.8596	<.0001
Number of Chronic Conditions (from hospital data)=1	0.1570	0.0049
Number of Chronic Conditions (from hospital data)=2	0.1638	0.0317
Number of Chronic Conditions (from hospital data)=3	0.1457	0.1301
Number of Chronic Conditions (from hospital data)=4 or more	0.2159	0.0857
Number of Chronic Conditions (from home health prescription)=1	0.3898	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.2645	0.0832
Number of Chronic Conditions (from drug prescriptions)=1	-0.2228	0.0042
Number of Chronic Conditions (from drug prescriptions)=2	-0.2887	0.0012
Number of Chronic Conditions (from drug prescriptions)=3	-0.2736	0.0079
Number of Chronic Conditions (from drug prescriptions)=4	-0.3649	0.0023
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.3187	0.0260
Reside in Mountain area on 12/ 31/ 2012	-0.0273	0.4157
Reside in Hill area on 12/ 31/ 2012	0.0419	0.0161
Age on 12/ 31/ 2012	-0.0638	0.0170
Cancer (from home health prescription)	0.6146	<.0001
Genitourinary Disease (from home health prescription)	-0.2349	0.2216
Blood Diseases (from home health prescription)	0.5082	0.1476
Infectious Disease (from home health prescription)	0.3613	0.1489
Neurologic Diseases (from home health prescription)	0.2858	0.0007
Mental Disorders (from home health prescription)	0.1399	0.1595
Respiratory Diseases (from home health prescription)	0.1985	0.1585
Cancer (chemo or radiation)	0.2160	0.0058
Genitourinary (dialysis)	0.1829	0.1227
Aneurysm, Thoracic-all stages	-0.1892	0.2250
Aortic Stenosis-Stage 1	0.1531	0.1625
Aortic Stenosis-Stage 3	0.1564	0.2244
Arrhythmias-Stage 2	-0.1538	0.0208
Cardiomyopathies-Stage 2	-0.1466	0.2544
Cardiomyopathies-Stage 3	0.3133	0.0224
Conduction Disorders-all stages	-0.1835	0.0881
Congestive Heart Failure-Stage3	0.2137	0.0021
Coronary Artery Disease-Stage 2	0.0566	0.4552
Coronary Artery Disease-Stage 3	0.1414	0.0418
Essential Hypertension-Stage 1	-0.1169	0.0073
Essential Hypertension-Stage 2	-0.2785	0.0001
Essential Hypertension-Stage 3	0.1298	0.0362
Infective Endocarditis-Stage 3	0.9736	0.0197

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Mitral Stenosis-Stage 2	-0.3114	0.1333
Pericarditis: Viral or Traumatic-Stage 2 or 3	-0.8469	0.0349
Thrombophlebitis-Stage 1	0.2275	0.1472
Thrombophlebitis-Stage 2 or 3	-0.1048	0.4208
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 1	0.1183	0.1409
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3	0.1645	0.2385
Diabetes Mellitus Type 1 or Type 2-Stage 2	-0.1939	0.0562
Diabetes Mellitus Type 1 or Type 2-Stage 3	-0.4025	0.0333
Hyperthyroidism-Stage 2 or 3	-0.2385	0.4296
Hypothyroidism-Stage 1	-0.1541	0.2322
Crohns Disease-Stage 1	0.8307	0.0454
Functional Digestive Disorders-Stage 1	0.1566	0.1709
Hernia, Hiatal or Reflux Esophagitis-Stage 2 or 3	0.7272	0.0040
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.4353	0.0125
Neoplasm, Malignant: Colon and Rectum-Stage 3	-0.2956	0.1051
Neoplasm, Malignant: Stomach-Stage 1	0.5695	0.0019
Neoplasm, Malignant: Stomach-Stage 3	0.3811	0.1781
Calculus of the Urinary Tract-Stage 1	-0.1383	0.3085
Neoplasm, Malignant: Bladder, Urinary-Stage 3	0.2547	0.4352
Neoplasm, Malignant: Kidneys-Stage 1	0.1778	0.2881
Renal Failure-Stage 2 or 3	0.3215	<.0001
Anemia: Aplastic, Acquired-Stage 2 or 3	0.2550	0.2048
Neoplasm, Malignant Hematologic-Stage 1	0.1501	0.1974
Neoplasm, Malignant Hematologic-Stage 3	-0.8912	0.0050
Cholecystitis and Cholelithiasis-Stage 2	0.4441	0.0408
Cirrhosis of the Liver-Stage 2 or 3	0.5304	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	1.1048	<.0001
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.2846	0.0009
Rheumatic Fever- Stage 2	0.2675	0.0866
Rheumatic Fever- Stage 3	0.2708	0.3136
Neoplasm, Malignant: Prostate-Stage 2	-0.3625	0.0290
Neoplasm, Malignant: Prostate-Stage 3	-0.2814	0.1619
Progressive Systemic Sclerosis-Stage 1	1.1059	0.1109
Cerebrovascular Disease-Stage 1	0.1684	0.0636
Cerebrovascular Disease-Stage 2	-0.2728	0.0094
Cerebrovascular Disease-Stage 3	-0.2874	0.0742
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.4422	0.0018
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.7034	0.0002
Epilepsy-all stages	-0.3392	0.0337
Bipolar Disorder - Manic Episode-Stage 1	0.6002	0.4877
Depression-Stage 1 or 2	-0.1468	0.3176
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	1.2556	0.0028
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.1890	0.3351

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.0777	0.0884
Chronic Obstructive Pulmonary Disease-Stage 3	0.1545	0.2047
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.7063	<.0001
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	0.3644	0.2398
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.3851	0.0755
Pulmonary Embolism-Stage 3	-0.2243	0.2238
Any Cancer - Stage 2 (from hospital data)	0.4412	0.0019
Any Cancer - Stage 3 (from hospital data)	1.2233	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.0637	0.2223
Any Cardiovascular Disease - Stage 2 (from hospital data)	0.1167	0.1053
Any Endocrine Disease - Stage 1 (from hospital data)	-0.1650	0.0986
Any Endocrine Disease - Stage 3 (from hospital data)	0.3361	0.0522
Any Ear,Nose,Throat Disease - Stage 1 (from hospital data)	-0.3128	0.1174
Any Ear,Nose,Throat Disease - Stage 2 (from hospital data)	-1.4375	0.0217
Any Eye Disease - All stages (from hospital data)	-0.2230	0.0012
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.1736	0.0372
Any Genitourinary Disease - Stage 1 (from hospital data)	0.2720	0.1932
Any Genitourinary Disease - Stage 2 (from hospital data)	0.3593	0.0553
Any Genitourinary Disease - Stage 3 (from hospital data)	0.4935	0.0195
Any Hemotologic Disease - Stage 3 (from hospital data)	0.3202	0.1946
Any Hepatobiliary Disease - Stage 2 (from hospital data)	-0.4870	0.0127
Any Immunologic Disease - All stages (from hospital data)	0.3766	0.3824
Any Infectious Disease - Stage 3 (from hospital data)	0.2171	0.2273
Any Male Genital System - All stages (from hospital data)	-0.1061	0.0692
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.1530	0.0217
Any Neurologic Disease - Stage 1 (from hospital data)	-0.1396	0.1837
Any Neurologic Disease - Stage 2 (from hospital data)	0.2235	0.0339
Any Neurologic Disease - Stage 3 (from hospital data)	0.3632	0.0284
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2738	0.0487
Any Psychologic Disease - Stage 3 (from hospital data)	-0.3812	0.0297
Any Respiratory Disease - Stage 2 (from hospital data)	0.2521	0.0012
Any Respiratory Disease - Stage 3 (from hospital data)	0.4709	<.0001
Any Skin Disease - Stage 1 (from hospital data)	-0.2014	0.1407
Any Skin Disease - Stage 2 (from hospital data)	-0.4164	0.0263
Neoplasm, Malignant: Melanoma-Stage 2	-0.5839	0.1101
Neoplasm, Malignant: Melanoma-Stage 3	0.6833	0.1751
Cancer (from drug prescriptions)	0.1681	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.2349	<.0001
Endocrine Disease (from drug prescriptions)	-0.0963	0.2515
Eye Disease (from drug prescriptions)	-0.1283	<.0001
Genitourinary Disease (from drug prescriptions)	0.4556	<.0001
Hematologic Disease (from drug prescriptions)	0.1419	0.0185
Hepatobiliary Disease (from drug prescriptions)	0.5789	0.0029

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Musculoskeletal Disease (from drug prescriptions)	0.0610	0.4251
Neurologic Diseases (from drug prescriptions)	0.0903	0.0217
Psychological Disease (from drug prescriptions)	-0.0954	0.2745
Respiratory Disease (from drug prescriptions)	0.1867	<.0001
Skin Disease (from drug prescriptions)	-0.3674	0.0144
Number of day hospitalizations	-0.0412	0.1517
Oral anti-coagulants	0.1213	0.0014
Alpha-blockers	-0.0742	0.0360
Statins	-0.1293	<.0001
ACE/ARB	-0.1708	<.0001
Anti-platelets	0.0372	0.1632
Calcium channel blockers	-0.0590	0.0133
Digitalis glycosides	0.1134	0.0086
Nitrates	0.2066	<.0001
Diuretics	0.0565	0.0325
Endocrine Disease (from any data source)	0.0742	0.3637
Number of ER visits labeled 'Yellow'	-0.1352	0.0027
Total number of ER visits	0.3208	<.0001
Genitourinary Disease (from any data source)	0.1205	0.3247
History of Endocrine Disease (from drug prescriptions) *	0.0680	0.0589
History of Neurological Disease (from drug prescriptions) *	0.0426	0.2080
History of Psychological Disease (from drug prescriptions) *	0.0262	0.3867
History of Respiratory Disease (from drug perscriptions) *	0.0918	0.0003
History of Aortic Stenosis-Stage 1 *	-0.1295	0.0581
History of Aortic Stenosis-Stage 3 *	0.0889	0.2974
History of Arrhythmias-Stage 2 *	0.0850	0.0061
History of Arrhythmias-Stage 3 *	0.1402	0.0717
History of Cardiomyopathies-Stage 2 *	0.1934	0.0028
History of Cardiomyopathies-Stage 3 *	0.2510	0.0033
History of Congestive Heart Failure-Stage 3 *	0.2257	<.0001
History of Coronary Artery Disease-Stage 1 *	0.1133	<.0001
History of Coronary Artery Disease-Stage 2 *	0.1341	0.0004
History of Coronary Artery Disease-Stage 3 *	0.0377	0.3346
History of Essential Hypertension-Stage 2 *	-0.0501	0.1298
History of Essential Hypertension-Stage 3 *	0.1584	<.0001
History of Infective Endocarditis-Stage 3 *	-0.3815	0.1951
History of Mitral Stenosis-Stage 3 *	0.1789	0.1632
History of Pericarditis: Chronic-Stage 2 or 3 *	-0.1570	0.3956
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	0.2146	0.2431
History of Thrombophlebitis-Stage 2 or 3 *	-0.1115	0.1430
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.0900	0.3217
hx_Drug-Drug interactions	0.0611	0.0593

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.0914	0.0047
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.3099	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.1290	0.0545
History of Hypothyroidism-Stage 2 or 3 *	0.2294	0.1961
History of Crohns Disease-Stage 2 or 3 *	0.8327	0.0190
History of Diverticular Disease-Stage 1 *	0.1541	0.0060
History of Diverticular Disease-Stage 2 or 3 *	0.1222	0.1954
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.4265	<.0001
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.6959	0.0011
History of Neoplasm, Malignant: Bladder, Urinary-Stage 1*	-0.0471	0.3540
History of Neoplasm, Malignant: Bladder, Urinary-Stage 3 *	0.3138	0.3404
History of Neoplasm, Malignant: Kidneys-Stage 3 *	0.3046	0.3079
History of Renal Failure-Stage 2 or 3 *	0.1469	<.0001
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.2973	0.0315
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.2804	0.0030
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.1374	0.3846
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.1693	0.0044
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.1182	0.0785
History of Cholecystitis and Cholelithiasis-Stage 3 *	0.1319	0.2852
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.2384	0.0009
History of Neoplasm, Malignant: Pancreas-Stage 1 *	-0.5740	0.0601
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	0.5440	0.3348
History of Rheumatic Fever-Stage 2 *	0.0840	0.3429
History of Neoplasm, Malignant: Prostate-Stage 3 *	0.1568	0.3460
History of Progressive Systemic Sclerosis-Stage 1 *	0.4957	0.3135
History of Cerebrovascular Disease-Stage 1 *	0.1907	<.0001
History of Cerebrovascular Disease-Stage 2 *	0.1589	<.0001
History of Cerebrovascular Disease-Stage 3 *	0.1081	0.0017
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.3414	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.6441	<.0001
History of Potentially inappropriate prescribing - always to be avoided drugs *	-0.0564	0.0053
History of polypharmacy *	0.1075	<.0001
History of Depression-Stage 1 or 2 *	0.0912	0.1498
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.4811	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.2667	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.5780	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.3130	0.0002
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	0.3404	0.1018

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Pneumonia: Bacterial-Stage 2 *	0.2648	0.1229
History of Pneumonia: Bacterial-Stage 3 *	0.2801	0.0001
History of Pulmonary Embolism-Stage 3 *	-0.1687	0.1666
History of Neoplasm, Malignant: Melanoma-Stage 3 *	-0.6881	0.1493
History of Oral Anti-coagulants *	0.0418	0.2647
History of Other Cardiovascular drugs *	0.1298	<.0001
Male Genital System (from any data source)	-0.2023	<.0001
Musculoskeletal Disease (from any data source)	-0.1057	0.1679
Hospitalization	0.2060	<.0001
Number of hospitalizations	-0.0217	0.2903
Polypharmacy	0.0913	0.0005
Psychological Disease (from any data source)	0.2569	0.0032
Cardiovascular Disease (from hospital data)	0.0736	0.2121
Endocrine Disease (from hospital data)	0.1978	0.0758
Gastrointestinal Disease (from hospital data)	-0.0878	0.2559
Genitourinary Disease (from hospital data)	-0.4235	0.0770
Hepatobiliary (from hospital data)	-0.1192	0.0959
Neurologic Disease (from hospital data)	0.0927	0.3439
Skin Disease (from any data source)	0.2454	0.0506
Any of the other 9 Cardiovascular drugs	0.0558	0.2144
Number of the other 9 Cardiovascular drugs	0.0680	<.0001

Females 85 and over

Variable	Coefficient	p-value
Intercept	-4.4541	<.0001
Number of Chronic Conditions (from any data source)=1	0.1340	0.0095
Number of Chronic Conditions (from any data source)=2	0.2527	<.0001
Number of Chronic Conditions (from any data source)=3	0.3482	<.0001
Number of Chronic Conditions (from any data source)=4	0.3660	<.0001
Number of Chronic Conditions (from any data source)=5	0.3330	0.0004
Number of Chronic Conditions (from any data source)=6 or more	0.3306	0.0032
Number of Chronic Conditions (from hospital data)=1	-0.0564	0.3573
Number of Chronic Conditions (from hospital data)=2	-0.0127	0.8565
Number of Chronic Conditions (from hospital data)=3	0.1268	0.0996
Number of Chronic Conditions (from hospital data)=4 or more	0.1910	0.0340
Number of Chronic Conditions (from home health prescription)=1	0.3946	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.6315	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.0973	0.0439
Number of Chronic Conditions (from drug prescriptions)=2	-0.1674	0.0036
Number of Chronic Conditions (from drug prescriptions)=3	-0.1843	0.0057

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Number of Chronic Conditions (from drug prescriptions)=4	-0.0790	0.3133
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.0275	0.7718
Age on 12 /31/ 2012	-0.0002	<.0001
Cardiovascular Disease (from home health prescription)	-0.0967	0.0033
Blood Diseases (from home health prescription)	-0.3867	0.0183
Respiratory Diseases (from home health prescription)	-0.3093	0.0020
Skin Disease (from home health prescription)	0.2047	0.0023
Genitourinary (dialysis)	0.8709	0.0002
Aortic Stenosis-Stage 3	0.2778	0.0107
Conduction Disorders-all stages	-0.2408	0.0266
Congestive Heart Failure-Stage3	0.1248	0.0454
Coronary Artery Disease-Stage 1	-0.1922	0.0014
Coronary Artery Disease-Stage 2	0.2265	0.0047
Essential Hypertension-Stage 1	-0.3230	<.0001
Essential Hypertension-Stage 2	-0.1393	0.0052
Mitral Stenosis-Stage 2	-0.4336	0.0440
Diabetes Mellitus Type 1 or Type 2-Stage 2	-0.3893	0.0295
Functional Digestive Disorders-Stage 1	0.2998	0.0008
Hernia, Hiatal or Reflux Esophagitis-Stage 1	-0.2273	0.0578
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.9583	0.0003
Renal Failure-Stage 2 or 3	0.1598	0.0010
Neoplasm, Malignant: Breast, Female-Stage 1	-0.4438	0.0011
Anemia: Aplastic, Acquired-Stage 2 or 3	0.9430	<.0001
Rheumatic Fever- Stage 2	0.2483	0.0251
Rheumatic Fever- Stage 3	0.7922	<.0001
Cerebrovascular Disease-Stage 3	0.7130	0.0005
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.2095	0.0381
Bipolar Disorder - Manic Episode-Stage 2	1.3679	0.0661
Any Cancer - Stage 2 (from hospital data)	1.1401	<.0001
Any Cancer - Stage 3 (from hospital data)	0.9674	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	0.2126	0.0002
Any Cardiovascular - Stage 3 (from hospital data)	0.1855	<.0001
Any Endocrine - Stage 2 (from hospital data)	0.4821	0.0015
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.2081	0.0003
Any Hemotologic Disease - Stage 3 (from hospital data)	0.3382	0.0529
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.0783	0.0438
Any Neurologic Disease - Stage 3 (from hospital data)	-0.4771	0.0171
Any Respiratory Disease - Stage 3 (from hospital data)	0.2027	0.0040
Neoplasm, Malignant: Melanoma-Stage 2	-1.4509	0.0019
Cardiovascular Disease (from drug prescriptions)	-0.2006	<.0001
Endocrine Disease (from drug prescriptions)	-0.1009	0.0041
Genitourinary Disease (from drug prescriptions)	0.2622	0.0004
Hematologic Disease (from drug prescriptions)	0.1242	0.0087
Hepatobiliary Disease (from drug prescriptions)	0.8195	0.0083

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Oral anti-coagulants	-0.0905	0.0222
Statins	-0.2123	<.0001
ACE/ARB	-0.1098	<.0001
Calcium channel blockers	-0.0802	0.0002
Anti-arrhythmics	0.0988	0.0106
Digitalis glycosides	0.1114	0.0002
Nitrates	0.1279	<.0001
Diuretics	0.0611	0.0088
Number of ER visits labeled 'Yellow'	-0.2129	0.0004
Total number of ER visits	0.4649	<.0001
Eye Disease (from any data source)	-0.1404	<.0001
History of Cancer (from drug prescriptions) *	0.1381	0.0002
History of Endocrine Disease (from drug prescriptions) *	0.1108	0.0005
History of Psychological Disease (from drug prescriptions) *	0.0824	<.0001
History of Respiratory Disease (from drug prescriptions) *	0.0686	0.0043
History of Aortic Stenosis-Stage 1 *	0.1279	0.0492
History of Aortic Stenosis-Stage 3 *	0.3272	<.0001
History of Arrhythmias-Stage 2 *	0.1359	<.0001
History of Cardiomyopathies-Stage 3 *	0.4752	0.0004
History of Congestive Heart Failure-Stage 3 *	0.1905	<.0001
History of Coronary Artery Disease-Stage 1 *	0.0616	0.0342
History of Coronary Artery Disease-Stage 2 *	0.1270	0.0059
History of Essential Hypertension-Stage 3 *	0.1426	<.0001
History of Mitral Stenosis-Stage 2 *	0.2121	0.0393
History of Mitral Stenosis-Stage 3 *	0.2109	0.0254
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.1684	0.0609
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.0983	0.0030
History of Neoplasm, Malignant: Kidneys-Stage 1 *	0.3075	0.0501
History of Renal Failure-Stage 2 or 3 *	0.1614	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.5190	0.0032
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.3169	0.0038
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.2489	0.0263
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.1278	0.0084
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.1977	0.0150
History of Rheumatic Fever-Stage 2 *	0.1641	0.0171
History of Rheumatic Fever-Stage 3 *	0.4295	<.0001
History of Cerebrovascular Disease-Stage 1 *	0.1780	<.0001
History of Cerebrovascular Disease-Stage 2 *	0.1248	0.0001
History of Cerebrovascular Disease-Stage 3 *	0.2010	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.2327	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.3112	<.0001
History of polypharmacy *	0.0412	0.0637

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.1641	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.2718	0.0010
History of Pneumonia: Bacterial-Stage 3 *	0.2305	0.0024
History of Oral Anti-coagulants *	0.1412	0.0001
History of Other Cardiovascular drugs *	0.0945	0.0002
Musculoskeletal Disease (from any data source)	-0.1212	<.0001
Hospitalization	0.4053	<.0001
Number of hospitalizations	-0.0572	0.0152
Polypharmacy	0.0686	0.0047
Psychological Disease (from any data source)	0.0241	0.3329
Respiratory Disease (from any data source)	0.1452	<.0001
Cancer (from hospital data)	0.1946	0.0015
Number of the other 9 Cardiovascular drugs	0.0343	0.0067

Males 85 and over

Variable	Coefficient	p-value
Intercept	-6.5943	<.0001
Number of Chronic Conditions (from any data source)=1	0.0148	0.7836
Number of Chronic Conditions (from any data source)=2	0.0567	0.3273
Number of Chronic Conditions (from any data source)=3	0.1133	0.0832
Number of Chronic Conditions (from any data source)=4	0.1108	0.1382
Number of Chronic Conditions (from any data source)=5	0.1027	0.2348
Number of Chronic Conditions (from any data source)=6 or more	-0.00058	0.9956
Number of Chronic Conditions (from hospital data)=1	0.2260	<.0001
Number of Chronic Conditions (from hospital data)=2	0.3490	<.0001
Number of Chronic Conditions (from hospital data)=3	0.3454	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.2443	0.0024
Number of Chronic Conditions (from home health prescription)=1	0.5620	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.9112	<.0001
Age on 12/ 31/ 2012	0.0552	<.0001
Cardiovascular Disease (from home health prescription)	-0.1808	0.0009
Neurologic Diseases (from home health prescription)	-0.2313	0.0125
Cancer (chemo or radiation)	0.4807	0.0028
Aortic Stenosis-Stage 3	0.3770	0.0178
Coronary Artery Disease-Stage 2	0.1810	0.0345
Essential Hypertension-Stage 1	-0.1334	0.0049
Mitral Stenosis-Stage 3	0.7232	0.0092
Crohns Disease-Stage 2 or 3	2.0123	0.0823
Renal Failure-Stage 2 or 3	0.2166	<.0001
Neoplasm, Malignant Hematologic-Stage 2	0.9737	0.0279
Cerebrovascular Disease-Stage 2	-0.3039	0.0005

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.4636	0.0033
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	1.0085	<.0001
Any Cancer - Stage 3 (from hospital data)	1.2151	<.0001
Any Cardiovascular - Stage 3 (from hospital data)	0.1683	0.0006
Any Eye Disease – All stages (from hospital data)	-0.3186	0.0024
Any Psychologic Disease - Stage 1 (from hospital data)	-0.3000	0.0483
Any Psychologic Disease - Stage 2 (from hospital data)	0.3406	0.0468
Cancer (from any data source)	0.0881	0.0255
Cardiovascular Disease (from drug prescriptions)	-0.0958	0.0203
Eye Disease (from drug prescriptions)	-0.1138	0.0019
Genitourinary Disease (from drug prescriptions)	0.3297	0.0004
Hepatobiliary Disease (from drug prescriptions)	1.1786	0.0031
Respiratory Disease (from drug prescriptions)	0.2047	<.0001
Day hospitalization	-0.2140	0.0003
Statins	-0.1823	<.0001
ACE/ARB	-0.1220	<.0001
Digitalis glycosides	0.1380	0.0011
Nitrates	0.1895	<.0001
Diuretics	0.1133	<.0001
Number of ER visits labeled 'Yellow'	-0.3232	<.0001
Total number of ER visits	0.5087	<.0001
Genitourinary Disease (from any data source)	0.1686	0.0037
Hematologic Disease (from any data source)	0.1699	0.0006
History of Endocrine Disease (from drug prescriptions) *	0.0877	0.0078
History of Psychological Disease (from drug prescriptions) *	0.1168	0.0009
History of Aortic Stenosis-Stage 3 *	0.3893	0.0002
History of Arrhythmias-Stage 2 *	0.1078	0.0013
History of Congestive Heart Failure-Stage 3 *	0.2010	0.0002
History of Coronary Artery Disease-Stage 1 *	0.1312	0.0002
History of Coronary Artery Disease-Stage 2 *	0.1103	0.0392
History of Coronary Artery Disease-Stage 3 *	0.1075	0.0476
History of Essential Hypertension-Stage 3 *	0.1204	0.0079
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	-0.7456	0.0101
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.1159	0.0141
History of Neoplasm, Malignant: Bladder, Urinary-Stage 1*	0.1311	0.0546
History of Renal Failure-Stage 2 or 3 *	0.1660	<.0001
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.3476	0.0069
History of Neoplasm, Malignant Hematologic-Stage 2 *	-1.1984	0.0057
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.2949	0.0083
History of Progressive Systemic Sclerosis-Stage 1 *	2.4748	0.0284
History of Cerebrovascular Disease-Stage 1 *	0.2552	<.0001
History of Cerebrovascular Disease-Stage 3 *	0.1474	0.0006
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.2598	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.2971	0.0457
History of Potentially inappropriate prescribing - always to be avoided drugs *	-0.0705	0.0104
History of polypharmacy *	0.1276	0.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.2083	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.2686	0.0030
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.4042	0.0142
History of Pneumonia: Bacterial-Stage 2 *	0.4137	0.0392
History of Other Cardiovascular drugs *	0.1277	0.0005
Immunologic Disease (from any data source)	0.9577	0.0461
Male Genital System (from any data source)	-0.1301	<.0001
Neurologic Disease (from any data source)	0.1639	<.0001
Polypharmacy	0.0778	0.0206
Psychological Disease (from any data source)	0.1247	0.0017
Endocrine Disease (from hospital data)	0.1413	0.0154
Respiratory Disease (from hospital data)	0.3324	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only

Predicting Risk Using a Population Based Longitudinal Database
Louis et al
STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	See abstract section: Design.
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	The outcome measures used and what was found are summarized in the sections of the abstract: Main outcome measures and Results
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	See the Introduction section pages 4 and 5
Objectives	3	State specific objectives, including any prespecified hypotheses	Our objectives are described in the last two paragraphs of the Introduction
Methods			
Study design	4	Present key elements of study design early in the paper	See last paragraph of the Introduction and the Methods section
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	See Study Data and Study Population at beginning of Methods section
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	Our study includes 100% of the adult population See Study Data and Study Population at beginning of Methods section
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	See the Dependent variable and Independent variable sections in the Methods section
Data sources/ measurement	8*	For each variable of interest, give sources of data	See the Dependent variable

For Peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	and Independent variable sections in the Methods section. In addition we have included an Appendix with detailed mapping to independent variable.
Bias	9	Describe any efforts to address potential sources of bias	See Evaluation of the Models (page 10) in the Methods section
Study size	10	Explain how the study size was arrived at	Our study includes 100% of the adult population
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	See the Dependent variable and Independent variable sections in the Methods section. In addition we have included an Appendix with detailed mapping to independent variable.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	See the Modelling section and Evaluation of the Models section (pages 9 and 10) of the Methods
		(*b*) Describe any methods used to examine subgroups and interactions	See the Modelling section (pages 9 and 10) of the Methods
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	Our study includes 100% of the adult population
		(*e*) Describe any sensitivity analyses	See Evaluation of the Models last paragraph of Methods.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	See Page 11 Results section
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	See Table 1 and Results section (page 11-14)
		(b) Indicate number of participants with missing data for each variable of interest	

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	See Table 2 and figure and pages 14 and 15 of Results section
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	See pages 13
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	See Table 2 and page 13 and 14.
Discussion			
Key results	18	Summarise key results with reference to study objectives	See Discussion section page 16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	See Discussion section pages 17 and 18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	See Discussion section page 18
Generalisability	21	Discuss the generalisability (external validity) of the study results	See Discussion section pages 18 and 19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	The study was funded by the Emilia-Romagna, region of Italy. See page 3.

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open

Predicting Risk of Hospitalization or Death: A Retrospective Population Based Analysis

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-005223.R2
Article Type:	Research
Date Submitted by the Author:	30-Aug-2014
Complete List of Authors:	Louis, Daniel; Sidney Kimmel Medical College at Thomas Jefferson University, Center for Research in Medical Education and Health Care Robeson, Mary; Sidney Kimmel Medical College at Thomas Jefferson University, Center for Research in Medical Education and Health Care McAna, John; Thomas Jefferson University, Jefferson School of Population Health Maio, Vittorio; Thomas Jefferson University, Jefferson School of Population Health Keith, Scott; Thomas Jefferson University, Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics Liu, Mengdan; Sidney Kimmel Medical College at Thomas Jefferson University, Center for Research in Medical Education and Health Care Gonnella, Joseph; Sidney Kimmel Medical College at Thomas Jefferson University, Center for Research in Medical Education and Health Care Grilli, Roberto; Agenzia Sanitaria e Sociale Regione Emilia-Romagna,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Patient-centred medicine
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

Predicting Risk of Hospitalization or Death: A Retrospective Population Based Analysis

Daniel Z. Louis[1], Mary Robeson[1], John McAna[2], Vittorio Maio[2], Scott W. Keith[3], Mengdan Liu[1], Joseph S. Gonnella[1], Roberto Grilli[4]

[1]Center for Research in Medical Education and Health Care, Sidney Kimmel Medical College at Thomas Jefferson University. 1015 Walnut Street, Suite 319. Philadelphia, PA 19107, USA

[2]Jefferson School of Population Health, Thomas Jefferson University. 901 Walnut Street, 10th Floor. Philadelphia, PA 19107, USA

[3]Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics, Thomas Jefferson University. 1015 Chestnut Street, Suite M-100. Philadelphia, PA 19107, USA

[4]Agenzia Sanitaria e Sociale Regione Emilia-Romagna, Viale Aldo Moro, 21. Bologna, 40128 Italy.

Corresponding author: Daniel Z. Louis, Sidney Kimmel Medical College, 1015 Walnut Street, Suite 319, Philadelphia, PA 19107. USA; Tel: 215-955-6959, daniel.louis@jefferson.edu

Copyright: The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (non-exclusive for Dr. Grilli as a government employee) on a worldwide basis to the BMJ publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license.

Ethical approval: This study was approved by the Institutional Review Board of Thomas Jefferson University as an expedited retrospective database/record review. The IRB granted a waiver of informed consent.

Keywords: hospitalization, risk, medical home, patient-centered care

For peer review only

Abstract

Objectives: Develop predictive models using an administrative health care database that provide information for Patient Centered Medical Homes to proactively identify patients at risk of hospitalization for conditions that may be impacted through improved patient care.

Design: Retrospective health care utilization analysis with multivariate logistic regression models.

Data: A population-based longitudinal database of residents served by the Emilia-Romagna, Italy health service in the years 2004-2012 including demographic information and utilization of health services by 3,726,380 people age ≥ 18 years.

Outcome measures: Models designed to predict risk of hospitalization or death in 2012 for problems that are potentially avoidable were developed and evaluated using the area under the receiver operating curve C-statistic, in terms of their sensitivity, specificity, and positive predictive value, and for calibration to assess performance across levels of predicted risk.

Results: Among the 3,726,380 adult residents of Emilia-Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions or died in 2012 (3.6% hospitalized, 1.3% died). The C-statistic for predicting 2012 outcomes was 0.856. The model was well calibrated across categories of predicted risk. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%.

Conclusions: We have developed a population-based model using a longitudinal administrative database that identifies the risk of hospitalization for residents of the Emilia-Romagna region

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

with a level of performance as high as, or higher than, similar models. The results of this model, along with profiles of patients identified as high risk are being provided to the physicians and other health care professionals associated with the Patient Centered Medical Homes to aid in planning for care management and interventions that may reduce their patients' likelihood of a preventable, high-cost hospitalization.

Strengths and Limitations of this study:

- This study included the entire adult population of the Emilia-Romagna Region of Italy, over 3.7 million people.
- The study used an existing longitudinal administrative health care database with both the advantage of much lower cost than new data collection and the disadvantage of potential errors in administrative data.
- The results of the study are being used to assist in the development of newly formed Patient Centered Medical Homes.

Word Count: 3,952 (excluding title page, abstract, references, figure 1, table1 and table 2.)

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

The predominant healthcare delivery system, which has been a passive model, reacting to patients' problems, is shifting to a more proactive model designed to take the initiative in providing care for an increasingly older population that has a greater prevalence of chronic conditions, often with multiple medical and social needs. These changes are driving the reorganization of the primary care system, emphasizing coordination and cooperation among healthcare professionals.[1,2,3] Among the approaches to addressing this need has been the establishment of Patient Centered Medical Homes, organizations in which teams of healthcare providers are engaged in delivering comprehensive, coordinated, patient-centered care to patient defined populations.

Primary care has a central role in the Italian National Health Service (NHS). Twenty one regional governments are responsible for ensuring the delivery of a health benefits package through a network of geographically defined, population-based Local Health Authorities. Primary care physicians work for these authorities as independent contractors and act as "gatekeepers" for specialty and other referral services for their patients.[4]

With the belief that a strong primary care system is conducive to improving population health, the NHS initiated reforms that encouraged primary care physicians to organize into collaborative arrangements. To this end, the Regione Emilia-Romagna (RER), a large northern region with a population of about 4.5 million, has recently launched a plan in its 11 Local Health Authorities to establish Patient-Centered Medical Homes to better coordinate patient care and help patients avoid unnecessary hospitalizations.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The identification of those patients who would benefit most from outreach efforts is fundamental to achieving these goals of promoting and practicing population health in Patient Centered Medical Homes. The RER has established three objectives for this project: 1) develop predictive models to identify patients at high risk of hospitalization or death, 2) create "risk of hospitalization" patient profiles that provide information about their high-risk patients to the general practitioners in the newly formed Patient Centered Medical Homes, and 3) assess the extent to which these models and reports provide additional information useful in the identification of patients who may benefit from case management or disease management.

This paper will address the first of the three goals. We describe the development of a predictive model using the RER's regional longitudinal administrative health care database to help identify patients who are most at risk of hospitalization for conditions that may be impacted through improved patient care. This model will then be used to inform the providers associated with the Patient Centered Medical Homes and aid in their planning for care management and interventions that can reduce their patients' likelihood of a preventable, high-cost hospitalization.

Methods

Study Data and Study Population

The model was developed using the population-based longitudinal health care database of the residents served by the RER Health Service in the years 2004 through 2012.[5] This administrative database includes demographic information for all residents (gender, birth and death dates, location of current residence, and primary care physician), hospital discharge abstract data (ICD-9-CM diagnosis and procedure codes, and admission and discharge dates),

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

emergency room utilization information, outpatient pharmacy data at the individual prescription level, specialty care (laboratory, diagnostics, therapeutic procedures, rehabilitation, and specialist visits), home health data, and information on each primary care physician in the region. Each patient has an anonymous identifier assigned by the RER so that an individual's utilization can be tracked over time without jeopardizing patient privacy.

The study population consisted of all residents of the RER who were at least 18 years of age and still alive as of 31st December 2011. Health care utilization data from 2011 and history variables using data from 2004 through 2010 were used to predict outcomes in 2012.

Dependent Variable

The dependent variable was defined as the occurrence of a hospitalization for problems that are potentially avoidable, or whose progression may have been avoided or delayed through appropriate patient care, or the death of the individual, either in or out of the hospital, for any reason in 2012. We included deaths in the dependent variable since we believe that, for example, a patient with coronary artery disease who dies secondary to an acute myocardial infarction, should be included in the dependent variable even if the death is out of the hospital. We decided to not limit the hospitalization to emergency admissions, since a planned admission may also be an indicator of a worsening medical problem. In order to operationally define the dependent variable, we (authors JSG and DZL) reviewed the Disease Staging[6, 7] primary diagnostic category and severity stage of all day and inpatient hospital admissions (for adults age 18+) in RER for one year, to select those admissions that should be included in the dependent variable.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Admissions to deliver a baby, admission for dental diseases or admissions for vague signs or symptoms with no identified etiology were excluded. Admissions for problems that are not predictable/preventable were excluded while those where screening may identify problems that can potentially be treated to avoid progression were included. For example, admissions for stage 1, chronic cholecystitis or cholelithiasis were excluded, but admissions for advanced stage 2 or 3 complications such as ascending cholangitis or pancreatitis were included.

We felt that inclusion of hospitalization for cancer in the dependent variable should depend on the ability to either prevent or avoid progression of the disease. We therefore included colon cancer and cervical cancer in the definition because they are potentially preventable but excluded all other cancers where prevention/prediction is not currently possible.

Inclusion of injuries, burns, or toxic reaction to prescription or non-prescription drugs would ideally be based on the cause of these problems. Since the etiology of these problems is typically not available in the administrative data being used in this project, we made the decision to include or exclude based on our subjective judgment of the likelihood of preventability. For example, adverse drug reactions were included but burns were excluded from the definition of the dependent variable.

There is no obvious medical reason for a hospital admission for patients with stage 1 diabetes mellitus or stage 1 essential hypertension without complications. These problems are typically treatable in the outpatient setting. A hospitalization implies a potential problem in the care of these patients, so we decided to include these admissions as a part of the dependent variable.[8]

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Independent Variables

A broad range of candidate predictor variables was developed taking advantage of the RER administrative data. The independent variables used for modeling were defined from the RER administrative data for the years 2004 through 2011. Demographic data included patient age, sex, and geographic location of residence. We developed a mapping to broad disease categories defined primarily in terms of the affected body system from home health care data, pharmacy data, and hospital discharge abstract data. (See Appendix 1.)

For those patients who had been hospitalized, more specific diagnostic data were available. We reviewed the classification of patients hospitalized historically using the Disease Staging diagnostic category and disease severity stages.[6,7] Based on the frequencies specific diagnostic category/stage predictor variables were defined for either specific stages of frequent diseases, or by combinations across similar categories. Predictor variables were defined based on the number of emergency room visits using the RER classifications system for the urgency of the visit.

Pharmacy data were used to identify polypharmacy[9] (defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days), potential drug-drug interactions (DDI)[10] and potentially inappropriate medication use in patients[11] 65 years and older. Since cardiovascular disease is highly prevalent, we reviewed the use of cardiovascular drugs and created a variable for each of the following 11 classes of drugs (oral anti-coagulants, beta-blockers, angiotensin converting enzyme / angiotensin II receptor blockers, anti-platelets,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

calcium channel blockers, anti-arrhythmics, digitalis glycosides, nitrates, diuretics, alpha-blockers, statins) to account for the complexity of therapeutic regimen at the patient level.

To take advantage of the fact that the RER database includes multiple years of data, we created history variables using the utilization for each year of data available. Since we were working with the 2011 data to predict hospitalization or death in 2012, we created history variables based on 2004-2010 data. This set included 83 of the diagnostic category/stage variables as well as 11 variables based on pharmacy utilization such as exposure to polypharmacy and use of cardiovascular drugs. If the individual had a history of a disease in any of the years from 2004 to 2010 they were flagged as having a history of that disease and this was used as a potential predictor variable.

Modeling

Logistic regression models were used to estimate predicted probabilities for the occurrence of an inpatient hospital stay for the selected conditions or death for individual patients. Risk of hospitalization or death, and the variables that relate to those risks are highly dependent on age and gender. Regression models were fit in each of 14 gender and age strata using SAS Version 9.2 (SAS Institute, Cary NC). A stepwise process with relaxed covariate entry and retention criteria (inclusion p-value <=0.8, retention <=0.5) was used. At each step in this process, an attempt is made to remove any unimportant variables from the model before adding a potentially important variable. Each addition or deletion of a variable to or from a potential model is a separate step and, at each step, a new model is fitted. This process results in a reduced, but robust

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

set of independent variables that predict outcome or that might have importance as adjustment terms for the model in each age/gender stratum.

Evaluation of the Models

The predictive accuracy of the modelling was evaluated using C-statistics (the area under the receiver operating characteristics curve), along with three measures traditionally used with clinical screening tools: sensitivity, specificity and positive predictive value.

C-statistics were used to evaluate the models in two ways. The first evaluation consisted of fitting the model developed using utilization and demographic data from 2011, along with historical variables based on 2004-2010 data, and outcomes (hospitalization or death) from 2012 and then computing a C-statistic to evaluate how the models performed at predicting those outcomes on which the models were conditioned. However, this evaluation is not consistent with evaluating how the data are used in practice. In practice, we have current predictor information, but the outcomes have not been realized. To better estimate how the models are likely to perform in this setting, we fit models to outcomes data up to a year prior to the most current available (e.g., 2011 outcomes modelled with predictors from 2010, along with historical variables based on 2004-2009 data). We then computed a C-statistic for projections made on the risk of hospitalization or death outcomes (in 2012) using the next year's predictor information (in 2011). This way, the models are forced to make projections into the future, but we have the actual observed outcomes data to evaluate the modelling process as it would be used in practice. The resulting C-statistics obtained from these two model runs were compared.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

In order to evaluate the performance of the model across different risk thresholds we classified predicted risk scores. "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%. These risk thresholds were selected after consultation with physicians practicing in the medical homes to yield a total of about 10% of the 1,500 patients enrolled with a typical primary care physician.

Results

Among the 3,726,380 adult residents of Emilia Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions defined earlier or died in 2012 (3.6% hospitalized, 1.3% died).

Table 1 shows the distribution of the demographics (age and gender), number of chronic conditions, body systems impacted by the selected chronic conditions, polypharmacy and inappropriate prescribing among the eligible RER residents, as of December 31, 2011. Table 1 also compares these characteristics of the total adult population of the region to the subgroups of the population classified in the "very high risk" and "high risk" categories. Based on the model results, 114,255 individuals were identified as having a predicted risk of hospitalization or death in 2012 of ≥ 25% and classified as "very high risk." An additional 134,610 individuals had a predicted risk of hospitalization or death in 2012 of 15-24% and were classified as "high risk."

Table 1. Demographic and clinical characteristics of the Regione Emilia-Romagna population, overall and by risk category

Total Population*	Very High Risk**	High Risk**

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	3,726,380		114,255		134,610	
Gender	N.	%	N.	%	N.	%
Male	1,788,048	48.0%	54,357	47.6%	61,803	45.9%
Female	1,938,332	52.0%	59,898	52.4%	72,807	54.1%
Age groups						
18-24	258,338	6.9%	76	0.1%	105	0.1%
25-34	499,786	13.4%	302	0.3%	391	0.3%
35-44	732,626	19.7%	1,137	1.0%	1,198	0.9%
45-54	676,047	18.1%	2,612	2.3%	2,485	1.8%
55-64	550,689	14.8%	5,391	4.7%	5,287	3.9%
65-74	482,346	12.9%	13,154	11.5%	14,471	10.8%
74-85	364,369	9.8%	33,430	29.3%	44,857	33.3%
85+	162,179	4.4%	58,153	50.9%	65,816	48.9%
Number of Chronic Conditions						
0-1	2,775,888	74.5%	8,176	7.2%	24,618	18.3%
2 or more	950,492	25.5%	106,079	92.8%	109,992	81.7%
5 or more	99,337	2.7%	45,445	39.8%	20,576	15.3%
Selected Conditions/Body Systems						
Cancer	99,328	2.7%	23,872	20.9%	14,305	10.6%
Cardiovascular	967,796	26.0%	96,157	84.2%	103,749	77.1%
Male Genitourinary[#]	130,609	7.3%	14,616	26.9%	16,776	27.1%
Ear, Nose, Throat	5,364	0.1%	240	0.2%	242	0.2%
Endocrine	429,528	11.5%	40,653	35.6%	37,471	27.8%
Eye	114,117	3.1%	9,558	8.4%	13,478	10.0%
Gastrointestinal	580,946	15.6%	74,718	65.4%	66,305	49.3%
Gynecologic[##]	21,806	1.1%	333	0.6%	405	0.6%
Hematologic	45,022	1.2%	15,353	13.4%	6,591	4.9%
Hepatobiliary	24,785	0.7%	6,477	5.7%	3,306	2.5%
Immunologic	3,281	0.1%	464	0.4%	273	0.2%
Infectious Disease	4,723	0.1%	2,207	1.9%	727	0.5%
Musculoskeletal	419,184	11.2%	43,436	38.0%	41,000	30.5%
Neurologic	173,751	4.7%	34,494	30.2%	24,838	18.5%
Psychological	291,308	7.8%	43,387	38.0%	33,715	25.0%
Respiratory	176,830	4.7%	39,082	34.2%	21,763	16.2%
Skin	28,339	0.8%	7,645	6.7%	3,008	2.2%
Urogenital	37,728	1.0%	16,501	14.4%	5,740	4.3%
Polypharmacy[^]	609,278	16.4%	92,153	80.7%	92,156	68.5%
Any potentially inappropriate medications (age 65 years or older)[^^]	257,033	25.5%	51,055	48.7%	49,003	39.2%

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

* Adults (age 18 or older) and alive at 31 December 2011.

** "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%.

Men only.

Women only.

^ Polypharmacy is defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days.

^^ The list of potentially inappropriate medications can be found in: Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *Journal of Clinical Pharmacy and Therapeutics* 2010;35:219-229.

There was little difference across the risk categories by gender. Age distributions for the "very high risk" and "high risk" groups were shifted more towards the older age groups than those in the overall study population. Residents age 85 or older represented about 4.5% of the RER population, but about 50% of the "very high" and "high" predicted risk groups. More than 75% of the residents over age 85 were classified as "very high" or "high" risk. However, age alone was not sufficient to predict their risk. For example, residents between 75 and 84 years of age made up 23% of the "very high" risk group and 41% of the "high" risk group, but over 85% of the residents in this age category had neither "very high" nor "high" predicted risk.

Across age and gender strata, demographics and heath care utilization experience in 2011 were the most commonly used independent variables for predicting hospitalization or death in 2012. Selected history variables flagging chronic problems such as cardiovascular disease, diabetes mellitus and chronic renal failure and a history of prescriptions for cardiovascular medications and polypharmacy were also significant predictors.

The residents in the two higher risk groups were more likely than others to have multiple chronic diseases and to experience polypharmacy and inappropriate medication use. The residents identified as "very high risk" or "high risk" by the model also showed a number of striking

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

differences from others in terms of the occurrence of some of the most prevalent health conditions by type and body system. Although cardiovascular conditions were not uncommon in the total adult population (26.0%), they were far more common among those classified as "very high risk" and "high risk" (84.2% and 77.1%, respectively). Similarly, gastrointestinal conditions affected 15.6% of the total population, but were diagnosed in 65.4% of the "very high risk" and 49.3% of the "high risk" patients. Cancer occurred in 2.7% of the total population, but 20.9% of the "very high risk" and 10.6% of the "high risk" patients had a cancer diagnosis. Mental health problems were identified in 7.8% of the adult population, but in 34.2% of the "very high risk" and 25.0% of the "high risk" patients.

The C-statistic for the model of 2012 outcomes developed using 2011 predictors and the C-statistic based on the parameters from the model of 2011 outcomes regressed on 2010 predictors applied to the 2011 predictors and 2012 outcomes were very similar (0.856 and 0.853, respectively). These results suggest that the relationship between predictors and risk of hospitalization changed little in one year and that model parameters developed in a prior year can be used reliably with the most current year's data to predict unknown outcomes in the next year with only a minimal loss in performance in this population.

Table 2 shows the sensitivity, specificity, positive predictive value and number of true positives for the model at the two selected cut-off points. The sensitivity (percentage of patients actually hospitalized who had been identified by the model as having a predicted risk higher than the cut-off point) was 29.8% for those with the "very high" risk scores. This percentage represents 46,950 of the 157,550 residents of the region who were hospitalized for a selected condition or died in 2012. If we modify the risk score threshold to include individuals with a predicted risk of

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

hospitalization for selected conditions or death of ≥ 15% (i.e., both the "very high risk" and the "high risk" patients) the sensitivity is .471. The true negative rate (specificity) is very high for both risk thresholds (.981 and .951, respectively).

Table 2. Performance of the "Risk of Hospitalization" model for residents identified as "Very High Risk" and "High or Very High Risk"

	Cut-off points for comparison	
Measure	**"Very high risk"***	**"Very high risk"* + "High risk"****
Sensitivity#	0.298	0.471
Specificity##	0.981	0.951
Positive Predictive Value^	0.411	0.298
True positives^^	46,950	74,196

* "Very high risk" is defined as patients with a predicted risk of hospitalization of ≥ 25%.
** "Very high risk" + "High risk", is defined as patients with a predicted risk of hospitalization of ≥15%.
# Sensitivity is defined as the proportion of those hospitalized who were predicted to be hospitalized (true positive rate).
Specificity is the proportion of those not hospitalized who were not predicted to be hospitalized (true negative rate).
^ Positive Predictive Value is the proportion of those predicted to be hospitalized who were actually hospitalized.
^^ True positives are the number of residents who were predicted to be at risk for hospitalization at the predicted risk threshold and were actually hospitalized.

The model appears to be well calibrated across levels of risk. Figure 1 depicts the RER population divided into groups by deciles of predicted risk of hospitalization or death from the models. The observed prevalence of hospitalization or death is compared to the average predicted risk among individuals in each of the ten predicted risk groups. For example, the overall rate of hospitalization for the selected conditions or death in 2012 was 4.2%. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%. (Regression coefficients and

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

significance levels of independent variables for models for each of 14 age and gender strata are displayed in Appendix 2.)

[Insert Figure 1 about here]

Discussion

We have developed a population-based model that identifies the risk of hospitalization for all adult RER residents and does so with a level of performance (c=0.85) as high as, or higher than, similar models. In addition, we believe that the definition of the dependent variable chosen for our models increases the probability that they are identifying patients who risk can potentially be improved by appropriate care. A systematic review by Kansagara[12] of models designed to predict readmissions, showed C-statistic results in the range of 0.55 to 0.83. Recent work by Billings et al[13] to develop models predictive of emergent admissions in the UK had results ranging from 0.73 to 0.78. Li Wang, et al. (2013),[14] using information available through the United States Department of Veterans Affairs that also included lab data, demonstrated C-statistics of 0.81 and 0.79, respectively, for their models of 90-day or 12-month hospitalization or death outcomes. At a predicted risk of ≥25% our model had a Positive Predictive Value (PPV) of .411. Billings et al[14] reported a PPV of .417 at a risk threshold of 30. There is a trade-off in using our model, or any predictive model, between the threshold for follow-up and predictive accuracy. A lower risk threshold would identify more patients but with a lower prevalence of hospitalization or death.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Although previous studies have developed models predictive of hospital care, these models fall short of the needs of the Patient Centered Medical Homes being implemented in RER. Typically, these models have focused on specific age groups,[15] conditions, or types of admissions, such as emergent[14] or unplanned admissions or rehospitalizations, or health insurance plans in the United States, including private insurance plans, Medicare and Medicaid plans.[16,17] The models we have developed are applied to the entire adult population of RER. They use existing administrative data, which makes them cost effective to apply.

Patient Centered Medical Homes, including those instituted in RER, are responsible for addressing the needs of their population and making the best use of their finite resources to accomplish this. Preventing unnecessary admissions could improve both the quality of care and health status of the enrolled population, and result in a substantial savings. To accomplish this, predictive models and risk stratification tools such as those developed for this project are needed to identify patients at risk of preventable admissions and provide information that can be used by the medical homes to help manage care.

There are some limitations to the model. The model is developed from administrative data. Administrative data are collected for reimbursement and tracking utilization and not for medical research. They lack the clinical specificity that would be desirable in assessing an individual's medical problems. Patients with medical problems who have not used the services included in the database cannot be identified. While the hospital discharge abstract data do include diagnostic information coded using ICD-9-CM, no similar data are available for outpatient encounters in the RER database. The mortality data available for this project did not include information about cause of death. Therefore, some proportion of patients whose death was not

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

predictable were included, limiting model performance. In addition, our models use prior utilization among the predictor variables. With the administrative data we cannot distinguish between appropriate and inappropriate treatment which may bias our results.

Despite the limitations of administrative data, they have many advantages for this project: they are readily available, relatively inexpensive to analyze and cover large populations over many years. They are ideal for uncovering patterns of care. If information from the medical records is needed, the results of these analyses can then be supplemented by focused clinical reviews at the local level. Also, The RER has a system in place to monitor the quality of diagnosis and procedure coding in their hospital discharge abstract data. Controls at both the hospital and regional level assess the validity of coding and the consistency of codes assigned such as congruity between sex, age and diagnosis and between diagnosis and procedure. The existence of the RER administrative database made it feasible to develop the models described in this article at relatively low cost and to update the models over time without additional data collection that others have found necessary.[13]

Currently, these risk scores are being integrated with other information in profiles of high-risk patients furnished to providers in 12 newly formed medical homes, including 83 primary care physicians serving a total of about 100,000 patients, in the Parma Local Health Authority located in RER. Along with the risk scores, this information includes data about previous hospitalizations, use of referrals, medications, long-term care and home care services, and a number of process-like quality indicators for diabetic and cardiovascular patients, and for appropriate medication use in older patients.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

We believe that the Italian health care system offers a number of advantages in the goal of reducing potentially avoidable hospitalization. Every Italian must enroll with a primary care physician. The population is quite stable with modest levels of change of residence or change of primary care physician. Every Italian is entitled to health care with little or no cost at the point of service. There is no problem with loss of, or change in, insurance coverage. Primary care physicians are primarily paid on a per capita basis. But the Emilia-Romagna region has the ability to negotiate incentive payments designed to address and monitor improvements in medical management such as that addressed in our study.

Of course, model results need to lead to effective interventions to have a positive impact on patient care. To this end, we are working with the physicians, nurses, and other health care professionals as well as the administration of the newly formed Medical Homes in Parma to assist them in understanding how to use the results of these models and in developing potentially effective interventions. The individual profiles of high risk patients provided to the health care team in the Medical Homes allow them to trigger specific actions such as inviting patients to enroll in disease management programs for chronic problems such as heart failure, chronic obstructive pulmonary disease, or diabetes mellitus, activating home health assistance, initiating a medication review, or recommending that the patient come in for an office visit. An evaluation of the use and usefulness of the profiles and intervention is under way.

In summary, these models provide a means of identifying patients at high risk for hospitalization. The risk predictions, in conjunction with the risk profile, show promise as a useful organizational tool for the regional Patient Centered Medical Homes to develop and implement proactive case management and disease management programs. The RER is reviewing the results of the Parma

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Local Health Authority pilot project of the profiles. Once their usefulness has been further evaluated, their use will be expanded to other Medical Homes in development in the other Local Health Authorities in the Emilia-Romagna region. If similar data are available, these models can be applied in other Italian regions and other countries investing in organization similar to the Patient Centered Medical Home.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1 Legend:

Figure 1. Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups.

Contributorship statement: DZL, RG, VM, and JSG were responsible for the conceptualization of this project. MR, JM, and ML were responsible for creation of the datasets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form and declare grants from the Regione Emilia-Romagna (Louis, Robeson, McAna, Maio, Keith, Liu, Gonnella) during the conduct of the study. Louis and Gonnella declare personal fees from Truven Health Analytics, The National School of Public Health, Portugal, INSIEL Mercato.

Funding: This research was funded by the Agenzia Sanitaria e Sociale of the Emilia-Romagna Region, Italy. The resulting models are being used by the Local Health Authority of Parma, Italy and the Patient Centered Medical Homes located in the Parma Local Health Authority.

Data Sharing: No additional data available.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1 Saltman RB, Rico A, Boerma W, eds. Primary care in the driver's seat? Organizational reform in European primary care. New York, New York: World Health Organization European Observatory on Health Systems and Policies Series 2006.

2 Smith J, Goodwin N. Towards managed primary care: The role and experience of primary care organizations. Burlington, VT: Ashgate Publishing Company 2006.

3 Defining the PCMH. Agency for Healthcare Research and Quality Web site. http://www.pcmh.ahrq.gov/portal/server.pt/community/pcmh__home/1483/pcmh_defining_the_pcmh_v2 (accessed June 30, 2013).

4 Lo Scalzo A, Donatini A, Orzella l, et al. Italy: Health system review. *Health Syst Transit*. 2009;**11**(6):1-216.

5 The Emilia-Romagna Regional Health Service and the new welfare system: Facilities, expenditure, and activities as of 31.12.2010. Programs, agreements and organisational models. Regione Emilia-Romagna Assessorato Politiche per la Salute. Bologna, Italy 2012.

6 Gonnella JS, Hornbrook MC, Louis DZ. Staging of Disease: A Case-Mix Measurement. *JAMA* 1984;**251**:637–44.

7 Gonnella JS, Louis DZ, Gozum ME, et al, eds. Clinical Criteria for Disease Staging. 6th edition. Ann Arbor, MI: Truven Health Analytics 2012.

8 Louis DZ, F Taroni, R Melotti, et al. Increasing appropriateness of hospital admissions in the Emilia-Romagna region of Italy. *J Health Serv Res Policy* 2008;**13**(4):202-8.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

9 Slabaugh L, Maio V, Templin M, et al. Prevalence and risk of polypharmacy amongst elderly primary care patients in Regione Emilia-Romagna, Italy. *Drug Aging* 2010;**27**:1019-28.

10 Gagne JJ, Maio V, Rabinowitz C. Prevalence and Predictors of Potential Drug-Drug Interactions among Ambulatory Patients in Regione Emilia-Romagna, Italy. *J Clin Pharm Ther* 2008;**33**:141-51.

11 Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *J Clin Pharm Ther* 2010;**35**:219-29.

12 Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: A systematic review. Washington, DC: Department of Veterans Affairs US 2011.

13 Billings J, Georghiou T, Blunt I, et al. Choosing a model to predict hospital admission: An observational study of new variants of predictive models for case finding. *BMJ Open* 2013;**3**(8):e003352. doi: 10.1136/bmjopen-2013-003352 [published Online First: 26 August 2013].

14 Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care* 2013;**51**(4):368-73.

15 Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. *Med Care* 2008;**46**(7):726-31.

16 Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care* 2012;**50**(2):131-9.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

17 McAna JF, Crawfrod AG, Novinger BW, et al. A predictive model of hospitalization risk among diabled Medicaid enrollees. *Am J Manag Care* 2013;**19**(5):166-74.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Predicting Risk of Hospitalization or Death: A Retrospective -Population Based Analysis

Daniel Z. Louis[1], Mary Robeson[1], John McAna[2], Vittorio Maio[2], Scott W. Keith[3], Mengdan Liu[1], Joseph S. Gonnella[1], Roberto Grilli[4]

[1]Center for Research in Medical Education and Health Care, ~~Jefferson~~ Sidney Kimmel Medical College ~~of~~ at Thomas Jefferson University. 1015 Walnut Street, Suite 319. Philadelphia, PA 19107, USA

[2]Jefferson School of Population Health, Thomas Jefferson University. 901 Walnut Street, 10th Floor. Philadelphia, PA 19107, USA

[3]Department of Pharmacology & Experimental Therapeutics, Division of Biostatistics, Thomas Jefferson University. 1015 Chestnut Street, Suite M-100. Philadelphia, PA 19107, USA

[4]Agenzia Sanitaria e Sociale Regione Emilia-Romagna, Viale Aldo Moro, 21. Bologna, 40128 Italy.

Corresponding author: Daniel Z. Louis, ~~Jefferson~~ Sidney Kimmel Medical College, 1015 Walnut Street, Suite 319, Philadelphia, PA 19107. USA; Tel: 215-955-6959, daniel.louis@jefferson.edu

Copyright: The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (non-exclusive for Dr. Grilli as a government employee) on a worldwide basis to the BMJ publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license.

Ethical approval: This study was approved by the Institutional Review Board of Thomas Jefferson University as an expedited retrospective database/record review. The IRB granted a waiver of informed consent.

For peer review only

Abstract

Objectives: Develop predictive models using an administrative health care database that provide information for Patient Centered Medical Homes to proactively identify patients at risk of hospitalization for conditions that may be impacted through improved patient care.

Design: Retrospective health care utilization analysis with multivariate logistic regression models.

Data: A population-based longitudinal database of residents served by the Emilia-Romagna, Italy health service in the years 2004-2012 including demographic information and utilization of health services by 3,726,380 people age ≥ 18 years.

Outcome measures: Models designed to predict risk of hospitalization or death in ~~calendar year~~ 2012 for problems that are potentially avoidable were developed and evaluated using the area under the receiver operating curve C-statistic, in terms of their sensitivity, specificity, and positive predictive value, and for calibration to assess performance across levels of predicted risk.

Results: Among the 3,726,380 adult residents of Emilia-Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions or died in 2012 (3.6% hospitalized, 1.3% died). The C-statistic for ~~the model~~ predicting 2012 outcomes was 0.856. The model was well calibrated across categories of predicted risk. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Conclusions: We have developed a population-based model using a longitudinal administrative database that identifies the risk of hospitalization for residents of the Emilia-Romagna region with a level of performance as high as, or higher than, similar models. The results of this model, along with profiles of patients identified as high risk are being provided to the physicians and other health care professionals associated with the Patient Centered Medical Homes to aid in planning for care management and interventions that may reduce their patients' likelihood of a preventable, high-cost hospitalization.

Strengths and Limitations of this study:

- This study included the entire adult population of the Emilia-Romagna Region of Italy, over 3.7 million people.
- The study used an existing longitudinal administrative health care database with both the advantage of much lower cost than new data collection and the disadvantage of potential errors in administrative data.
- The results of the study are being used to assist in the development of newly formed Patient Centered Medical Homes.

Keywords: hospitalization, risk, medical home, patient-centered care

Word Count: 3,952~~861~~ (excluding title page, abstract, references, figure 1, ~~and~~ table1 and table 2~~s~~.)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

The predominant healthcare delivery system, which has been a passive model, reacting to patients' problems, is shifting to a more proactive model designed to take the initiative in providing care for an increasingly older population that has a greater prevalence of chronic conditions, often with multiple medical and social needs. These changes are driving the reorganization of the primary care system, emphasizing coordination and cooperation among healthcare professionals.[1,2,3] Among the approaches to addressing this need has been the establishment of ~~primary care organizations incorporating integrated teams of physicians and other healthcare professionals that "seek to increase the influence of primary care professionals, and in particular general practitioners (GPs), in health planning and resource allocation." Prominent among these new models of primary care is the~~ Patient Centered Medical Homes, ~~an~~ organizations in which ~~a~~ teams of healthcare providers ~~is~~ are engaged in delivering comprehensive, coordinated, patient-centered care to patient defined populations.

Primary care has a central role in the Italian National Health Service (NHS). Twenty one regional governments are responsible for ensuring the delivery of a health benefits package through a network of geographically defined, population-based Local Health Authorities. Primary care physicians work for these authorities as independent contractors and act as "gatekeepers" for specialty and other referral services for their patients.[4]

With the belief that a strong primary care system is conducive to improving population health, the NHS initiated reforms that encouraged primary care physicians to organize into collaborative arrangements. To this end, the Regione Emilia-Romagna (RER), a large northern region with a population of about 4.5 million, has recently launched a plan in its 11 Local Health Authorities to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

establish Patient-Centered Medical Homes to better coordinate patient care and help patients avoid unnecessary hospitalizations.

~~The identification of those patients who would benefit most from outreach efforts is fundamental to achieving these goals of promoting and practicing population health in Patient-Centered Medical Homes. To accomplish this, predictive models and risk stratification tools are needed to identify patients at risk of a worsening of their health status. According to Knutson and Bella,[5] "Predictive models are data-driven, decision-support tools that estimate an individual's future potential health care costs and/or opportunities for care management." A good model will identify as many of these patients as possible while excluding those for whom these interventions would be unnecessary or ineffective.~~

The identification of those patients who would benefit most from outreach efforts is fundamental to achieving these goals of promoting and practicing population health in Patient-Centered Medical Homes. The RER has established three objectives for this project: 1) develop predictive models to identify patients at high risk of hospitalization or death, 2) create "risk of hospitalization" patient profiles that provide information about their high-risk patients to the general practitioners in the newly formed Patient-Centered Medical Homes, and 3) assess the extent to which these models and reports provide additional information useful in the identification of patients who may benefit from case management or disease management.

This paper will address the first of the three goals. We describe the development of a predictive model using the RER's regional longitudinal administrative health care database to help identify patients who are most at risk of hospitalization for conditions that may be impacted through improved patient care. This model will then be used to inform the providers associated with the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Patient -Centered Medical Homes and aid in their planning for care management and interventions that can reduce their patients' likelihood of a preventable, high-cost hospitalization.

Methods

Study Data and Study Population

The model was developed using the population-based longitudinal health care database of the residents served by the RER Health Service in the years 2004 through 2012.[6] This administrative database includes demographic information for all residents (gender, birth and death dates, location of current residence, and primary care physician), hospital discharge abstract data (ICD-9-CM diagnosis and procedure codes, and admission and discharge dates), emergency room utilization information, outpatient pharmacy data at the individual prescription level, specialty care (laboratory, diagnostics, therapeutic procedures, rehabilitation, and specialist visits), home health data, and information on each primary care physician in the region. Each patient has an anonymous identifier assigned by the RER so that an individual's utilization can be tracked over time without jeopardizing patient privacy.

The study population consisted of all residents of the RER who were at least 18 years of age and still alive as of 31st December 2011. Health care utilization data from 2011 and history variables using data from 2004 through 2010 were used to predict outcomes in 2012.

Dependent Variable

The dependent variable was defined as the occurrence of a hospitalization for problems that are

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

potentially avoidable, or whose progression may have been avoided or delayed through appropriate patient care, or the death of the individual, either in or out of the hospital, for any reason in 2012. We included deaths in the dependent variable since we believe that, for example, a patient with coronary artery disease who dies secondary to an acute myocardial infarction, should be included in the dependent variable even if the death is out of the hospital. We decided to not limit the hospitalization to emergency admissions, since a planned admission may also be an indicator of a worsening medical problem. In order to operationally define the dependent variable, we (authors JSG and DZL) reviewed the Disease Staging[7, 8] primary diagnostic category and severity stage of all day and inpatient hospital admissions (for adults age 18+) in RER for one year, to select those admissions that should be included in the dependent variable.

~~Deliveries~~Admissions to deliver a baby, admission for dental diseases or admissions for vague signs or symptoms with no identified etiology were excluded. Admissions for problems that are not predictable/preventable were excluded while those where screening may identify problems that can potentially be treated to avoid progression were included. For example, admissions for stage 1, chronic cholecystitis or cholelithiasis were excluded, but admissions for advanced stage 2 or 3 complications such as ascending cholangitis or pancreatitis were included.

We felt that inclusion of hospitalization for cancer in the dependent variable should depend on the ability to either prevent or avoid progression of the disease. We therefore included colon cancer and cervical cancer in the definition because they are potentially preventable but excluded all other cancers where prevention/prediction is not currently possible.

Inclusion of injuries, burns, or toxic reaction to prescription or non-prescription drugs would

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ideally be based on the cause of these problems. Since the etiology of these problems is typically not available in the administrative data being used in this project, we made the decision to include or exclude based on our subjective judgment of the likelihood of preventability. For example, adverse drug reactions were included but burns were excluded from the definition of the dependent variable.

There is no obvious medical reason for a hospital admission for patients with stage 1 diabetes mellitus or stage 1 essential hypertension without complications. These problems are typically treatable in the outpatient setting. A hospitalization implies a potential problem in the care of these patients, so we decided to include these admissions as a part of the dependent variable.[9]

Independent Variables

A broad range of candidate predictor variables was developed taking advantage of the RER administrative data. The independent variables used for modeling were defined from the RER administrative data for the years 2004 through 2011. Demographic data included patient age, sex, and geographic location of residence. We developed a mapping to broad disease categories defined primarily in terms of the affected body system from home health care data, pharmacy data, and hospital discharge abstract data. (See Appendix 1.)

For those patients who had been hospitalized, more specific diagnostic data were available. We reviewed the classification of patients hospitalized historically using the Disease Staging diagnostic category and disease severity stages.[7,8] Based on the frequencies specific diagnostic

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

category/stage predictor variables were defined for either specific stages of frequent diseases, or by combinations across similar categories. Predictor variables were defined based on the number of emergency room visits using the RER classifications system for the urgency of the visit.

Pharmacy data were used to identify polypharmacy[10] (defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days), potential drug-drug interactions (DDI)[11] and potentially inappropriate medication use in patients[12] 65 years and older. Since cardiovascular disease is highly prevalent, we reviewed the use of cardiovascular drugs and created a variable for each of the following 11 classes of drugs (oral anti-coagulants, beta-blockers, angiotensin converting enzyme / angiotensin II receptor blockers, anti-platelets, calcium channel blockers, anti-arrhythmics, digitalis glycosides, nitrates, diuretics, ~~alfa~~alpha-blockers, statins) to account for the complexity of therapeutic regimen at the patient level.

To take advantage of the fact that the RER database includes multiple years of data, we created history variables using the utilization for each year of data available. Since we were working with the 2011 data to predict hospitalization or death in 2012, we created history variables based on 2004-2010 data. This set included 83 of the diagnostic category/stage variables as well as 11 variables based on pharmacy utilization such as exposure to polypharmacy and use of cardiovascular drugs. If the individual had a history of a disease in any of the years from 2004 to 2010 they were flagged as having a history of that disease and this was used as a potential predictor variable.

Modeling

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Logistic regression models were used to estimate predicted probabilities for the occurrence of an inpatient hospital stay for the selected conditions or death for individual patients. Risk of hospitalization or death, and the variables that relate to those risks are highly dependent on age and gender. Regression models were fit in each of 14 gender and age strata using SAS Version 9.2 (SAS Institute, Cary NC). A stepwise process with relaxed covariate entry and retention criteria (inclusion p-value <=0.8, retention <=0.5) was used. At each step in this process, an attempt is made to remove any unimportant variables from the model before adding a potentially important variable. Each addition or deletion of a variable to or from a potential model is a separate step and, at each step, a new model is fitted. This process results in a reduced, but robust set of independent variables that predict outcome or that might have importance as adjustment terms for the model in each age/gender stratum.

Evaluation of the Models

The predictive accuracy of the modelling was evaluated using C-statistics (the area under the receiver operating characteristics curve), along with three measures traditionally used with clinical screening tools: sensitivity, specificity and positive predictive value.

C-statistics were used to evaluate the models in two ways. The first evaluation consisted of fitting the model developed using utilization and demographic data from 2011, along with historical variables based on 2004-2010 data, and outcomes (hospitalization or death) from 2012 and then computing a C-statistic to evaluate how the models performed at predicting those outcomes on which the models were conditioned. However, this evaluation is not consistent with evaluating how the data are used in practice. In practice, we have current predictor information,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

but the outcomes have not been realized. To better estimate how the models are likely to perform in this setting, we fit models to outcomes data up to a year prior to the most current available (e.g., 2011 outcomes modelled with predictors from 2010, along with historical variables based on 2004-2009 data). We then computed a C-statistic for projections made on the risk of hospitalization or death outcomes (in 2012) using the next year's predictor information (in 2011). This way, the models are forced to make projections into the future, but we have the actual observed outcomes data to evaluate the modelling process as it would be used in practice. The resulting C-statistics obtained from these two model runs were compared.

In order to evaluate the performance of the model across different risk thresholds we classified predicted risk scores. "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%. These risk thresholds were selected after consultation with physicians practicing in the medical homes to yield a total of about 10% of the 1,500 patients enrolled with a typical primary care physician.

Results

Among the 3,726,380 adult residents of Emilia Romagna at the end of 2011, 449,163 (12.1%) were hospitalized in 2012; 4.2% were hospitalized for the selected conditions defined earlier or died in 2012 (3.6% hospitalized, 1.3% died).

Table 1 shows the distribution of the demographics (age and gender), number of chronic conditions, body systems impacted by the selected chronic conditions, polypharmacy and inappropriate prescribing among the eligible RER residents, as of December 31, 2011. T~~The~~

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

table 1 also compares these characteristics of the total adult population of the region to the subgroups of the population classified in the "very high risk" and "high risk" categories. Based on the model results, 114,255 individuals were identified as having a predicted risk of hospitalization or death in 2012 of ≥ 25% and classified as "very high risk." An additional 134,610 individuals had a predicted risk of hospitalization or death in 2012 of 15-24% and were classified as "high risk."

Table 1. Demographic and clinical characteristics of the Regione Emilia-Romagna population, overall and by risk category

	Total Population*		Very High Risk**		High Risk**	
	3,726,380		114,255		134,610	
Gender	N.	%	N.	%	N.	%
Male	1,788,048	48.0%	54,357	47.6%	61,803	45.9%
Female	1,938,332	52.0%	59,898	52.4%	72,807	54.1%
Age groups						
18-24	258,338	6.9%	76	0.1%	105	0.1%
25-34	499,786	13.4%	302	0.3%	391	0.3%
35-44	732,626	19.7%	1,137	1.0%	1,198	0.9%
45-54	676,047	18.1%	2,612	2.3%	2,485	1.8%
55-64	550,689	14.8%	5,391	4.7%	5,287	3.9%
65-74	482,346	12.9%	13,154	11.5%	14,471	10.8%
74-85	364,369	9.8%	33,430	29.3%	44,857	33.3%
85+	162,179	4.4%	58,153	50.9%	65,816	48.9%
Number of Chronic Conditions						
0-1	2,775,888	74.5%	8,176	7.2%	24,618	18.3%
2 or more	950,492	25.5%	106,079	92.8%	109,992	81.7%
5 or more	99,337	2.7%	45,445	39.8%	20,576	15.3%
Selected Conditions/Body Systems						
Cancer	99,328	2.7%	23,872	20.9%	14,305	10.6%
Cardiovascular	967,796	26.0%	96,157	84.2%	103,749	77.1%
Male Genitourinary#	130,609	7.3%	14,616	26.9%	16,776	27.1%
Ear, Nose, Throat	5,364	0.1%	240	0.2%	242	0.2%
Endocrine	429,528	11.5%	40,653	35.6%	37,471	27.8%

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Eye	114,117	3.1%	9,558	8.4%	13,478	10.0%
Gastrointestinal	580,946	15.6%	74,718	65.4%	66,305	49.3%
Gynecologic##	21,806	1.1%	333	0.6%	405	0.6%
Hematologic	45,022	1.2%	15,353	13.4%	6,591	4.9%
Hepatobiliary	24,785	0.7%	6,477	5.7%	3,306	2.5%
Immunologic	3,281	0.1%	464	0.4%	273	0.2%
Infectious Disease	4,723	0.1%	2,207	1.9%	727	0.5%
Musculoskeletal	419,184	11.2%	43,436	38.0%	41,000	30.5%
Neurologic	173,751	4.7%	34,494	30.2%	24,838	18.5%
Psychological	291,308	7.8%	43,387	38.0%	33,715	25.0%
Respiratory	176,830	4.7%	39,082	34.2%	21,763	16.2%
Skin	28,339	0.8%	7,645	6.7%	3,008	2.2%
Urogenital	37,728	1.0%	16,501	14.4%	5,740	4.3%
Polypharmacy^	609,278	16.4%	92,153	80.7%	92,156	68.5%
Any potentially inappropriate medications (age 65 years or older)^^	257,033	25.5%	51,055	48.7%	49,003	39.2%

* Adults (age 18 or older) and alive at 31 December 2011.

** "Very high risk" was defined as patients with a predicted risk of hospitalization or death in the following year of ≥ 25% while "high risk" was defined as patients with a predicted risk of hospitalization of 15-24%.

Men only.

Women only.

^ Polypharmacy is defined as the simultaneous use of five or more active ingredients for at least 15 consecutive days.

^^ The list of potentially inappropriate medications can be found in: Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *Journal of Clinical Pharmacy and Therapeutics* 2010;35:219-229.

There was little difference across the risk categories by gender. Age distributions for the "very high risk" and "high risk" groups were shifted more towards the older age groups than those in the overall study population. Residents age 85 or older represented about 4.5% of the RER population, but about 50% of the "very high" and "high" predicted risk groups. More than 75% of the residents over age 85 were classified as "very high" or "high" risk. However, age alone was not sufficient to ~~their~~ predict their risk. For example, residents between 75 and 84 years of

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

age made up 23% of the "very high" risk group and 41% of the "high" risk group, but over 85% of the residents in this age category had neither "very high" nor "high" predicted risk.

Across age and gender strata, demographics and heath care utilization experience in 2011 were the most commonly used independent variables for predicting hospitalization or death in 2012. Selected history variables flagging chronic problems such as cardiovascular disease, diabetes mellitus and chronic renal failure and a history of prescriptions for cardiovascular medications and polypharmacy were also significant predictors.

The residents in the two higher risk groups were more likely than others to have multiple chronic diseases and to experience polypharmacy and inappropriate medication use. The residents identified as "very high risk" or "high risk" by the model also showed a number of striking differences from others in terms of the occurrence of some of the most prevalent health conditions by type and body system. Although cardiovascular conditions were not uncommon in the total adult population (26.0%), they were far more common among those classified as "very high risk" and "high risk" (84.2% and 77.1%, respectively). Similarly, gastrointestinal conditions affected 15.6% of the total population, but were diagnosed in 65.4% of the "very high risk" and 49.3% of the "high risk" patients. Cancer occurred in 2.7% of the total population, but 20.9% of the "very high risk" and 10.6% of the "high risk" patients had a cancer diagnosis. Mental health problems were identified in 7.8% of the adult population, but in 34.2% of the "very high risk" and 25.0% of the "high risk" patients.

The C-statistic for the model of 2012 outcomes developed using 2011 predictors and the C-statistic based on the parameters from the model of 2011 outcomes regressed on 2010 predictors applied to the 2011 predictors and 2012 outcomes were very similar (0.856 and 0.853,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

respectively). These results suggest that the relationship between predictors and risk of hospitalization changed little in one year and that model parameters developed in a prior year can be used reliably with the most current year's data to predict unknown outcomes in the next year with only a minimal loss in performance in this population.

Table 2 shows the sensitivity, specificity, positive predictive value and number of true positives for the model at the two selected cut-off points. The sensitivity (percentage of patients actually hospitalized who had been identified by the model as having a predicted risk higher than the cut-off point) was 29.8% for those with the "very high" risk scores. This percentage represents 46,950 of the 157,550 residents of the region who were hospitalized for a selected condition or died in 2012. If we modify the risk score threshold to include individuals with a predicted risk of hospitalization for selected conditions or death of $\geq$ 15% (i.e., both the "very high risk" and the "high risk" patients) the sensitivity is .471. The true negative rate (specificity) is very high for both risk thresholds (.981 and .951, respectively).

Table 2. Performance of the "Risk of Hospitalization" model for residents identified as "Very High Risk" and "High or Very High Risk"

	Cut-off points for comparison	
Measure	**"Very high risk"***	**"Very high risk"* + "High risk"****
Sensitivity#	0.298	0.471
Specificity##	0.981	0.951
Positive Predictive Value^	0.411	0.298
True positives^^	46,950	74,196

* "Very high risk" is defined as patients with a predicted risk of hospitalization of $\geq$ 25%.
** "Very high risk" + "High risk", is defined as patients with a predicted risk of hospitalization of $\geq$15%.
# Sensitivity is defined as the proportion of those hospitalized who were predicted to be hospitalized (true positive rate).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Specificity is the proportion of those not hospitalized who were not predicted to be hospitalized (true negative rate).
^ Positive Predictive Value is the proportion of those predicted to be hospitalized who were actually hospitalized.
^^ True positives are the number of residents who were predicted to be at risk for hospitalization at the predicted risk threshold and were actually hospitalized.

The model appears to be well calibrated across levels of risk. ~~The~~ Figure 1 depicts the RER population divided into groups by deciles of predicted risk of hospitalization or death from the models. The observed prevalence of hospitalization or death is compared to the average predicted risk among individuals in each of the ten predicted risk groups. For example, the overall rate of hospitalization for the selected conditions or death in 2012 was 4.2%. For those patients in the highest predicted risk decile group, the average predicted risk was 23.9% and the actual prevalence of hospitalization or death was 24.2%. (Regression coefficients and significance levels of independent variables for models for each ~~ot~~of 14 age and gender strata are displayed in Appendix 2.)

[Insert Figure 1 about here]

Discussion

We have developed a population-based model that identifies the risk of hospitalization for all adult RER residents and does so with a level of performance (c=0.85) as high as, or higher than, similar models. In addition, we believe that the definition of the dependent variable chosen for our models increases the probability that they are identifying patients who risk can potentially be improved by appropriate care. A systematic review by Kansagara[13] of models designed to predict readmissions, showed C-statistic results in the range of 0.55 to 0.83. Recent work by

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Billings et al[14] to develop models predictive of emergent admissions in the UK had results ranging from 0.73 to 0.78. Li Wang, et al. (2013),[15] using information available through the United States Department of Veterans Affairs ~~US Veteran's Administration~~ that also included lab data, demonstrated C~~c~~-statistics of 0.81 and 0.79, respectively, for their models of 90-day or 12-month hospitalization or death outcomes. At a predicted risk of ≥25% our model had a Positive Predictive Value (PPV) of .411. Billings et al[14] reported a PPV of .417 at a risk threshold of 30. There is a trade-off in using our model, or any predictive model, between the threshold for follow-up and predictive accuracy. A lower risk threshold would identify more patients but with a lower prevalence of hospitalization or death.

Although previous studies have developed models predictive of hospital care, these models fall short of the needs of the Patient Centered Medical Homes being implemented in RER. Typically, these models have focused on specific age groups,[16] conditions, or types of admissions, such as emergent[14] or unplanned admissions or rehospitalizations, or health insurance plans in the United States, including private insurance plans, Medicare and Medicaid plans.[17,18] The models we have developed are applied to the entire adult population of RER. They use existing administrative data, which makes them cost effective to apply.

Patient Centered Medical Homes, including those instituted in RER, are responsible for addressing the needs of their population and making the best use of their finite resources to accomplish this. Preventing unnecessary admissions could improve both the quality of care and health status of the enrolled population, and result in a substantial savings. To accomplish this, predictive models and risk stratification tools such as those developed for this project are needed

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

to identify patients at risk of preventable admissions and provide information that can be used by the medical homes to help manage care.

There are some limitations to the model. The model is developed from administrative data. Administrative data are collected for reimbursement and tracking utilization and not for medical research. They lack the clinical specificity that would be desirable in assessing an individual's medical problems. Patients with medical problems who have not used the services included in the database cannot be identified. While the hospital discharge abstract data do include diagnostic information coded using ICD-9-CM, no similar data are available for outpatient encounters in the RER database. The mortality data available for this project did not include information about cause of death. Therefore, some proportion of patients whose death was not predictable were included, limiting model performance. In addition, our models use prior utilization among the predictor variables. With the administrative data we cannot distinguish between appropriate and inappropriate treatment which may bias our results.

Despite the limitations of administrative data, they have many advantages for this project: they are readily available, relatively inexpensive to analyze and cover large populations over many years. They are ideal for uncovering patterns of care. If information from the medical records is needed, the results of these analyses can then be supplemented by focused clinical reviews at the local level. Also, The RER has a system in place to monitor the quality of diagnosis and procedure coding in their hospital discharge abstract data. Controls at both the hospital and regional level assess the validity of coding and the consistency of codes assigned such as congruity between sex, age and diagnosis and between diagnosis and procedure. The existence of the RER administrative database made it feasible to develop the models described in this

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

article at relatively low cost and to update the models over time without additional data collection that others have found necessary.[14]

Currently, these risk scores are being integrated with other information in profiles of high-risk patients furnished to providers in 12 newly formed medical homes, including 83 primary care physicians serving a total of about 100,000 patients, in the Parma Local Health Authority located in RER. Along with the risk scores, this information includes data about previous hospitalizations, use of referrals, medications, long-term care and home care services, and a number of process-like quality indicators for diabetic and cardiovascular patients, and for appropriate medication use in older patients.

We believe that the Italian health care system offers a number of advantages in the goal of reducing potentially avoidable hospitalization. Every Italian must enroll with a primary care physician. The population is quite stable with modest levels of change of residence or change of primary care physician. Every Italian is entitled to health care with little or no cost at the point of service. There is no problem with loss of, or change in, insurance coverage. Primary care physicians are primarily paid on a per capita basis. But the Emilia-Romagna region has the ability to negotiate incentive payments designed to address and monitor improvements in medical management such as that addressed in our study.

Of course, model results need to lead to ~~an~~ effective interventions to have a positive impact on patient care. To this end, we are working with the physicians, nurses~~, and~~, and other health care professionals as well as the administration of the newly formed Medical Homes in Parma to assist them in understanding how to use the results of these models and in developing potentially effective interventions. The individual profiles of high risk patients provided to the health care

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

team in the Medical Homes allow them to trigger specific actions such as inviting patients to enroll in disease management programs for chronic problems such as heart failure, chronic obstructive pulmonary disease, or diabetes mellitus, activating home health assistance, initiating a medication review, or recommending that the patient come in for an office visit. An evaluation of the use and usefulness of the profiles and intervention is under way.

In summary, these models provide a means of identifying patients at high risk for hospitalization. The risk predictions, in conjunction with the risk profile, show promise as a useful organizational tool for the regional Patient Centered Medical Homes to develop and implement proactive case management and disease management programs. The RER is reviewing the results of the Parma Local Health Authority pilot project of the profiles. Once their usefulness has been further evaluated, their use will be expanded to other Medical Homes in development in the other Local Health Authorities in the Emilia-Romagna region. If similar data are available, these models can be applied in other Italian regions and other countries investing in organization similar to the Patient Centered Medical Home.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1 Legend:

Figure 1. Model calibration: Predicted risk and observed prevalence of hospitalization or death in 2012 by predicted risk decile groups.

Contributorship statement: DZL, RG, VM, and JSG were responsible for the conceptualization of this project. MR, JM, and ML were responsible for creation of the datasets used in this project. DZL,VM, MR, and JSG were responsible for the definition of analytical variables. SWK, MR, ML and JM were responsible for modeling and statistical analysis. DZL managed the research team. RG and JSG advised on the analyses and results. All authors contributed to the preparation of the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form and declare grants from the Regione Emilia-Romagna (Louis, Robeson, McAna, Maio, Keith, Liu, Gonnella) during the conduct of the study. Louis and Gonnella declare personal fees from Truven Health Analytics, The National School of Public Health, Portugal, INSIEL Mercato.

Funding: This research was funded by the Agenzia Sanitaria e Sociale of the Emilia-Romagna Region, Italy. The resulting models are being used by the Local Health Authority of Parma, Italy and the Patient Centered Medical Homes located in the Parma Local Health Authority.

Data Sharing: No additional data available.

1 Saltman RB, Rico A, Boerma W, eds. Primary care in the driver's seat? Organizational reform in European primary care. New York, New York: World Health Organization European Observatory on Health Systems and Policies Series 2006.

2 Smith J, Goodwin N. Towards managed primary care: The role and experience of primary care organizations. Burlington, VT: Ashgate Publishing Company 2006.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

3 Defining the PCMH. Agency for Healthcare Research and Quality Web site. http://www.pcmh.ahrq.gov/portal/server.pt/community/pcmh__home/1483/pcmh_defining_the_pcmh_v2 (accessed June 30, 2013).

4 Lo Scalzo A, Donatini A, Orzella l, et al. Italy: Health system review. *Health Syst Transit*. 2009;**11**(6):1-216.

~~5 Knutson D, Bella M, Llanos K. Predictive Modeling: A Guide for State Medicaid Purchasers. Center for Health Care Strategies, Inc. 2009.~~

6 The Emilia-Romagna Regional Health Service and the new welfare system: Facilities, expenditure, and activities as of 31.12.2010. Programs, agreements and organisational models. Regione Emilia-Romagna Assessorato Politiche per la Salute. Bologna, Italy 2012.

7 Gonnella JS, Hornbrook MC, Louis DZ. Staging of Disease: A Case-Mix Measurement. *JAMA* 1984;**251**:637–44.

8 Gonnella JS, Louis DZ, Gozum ME, et al, eds. Clinical Criteria for Disease Staging. 6th edition. Ann Arbor, MI: Truven Health Analytics 2012.

9 Louis DZ, F Taroni, R Melotti, et al. Increasing appropriateness of hospital admissions in the Emilia-Romagna region of Italy. *J Health Serv Res Policy* 2008;**13**(4):202-8.

10 Slabaugh L, Maio V, Templin M, et al. Prevalence and risk of polypharmacy amongst elderly primary care patients in Regione Emilia-Romagna, Italy. *Drug Aging* 2010;**27**:1019-28.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

11 Gagne JJ, Maio V, Rabinowitz C. Prevalence and Predictors of Potential Drug-Drug Interactions among Ambulatory Patients in Regione Emilia-Romagna, Italy. *J Clin Pharm Ther* 2008;**33**:141-51.

12 Maio V, Del Canale S, Abouzaid S. Using Explicit Criteria to Evaluate the Quality of Prescribing in Elderly Italian Outpatients: A Cohort Study. *J Clin Pharm Ther* 2010;**35**:219-29.

13 Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: A systematic review. Washington, DC: Department of Veterans Affairs US 2011.

14 Billings J, Georghiou T, Blunt I, et al. Choosing a model to predict hospital admission: An observational study of new variants of predictive models for case finding. *BMJ Open* 2013;**3**(8):e003352. doi: 10.1136/bmjopen-2013-003352 [published Online First: 26 August 2013].

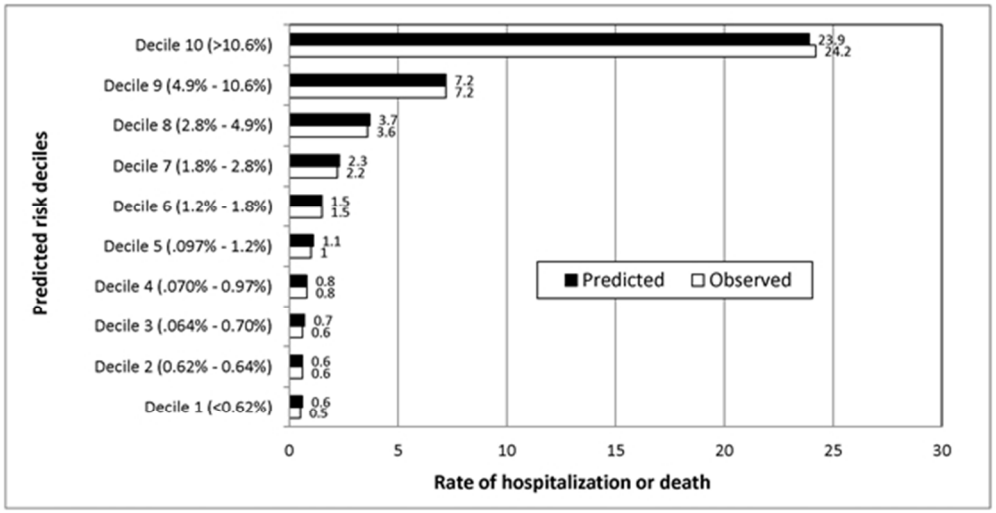
15 Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care* 2013;**51**(4):368-73.

16 Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. *Med Care* 2008;**46**(7):726-31.

17 Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care* 2012;**50**(2):131-9.

18 McAna JF, Crawfrod AG, Novinger BW, et al. A predictive model of hospitalization risk among diabled Medicaid enrollees. *Am J Manag Care* 2013;**19**(5):166-74.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



54x27mm (300 x 300 DPI)

For peer review only

Appendix 1: Mapping to Body System or Etiology Groups

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Cancer	Neoplasm, Malignant: Cardiovascular, Hypopharynx, Oral Cavity, Oropharynx, Salivary Glands and Mandible, Other Endocrine System, Larynx, Glottis, Larynx, Subglottic, Larynx, Supraglottic, Nasopharyngeal, Sinuses, Ocular Melanoma, Other Eye and Periocular, Colon and Rectum, Esophagus, Small Bowel, Stomach, Other Gastrointestinal System, Bladder, Urinary, Kidneys, Other Genitourinary System, Breast (Female), Cervix Uteri, Endometrium, Ovaries, Vagina, Vulva, Other Female Genitalia, Hodgkin's Lymphoma, Multiple Myeloma, Mastocytosis, Pancreas, Other Hepatobiliary Tract, Breast (Male), Penile, Prostate, Testicular, Primary Bone, Waldenstrom's Macroglobulinemia, Nonspecific Sites, Unspecified Primary Site, Lungs, Bronchi, or Mediastinum, Hodgkin's Disease Lymphocytic Depletion, Hodgkin's Disease Lymphocytic Predominance, Hodgkin's Disease Mixed Cellularity, Hodgkin's Disease Nodular Sclerosis, Lymphatic and Hematopoietic (Other Types), Lymphoma, Cutaneous T Cell (Mycosis Fungoides), Lymphoma (Diffuse Mixed Small and Large Cell), Lymphoma (Diffuse Large Cell), Lymphoma (Follicular Predominantly Large Cell), Lymphoma (Histiocytic Cell), Lymphoma (Lymphoblastic), Other Respiratory System, Carcinoma (Basal Cell stage 2/3), Carcinoma (Squamous Cell), Melanoma, Other Skin and Soft	Antineoplastics 5HT3 Antagonists	2005-2009: Visits prescribed due to the presence of cancer. Beginning in 2010, the following ICD-9-CM codes were in the record: 140-208, 235-239, V10, V16	Visits prescribed for radiation therapy, or for Injection or infusion of chemotherapeutic Substances for cancer treatment

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Tissue Neoplasm: Pheochromocytoma, Eyelid, Central Nervous System, Lymphatic or Hematopoietic Leukemia: Acute Lymphocytic, Acute Nonlymphocytic, Chronic Lymphocytic, Chronic Myelogenous, Other Types Encounter for: Chemotherapy, Radiation Therapy ICD-9-CM Procedure codes: 99.25, 99.28, 00.10,00.15,92.2x			
Cardiovascular	Aneurysm: Abdominal, Thoracic Anomaly: Patent Ductus Arteriosus, Atrial Septal Defect, Atrioventricular Defects, Coarctation of the Aorta, Other Congenital Heart Disease, Pulmonary Valve Stenosis, Tetralogy of Fallot (stage 3), Transposition of the Great Arteries, Ventricular Septal Defects, Other Circulatory System Aortic: Regurgitation, Stenosis Mitral: Regurgitation, Stenosis Neoplasm: Benign of the Cardiovascular System Arrhythmias, Cardiomyopathies, Conduction Disorders, Congestive Heart Failure, Coronary Artery Disease Prior Coronary Revascularization, Coronary Artery Disease w/o Prior, Coronary Revascularization, Essential Hypertension, Infective, Endocarditis, Pericarditis: Chronic (stage 2/3), Viral or Traumatic (stage 2/3) Periarteritis Nodosa, Raynaud's Disease, Thromboangiitis,	Oral anti-coagulants beta-blockers ACE/ARB anti-platelets calcium channel blockers anti-arrhythmics digitalis glycosides nitrates diuretics alfa-blockers statins	2005-2009: Visits prescribed due to the presence of Congestive Heart Failure or not-well defined cardiopathy, and other diseases of cardiovascular system. Beginning in 2010, the following ICD-9-CM-CM codes were in the record: 390-454,456-459	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Obliterans, Thrombophlebitis, Tibial, lliac, Femoral, or Popliteal Artery Disease, Varicose Veins of Lower Extremities, Secondary Hypertension, Budd Chiari Syndrome, Rheumatic Fever (stage 2/3) Vasculitis Other: Atherosclerosis, Cardiac Conditions, Cardiovascular Symptoms, Circulatory Disorders, Diseases of Arteries, Diseases of Veins, Disorders of Pulmonary Circulation, Lymphatic Disorders			
Endocrine	Adrenal Insufficiency, Cushing's Syndrome, Diabetes insipidus, Diabetes Mellitus Type 1, Diabetes Mellitus Type 2 and Hyperglycemic States,Hyperthyroidism, Hypoglycemia, Hypothyroidism, Monotropic Hormone Deficiency, Primary Amyloidosis, Thyroiditis, Klinefelter's Syndrome, Turner's or Noonan's Syndrome, Obesity Goiter: Nontoxic or Euthyroid (stage 2/3) Neoplasm, Benign: Acromegaly, Adenoma, Parathyroid, Hyperparathyroidism, Primary Hyperaldosteronism, Other Endocrine System Neoplasm, Malignant:Thyroid Other: Endocrine Disorders, Electrolyte Disorders, Nutritional and Metabolic Disorders Anomaly: Adrenal Hyperplasia	Insulins biguanides sulfonylureas vasopressin thyroid replacement antithyroid agents	2005-2009: Visits prescribed due to the presence of diabetes mellitus Beginning in 2010, the following ICD-9-CM codes were in the record: 240-278	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Ear, Nose, Throat	Diseases of Salivary Gland, Incl. Parotitis or Benign Tumors, Other Disorders of Oral Cavity (stage 2), Cholesteatoma, Meniere's Disease, Otitis Media, Sinusitis Hearing Loss due to: Acoustic Trauma, Otosclerosis Neoplasm, Benign: Larynx, Sinuses, Oral Cavity and Pharyngeal Structures Pharyngitis: Non-Streptococcal (stage 2)			
Eye	Cataract, Conjunctivitis: Bacterial, Contusion or Ruptured Globe, Dacryostenosis or Dacryocystitis, Detachment of the Retina, Ectropion or Entropion (Abnormal Lower Lid Position), Endophthalmitis, Foreign Body: Orbit, Fracture: Orbit, Blow-Out, Fungal Infection of the Eye, Glaucoma, Hypovitaminosis A, Laceration: Cornea, Macular Degeneration, Orbital Infection, Prematurity: Retinopathy, Ptosis of Upper Lid, Retrobulbar Orbital Hemorrhage, Trachoma, Other Eye Disorders Injury or Laceration: Eyelid, Periocular, Cornea, Conjunctiv Injury: Eyes, Nonionizing Radiation Keratitis: Acanthamoeba, Bacterial Neoplasm, Benign: Eye	Sympaticomimetic agents parasympaticomimetic agents anhydrase inhibitors ophthalmic beta blockers		
Gastrointestinal	Anorectal Suppuration, Celiac Disease, Clostridium difficile Colitis, Crohn's Disease, Diverticular Disease, Food Poisoning: Other Organisms (stage 3), Functional Digestive Disorders, Gastritis, Hemorrhoids, Hernia (External), Hernia (Hiatal or Reflux Esophagitis), Intussusception (stage 2), Irritable Bowel Syndrome, Gastroenteritis	Intestinal corticosteroids agents H2 antagonists prostaglandins proton pump inhibitors	2005-2009: Visits prescribed due to the presence of Gastrointestinal Diseases Beginning in 2010,	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Neoplasm, Benign: Adenomatous Polyps, Colon, Small Bowel, Other Gastrointestinal System Peptic Ulcer Disease, Salmonellosis (stage 3), Ulcerative Colitis, Vascular Insufficiency of the Bowels, Complications of Gastrointestinal Treatment, Gastroenteritis (stage 2/3) Other Diseases of Esophagus, Stomach, and Duodenum Other Gastrointestinal Disorders, Other Gastrointestinal Infections (stage 2), Other Gastrointestinal or Abdominal Symptoms Anomaly: Congenital Megacolon, Other Digestive or Hepatobiliary System Burns, Chemical: Esophagus, Stomach, or Small Intestine, Laceration: Esophagus		the following ICD-9-CM codes were in the record: 520-539,550-579	
Genitourinary	Bladder Disorders, Calculus of the Urinary Tract, Glomerulonephritis, Acute, Injury: Urinary Tract, Nephrotic Syndrome (stage 2/3), Renal Failure (stage 2/3), Urethritis, Urinary Tract Infections, Neoplasm, Benign: Urinary Tract, Other Disorders of Kidney or Ureter, Other Urinary Symptoms, Encounter for Dialysis, Anomaly: Defects of Kidney, Defects of Lower Genitourinary Tract, Syphilis: Congenital	Agents for hyperkalemia and hyperphosphatemia	2005-2009: Visits prescribed due to the presence of renal failure and Other diseases of the genito-urinary system Beginning in 2010, the following ICD-9-CM codes were in the record: 580-629	Visits prescribed for dialysis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Gynecological and Obstetrics	Anomaly: External Female Genitalia, Anomaly: Uterus, Dysfunctional Uterine Bleeding, Endometriosis, Neoplasm, Benign: Ovary (stage 2), Pelvic Inflammatory Disease, Uterine Infection, Uterovaginal Prolapse, Vulvovaginitis, Other Disorders of Female Genital System			
Hematological	Agranulocytosis, Anemia: Aplastic, Acquired (stage 2/3), Folic Acid Deficiency, Hemolytic (stage 2/3), Iron Deficiency, Sickle Cell, Thalassemia, Vitamin B-12 Deficiency, Other Graft versus Host reaction, Hemophilia A or B, Polycythemia Vera, Other Disorders of Blood and Blood-Forming Organs	Iron vitamin B12 folic acids	Beginning in 2010, the following ICD-9-CM codes were in the record: 280-289	
Hepatobiliary	Cholecystitis and Cholelithiasis, Cirrhosis of the Liver (stage 2/3), Disorders of Bilirubin Excretion, Hepatitis A, Hepatitis B, Hepatitis C, Hepatitis D, Hepatitis E, Hepatitis G, Hepatitis (Chemical), Pancreatitis, Wilson's Disease, Neoplasm, Benign: Hepatobiliary System, Other Hepatobiliary and Pancreatic Disorders, Other Hepatobiliary Infections, Other Pancreatic Disorders	Interferons blood substitutives and plasmatic protein fractions		
Immunologic Diseases	Human Immunodeficiency Virus Type I (HIV) Infection, Other Immunodeficient Disorders, Pneumonia: Pneumocystis carinii	Nucleosides and nucleotides reverse transcriptase inhibitors	2005-2009: Visits prescribed due to HIV Infections Beginning in 2010, the following ICD-9-CM codes were in the record: 279	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Infectious Diseases	Aspergillosis, Chlamydial Infection Except Trachoma or Pneumonia, Cryptococcosis, Cytomegalovirus Disease (Acquired), Infectious Mononucleosis (stage 2), Mucormycosis, Reye's Syndrome (stage 3), Rubella: Acquired (stage 3), Schistosomiasis, Other Bacterial Infections, Other Fungal Infections, Other Infectious and Parasitic Infections, Other Viral Infections, Cytomegalovirus Disease (Congenital), Parainfluenza Virus Infection, Pneumonia: Chlamydial, Sarcoidosis, Other Respiratory Infections, Scabies		Beginning in 2010, the following ICD-9-CM codes were in the record: 001-139	
Male Genital	Benign Prostatic Hypertrophy, Gonorrhea: Male, Prostatitis	Alfa-adrenoreceptor antagonists testosterone 5-alfa reductase inhibitors		
Musculoskeletal	Vitamin D Deficiency, Dislocation: Knee, Eosinophilia Myalgia Syndrome, Fracture: Acetabulum, Fracture: Calcaneus (stage 2), Fracture: Femur, Except Head or Neck, Fracture: Femur, Head or Neck, Fracture: Fibula (stage 2), Fracture: Humerus (Shaft), Fracture: Humerus (Supracondylar) (stage 2), Fracture: Radial Shaft, Ulna or Olecranon (stage 2), Fracture: Radius, Lower End (stage 2), Fracture: Tibia (stage 2/3), Fracture or Dislocation: Patella (stage 2), Fracture or Sprain: Ankle (stage 2), Fracture, Dislocation, or Sprain: Facial Bones (stage 2/3), Fracture, Dislocation, or Sprain: Foot (stage 2), Fracture, Dislocation, or Sprain: Hip or Pelvis (stage 2/3), Fracture, Dislocation, or	Colchicine uric acid inhibitors antiinflammatory non-steroids gold salts aminoquinolines bisphosphonates calcitonin	2005-2009: Visits prescribed due to the presence of Arthrosis, Arthritis and other osteo-muscular and connective diseases, and Fractures of femurs and other consequences of fractures.	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Sprain: Humerus (Head) or Shoulder (stage 2), Fracture, Dislocation, or Sprain: Wrist or Hand or Fingers (stage 2), Gout, Herniated Intervertebral Disc, Infectious Arthritis (stage 2/3), Injury, Chest Wall, Injury, Knee, Semilunar Cartilages (stage 2), Injury, Open Wound, or Blunt Trauma: Lower Extremity (stage 2), Injury, Open Wound, or Blunt Trauma: Upper Extremity (stage 2/3), Muscular Dystrophy, Osteoarthritis, Osteochondrodysplasia, Osteomalacia, Osteomyelitis (stage 2/3), Osteoporosis, Progressive Systemic Sclerosis, Rheumatoid Arthritis, Scoliosis of the Thoracic Spine, Spondylitis, Ankylosing, Systemic Lupus Erythematosus, Anomaly: Musculoskeletal System, Injury: Other and Ill-Defined Musculoskeletal Sites, Neoplasm, Benign: Musculoskeletal Syst. or Connective Tissue, Other Arthropathies, Bone and Joint Disorders, Other Disorders of Connective Tissue, Other Spinal and Back Disorders, Myasthenia Gravis, Complications of Surgical and Medical Care (stage 1), Injury, Open Wound, or Blunt Trauma: Abdomen or Trunk (stage 2/3), Injury: Other (stage 3)		Beginning in 2010, the following ICD-9-CM codes were in the record: 710-739	
Neurologic Diseases	Down's Syndrome, Herpes zoster, Poliomyelitis, Post-Polio Syndrome, Syphilis: Acquired, Tetanus (stage 1), Toxoplasmosis: Acquired (stage 3), Amyotrophic Lateral Sclerosis, Cerebral Palsy, Cerebrovascular Disease, Disease of Nervous System Secondary to Implants or Grafts, Epilepsy, Guillain-Barre Syndrome (stage 2), Headache (stage 2), Huntington's Chorea, Injury: Craniocerebral,	Anticholinesterase agents anticonvulsivant barbiturates and congeners alprostadil ergot alkaloids 5HT1 agonists dopamine	2005-2009: Visits prescribed due to the presence of Dementia and Alzheimer's syndrome, Parkinson's and	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
	Injury: Spine and spinal cord, Meningitis, Encephalitis, and Myelitis: Viral, Meningitis: Bacterial, Mental Retardation, Multiple Sclerosis, Neurofibromatosis Type I [Von Recklinghausen's Disease], Parkinson's Disease, Other CNS Inflammation, Infection, or Disorder, Other Cranial Nerve Disorders, Other Neurological Conditions, Other Peripheral Nerve Disorders, Other Spinal Lesions, Anomaly: Neural Tube Defects, Rubella: Congenital (stage 2), Anomaly: Other Nervous System, Injury: Other	MAO b inhibitors	other CNS degenerative disease, hemiplegia, monoplegia, and other associated syndroms, and acute and chronic cerebrovascular diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 320-389,797	
Psychological	Dementia: Primary Degenerative (Alzheimer's or Pick's), Antisocial Personality Disorder, Bipolar Disorder - Major Depressive Episode, Bipolar Disorder - Manic Episode, Depression, Generalized Anxiety Disorder, Obsessive-Compulsive Neurosis, Schizophrenia, Autism, Other Neuroses, Other Psychoses Drug Abuse, Dependence, Intoxication: Alcohol, Amphetamine, Barbiturate, Cannabis, Cocaine, Hallucinogen, Opioid, Other Eating disorders: Anorexia Nervosa, Bulimia Nervosa	Antidepressants antipsychotics agents	2005-2009: Visits prescribed due to the presence of psychoses, neuroses, and mental retardation Beginning in 2010, the following ICD-9-CM codes were in the record: 290-319	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Body System or Etiology Group	Hospital Discharge data	Outpatient pharmacy data	Home health care	Specialty visits
Respiratory	Coxsackie and ECHO Infections (stage 2/3), Anomaly: Tracheoesophageal Malformations, Asbestosis, Asthma, Berylliosis, Byssinosis, Chronic Obstructive Pulmonary Disease, Coal Miner's Pneumoconiosis, Croup, Cystic Fibrosis, Emphysema, Hypersensitivity Pneumonitis, Influenza, Mycoplasma pneumoniae Infection, Parainfluenza Virus Infection (stage 2), Pneumonia: Bacterial, Pneumonia: Legionella, Pulmonary Alveolar Proteinosis, Pulmonary Embolism (stage 3), Radiation Pneumonitis, Silicosis, Tuberculosis, Complications of Tracheostomy, Other Disorders of Respiratory System, Other Respiratory Disease Due to External Agents, Other Respiratory Symptoms, Pneumonia: Aspiration, Neoplasm, Benign: Respiratory System	Inhaled corticosteroids beta-2-adrenoreceptor agonists xanthines leucotrienies antagonists cromolyn pancreatic enzymes mucolytics antituberculosis antibiotics isoniazid	2005-2009: Visits prescribed due to the presence of respiratory diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 460-519	
Skin	Herpes Virus Ocular Infection (stage 1), Urticaria, Candida (Monilial) Infections, Clostridial Wound Infection (stage 2), Herpes Simplex Infections, Complications of Surgical and Medical Care (stage 2/3), Anomaly: Integument (Genodermatoses), Decubitus Ulcers, Erythema Multiforme, Erythroderma, Immunologically Mediated Blistering Skin Diseases, Infections of Skin and Subcutaneous Tissue, Neoplasm, Malignant: Carcinoma, Basal Cell (stage 1), Neoplasm: Atypical Nevus (stage 1), Psoriasis Vulgaris, Other Inflammations & Infections of Skin & SubQ Tissue, Burns, Neoplasm, Benign: Skin or Subcutaneous Tissue (stage 1)	Oral and topical antipsoriasis agents	2005-2009: Visits prescribed due to the presence of decubitus ulcers and othere skin diseases Beginning in 2010, the following ICD-9-CM codes were in the record: 680-709	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Appendix: Regression coefficients and significance levels

Females 18-34

Variable	Coefficient	p-value
Intercept	-5.0771	<.0001
Number of Chronic Conditions (from any data source = 1)	0.4347	<.0001
Number of Chronic Conditions (from any data source = 2)	0.8614	<.0001
Number of Chronic Conditions (from any data source = 3)	1.1305	<.0001
Number of Chronic Conditions (from any data source = 4 or more)	1.7194	<.0001
Number of Chronic Conditions (from hospital data = 1)	0.9074	<.0001
Number of Chronic Conditions (from hospital data = 2 or more)	0.8834	<.0001
Total number of ER visits	0.2634	<.0001
History of Obesity-Stage 2 or 3 *	1.6342	<.0001
History of polypharmacy *	0.5968	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	2.1087	<.0001

Males 18-34

Variable	Coefficient	p-value
Intercept	-5.2835	<.0001
Number of Chronic Conditions (from any data source) = 1	0.6534	<.0001
Number of Chronic Conditions (from any data source) = 2	1.2390	<.0001
Number of Chronic Conditions (from any data source) = 3	1.5240	<.0001
Number of Chronic Conditions (from any data source) = 4 or more	2.0556	<.0001
Neurologic Diseases (from home health prescription)	1.6802	<.0001
Renal Failure-Stage 2 or 3	1.3802	<.0001
Any Gastrointestinal Disease - Stage 2 (from hospital data)	1.1512	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	0.8614	<.0001
Any Psychologic Disease - Stage 2 (from hospital data)	1.0198	<.0001
Any Respiratory Disease - Stage 2 (from hospital data)	0.9451	<.0001
Anti-arrhythmics	1.5415	<.0001
Total number of ER visits	0.2371	<.0001
History of Neurologic Diseases (from drug prescriptions) *	0.4880	<.0001
History of Crohns Disease-Stage 2 or 3 *	1.4684	0.0004
History of Neoplasm, Malignant: Colon and Rectum-Stage 2 *	2.9037	0.0063
History of Calculus of the Urinary Tract-Stage 1 *	1.0806	<.0001
History of Cirrhosis of the Liver-Stage 2 or 3 *	1.2212	<.0001
History of Pancreatitis-any stage *	1.7777	<.0001
History of Cerebrovascular Disease-Stage 2 *	2.0588	0.0004
History of Obesity-Stage 2 or 3 *	1.6569	<.0001
History of polypharmacy *	0.4747	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	1.6103	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Hospitalization	0.6338	<.0001

Females 35-44

Variable	Coefficient	p-value
Intercept	-4.9905	<.0001
Number of Chronic Conditions (from any data source) = 1	0.5265	<.0001
Number of Chronic Conditions (from any data source)= 2	0.8446	<.0001
Number of Chronic Conditions (from any data source) = 3	0.8519	<.0001
Number of Chronic Conditions (from any data source)= 4 or more	0.6525	0.0155
Number of Chronic Conditions (from any data source) = 5 or more	0.5419	0.0934
Number of Chronic Conditions (from hospital data) = 1	0.8319	<.0001
Number of Chronic Conditions (from hospital data) = 2	1.0387	<.0001
Number of Chronic Conditions (from hospital data) = 3 or more	1.3840	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.5128	0.0390
Number of Chronic Conditions (from drug prescriptions)=1	-0.0948	0.3171
Number of Chronic Conditions (from drug prescriptions)=2	0.0670	0.6542
Number of Chronic Conditions (from drug prescriptions)=3 or more	0.2362	0.2746
Reside in Mountain area on 12/ 31/ 2012	0.1865	0.0196
Reside in Hill area on 12/ 31/ 2012	-0.0128	0.7341
Cardiovascular Disease (from home health prescription)	1.7641	0.0006
Endocrine Disease (from home health prescription)	1.5904	0.0930
Infectious Disease (from home health prescription)	1.6836	0.0468
Genitourinary (dialysis)	0.7208	0.0081
Aortic Stenosis-Stage 1	2.2652	0.0004
Arrhythmias-Stage 2	1.0050	0.0016
Neoplasm, Malignant: Stomach-Stage 3	1.8592	0.0204
Neoplasm, Malignant: Breast, Female-Stage 3	0.7628	0.0204
Progressive Systemic Sclerosis-Stage 1	0.9852	0.0321
Progressive Systemic Sclerosis-Stage 2 or 3	1.6206	0.0087
Obesity-Stage 2 or 3	0.4604	0.0039
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	0.6965	0.0223
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.7457	0.0325
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	1.6501	0.0187
Any Cancer - Stage 3 (from hospital data)	1.2094	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.5051	0.0014
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.4724	0.0065
Any Genitourinary Disease - Stage 2 (from hospital data)	0.3895	0.0581
Any Gynecologic Disease - Stage 1 (from hospital data)	-0.5549	<.0001
Any Hepatobiliary Disease - Stage 1 (from hospital data)	-0.5393	0.0029
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.4858	0.0002

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any Neurologic Disease - Stage 3 (from hospital data)	0.8922	<.0001
Any Psychologic Disease - Stage 2 (from hospital data)	0.5166	0.0038
Any Respiratory Disease - Stage 2 (from hospital data)	0.5162	0.0436
Endocrine Disease (from drug prescriptions)	-0.4482	<.0001
Genitourinary Disease (from drug prescriptions)	0.9552	0.0001
Respiratory Disease (from drug prescriptions)	-0.2258	0.0107
Cardiovascular Disease (from any data source)	-0.2700	0.0118
Day hospitalization	-0.2627	0.0004
ACE/ARB	0.2465	0.0266
Digitalis glycosides	0.9678	0.0285
Number of ER visits labeled 'Yellow'	-0.2515	0.0378
Total number of ER visits	0.4548	0.0001
Eye Disease (from any data source)	-0.5158	0.0174
History of Cancer (from drug prescriptions) *	0.2289	0.0247
History of Endocrine Disease (from drug prescriptions) *	0.1529	0.0624
History of Psychological Disease (from drug prescriptions) *	0.2272	<.0001
History of Arrhythmias-Stage 2 *	0.4871	0.0794
History of Cardiomyopathies-Stage 3 *	1.1771	0.0262
History of Thrombophlebitis-Stage 2 or 3 *	0.9344	0.0030
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	1.1944	<.0001
History of Crohns Disease-Stage 2 or 3 *	0.7513	0.0377
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	2.0723	<.0001
History of Calculus of the Urinary Tract-Stage 1 *	0.4914	0.0146
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.8776	0.0003
History of Neoplasm, Malignant: Kidneys-Stage 1 *	1.3977	0.0102
History of Neoplasm, Malignant: Kidneys-Stage 3 *	3.2491	0.0100
History of Pancreatitis-any stage *	0.8084	0.0199
History of Cerebrovascular Disease-Stage 1 *	0.8172	0.0057
History of Cerebrovascular Disease-Stage 2 *	1.1315	0.0013
History of Obesity-Stage 2 or 3 *	1.2321	<.0001
History of polypharmacy *	0.3345	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.6431	0.0246
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.8255	<.0001
History of Depression-Stage 1 or 2 *	0.2207	0.0716
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.5358	0.0330
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.9824	0.0003
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	1.8886	0.0964
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	1.6543	0.0122
History of Pneumonia: Bacterial-Stage 3 *	1.3040	0.0015
Immunologic Disease (from any data source)	0.8590	0.0001
Polypharmacy	0.2838	<.0001
Number of the other 9 Cardiovascular drugs	0.1330	0.0036

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Males 35-44

Variable	Coefficient	p-value
Intercept	-4.8083	<.0001
Number of Chronic Conditions (from any data source)=1	0.5439	<.0001
Number of Chronic Conditions (from any data source)=2	0.7994	<.0001
Number of Chronic Conditions (from any data source)=3	0.7949	<.0001
Number of Chronic Conditions (from any data source)=4	1.1832	<.0001
Number of Chronic Conditions (from any data source)=5 or more	0.9615	<.0001
Age on 12/ 31/ 2012	-0.00005	0.0031
Cancer (from home health prescription)	1.6698	0.0023
Blood Diseases (from home health prescription)	2.7427	0.0385
Infectious Disease (from home health prescription)	3.1524	0.0129
Neurologic Diseases (from home health prescription)	2.0454	<.0001
Genitourinary (dialysis)	1.0696	<.0001
Cardiomyopathies-Stage 3	1.7220	0.0016
Infective Endocarditis-Stage 3	3.7783	0.0015
Mitral Stenosis-Stage 3	3.0377	0.0345
Pericarditis: Chronic-Stage 2 or 3	1.2938	0.0240
Crohns Disease-Stage 1	1.6408	<.0001
Neoplasm, Malignant: Colon and Rectum-Stage 2	1.4659	0.0030
Cirrhosis of the Liver-Stage 2 or 3	0.6646	0.0066
Neoplasm, Malignant: Pancreas-Stage 1	2.4864	0.0143
Pancreatitis-all stages	0.8241	0.0053
Cerebrovascular Disease-Stage 3	0.9540	<.0001
Epilepsy-all stages	0.5515	0.0247
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	0.6065	0.0050
Any Cancer - Stage 3 (from hospital data)	1.7602	<.0001
Any Endocrine Disease - Stage 1 (from hospital data)	-0.5267	0.0028
Any Endocrine - Stage 2 (from hospital data)	0.5688	0.0028
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-1.5927	<.0001
Any Immunologic Disease - All stages (from hospital data)	0.6498	0.0068
Any Psychologic Disease - Stage 2 (from hospital data)	0.7963	<.0001
Any Psychologic Disease - Stage 3 (from hospital data)	1.4603	0.0014
Endocrine Disease (from drug prescriptions)	0.2204	0.0058
Number of day hospitalizations	0.2078	0.0003
Oral anti-coagulants	0.4822	0.0028
Anti-arrhythmics	0.7069	0.0027
Total number of ER visits	0.2455	<.0001
Eye Disease (from any data source)	-0.4757	0.0200
History of Neurological Disease (from drug prescriptions) *	0.3728	<.0001
History of Psychological Disease (from drug prescriptions) *	0.2678	<.0001
History of Arrhythmias-Stage 2 *	0.7259	0.0008
History of Cardiomyopathies-Stage 2 *	0.8655	0.0009

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Essential Hypertension-Stage 3 *	0.8004	0.0062
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	1.0253	0.0008
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.7425	0.0034
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.7380	0.0071
History of Crohns Disease-Stage 2 or 3 *	1.3269	<.0001
History of Calculus of the Urinary Tract-Stage 1 *	0.7919	<.0001
History of Renal Failure-Stage 2 or 3 *	0.5494	0.0107
History of Cholecystitis and Cholelithiasis-Stage 3 *	1.4633	0.0012
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.5857	0.0009
History of Pancreatitis-any stage *	1.2530	<.0001
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	3.8605	0.0026
History of Obesity-Stage 2 or 3 *	0.8764	<.0001
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.8467	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.5825	0.0005
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.7745	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.7316	0.0018
History of Pneumonia: Bacterial-Stage 3 *	1.0518	0.0006
History of Other Cardiovascular drugs *	0.2342	0.0043
Hospitalization	0.7414	<.0001
Gastrointestinal Disease (from hospital data)	0.9745	<.0001
Respiratory Disease (from hospital data)	0.4067	0.0025
Any of the other 9 Cardiovascular drugs	-0.5914	<.0001
Number of the other 9 Cardiovascular drugs	0.2767	<.0001

Females 45-54

Variable	Coefficient	p-value
Intercept	-4.9051	<.0001
Number of Chronic Conditions (from any data source)=1	0.3066	<.0001
Number of Chronic Conditions (from any data source)=2	0.4393	<.0001
Number of Chronic Conditions (from any data source)=3	0.4533	<.0001
Number of Chronic Conditions (from any data source)=4	0.3924	0.0002
Number of Chronic Conditions (from any data source)=5 or more	0.3544	0.0079
Number of Chronic Conditions (from hospital data)=1	0.4819	<.0001
Number of Chronic Conditions (from hospital data)=2	0.6828	<.0001
Number of Chronic Conditions (from hospital data)=3 or more	0.8788	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.9174	<.0001
Reside in Mountain area on 12/ 31/ 2012	0.2243	0.0008
Reside in Hill area on 12/ 31/ 2012	-0.0104	0.7596
Essential Hypertension-Stage 1	-0.3166	0.0082
Pericarditis: Chronic-Stage 2 or 3	2.0018	0.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Neoplasm, Malignant: Colon and Rectum-Stage 2	0.8110	0.0074
Cholecystitis and Cholelithiasis-Stage 2	0.8570	0.0043
Cirrhosis of the Liver-Stage 2 or 3	0.4736	0.0176
Pancreatitis-all stages	0.8873	0.0217
Progressive Systemic Sclerosis-Stage 2 or 3	0.9823	0.0173
Cerebrovascular Disease-Stage 1	0.5623	0.0340
Cerebrovascular Disease-Stage 3	-0.6910	0.0224
Obesity-Stage 2 or 3	0.4054	0.0063
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	1.3912	0.0132
Any Cancer - Stage 3 (from hospital data)	1.7066	<.0001
Any Ear,Nose,Throat Disease - Stage 1 (from hospital data)	-0.6186	0.0401
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.4730	<.0001
Any Gynecologic Disease - Stage 1 (from hospital data)	-0.5578	<.0001
Any Infectious Disease - Stage 3 (from hospital data)	-1.8556	0.0017
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.4338	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	1.0267	<.0001
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2591	0.0256
Any Psychologic Disease - Stage 2 (from hospital data)	0.4123	0.0034
Any Psychologic Disease - Stage 3 (from hospital data)	1.3041	0.0042
Any Respiratory Disease - Stage 2 (from hospital data)	0.4618	0.0149
Any Respiratory Disease - Stage 3 (from hospital data)	0.7372	0.0009
Any Skin Disease - Stage 1 (from hospital data)	-0.3787	0.0245
Gastrointestinal Disease (from drug prescriptions)	0.2535	<.0001
Genitourinary Disease (from drug prescriptions)	1.1262	<.0001
Oral anti-coagulants	0.3753	0.0063
Anti-arrhythmics	0.7321	<.0001
Digitalis glycosides	0.9270	0.0003
Total number of ER visits	0.2068	<.0001
History of Cancer (from drug prescriptions) *	0.1500	0.0286
History of Psychological Disease (from drug prescriptions) *	0.2010	<.0001
History of Aortic Stenosis-Stage 3 *	1.6666	0.0170
History of Coronary Artery Disease-Stage 3 *	0.5560	0.0019
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	1.0444	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.6990	0.0012
History of Crohns Disease-Stage 2 or 3 *	1.3534	<.0001
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.9831	0.0009
History of Neoplasm, Malignant: Stomach-Stage 3 *	2.1435	0.0007
History of Ulcerative Colitis-any stage *	0.6736	0.0177
History of Calculus of the Urinary Tract-Stage 1 *	0.7453	<.0001
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.5209	0.0218
History of Renal Failure-Stage 2 or 3 *	0.5702	0.0006
History of Neoplasm, Malignant: Ovaries-Stage 2 or 3 *	0.7030	0.0286
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.7484	0.0026
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.4895	0.0014

*History variables are calculated on previous 5 years of exposure data

For peer review only

History of Pancreatitis-any stage *	0.7645	0.0043
History of Cerebrovascular Disease-Stage 3 *	0.3947	0.0087
History of Obesity-Stage 2 or 3 *	0.8330	<.0001
History of polypharmacy *	0.2928	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.6454	0.0038
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.6630	0.0001
History of Depression-Stage 1 or 2 *	0.4957	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.6357	0.0005
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	1.1324	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.4127	0.0104
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	1.3247	0.0050
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	1.6041	<.0001
History of Neoplasm, Malignant: Melanoma-Stage 3 *	1.6975	0.0026
History of Other Cardiovascular drugs *	0.1255	0.0236
Immunologic Disease (from any data source)	0.7145	0.0003
Infectious Disease (from any data source)	0.5052	0.0335
Neurologic Disease (from any data source)	0.2075	0.0003
Hospitalization	0.3975	<.0001
Polypharmacy	0.2655	<.0001
Any of the other 9 Cardiovascular drugs	-0.2999	0.0008
Number of the other 9 Cardiovascular drugs	0.1482	<.0001

Males 45-54

Variable	Coefficient	p-value
Intercept	-4.4469	<.0001
Number of Chronic Conditions (from any data source)=1	0.3859	<.0001
Number of Chronic Conditions (from any data source)=2	0.6634	<.0001
Number of Chronic Conditions (from any data source)=3	0.7465	<.0001
Number of Chronic Conditions (from any data source)=4	0.7901	<.0001
Number of Chronic Conditions (from any data source)=5 or more	0.5246	<.0001
Number of Chronic Conditions (from hospital data)=1	0.2577	0.0009
Number of Chronic Conditions (from hospital data)=2	0.3237	0.0015
Number of Chronic Conditions (from hospital data)=3 or more	0.4067	0.0021
Number of Chronic Conditions (from home health prescription)=1 or more	0.8811	<.0001
Age on 12 /31/ 2012	-0.00008	<.0001
Cancer (chemo or radiation)	0.5498	0.0011

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Genitourinary (dialysis)	0.9242	<.0001
Aortic Stenosis-Stage 3	2.0591	0.0014
Arrhythmias-Stage 2	0.5607	0.0005
Essential Hypertension-Stage 2	-0.6186	0.0030
Neoplasm, Malignant Hematologic-Stage 3	-1.5129	0.0096
Cirrhosis of the Liver-Stage 2 or 3	0.8760	<.0001
Pancreatitis-all stages	0.9702	<.0001
Chronic Obstructive Pulmonary Disease-Stage 3	1.2772	0.0051
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	1.5413	<.0001
Any Cancer - Stage 3 (from hospital data)	1.5337	<.0001
Any Cardiovascular - Stage 3 (from hospital data)	0.4528	<.0001
Any Endocrine Disease - Stage 1 (from hospital data)	-0.2681	0.0085
Any Endocrine - Stage 2 (from hospital data)	0.3403	0.0055
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.7672	<.0001
Any Immunologic Disease - All stages (from hospital data)	0.7049	<.0001
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.3976	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	0.4660	0.0008
Any Psychologic Disease - Stage 2 (from hospital data)	0.8976	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.1612	0.0030
Eye Disease (from drug prescriptions)	-0.5308	<.0001
Genitourinary Disease (from drug prescriptions)	0.4405	0.0099
Hematologic Disease (from drug prescriptions)	0.5097	0.0070
Hepatobiliary Disease (from drug prescriptions)	0.3691	0.0179
Number of day hospitalizations	0.1160	0.0093
Statins	-0.1389	0.0112
Anti-platelets	0.2288	<.0001
Anti-arrhythmics	0.3517	0.0063
Nitrates	0.4390	<.0001
Total number of ER visits	0.1627	<.0001
History of Cancer (from drug prescriptions) *	0.3251	0.0004
History of Aortic Stenosis-Stage 1 *	-0.9794	0.0150
History of Arrhythmias-Stage 2 *	0.3919	0.0019
History of Cardiomyopathies-Stage 3 *	0.7836	<.0001
History of Coronary Artery Disease-Stage 1 *	0.3743	<.0001
History of Thrombophlebitis-Stage 2 or 3 *	0.7954	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.3220	0.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.8677	<.0001
History of Calculus of the Urinary Tract-Stage 1 *	0.3374	0.0033
History of Renal Failure-Stage 2 or 3 *	0.4042	0.0033
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.5888	0.0001
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.4954	<.0001
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	1.9346	0.0029
History of Pancreatitis-any stage *	0.5981	0.0009
History of Obesity-Stage 2 or 3 *	0.5126	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only

History of polypharmacy *	0.2306	<.0001
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.8192	<.0001
History of Depression-Stage 1 or 2 *	0.2814	0.0098
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.4287	0.0020
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.9474	<.0001
History of Pneumonia: Bacterial-Stage 3 *	1.1780	<.0001
History of Other Cardiovascular drugs *	0.2404	<.0001
Male Genital System (from any data source)	-0.3177	0.0021
Neurologic Disease (from any data source)	0.2173	0.0002
Hospitalization	0.4249	<.0001
Number of hospitalizations	0.0445	0.1777
Polypharmacy	0.2976	<.0001
Gastrointestinal Disease (from hospital data)	-0.2445	0.0085

Females 55-64

Variable	Coefficient	p-value
Intercept	0.9467	0.6785
Number of Chronic Conditions (from any data source)=1	0.5017	<.0001
Number of Chronic Conditions (from any data source)=2	0.6666	<.0001
Number of Chronic Conditions (from any data source)=3	0.7010	<.0001
Number of Chronic Conditions (from any data source)=4	0.7868	0.0001
Number of Chronic Conditions (from any data source)=5	0.7545	0.0024
Number of Chronic Conditions (from any data source)=6 or more	0.5597	0.0587
Number of Chronic Conditions (from hospital data)=1	0.5017	<.0001
Number of Chronic Conditions (from hospital data)=2	0.6365	<.0001
Number of Chronic Conditions (from hospital data)=3	0.7653	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.7953	0.0007
Number of Chronic Conditions (from home health prescription)=1 or more	0.4889	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.3395	0.0003
Number of Chronic Conditions (from drug prescriptions)=2	-0.3996	0.0020
Number of Chronic Conditions (from drug prescriptions)=3	-0.4436	0.0073
Number of Chronic Conditions (from drug prescriptions)=4	-0.5404	0.0083
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.4198	0.0955
Age on 12/ 31/ 2012	-0.1094	0.0134
Endocrine Disease (from home health prescription)	0.9016	0.0026
Gastrointestinal Disease (from home health prescription)	-0.6807	0.1725
Genitourinary Disease (from home health prescription)	1.1277	0.1537
Blood Diseases (from home health prescription)	-1.8804	0.1597
Infectious Disease (from home health prescription)	1.2458	0.0624
Musculoskeletal Disease (from home health prescription)	0.7627	0.0235

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Neurologic Diseases (from home health prescription)	0.7672	0.0003
Respiratory Diseases (from home health prescription)	1.0350	0.0211
Skin Disease (from home health prescription)	0.4575	0.2643
Cancer (chemo or radiation)	0.2011	0.0741
Genitourinary (dialysis)	-0.2335	0.2676
Aneurysm, Thoracic-all stages	0.7181	0.1737
Arrhythmias-Stage 1	-0.8531	0.1853
Arrhythmias-Stage 3	-0.4355	0.3839
Congestive Heart Failure-Stage3	0.5326	0.0248
Essential Hypertension-Stage 2	-0.4672	0.0156
Mitral Stenosis-Stage 1	0.3593	0.1871
Mitral Stenosis-Stage 2	0.4834	0.2609
Pericarditis: Viral or Traumatic-Stage 2 or 3	-1.2494	0.1428
Thrombophlebitis-Stage 2 or 3	0.3434	0.2209
Tibial/lliac/Femoral/Popliteal Artery Disease-Stage 1	0.2126	0.4831
Diabetes Mellitus Type 1 or Type 2-Stage 3	-1.2846	0.0001
Hyperthyroidism-Stage 1	-0.6448	0.1376
Hypothyroidism-Stage 1	-0.1514	0.2913
Hypothyroidism-Stage 2 or 3	-0.4319	0.2665
Crohns Disease-Stage 2 or 3	-1.3093	0.2190
Diverticular Disease-Stage 1	-0.3203	0.2326
Diverticular Disease-Stage 2 or 3	-1.7244	0.0142
Gastritis-Stage 2 or 3	-0.4282	0.3525
Hernia, Hiatal or Reflux Esophagitis-Stage 1	-0.4908	0.0465
Hernia, Hiatal or Reflux Esophagitis-Stage 2 or 3	-0.6378	0.2767
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.6795	0.0391
Neoplasm, Malignant: Colon and Rectum-Stage 3	0.9264	<.0001
Neoplasm, Malignant: Stomach-Stage 1	0.5593	0.2519
Neoplasm, Malignant: Stomach-Stage 3	0.6181	0.1650
Ulcerative Colitis-all stages	0.5059	0.1961
Neoplasm, Malignant: Bladder, Urinary-Stage 1	-0.3433	0.3353
Neoplasm, Malignant: Kidneys-Stage 1	0.3389	0.3869
Renal Failure-Stage 2 or 3	0.3465	0.0464
Neoplasm, Malignant: Breast, Female-Stage 1	-0.9369	<.0001
Neoplasm, Malignant: Ovaries-Stage 1	-0.7428	0.0657
Anemia: Aplastic, Acquired-Stage 2 or 3	0.3219	0.2666
Neoplasm, Malignant Hematologic-Stage 1	-0.4013	0.0971
Neoplasm, Malignant Hematologic-Stage 2	-0.6307	0.0974
Neoplasm, Malignant Hematologic-Stage 3	-0.5380	0.2971
Cholecystitis and Cholelithiasis-Stage 1	0.7769	0.0033
Cholecystitis and Cholelithiasis-Stage 2	1.3013	0.0057
Cirrhosis of the Liver-Stage 2 or 3	0.7679	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.8360	0.0386
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.3469	0.0051

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Pancreatitis-all stages	1.0560	0.0079
Progressive Systemic Sclerosis-Stage 2 or 3	1.0431	0.0155
Cerebrovascular Disease-Stage 1	0.2701	0.2083
Cerebrovascular Disease-Stage 2	-0.2745	0.2425
Cerebrovascular Disease-Stage 3	-0.2845	0.3180
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.6421	0.0743
Epilepsy-all stages	-0.4298	0.1025
Bipolar Disorder - Major Depressive Episode-Stage 2 or 3	-0.5924	0.2896
Depression-Stage 1 or 2	0.1495	0.4002
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	-0.6312	0.2187
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.4370	0.1568
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.1433	0.4617
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.4562	0.0554
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.4995	0.0489
Pneumonia: Bacterial-Stage 1	0.3737	0.0753
Pneumonia: Bacterial-Stage 3	-0.4891	0.1195
Pulmonary Embolism-Stage 3	0.8005	0.0250
Any Cancer - Stage 1 (from hospital data)	0.2989	0.0931
Any Cancer - Stage 2 (from hospital data)	0.2308	0.2732
Any Cancer - Stage 3 (from hospital data)	0.8552	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.2364	0.0023
Any Cardiovascular Disease - Stage 2 (from hospital data)	0.0953	0.3950
Any Cardiovascular - Stage 3 (from hospital data)	0.1844	0.1691
Any Endocrine - Stage 2 (from hospital data)	0.4515	<.0001
Any Endocrine Disease - Stage 3 (from hospital data)	0.7544	0.0013
Any Ear,Nose,Throat Disease - Stage 2 (from hospital data)	-0.8908	0.3948
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.3647	0.0445
Any Hepatobiliary Disease - Stage 1 (from hospital data)	-0.6278	0.0950
Any Hepatobiliary Disease - Stage 2 (from hospital data)	-0.4898	0.3504
Any Infectious Disease - Stage 3 (from hospital data)	-0.3543	0.3858
Any Neurologic Disease - Stage 2 (from hospital data)	0.5032	0.0016
Any Neurologic Disease - Stage 3 (from hospital data)	0.4457	0.0740
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2030	0.1733
Any Psychologic Disease - Stage 2 (from hospital data)	0.1866	0.3379
Any Psychologic Disease - Stage 3 (from hospital data)	0.9858	0.0302
Any Respiratory Disease – Stage 1 (from hospital data)	-0.3063	0.0617
Any Respiratory Disease - Stage 3 (from hospital data)	0.9969	<.0001
Any Skin Disease - Stage 2 (from hospital data)	-0.5257	0.1185
Neoplasm, Malignant: Melanoma-Stage 2	-1.2688	0.1880
Neoplasm, Malignant: Melanoma-Stage 3	1.5149	0.0037
Cancer (from any data source)	-0.3957	0.0036
Cancer (from drug prescription)	0.4367	0.0002
Eye Disease (from drug prescriptions)	0.4615	0.0460
Gastrointestinal Disease (from drug prescriptions)	0.0544	0.1584

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Genitourinary Disease (from drug prescriptions)	0.2762	0.1949
Immunologic Disease (from drug prescriptions)	-1.0597	0.0780
Psychological Disease (from drug prescriptions)	-0.1619	0.3523
Respiratory Disease (from drug prescriptions)	0.2058	0.0005
Skin Disease (from drug prescriptions)	0.1296	0.4503
Day hospitalization	-0.0504	0.4576
Drug-Drug interactions	0.2803	0.0229
Statins	-0.2104	<.0001
Beta-blockers	0.0461	0.1974
Anti-platelets	0.1061	0.0107
Calcium channel blockers	0.0933	0.0329
Anti-arrhythmics	0.2835	0.0060
Digitalis glycosides	-0.2094	0.2138
Nitrates	0.4632	<.0001
Diuretics	0.2687	<.0001
Ear,Nose,Throat Disease (from any data source)	-0.3916	0.1154
Number of ER visits labeled 'Yellow'	0.0881	0.2672
Total number of ER visits	0.1621	0.0341
Eye Disease (from any data source)	-0.6320	0.0043
Genitourinary Disease (from any data source)	0.7024	0.0014
Gynecologic Disease (from any data source)	-0.6085	<.0001
Hematologic Disease (from any data source)	0.2896	0.0019
Hepatobiliary Disease (from any data source)	0.4386	0.0713
History of Cancer (from drug prescriptions) *	0.2544	<.0001
History of Neurological Disease (from drug prescriptions) *	0.0571	0.2898
History of Respiratory Disease (from drug prescriptions) *	0.0775	0.0855
History of Aortic Stenosis-Stage 1 *	0.1809	0.4275
History of Arrhythmias-Stage 2 *	0.3253	0.0041
History of Arrhythmias-Stage 3 *	0.3799	0.2085
History of Cardiomyopathies-Stage 2 *	0.4667	0.0155
History of Congestive Heart Failure-Stage 3 *	0.1210	0.4597
History of Coronary Artery Disease-Stage 1 *	0.3581	0.0002
History of Coronary Artery Disease-Stage 2 *	0.1121	0.4469
History of Coronary Artery Disease-Stage 3 *	0.3913	0.0023
History of Essential Hypertension-Stage 1 *	0.0703	0.1748
History of Infective Endocarditis-Stage 3 *	-0.5804	0.3469
History of Mitral Stenosis-Stage 2 *	0.4323	0.1047
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	-0.5354	0.2771
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.5543	0.0659
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.3017	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.3810	0.0005
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.4846	0.0020
History of Hypothyroidism-Stage 2 or 3 *	0.2600	0.3833

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Crohns Disease-Stage 2 or 3 *	0.7668	0.0458
History of Diverticular Disease-Stage 1 *	0.3579	0.0123
History of Diverticular Disease-Stage 2 or 3 *	0.7500	0.0026
History of Neoplasm, Malignant: Stomach-Stage 3 *	1.3562	0.0009
History of Calculus of the Urinary Tract-Stage 1 *	0.1328	0.4254
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.1600	0.4523
History of Neoplasm, Malignant: Kidneys-Stage 1 *	-0.2475	0.4638
History of Neoplasm, Malignant: Kidneys-Stage 3 *	1.1246	0.0296
History of Renal Failure-Stage 2 or 3 *	0.5356	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 1 *	-0.2892	0.0025
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.5799	0.0001
History of Neoplasm, Malignant: Ovaries-Stage 2 or 3 *	0.4641	0.0675
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.2917	0.1730
History of Neoplasm, Malignant Hematologic-Stage 2 *	0.2692	0.3031
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.6343	0.2195
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.4606	0.0057
History of Cholecystitis and Cholelithiasis-Stage 3 *	-0.3862	0.3756
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.1037	0.4226
History of Neoplasm, Malignant: Pancreas-Stage 1 *	0.9151	0.0222
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	-1.0822	0.1459
History of Pancreatitis-any stage *	0.2106	0.4490
History of Progressive Systemic Sclerosis-Stage 1 *	0.2697	0.2504
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	-0.4137	0.2802
History of Cerebrovascular Disease-Stage 1 *	0.2919	0.0479
History of Cerebrovascular Disease-Stage 2 *	0.1387	0.2983
History of Cerebrovascular Disease-Stage 3 *	0.3321	0.0030
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.2370	0.3913
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	1.0816	0.1258
History of Obesity-Stage 2 or 3 *	0.2556	0.0066
History of polypharmacy *	0.1792	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.5506	0.0209
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.6414	0.0002
History of Depression-Stage 1 or 2 *	0.3655	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.6588	0.0024
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.6840	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.4496	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.9085	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2 *	1.1063	0.0136

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	0.4570	0.1957
History of Pneumonia: Bacterial-Stage 2 *	0.7303	0.1047
History of Oral Anti-coagulants *	0.4605	<.0001
History of Other Cardiovascular drugs *	0.0431	0.3044
History of Statins *	0.0710	0.1377
Immunologic Disease (from any data source)	0.7304	0.0366
Neurologic Disease (from any data source)	0.2104	0.0008
Hospitalization	0.2413	0.0005
Polypharmacy	0.3614	<.0001
Psychological Disease (from any data source)	0.2990	0.0875
Cancer (from hospital data)	0.2380	0.2926
Gastrointestinal Disease (from hospital data)	-0.2031	0.0659
Genitourinary Disease (from hospital data)	-0.4981	0.0234
Hepatobiliary (from hospital data)	-0.5335	0.2235
Musculoskeletal Disease (from hospital data)	-0.4348	<.0001
Any of the other 9 Cardiovascular drugs	-0.1434	0.0142

Males 55-64

Variable	Coefficient	p-value
Intercept	-4.2367	<.0001
Number of Chronic Conditions (from any data source)=1	0.3641	<.0001
Number of Chronic Conditions (from any data source)=2	0.7530	<.0001
Number of Chronic Conditions (from any data source)=3	1.0181	<.0001
Number of Chronic Conditions (from any data source)=4	1.2055	<.0001
Number of Chronic Conditions (from any data source)=5	1.4339	<.0001
Number of Chronic Conditions (from any data source)=6 or more	1.4674	<.0001
Number of Chronic Conditions (from hospital data)=1	0.4141	<.0001
Number of Chronic Conditions (from hospital data)=2	0.5725	<.0001
Number of Chronic Conditions (from hospital data)=3	0.7463	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.6436	0.0066
Number of Chronic Conditions (from home health prescription)=1 or more	-0.5216	0.0939
Number of Chronic Conditions (from drug prescriptions)=1	-0.0825	0.3624
Number of Chronic Conditions (from drug prescriptions)=2	-0.3715	0.0074
Number of Chronic Conditions (from drug prescriptions)=3	-0.5199	0.0060
Number of Chronic Conditions (from drug prescriptions)=4	-0.7343	0.0025
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.8378	0.0069
Age on 12 /31/ 2012	-0.00009	<.0001
Cancer (from home health prescription)	1.5149	<.0001
Cardiovascular Disease (from home health prescription)	0.6241	0.0555
Endocrine Disease (from home health prescription)	1.2243	0.0016
Genitourinary Disease (from home health prescription)	0.5007	0.4113

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Blood Diseases (from home health prescription)	1.2676	0.2946
Musculoskeletal Disease (from home health prescription)	1.2147	0.0136
Neurologic Diseases (from home health prescription)	1.5200	<.0001
Mental Disorders (from home health prescription)	0.9790	0.0086
Respiratory Diseases (from home health prescription)	0.5207	0.3572
Skin Disease (from home health prescription)	1.2618	0.0224
Cancer (chemo or radiation)	0.4095	0.0006
Genitourinary (dialysis)	0.4977	0.0006
Aneurysm, Abdominal-all stages	-0.4983	0.0674
Aneurysm, Thoracic-all stages	-1.0860	0.0027
Aortic Stenosis-Stage 3	0.3053	0.4855
Arrhythmias-Stage 1	0.3575	0.2891
Arrhythmias-Stage 3	0.4253	0.1234
Cardiomyopathies-Stage 2	0.1958	0.2233
Cardiomyopathies-Stage 3	0.3422	0.0864
Congestive Heart Failure-Stage3	0.2219	0.2411
Coronary Artery Disease-Stage 1	-0.0683	0.3883
Coronary Artery Disease-Stage 2	0.0967	0.3847
Essential Hypertension-Stage 2	-0.1509	0.2134
Mitral Stenosis-Stage 2	0.4867	0.0815
Thrombophlebitis-Stage 1	0.3111	0.1622
Thrombophlebitis-Stage 2 or 3	0.2554	0.2361
Tibial/lliac/Femoral/Popliteal Artery Disease-Stage 1	0.1621	0.2869
Crohns Disease-Stage 1	1.2423	0.0002
Diverticular Disease-Stage 1	0.1565	0.4435
Diverticular Disease-Stage 2 or 3	-0.6513	0.1725
Functional Digestive Disorders-Stage 1	0.2423	0.3863
Hernia, Hiatal or Reflux Esophagitis-Stage 1	0.4054	0.0335
Hernia, Hiatal or Reflux Esophagitis-Stage 2 or 3	0.3879	0.3142
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.3750	0.1667
Neoplasm, Malignant: Colon and Rectum-Stage 3	0.1798	0.4389
Neoplasm, Malignant: Stomach-Stage 1	0.3938	0.2408
Neoplasm, Malignant: Stomach-Stage 3	0.5800	0.1449
Calculus of the Urinary Tract-Stage 2 or 3	0.2830	0.3339
Neoplasm, Malignant: Bladder, Urinary-Stage 3	0.8356	0.0617
Neoplasm, Malignant: Kidneys-Stage 3	0.9161	0.0054
Renal Failure-Stage 2 or 3	0.2518	0.0215
Anemia: Aplastic, Acquired-Stage 2 or 3	0.3357	0.2123
Neoplasm, Malignant Hematologic-Stage 2	-0.4371	0.1590
Neoplasm, Malignant Hematologic-Stage 3	-0.8421	0.0510
Cholecystitis and Cholelithiasis-Stage 1	0.1946	0.3670
Cholecystitis and Cholelithiasis-Stage 2	0.7374	0.0003
Cirrhosis of the Liver-Stage 2 or 3	0.6437	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	1.1672	0.0009

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Neoplasm, Malignant: Pancreas-Stage 2 or 3	0.4115	0.3686
Pancreatitis-all stages	0.6998	0.0093
Rheumatic Fever- Stage 2	0.2881	0.3795
Rheumatic Fever- Stage 3	0.7440	0.2101
Neoplasm, Malignant: Prostate-Stage 2	-0.9665	0.0008
Progressive Systemic Sclerosis-Stage 1	1.5792	0.0739
Cerebrovascular Disease-Stage 3	-0.3096	0.1750
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.3701	0.2504
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	-0.8789	0.3960
Bipolar Disorder - Major Depressive Episode-Stage 2 or 3	0.4031	0.3584
Bipolar Disorder - Manic Episode-Stage 2	0.3873	0.2024
Depression-Stage 1 or 2	-0.1959	0.3188
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.5678	0.0022
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.4929	0.0002
Chronic Obstructive Pulmonary Disease-Stage 3	0.7526	0.0053
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.4482	0.0104
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.7027	0.0037
Pneumonia: Bacterial-Stage 1	0.4147	0.0065
Pneumonia: Bacterial-Stage 2	-1.0459	0.0841
Any Cancer - Stage 1 (from hospital data)	0.1447	0.3765
Any Cancer - Stage 2 (from hospital data)	0.4529	0.0235
Any Cancer - Stage 3 (from hospital data)	0.8561	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.1085	0.2067
Any Cardiovascular - Stage 3 (from hospital data)	0.1646	0.0649
Any Endocrine - Stage 2 (from hospital data)	0.1840	0.0457
Any Endocrine Disease - Stage 3 (from hospital data)	0.1092	0.4222
Any Eye Disease - All stages (from hospital data)	0.1685	0.4757
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.3016	0.0049
Any Gastrointestinal Disease - Stage 2 (from hospital data)	0.2142	0.0818
Any Genitourinary Disease - Stage 2 (from hospital data)	0.1929	0.4584
Any Genitourinary Disease - Stage 3 (from hospital data)	-0.3144	0.2654
Any Male Genital System - All stages (from hospital data)	-0.4482	0.0010
Any Neurologic Disease - Stage 1 (from hospital data)	0.2517	0.1493
Any Neurologic Disease - Stage 2 (from hospital data)	0.3684	0.0263
Any Neurologic Disease - Stage 3 (from hospital data)	0.6947	0.0025
Any Respiratory Disease - Stage 2 (from hospital data)	0.3146	0.0525
Any Respiratory Disease - Stage 3 (from hospital data)	0.5118	0.0036
Neoplasm, Malignant: Melanoma-Stage 2	-1.0279	0.1274
Neoplasm, Malignant: Melanoma-Stage 3	0.8045	0.0919
Cancer (from any data source)	-0.5680	<.0001
Cancer (from drug prescription)	0.7548	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.2032	0.0068
Eye Disease (from drug prescriptions)	0.4930	0.0727
Gastrointestinal Disease (from drug prescriptions)	0.2771	0.0301

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Genitourinary Disease (from drug prescriptions)	0.5082	0.0008
Hematologic Disease (from drug prescriptions)	0.7320	<.0001
Hepatobiliary Disease (from drug prescriptions)	0.7078	<.0001
Musculoskeletal Disease (from drug prescriptions)	0.2658	0.0704
Neurologic Diseases (from drug prescriptions)	-0.1106	0.4344
Respiratory Disease (from drug prescriptions)	0.3397	0.0140
Skin Disease (from drug prescriptions)	-0.8534	0.0737
Day hospitalization	-0.3167	0.0005
Number of day hospitalizations	0.1936	0.0009
Oral anti-coagulants	0.2097	0.0073
Alpha-blockers	-0.0815	0.1273
Statins	-0.1195	0.0067
Beta-blockers	0.0758	0.0200
ACE/ARB	-0.1067	0.0015
Anti-platelets	0.2416	<.0001
Anti-arrhythmics	0.2796	<.0001
Nitrates	0.3750	<.0001
Diuretics	0.1841	<.0001
Ear,Nose,Throat Disease (from any data source)	-0.4299	0.0218
Total number of ER visits	0.1667	<.0001
Eye Disease (from any data source)	-0.5605	0.0459
Gastrointestinal Disease (from any data source)	-0.2670	0.0340
Genitourinary Disease (from any data source)	-0.2029	0.0436
Hematologic Disease (from any data source)	-0.2489	0.0209
History of Endocrine Disease (from drug prescriptions) *	0.1446	0.0036
History of Neurological Disease (from drug prescriptions) *	0.1136	0.0491
History of Aortic Stenosis-Stage 1 *	0.1778	0.2745
History of Aortic Stenosis-Stage 3 *	-0.3400	0.2752
History of Arrhythmias-Stage 2 *	0.1806	0.0180
History of Cardiomyopathies-Stage 2 *	0.2800	0.0099
History of Cardiomyopathies-Stage 3 *	0.2745	0.0450
History of Congestive Heart Failure-Stage 3 *	0.3791	0.0032
History of Coronary Artery Disease-Stage 1 *	0.2123	<.0001
History of Coronary Artery Disease-Stage 2 *	0.2602	0.0002
History of Coronary Artery Disease-Stage 3 *	0.1210	0.0479
History of Essential Hypertension-Stage 3 *	0.1309	0.1997
History of Mitral Stenosis-Stage 2 *	0.1638	0.4910
History of Pericarditis: Chronic-Stage 2 or 3 *	-0.6191	0.0980
History of Thrombophlebitis-Stage 2 or 3 *	0.2574	0.0966
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.3798	0.0282
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.0741	0.1711
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.2216	0.0034
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.2666	0.0234

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Hypothyroidism-Stage 1 *	-0.1031	0.4735
History of Crohns Disease-Stage 2 or 3 *	0.5055	0.1168
History of Diverticular Disease-Stage 1 *	0.1351	0.3093
History of Diverticular Disease-Stage 2 or 3 *	0.3463	0.1463
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.6400	0.0005
History of Neoplasm, Malignant: Stomach-Stage 1 *	0.7035	0.0151
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.7735	0.0405
History of Ulcerative Colitis-any stage *	0.1614	0.4858
History of Calculus of the Urinary Tract-Stage 1 *	0.1264	0.1702
History of Calculus of the Urinary Tract-Stage 2 or 3 *	0.3105	0.0119
History of Neoplasm, Malignant: Kidneys-Stage 1 *	-0.3428	0.0948
History of Renal Failure-Stage 2 or 3 *	0.3768	<.0001
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.4850	0.0332
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.2115	0.1691
History of Neoplasm, Malignant Hematologic-Stage 2 *	0.2152	0.3377
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.5625	0.1068
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.3497	0.0041
History of Cholecystitis and Cholelithiasis-Stage 3 *	0.1889	0.4655
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.3937	<.0001
History of Neoplasm, Malignant: Pancreas-Stage 1 *	0.3556	0.3596
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	-0.7080	0.2534
History of Pancreatitis-any stage *	0.5502	0.0003
History of Rheumatic Fever-Stage 3 *	0.6045	0.2213
History of Neoplasm, Malignant: Prostate-Stage 2 *	-0.3042	0.0181
History of Progressive Systemic Sclerosis-Stage 1 *	-1.8861	0.1260
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	1.2042	0.1939
History of Cerebrovascular Disease-Stage 1 *	0.1150	0.3529
History of Cerebrovascular Disease-Stage 2 *	0.1580	0.0735
History of Cerebrovascular Disease-Stage 3 *	0.2798	0.0003
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.5172	0.0355
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.9102	0.1461
History of Obesity-Stage 2 or 3 *	0.2153	0.0147
History of polypharmacy *	0.1251	0.0003
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.4806	0.0162
History of Depression-Stage 1 or 2 *	0.4028	<.0001
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.3627	0.009
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.6949	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.3625	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.9574	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.1444	0.4290

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2 *	0.3890	0.3220
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	0.7587	0.0080
History of Pneumonia: Bacterial-Stage 2 *	-0.9199	0.0638
History of Pneumonia: Bacterial-Stage 3 *	0.4583	0.0128
History of Neoplasm, Malignant: Melanoma-Stage 2 *	0.2292	0.4962
History of Neoplasm, Malignant: Melanoma-Stage 3 *	0.6301	0.2916
History of Oral Anti-coagulants *	0.0666	0.3850
History of Other Cardiovascular drugs *	0.0303	0.4200
History of Statins *	0.0592	0.1378
Immunologic Disease (from any data source)	0.2942	0.1479
Male Genital System (from any data source)	-0.1980	0.0015
Musculoskeletal Disease (from any data source)	-0.3422	0.0208
Neurologic Disease (from any data source)	0.2313	0.1007
Hospitalization	0.2053	0.0038
Number of hospitalizations	0.0291	0.2995
Polypharmacy	0.2601	<.0001
Psychological Disease (from any data source)	0.1041	0.0984
Respiratory Disease (from any data source)	-0.1691	0.2366
Cancer (from hospital data)	0.2023	0.3581
Cardiovascular Disease (from hospital data)	-0.0893	0.3710
Hepatobiliary (from hospital data)	-0.6082	<.0001
Musculoskeletal Disease (from hospital data)	-0.4016	0.0016
Neurologic Disease (from hospital data)	-0.6060	0.0072
Psychological Disease (from hospital data)	-0.2465	0.1059
Respiratory Disease (from hospital data)	-0.2243	0.1884
Skin Disease (from hospital data)	-0.5825	0.2155
Skin Disease (from any data source)	0.6403	0.1830
Any of the other 9 Cardiovascular drugs	0.0866	0.1859

Females 65-74

Variable	Coefficient	p-value
Intercept	-4.5504	<.0001
Number of Chronic Conditions (from any data source)=1	0.3294	0.0005
Number of Chronic Conditions (from any data source)=2	0.7012	<.0001
Number of Chronic Conditions (from any data source)=3	1.0162	<.0001
Number of Chronic Conditions (from any data source)=4	1.2243	<.0001
Number of Chronic Conditions (from any data source)=5	1.3625	<.0001
Number of Chronic Conditions (from any data source)=6 or more	1.5014	<.0001
Number of Chronic Conditions (from hospital data)=1	0.3904	<.0001
Number of Chronic Conditions (from hospital data)=2	0.4301	<.0001
Number of Chronic Conditions (from hospital data)=3	0.4970	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Number of Chronic Conditions (from hospital data)=4 or more	0.5908	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.4758	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.1393	0.1297
Number of Chronic Conditions (from drug prescriptions)=2	-0.3041	0.0058
Number of Chronic Conditions (from drug prescriptions)=3	-0.3950	0.0024
Number of Chronic Conditions (from drug prescriptions)=4	-0.4051	0.0076
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.4003	0.0256
Age on 12 /31/ 2012	-0.00012	<.0001
Cancer (from home health prescription)	0.6069	<.0001
Cardiovascular Disease (from home health prescription)	-0.2636	0.0301
Blood Diseases (from home health prescription)	0.6193	0.1479
Neurologic Diseases (from home health prescription)	0.4376	0.0016
Mental Disorders (from home health prescription)	-0.2879	0.1224
Cancer (chemo or radiation)	0.1589	0.0714
Aortic Stenosis-Stage 1	0.5120	0.0026
Cardiomyopathies-Stage 3	0.8973	0.0003
Congestive Heart Failure-Stage3	0.3086	0.0103
Coronary Artery Disease-Stage 1	-0.1015	0.2148
Coronary Artery Disease-Stage 2	0.4617	0.0005
Coronary Artery Disease-Stage 3	0.2471	0.0565
Essential Hypertension-Stage 1	-0.1363	0.0191
Essential Hypertension-Stage 3	0.2124	0.0429
Mitral Stenosis-Stage 1	-0.4732	0.0120
Thrombophlebitis-Stage 2 or 3	0.7060	<.0001
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 1	0.3844	0.0228
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3	0.8107	0.0004
Diabetes Mellitus Type 1 or Type 2-Stage 3	-0.6668	0.0022
Crohns Disease-Stage 2 or 3	0.7852	0.1300
Hernia, Hiatal or Reflux Esophagitis-Stage 1	0.1799	0.1780
Neoplasm, Malignant: Stomach-Stage 3	0.4943	0.1142
Neoplasm, Malignant: Bladder, Urinary-Stage 3	0.8348	0.0780
Neoplasm, Malignant: Breast, Female-Stage 1	-0.7289	<.0001
Neoplasm, Malignant: Breast, Female-Stage 3	-0.2375	0.1392
Anemia: Aplastic, Acquired-Stage 2 or 3	0.2640	0.1999
Neoplasm, Malignant Hematologic-Stage 1	0.2313	0.1024
Neoplasm, Malignant Hematologic-Stage 3	-1.0860	0.0017
Cirrhosis of the Liver-Stage 2 or 3	0.6134	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.7125	0.0116
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.7605	<.0001
Rheumatic Fever- Stage 2	0.3447	0.0706
Cerebrovascular Disease-Stage 3	-0.6890	0.0007
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.3887	0.2397
Bipolar Disorder - Major Depressive Episode-Stage 2 or 3	-0.7551	0.1418

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Bipolar Disorder - Manic Episode-Stage 2	0.6386	0.0329
Depression-Stage 1 or 2	-0.1537	0.1428
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	1.4878	0.0036
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.1325	0.0980
Chronic Obstructive Pulmonary Disease-Stage 3	0.5302	0.0063
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.5790	0.0005
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	0.6907	0.0477
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.7408	0.0003
Any Cancer - Stage 1 (from hospital data)	0.3101	<.0001
Any Cancer - Stage 3 (from hospital data)	1.2060	<.0001
Any Endocrine Disease - Stage 3 (from hospital data)	0.4493	0.0062
Any Eye Disease - All stages (from hospital data)	-0.1582	0.1255
Any Genitourinary Disease - Stage 2 (from hospital data)	0.3230	0.0655
Any Genitourinary Disease - Stage 3 (from hospital data)	-0.4009	0.1324
Any Gynecologic Disease - Stage 2 or 3 (from hospital data)	-0.4173	0.0548
Any Hepatobiliary Disease - Stage 1 (from hospital data)	-0.3374	0.0002
Any Hepatobiliary Disease - Stage 3 (from hospital data)	0.3881	0.1255
Any Immunologic Disease - All stages (from hospital data)	-0.5745	0.1396
Any Infectious Disease - Stage 3 (from hospital data)	0.5105	0.0511
Any Neurologic Disease - Stage 3 (from hospital data)	0.8526	<.0001
Any Psychologic Disease - Stage 2 (from hospital data)	0.3067	0.0376
Any Psychologic Disease - Stage 3 (from hospital data)	0.6691	0.0261
Any Respiratory Disease - Stage 3 (from hospital data)	0.1787	0.1601
Any Skin Disease - Stage 1 (from hospital data)	-0.2893	0.0063
Cancer (from any data source)	-0.3239	<.0001
Cancer (from drug prescriptions)	0.3190	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.3086	<.0001
Genitourinary Disease (from drug prescriptions)	0.6302	<.0001
Hematologic Disease (from drug prescriptions)	0.3673	<.0001
Hepatobiliary Disease (from drug prescriptions)	0.6321	0.0005
Musculoskeletal Disease (from drug prescriptions)	0.1877	0.0283
Respiratory Disease (from drug prescriptions)	0.1731	<.0001
Oral anti-coagulants	0.2036	0.0012
Statins	-0.1126	<.0001
ACE/ARB	-0.1348	<.0001
Anti-platelets	0.1104	0.0002
Anti-arrhythmics	0.0912	0.0986
Digitalis glycosides	0.2082	0.0046
Nitrates	0.2180	<.0001
Diuretics	0.0387	0.2255
Endocrine Disease (from any data source)	-0.1114	0.0002
Ear,Nose,Throat Disease (from any data source)	-0.4077	0.0494
Number of ER visits labeled 'Yellow'	-0.3525	<.0001
Total number of ER visits	0.5514	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Eye Disease (from any data source)	-0.2411	<.0001
Gastrointestinal Disease (from any data source)	-0.0917	0.0025
Gynecologic Disease (from any data source)	-0.4752	<.0001
History of Cancer (from drug prescriptions) *	0.1368	0.0038
History of Neurological Disease (from drug prescriptions) *	0.1374	<.0001
History of Aortic Stenosis-Stage 3 *	0.2156	0.1868
History of Arrhythmias-Stage 2 *	0.1567	0.0059
History of Cardiomyopathies-Stage 2 *	0.4598	0.0002
History of Congestive Heart Failure-Stage 3 *	0.2601	0.0033
History of Coronary Artery Disease-Stage 1 *	0.1550	0.0040
History of Coronary Artery Disease-Stage 2 *	0.1778	0.0271
History of Coronary Artery Disease-Stage 3 *	0.2409	0.0019
History of Essential Hypertension-Stage 2 *	0.1060	0.0526
History of Essential Hypertension-Stage 3 *	0.1504	0.0375
History of Mitral Stenosis-Stage 2 *	-0.2678	0.0937
History of Mitral Stenosis-Stage 3 *	0.3696	0.0337
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.2128	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.3147	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.4937	<.0001
History of Crohns Disease-Stage 2 or 3 *	0.5581	0.0954
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.4016	0.0044
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.4353	0.1847
History of Calculus of the Urinary Tract-Stage 1 *	0.2185	0.0607
History of Neoplasm, Malignant: Kidneys-Stage 3 *	-0.7174	0.1125
History of Renal Failure-Stage 2 or 3 *	0.4347	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 1 *	-0.1897	0.0064
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.2713	0.0494
History of Neoplasm, Malignant: Ovaries-Stage 1 *	0.4490	0.0239
History of Neoplasm, Malignant Hematologic-Stage 2 *	0.5429	0.0348
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.1929	0.0880
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.3420	<.0001
History of Progressive Systemic Sclerosis-Stage 2 or 3 *	0.7431	<.0001
History of Cerebrovascular Disease-Stage 1 *	0.2276	0.0043
History of Cerebrovascular Disease-Stage 3 *	0.1967	0.0029
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.4871	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.7255	0.0080
History of Obesity-Stage 2 or 3 *	0.1359	0.0870
History of polypharmacy *	0.1622	<.0001
History of Bipolar Disorder - Major Depressive Episode-Stage 2 or 3 *	0.6969	0.0010
History of Depression-Stage 1 or 2 *	0.1646	0.0034
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.8459	0.0008

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.6589	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.4628	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.8534	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.2386	0.1126
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2 *	0.6212	0.0971
History of Pneumonia: Bacterial-Stage 2 *	0.4624	0.1267
History of Pneumonia: Bacterial-Stage 3 *	0.4924	0.0003
History of Neoplasm, Malignant: Melanoma-Stage 3 *	0.8606	0.1231
History of Oral Anti-coagulants *	0.2508	<.0001
History of Other Cardiovascular drugs *	0.1002	0.0015
Musculoskeletal Disease (from any data source)	-0.3863	<.0001
Hospitalization	0.2562	<.0001
Number of hospitalizations	-0.0651	0.0068
Polypharmacy	0.2370	<.0001
Cardiovascular Disease (from hospital data)	-0.0957	0.1172
Musculoskeletal Disease (from hospital data)	-0.2090	0.0016
Any of the other 9 Cardiovascular drugs	0.0806	0.1274
Number of the other 9 Cardiovascular drugs	0.0425	0.0129

Males 65-74

Variable	Coefficient	p-value
Intercept	-4.0290	<.0001
Number of Chronic Conditions (from any data source)=1	0.3090	<.0001
Number of Chronic Conditions (from any data source)=2	0.4055	<.0001
Number of Chronic Conditions (from any data source)=3	0.6026	<.0001
Number of Chronic Conditions (from any data source)=4	0.7813	<.0001
Number of Chronic Conditions (from any data source)=5	0.7088	<.0001
Number of Chronic Conditions (from any data source)=6 or more	0.8178	<.0001
Number of Chronic Conditions (from hospital data)=1	0.2460	<.0001
Number of Chronic Conditions (from hospital data)=2	0.3676	<.0001
Number of Chronic Conditions (from hospital data)=3	0.3746	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.4616	<.0001
Number of Chronic Conditions (from home health prescription)=1 or more	0.7247	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.0302	0.6751
Number of Chronic Conditions (from drug prescriptions)=2	0.0804	0.3543
Number of Chronic Conditions (from drug prescriptions)=3	0.0770	0.4505
Number of Chronic Conditions (from drug prescriptions)=4	0.1604	0.1774
Number of Chronic Conditions (from drug prescriptions)=5 or more	0.3823	0.0066
Age on 12 /31/ 2012	-0.0001	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Cardiovascular Disease (from home health prescription)	-0.2088	0.0342
Gastrointestinal Disease (from home health prescription)	-0.9964	0.0117
Cancer (chemo or radiation)	0.4296	<.0001
Cardiomyopathies-Stage 3	0.5047	0.0007
Coronary Artery Disease-Stage 2	0.2966	<.0001
Coronary Artery Disease-Stage 3	0.2001	0.0086
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 1	0.2752	0.0019
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3	0.3313	0.0330
Hyperthyroidism-Stage 1	-0.8506	0.0059
Hyperthyroidism-Stage 2 or 3	-0.8443	0.0397
Hypothyroidism-Stage 1	-0.4005	0.0102
Diverticular Disease-Stage 1	0.3805	0.0034
Gastritis-Stage 1	0.2944	0.0367
Hernia, Hiatal or Reflux Esophagitis-Stage 1	0.3404	0.0214
Neoplasm, Malignant: Stomach-Stage 3	0.6357	0.0021
Neoplasm, Malignant: Bladder, Urinary-Stage 1	-0.3858	0.0003
Neoplasm, Malignant: Bladder, Urinary-Stage 3	1.0573	0.0003
Renal Failure-Stage 2 or 3	0.2462	0.0001
Cholecystitis and Cholelithiasis-Stage 2	0.8383	0.0010
Cirrhosis of the Liver-Stage 2 or 3	0.5557	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.8289	0.0004
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.0295	0.0026
Rheumatic Fever- Stage 2	0.4941	0.0138
Rheumatic Fever- Stage 3	0.9219	0.0348
Neoplasm, Malignant: Prostate-Stage 2	-0.5899	<.0001
Cerebrovascular Disease-Stage 1	0.2014	0.0319
Cerebrovascular Disease-Stage 3	-0.3785	0.0234
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.2581	0.1704
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.6623	<.0001
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	1.0595	<.0001
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.9074	<.0001
Pulmonary Embolism-Stage 3	0.5962	0.0059
Any Cancer - Stage 1 (from hospital data)	0.1424	0.1477
Any Cancer - Stage 3 (from hospital data)	0.9400	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.0869	0.0242
Any Ear,Nose,Throat Disease - Stage 1 (from hospital data)	-0.4741	0.0100
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.2854	<.0001
Any Hepatobiliary Disease - Stage 2 (from hospital data)	-0.5419	0.0149
Any Hepatobiliary Disease - Stage 3 (from hospital data)	0.4751	0.0179
Any Immunologic Disease - All stages (from hospital data)	0.6987	0.0029
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.2629	<.0001
Any Neurologic Disease - Stage 3 (from hospital data)	0.5546	0.0004
Any Psychologic Disease - Stage 1 (from hospital data)	-0.4107	0.0146
Any Psychologic Disease - Stage 3 (from hospital data)	-0.7275	0.0035

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any Respiratory Disease - Stage 3 (from hospital data)	0.6359	<.0001
Cancer (from any data source)	-0.3827	<.0001
Cancer (from drug prescriptions)	0.2794	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.3439	<.0001
Eye Disease (from drug prescriptions)	-0.3279	<.0001
Genitourinary Disease (from drug prescriptions)	0.2042	0.0754
Hepatobiliary Disease (from drug prescriptions)	0.5858	0.0002
Respiratory Disease (from drug prescriptions)	0.1468	<.0001
Drug-Drug interactions	0.1473	0.0128
Oral anti-coagulants	0.1161	0.0182
Statins	-0.1331	<.0001
Beta-blockers	0.0499	0.0248
ACE/ARB	-0.0997	<.0001
Anti-platelets	0.1992	<.0001
Anti-arrhythmics	0.1064	0.0130
Digitalis glycosides	0.3471	<.0001
Nitrates	0.3349	<.0001
Diuretics	0.1464	<.0001
Total number of ER visits	0.1632	<.0001
Gastrointestinal Disease (from any data source)	-0.1029	<.0001
Genitourinary Disease (from any data source)	0.3139	0.0021
History of Cancer (from drug prescriptions) *	0.1477	0.0020
History of Neurological Disease (from drug prescriptions) *	0.1240	0.0001
History of Cardiomyopathies-Stage 2 *	0.3116	<.0001
History of Cardiomyopathies-Stage 3 *	0.4397	<.0001
History of Congestive Heart Failure-Stage 3 *	0.3374	<.0001
History of Coronary Artery Disease-Stage 1 *	0.2568	<.0001
History of Essential Hypertension-Stage 2 *	0.1079	0.0140
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.1006	0.0031
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.1970	0.0001
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.3887	0.0005
History of Renal Failure-Stage 2 or 3 *	0.3301	<.0001
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.2586	0.0009
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.3021	<.0001
History of Neoplasm, Malignant: Prostate-Stage 3 *	0.3601	0.0252
History of Cerebrovascular Disease-Stage 1 *	0.1414	0.0309
History of Cerebrovascular Disease-Stage 3 *	0.2819	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.6145	<.0001
History of polypharmacy *	0.1051	<.0001
History of Bipolar Disorder - Manic Episode-Stage 1 *	0.6800	0.0265
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1 *	0.3941	0.0071
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.4235	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.3669	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.7646	<.0001
History of Pneumonia: Bacterial-Stage 3 *	0.3809	0.0002
History of Oral Anti-coagulants *	0.1253	0.0059
History of Other Cardiovascular drugs *	0.1019	0.0003
Male Genital System (from any data source)	-0.3918	<.0001
Musculoskeletal Disease (from any data source)	-0.1502	<.0001
Hospitalization	0.1762	0.0003
Number of hospitalizations	0.0176	0.3845
Polypharmacy	0.2522	<.0001
Cancer (from hospital data)	0.3120	0.0130
Genitourinary Disease (from hospital data)	-0.4713	<.0001
Psychological Disease (from hospital data)	0.5748	0.0003

Females 75-84

Variable	Coefficient	p-value
Intercept	-2.1966	0.0983
Number of Chronic Conditions (from any data source)=1	0.3204	<.0001
Number of Chronic Conditions (from any data source)=2	0.5216	<.0001
Number of Chronic Conditions (from any data source)=3	0.6879	<.0001
Number of Chronic Conditions (from any data source)=4	0.8108	<.0001
Number of Chronic Conditions (from any data source)=5	0.8615	<.0001
Number of Chronic Conditions (from any data source)=6 or more	0.9003	<.0001
Number of Chronic Conditions (from hospital data)=1	0.2734	<.0001
Number of Chronic Conditions (from hospital data)=2	0.4760	<.0001
Number of Chronic Conditions (from hospital data)=3	0.4602	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.5605	<.0001
Number of Chronic Conditions (from home health prescription)=1	0.4539	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.5024	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.2927	<.0001
Number of Chronic Conditions (from drug prescriptions)=2	-0.4346	<.0001
Number of Chronic Conditions (from drug prescriptions)=3	-0.5230	<.0001
Number of Chronic Conditions (from drug prescriptions)=4	-0.5943	<.0001
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.6388	<.0001
Age on 12/ 31/ 2012	-0.0494	0.0537
Cancer (from home health prescription)	0.5074	<.0001
Cardiovascular Disease (from home health prescription)	-0.1174	0.0117
Respiratory Diseases (from home health prescription)	0.4899	<.0001
Cancer (chemo or radiation)	0.3093	0.0003
Genitourinary (dialysis)	0.3668	0.0018
Aneurysm, Thoracic-all stages	0.4180	0.0286

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Aortic Stenosis-Stage 1	0.2859	0.0055
Arrhythmias-Stage 3	-0.4014	0.0049
Coronary Artery Disease-Stage 2	0.3630	<.0001
Essential Hypertension-Stage 3	-0.1678	0.0129
Mitral Stenosis-Stage 2	-0.3215	0.0749
Thrombophlebitis-Stage 1	0.3956	0.0024
Thrombophlebitis-Stage 2 or 3	-0.3711	0.0013
Hyperthyroidism-Stage 1	-0.3290	0.0847
Crohns Disease-Stage 1	0.7565	0.0329
Functional Digestive Disorders-Stage 1	0.2509	0.0124
Gastritis-Stage 1	0.3782	0.0002
Neoplasm, Malignant: Kidneys-Stage 3	-0.7235	0.1071
Renal Failure-Stage 2 or 3	0.2070	0.0002
Neoplasm, Malignant: Breast, Female-Stage 1	-0.3991	0.0001
Neoplasm, Malignant: Breast, Female-Stage 3	-0.4149	0.0077
Neoplasm, Malignant: Ovaries-Stage 1	0.5296	0.0693
Anemia: Aplastic, Acquired-Stage 2 or 3	0.3538	0.0227
Cirrhosis of the Liver-Stage 2 or 3	0.6150	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	0.8098	0.0011
Neoplasm, Malignant: Pancreas-Stage 2 or 3	0.7321	0.0866
Progressive Systemic Sclerosis-Stage 1	0.5340	0.0203
Cerebrovascular Disease-Stage 1	0.1406	0.0941
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.3594	<.0001
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.2377	0.0716
Bipolar Disorder - Manic Episode-Stage 1	-0.9145	0.1201
Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.2567	0.0037
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.5431	0.0015
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	0.9029	0.0394
Pneumonia: Bacterial-Stage 1	0.3367	0.0003
Pneumonia: Bacterial-Stage 3	-0.3613	0.0110
Pulmonary Embolism-Stage 3	-0.3302	0.0333
Any Cancer - Stage 1 (from hospital data)	0.2294	0.0003
Any Cancer - Stage 2 (from hospital data)	0.1477	0.1008
Any Cancer - Stage 3 (from hospital data)	1.3077	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.1001	0.0026
Any Cardiovascular - Stage 3 (from hospital data)	0.3030	<.0001
Any Eye Disease - All stages (from hospital data)	-0.1491	0.0443
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.2093	<.0001
Any Hemotologic Disease - Stage 3 (from hospital data)	0.4744	0.0174
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.1645	<.0001
Any Neurologic Disease - Stage 1 (from hospital data)	-0.1809	0.0076
Any Neurologic Disease - Stage 3 (from hospital data)	0.3131	<.0001
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2321	0.0008
Any Respiratory Disease – Stage 1 (from hospital data)	-0.1657	0.0476

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Any Respiratory Disease - Stage 3 (from hospital data)	0.5038	<.0001
Cancer (from any data source)	-0.1228	0.0560
Cancer (from drug prescriptions)	0.2540	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.2097	0.0046
Endocrine Disease (from drug prescriptions)	-0.0828	0.0222
Genitourinary Disease (from drug prescriptions)	0.4347	<.0001
Hematologic Disease (from drug prescriptions)	0.1757	0.0011
Hepatobiliary Disease (from drug prescriptions)	0.6510	0.0006
Neurologic Diseases (from drug prescriptions)	0.1086	0.0014
Respiratory Disease (from drug prescriptions)	0.1865	<.0001
Cardiovascular Disease (from any data source)	-0.0770	0.2400
Day hospitalization	-0.1056	0.0250
Oral anti-coagulants	0.1976	<.0001
Alpha-blockers	-0.1014	0.0038
Statins	-0.1832	<.0001
Beta-blockers	-0.0628	0.0031
ACE/ARB	-0.1956	<.0001
Calcium channel blockers	-0.0715	0.0007
Nitrates	0.1676	<.0001
Number of ER visits labeled 'Red'	0.2174	<.0001
Total number of ER visits	0.2167	<.0001
Eye Disease (from any data source)	-0.1462	<.0001
Gynecologic Disease (from any data source)	-0.4041	<.0001
History of Cancer (from drug prescriptions) *	0.1108	0.0048
History of Endocrine Disease (from drug prescriptions) *	0.1042	0.0016
History of Neurological Disease (from drug prescriptions) *	0.0915	0.0011
History of Psychological Disease (from drug prescriptions) *	0.0506	0.0295
History of Respiratory Disease (from drug prescriptions) *	0.0905	0.0002
History of Aortic Stenosis-Stage 3 *	0.1616	0.0515
History of Arrhythmias-Stage 2 *	0.1000	0.0015
History of Cardiomyopathies-Stage 2 *	0.1681	0.0697
History of Cardiomyopathies-Stage 3 *	0.3151	0.0108
History of Congestive Heart Failure-Stage 3 *	0.3407	<.0001
History of Coronary Artery Disease-Stage 1 *	0.1753	<.0001
History of Coronary Artery Disease-Stage 2 *	0.1517	0.0017
History of Coronary Artery Disease-Stage 3 *	0.1935	<.0001
History of Essential Hypertension-Stage 1 *	0.0413	0.0434
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.1886	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.1483	0.0027
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.1335	0.0715
History of Hypothyroidism-Stage 2 or 3 *	0.1892	0.0534
History of Diverticular Disease-Stage 2 or 3 *	0.1638	0.0572
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.7712	0.0010
History of Neoplasm, Malignant: Bladder, Urinary-Stage 3 *	1.3084	0.0182

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Neoplasm, Malignant: Kidneys-Stage 3 *	0.7219	0.0444
History of Renal Failure-Stage 2 or 3 *	0.2398	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 1 *	-0.2279	0.0002
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.3479	0.0056
History of Neoplasm, Malignant: Ovaries-Stage 1 *	0.3945	0.0551
History of Neoplasm, Malignant: Ovaries-Stage 2 or 3 *	0.4745	0.0314
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.2900	0.0103
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.4900	0.0545
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.1010	0.0831
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.1793	0.0089
History of Cholecystitis and Cholelithiasis-Stage 3 *	0.2803	0.0315
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.2739	<.0001
History of Rheumatic Fever-Stage 3 *	0.4281	0.0005
History of Progressive Systemic Sclerosis-Stage 1 *	0.4962	0.0004
History of Cerebrovascular Disease-Stage 1 *	0.1710	<.0001
History of Cerebrovascular Disease-Stage 2 *	0.1318	0.0003
History of Cerebrovascular Disease-Stage 3 *	0.2019	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.3755	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.4162	<.0001
History of Obesity-Stage 2 or 3 *	0.1187	0.1089
History of polypharmacy *	0.1086	<.0001
History of Bipolar Disorder - Manic Episode-Stage 2 *	0.3557	0.0356
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.2343	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.6055	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.3311	0.0150
History of Pneumonia: Bacterial-Stage 3 *	0.2981	0.0013
History of Oral Anti-coagulants *	0.1008	0.0057
History of Other Cardiovascular drugs *	0.1223	<.0001
Musculoskeletal Disease (from any data source)	-0.0624	0.0065
Hospitalization	0.1280	0.0248
Number of hospitalizations	-0.0643	0.0011
Polypharmacy	0.1500	<.0001
Psychological Disease (from any data source)	0.1240	<.0001
Any of the other 9 Cardiovascular drugs	0.0679	0.1160
Number of the other 9 Cardiovascular drugs	0.1028	<.0001

Males 75-84

Variable	Coefficient	p-value
Intercept	-1.0190	0.4621
Number of Chronic Conditions (from any data source)=1	0.2984	0.0003

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Number of Chronic Conditions (from any data source)=2	0.5009	<.0001
Number of Chronic Conditions (from any data source)=3	0.5987	<.0001
Number of Chronic Conditions (from any data source)=4	0.7284	<.0001
Number of Chronic Conditions (from any data source)=5	0.7507	<.0001
Number of Chronic Conditions (from any data source)=6 or more	0.8596	<.0001
Number of Chronic Conditions (from hospital data)=1	0.1570	0.0049
Number of Chronic Conditions (from hospital data)=2	0.1638	0.0317
Number of Chronic Conditions (from hospital data)=3	0.1457	0.1301
Number of Chronic Conditions (from hospital data)=4 or more	0.2159	0.0857
Number of Chronic Conditions (from home health prescription)=1	0.3898	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.2645	0.0832
Number of Chronic Conditions (from drug prescriptions)=1	-0.2228	0.0042
Number of Chronic Conditions (from drug prescriptions)=2	-0.2887	0.0012
Number of Chronic Conditions (from drug prescriptions)=3	-0.2736	0.0079
Number of Chronic Conditions (from drug prescriptions)=4	-0.3649	0.0023
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.3187	0.0260
Reside in Mountain area on 12/ 31/ 2012	-0.0273	0.4157
Reside in Hill area on 12/ 31/ 2012	0.0419	0.0161
Age on 12/ 31/ 2012	-0.0638	0.0170
Cancer (from home health prescription)	0.6146	<.0001
Genitourinary Disease (from home health prescription)	-0.2349	0.2216
Blood Diseases (from home health prescription)	0.5082	0.1476
Infectious Disease (from home health prescription)	0.3613	0.1489
Neurologic Diseases (from home health prescription)	0.2858	0.0007
Mental Disorders (from home health prescription)	0.1399	0.1595
Respiratory Diseases (from home health prescription)	0.1985	0.1585
Cancer (chemo or radiation)	0.2160	0.0058
Genitourinary (dialysis)	0.1829	0.1227
Aneurysm, Thoracic-all stages	-0.1892	0.2250
Aortic Stenosis-Stage 1	0.1531	0.1625
Aortic Stenosis-Stage 3	0.1564	0.2244
Arrhythmias-Stage 2	-0.1538	0.0208
Cardiomyopathies-Stage 2	-0.1466	0.2544
Cardiomyopathies-Stage 3	0.3133	0.0224
Conduction Disorders-all stages	-0.1835	0.0881
Congestive Heart Failure-Stage3	0.2137	0.0021
Coronary Artery Disease-Stage 2	0.0566	0.4552
Coronary Artery Disease-Stage 3	0.1414	0.0418
Essential Hypertension-Stage 1	-0.1169	0.0073
Essential Hypertension-Stage 2	-0.2785	0.0001
Essential Hypertension-Stage 3	0.1298	0.0362
Infective Endocarditis-Stage 3	0.9736	0.0197

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Mitral Stenosis-Stage 2	-0.3114	0.1333
Pericarditis: Viral or Traumatic-Stage 2 or 3	-0.8469	0.0349
Thrombophlebitis-Stage 1	0.2275	0.1472
Thrombophlebitis-Stage 2 or 3	-0.1048	0.4208
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 1	0.1183	0.1409
Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3	0.1645	0.2385
Diabetes Mellitus Type 1 or Type 2-Stage 2	-0.1939	0.0562
Diabetes Mellitus Type 1 or Type 2-Stage 3	-0.4025	0.0333
Hyperthyroidism-Stage 2 or 3	-0.2385	0.4296
Hypothyroidism-Stage 1	-0.1541	0.2322
Crohns Disease-Stage 1	0.8307	0.0454
Functional Digestive Disorders-Stage 1	0.1566	0.1709
Hernia, Hiatal or Reflux Esophagitis-Stage 2 or 3	0.7272	0.0040
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.4353	0.0125
Neoplasm, Malignant: Colon and Rectum-Stage 3	-0.2956	0.1051
Neoplasm, Malignant: Stomach-Stage 1	0.5695	0.0019
Neoplasm, Malignant: Stomach-Stage 3	0.3811	0.1781
Calculus of the Urinary Tract-Stage 1	-0.1383	0.3085
Neoplasm, Malignant: Bladder, Urinary-Stage 3	0.2547	0.4352
Neoplasm, Malignant: Kidneys-Stage 1	0.1778	0.2881
Renal Failure-Stage 2 or 3	0.3215	<.0001
Anemia: Aplastic, Acquired-Stage 2 or 3	0.2550	0.2048
Neoplasm, Malignant Hematologic-Stage 1	0.1501	0.1974
Neoplasm, Malignant Hematologic-Stage 3	-0.8912	0.0050
Cholecystitis and Cholelithiasis-Stage 2	0.4441	0.0408
Cirrhosis of the Liver-Stage 2 or 3	0.5304	<.0001
Neoplasm, Malignant: Pancreas-Stage 1	1.1048	<.0001
Neoplasm, Malignant: Pancreas-Stage 2 or 3	1.2846	0.0009
Rheumatic Fever- Stage 2	0.2675	0.0866
Rheumatic Fever- Stage 3	0.2708	0.3136
Neoplasm, Malignant: Prostate-Stage 2	-0.3625	0.0290
Neoplasm, Malignant: Prostate-Stage 3	-0.2814	0.1619
Progressive Systemic Sclerosis-Stage 1	1.1059	0.1109
Cerebrovascular Disease-Stage 1	0.1684	0.0636
Cerebrovascular Disease-Stage 2	-0.2728	0.0094
Cerebrovascular Disease-Stage 3	-0.2874	0.0742
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.4422	0.0018
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.7034	0.0002
Epilepsy-all stages	-0.3392	0.0337
Bipolar Disorder - Manic Episode-Stage 1	0.6002	0.4877
Depression-Stage 1 or 2	-0.1468	0.3176
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 1	1.2556	0.0028
Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3	0.1890	0.3351

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Chronic Obstructive Pulmonary Disease-Stage 1 or 2	0.0777	0.0884
Chronic Obstructive Pulmonary Disease-Stage 3	0.1545	0.2047
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	0.7063	<.0001
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 2	0.3644	0.2398
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3	0.3851	0.0755
Pulmonary Embolism-Stage 3	-0.2243	0.2238
Any Cancer - Stage 2 (from hospital data)	0.4412	0.0019
Any Cancer - Stage 3 (from hospital data)	1.2233	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	-0.0637	0.2223
Any Cardiovascular Disease - Stage 2 (from hospital data)	0.1167	0.1053
Any Endocrine Disease - Stage 1 (from hospital data)	-0.1650	0.0986
Any Endocrine Disease - Stage 3 (from hospital data)	0.3361	0.0522
Any Ear,Nose,Throat Disease - Stage 1 (from hospital data)	-0.3128	0.1174
Any Ear,Nose,Throat Disease - Stage 2 (from hospital data)	-1.4375	0.0217
Any Eye Disease - All stages (from hospital data)	-0.2230	0.0012
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.1736	0.0372
Any Genitourinary Disease - Stage 1 (from hospital data)	0.2720	0.1932
Any Genitourinary Disease - Stage 2 (from hospital data)	0.3593	0.0553
Any Genitourinary Disease - Stage 3 (from hospital data)	0.4935	0.0195
Any Hemotologic Disease - Stage 3 (from hospital data)	0.3202	0.1946
Any Hepatobiliary Disease - Stage 2 (from hospital data)	-0.4870	0.0127
Any Immunologic Disease - All stages (from hospital data)	0.3766	0.3824
Any Infectious Disease - Stage 3 (from hospital data)	0.2171	0.2273
Any Male Genital System - All stages (from hospital data)	-0.1061	0.0692
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.1530	0.0217
Any Neurologic Disease - Stage 1 (from hospital data)	-0.1396	0.1837
Any Neurologic Disease - Stage 2 (from hospital data)	0.2235	0.0339
Any Neurologic Disease - Stage 3 (from hospital data)	0.3632	0.0284
Any Psychologic Disease - Stage 1 (from hospital data)	-0.2738	0.0487
Any Psychologic Disease - Stage 3 (from hospital data)	-0.3812	0.0297
Any Respiratory Disease - Stage 2 (from hospital data)	0.2521	0.0012
Any Respiratory Disease - Stage 3 (from hospital data)	0.4709	<.0001
Any Skin Disease - Stage 1 (from hospital data)	-0.2014	0.1407
Any Skin Disease - Stage 2 (from hospital data)	-0.4164	0.0263
Neoplasm, Malignant: Melanoma-Stage 2	-0.5839	0.1101
Neoplasm, Malignant: Melanoma-Stage 3	0.6833	0.1751
Cancer (from drug prescriptions)	0.1681	<.0001
Cardiovascular Disease (from drug prescriptions)	-0.2349	<.0001
Endocrine Disease (from drug prescriptions)	-0.0963	0.2515
Eye Disease (from drug prescriptions)	-0.1283	<.0001
Genitourinary Disease (from drug prescriptions)	0.4556	<.0001
Hematologic Disease (from drug prescriptions)	0.1419	0.0185
Hepatobiliary Disease (from drug prescriptions)	0.5789	0.0029

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Musculoskeletal Disease (from drug prescriptions)	0.0610	0.4251
Neurologic Diseases (from drug prescriptions)	0.0903	0.0217
Psychological Disease (from drug prescriptions)	-0.0954	0.2745
Respiratory Disease (from drug prescriptions)	0.1867	<.0001
Skin Disease (from drug prescriptions)	-0.3674	0.0144
Number of day hospitalizations	-0.0412	0.1517
Oral anti-coagulants	0.1213	0.0014
Alpha-blockers	-0.0742	0.0360
Statins	-0.1293	<.0001
ACE/ARB	-0.1708	<.0001
Anti-platelets	0.0372	0.1632
Calcium channel blockers	-0.0590	0.0133
Digitalis glycosides	0.1134	0.0086
Nitrates	0.2066	<.0001
Diuretics	0.0565	0.0325
Endocrine Disease (from any data source)	0.0742	0.3637
Number of ER visits labeled 'Yellow'	-0.1352	0.0027
Total number of ER visits	0.3208	<.0001
Genitourinary Disease (from any data source)	0.1205	0.3247
History of Endocrine Disease (from drug prescriptions) *	0.0680	0.0589
History of Neurological Disease (from drug prescriptions) *	0.0426	0.2080
History of Psychological Disease (from drug prescriptions) *	0.0262	0.3867
History of Respiratory Disease (from drug perscriptions) *	0.0918	0.0003
History of Aortic Stenosis-Stage 1 *	-0.1295	0.0581
History of Aortic Stenosis-Stage 3 *	0.0889	0.2974
History of Arrhythmias-Stage 2 *	0.0850	0.0061
History of Arrhythmias-Stage 3 *	0.1402	0.0717
History of Cardiomyopathies-Stage 2 *	0.1934	0.0028
History of Cardiomyopathies-Stage 3 *	0.2510	0.0033
History of Congestive Heart Failure-Stage 3 *	0.2257	<.0001
History of Coronary Artery Disease-Stage 1 *	0.1133	<.0001
History of Coronary Artery Disease-Stage 2 *	0.1341	0.0004
History of Coronary Artery Disease-Stage 3 *	0.0377	0.3346
History of Essential Hypertension-Stage 2 *	-0.0501	0.1298
History of Essential Hypertension-Stage 3 *	0.1584	<.0001
History of Infective Endocarditis-Stage 3 *	-0.3815	0.1951
History of Mitral Stenosis-Stage 3 *	0.1789	0.1632
History of Pericarditis: Chronic-Stage 2 or 3 *	-0.1570	0.3956
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	0.2146	0.2431
History of Thrombophlebitis-Stage 2 or 3 *	-0.1115	0.1430
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.0900	0.3217
hx_Drug-Drug interactions	0.0611	0.0593

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.0914	0.0047
History of Diabetes Mellitus Type 1 or Type 2-Stage 2 *	0.3099	<.0001
History of Diabetes Mellitus Type 1 or Type 2-Stage 3 *	0.1290	0.0545
History of Hypothyroidism-Stage 2 or 3 *	0.2294	0.1961
History of Crohns Disease-Stage 2 or 3 *	0.8327	0.0190
History of Diverticular Disease-Stage 1 *	0.1541	0.0060
History of Diverticular Disease-Stage 2 or 3 *	0.1222	0.1954
History of Neoplasm, Malignant: Colon and Rectum-Stage 3 *	0.4265	<.0001
History of Neoplasm, Malignant: Stomach-Stage 3 *	0.6959	0.0011
History of Neoplasm, Malignant: Bladder, Urinary-Stage 1*	-0.0471	0.3540
History of Neoplasm, Malignant: Bladder, Urinary-Stage 3 *	0.3138	0.3404
History of Neoplasm, Malignant: Kidneys-Stage 3 *	0.3046	0.3079
History of Renal Failure-Stage 2 or 3 *	0.1469	<.0001
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.2973	0.0315
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.2804	0.0030
History of Neoplasm, Malignant Hematologic-Stage 2 *	-0.1374	0.3846
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.1693	0.0044
History of Cholecystitis and Cholelithiasis-Stage 2 *	0.1182	0.0785
History of Cholecystitis and Cholelithiasis-Stage 3 *	0.1319	0.2852
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.2384	0.0009
History of Neoplasm, Malignant: Pancreas-Stage 1 *	-0.5740	0.0601
History of Neoplasm, Malignant: Pancreas-Stage 2 or 3 *	0.5440	0.3348
History of Rheumatic Fever-Stage 2 *	0.0840	0.3429
History of Neoplasm, Malignant: Prostate-Stage 3 *	0.1568	0.3460
History of Progressive Systemic Sclerosis-Stage 1 *	0.4957	0.3135
History of Cerebrovascular Disease-Stage 1 *	0.1907	<.0001
History of Cerebrovascular Disease-Stage 2 *	0.1589	<.0001
History of Cerebrovascular Disease-Stage 3 *	0.1081	0.0017
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.3414	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.6441	<.0001
History of Potentially inappropriate prescribing - always to be avoided drugs *	-0.0564	0.0053
History of polypharmacy *	0.1075	<.0001
History of Depression-Stage 1 or 2 *	0.0912	0.1498
History of Drug Abuse, Dependence, Intoxication: Alcohol-Stage 2 or 3 *	0.4811	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.2667	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.5780	<.0001
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.3130	0.0002
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 3 *	0.3404	0.1018

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Pneumonia: Bacterial-Stage 2 *	0.2648	0.1229
History of Pneumonia: Bacterial-Stage 3 *	0.2801	0.0001
History of Pulmonary Embolism-Stage 3 *	-0.1687	0.1666
History of Neoplasm, Malignant: Melanoma-Stage 3 *	-0.6881	0.1493
History of Oral Anti-coagulants *	0.0418	0.2647
History of Other Cardiovascular drugs *	0.1298	<.0001
Male Genital System (from any data source)	-0.2023	<.0001
Musculoskeletal Disease (from any data source)	-0.1057	0.1679
Hospitalization	0.2060	<.0001
Number of hospitalizations	-0.0217	0.2903
Polypharmacy	0.0913	0.0005
Psychological Disease (from any data source)	0.2569	0.0032
Cardiovascular Disease (from hospital data)	0.0736	0.2121
Endocrine Disease (from hospital data)	0.1978	0.0758
Gastrointestinal Disease (from hospital data)	-0.0878	0.2559
Genitourinary Disease (from hospital data)	-0.4235	0.0770
Hepatobiliary (from hospital data)	-0.1192	0.0959
Neurologic Disease (from hospital data)	0.0927	0.3439
Skin Disease (from any data source)	0.2454	0.0506
Any of the other 9 Cardiovascular drugs	0.0558	0.2144
Number of the other 9 Cardiovascular drugs	0.0680	<.0001

Females 85 and over

Variable	Coefficient	p-value
Intercept	-4.4541	<.0001
Number of Chronic Conditions (from any data source)=1	0.1340	0.0095
Number of Chronic Conditions (from any data source)=2	0.2527	<.0001
Number of Chronic Conditions (from any data source)=3	0.3482	<.0001
Number of Chronic Conditions (from any data source)=4	0.3660	<.0001
Number of Chronic Conditions (from any data source)=5	0.3330	0.0004
Number of Chronic Conditions (from any data source)=6 or more	0.3306	0.0032
Number of Chronic Conditions (from hospital data)=1	-0.0564	0.3573
Number of Chronic Conditions (from hospital data)=2	-0.0127	0.8565
Number of Chronic Conditions (from hospital data)=3	0.1268	0.0996
Number of Chronic Conditions (from hospital data)=4 or more	0.1910	0.0340
Number of Chronic Conditions (from home health prescription)=1	0.3946	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.6315	<.0001
Number of Chronic Conditions (from drug prescriptions)=1	-0.0973	0.0439
Number of Chronic Conditions (from drug prescriptions)=2	-0.1674	0.0036
Number of Chronic Conditions (from drug prescriptions)=3	-0.1843	0.0057

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Number of Chronic Conditions (from drug prescriptions)=4	-0.0790	0.3133
Number of Chronic Conditions (from drug prescriptions)=5 or more	-0.0275	0.7718
Age on 12 /31/ 2012	-0.0002	<.0001
Cardiovascular Disease (from home health prescription)	-0.0967	0.0033
Blood Diseases (from home health prescription)	-0.3867	0.0183
Respiratory Diseases (from home health prescription)	-0.3093	0.0020
Skin Disease (from home health prescription)	0.2047	0.0023
Genitourinary (dialysis)	0.8709	0.0002
Aortic Stenosis-Stage 3	0.2778	0.0107
Conduction Disorders-all stages	-0.2408	0.0266
Congestive Heart Failure-Stage3	0.1248	0.0454
Coronary Artery Disease-Stage 1	-0.1922	0.0014
Coronary Artery Disease-Stage 2	0.2265	0.0047
Essential Hypertension-Stage 1	-0.3230	<.0001
Essential Hypertension-Stage 2	-0.1393	0.0052
Mitral Stenosis-Stage 2	-0.4336	0.0440
Diabetes Mellitus Type 1 or Type 2-Stage 2	-0.3893	0.0295
Functional Digestive Disorders-Stage 1	0.2998	0.0008
Hernia, Hiatal or Reflux Esophagitis-Stage 1	-0.2273	0.0578
Neoplasm, Malignant: Colon and Rectum-Stage 2	-0.9583	0.0003
Renal Failure-Stage 2 or 3	0.1598	0.0010
Neoplasm, Malignant: Breast, Female-Stage 1	-0.4438	0.0011
Anemia: Aplastic, Acquired-Stage 2 or 3	0.9430	<.0001
Rheumatic Fever- Stage 2	0.2483	0.0251
Rheumatic Fever- Stage 3	0.7922	<.0001
Cerebrovascular Disease-Stage 3	0.7130	0.0005
Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3	0.2095	0.0381
Bipolar Disorder - Manic Episode-Stage 2	1.3679	0.0661
Any Cancer - Stage 2 (from hospital data)	1.1401	<.0001
Any Cancer - Stage 3 (from hospital data)	0.9674	<.0001
Any Cardiovascular Disease - Stage 1 (from hospital data)	0.2126	0.0002
Any Cardiovascular - Stage 3 (from hospital data)	0.1855	<.0001
Any Endocrine - Stage 2 (from hospital data)	0.4821	0.0015
Any Gastrointestinal Disease - Stage 1 (from hospital data)	-0.2081	0.0003
Any Hemotologic Disease - Stage 3 (from hospital data)	0.3382	0.0529
Any Musculoskeletal Disease - Stage 1 (from hospital data)	-0.0783	0.0438
Any Neurologic Disease - Stage 3 (from hospital data)	-0.4771	0.0171
Any Respiratory Disease - Stage 3 (from hospital data)	0.2027	0.0040
Neoplasm, Malignant: Melanoma-Stage 2	-1.4509	0.0019
Cardiovascular Disease (from drug prescriptions)	-0.2006	<.0001
Endocrine Disease (from drug prescriptions)	-0.1009	0.0041
Genitourinary Disease (from drug prescriptions)	0.2622	0.0004
Hematologic Disease (from drug prescriptions)	0.1242	0.0087
Hepatobiliary Disease (from drug prescriptions)	0.8195	0.0083

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Oral anti-coagulants	-0.0905	0.0222
Statins	-0.2123	<.0001
ACE/ARB	-0.1098	<.0001
Calcium channel blockers	-0.0802	0.0002
Anti-arrhythmics	0.0988	0.0106
Digitalis glycosides	0.1114	0.0002
Nitrates	0.1279	<.0001
Diuretics	0.0611	0.0088
Number of ER visits labeled 'Yellow'	-0.2129	0.0004
Total number of ER visits	0.4649	<.0001
Eye Disease (from any data source)	-0.1404	<.0001
History of Cancer (from drug prescriptions) *	0.1381	0.0002
History of Endocrine Disease (from drug prescriptions) *	0.1108	0.0005
History of Psychological Disease (from drug prescriptions) *	0.0824	<.0001
History of Respiratory Disease (from drug prescriptions) *	0.0686	0.0043
History of Aortic Stenosis-Stage 1 *	0.1279	0.0492
History of Aortic Stenosis-Stage 3 *	0.3272	<.0001
History of Arrhythmias-Stage 2 *	0.1359	<.0001
History of Cardiomyopathies-Stage 3 *	0.4752	0.0004
History of Congestive Heart Failure-Stage 3 *	0.1905	<.0001
History of Coronary Artery Disease-Stage 1 *	0.0616	0.0342
History of Coronary Artery Disease-Stage 2 *	0.1270	0.0059
History of Essential Hypertension-Stage 3 *	0.1426	<.0001
History of Mitral Stenosis-Stage 2 *	0.2121	0.0393
History of Mitral Stenosis-Stage 3 *	0.2109	0.0254
History of Tibial/Iliac/Femoral/Popliteal Artery Disease-Stage 2 or 3 *	0.1684	0.0609
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.0983	0.0030
History of Neoplasm, Malignant: Kidneys-Stage 1 *	0.3075	0.0501
History of Renal Failure-Stage 2 or 3 *	0.1614	<.0001
History of Neoplasm, Malignant: Breast, Female-Stage 3 *	0.5190	0.0032
History of Anemia: Aplastic, Acquired-Stage 2 or 3 *	0.3169	0.0038
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.2489	0.0263
History of Cholecystitis and Cholelithiasis-Stage 1 *	0.1278	0.0084
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.1977	0.0150
History of Rheumatic Fever-Stage 2 *	0.1641	0.0171
History of Rheumatic Fever-Stage 3 *	0.4295	<.0001
History of Cerebrovascular Disease-Stage 1 *	0.1780	<.0001
History of Cerebrovascular Disease-Stage 2 *	0.1248	0.0001
History of Cerebrovascular Disease-Stage 3 *	0.2010	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.2327	<.0001
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.3112	<.0001
History of polypharmacy *	0.0412	0.0637

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.1641	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.2718	0.0010
History of Pneumonia: Bacterial-Stage 3 *	0.2305	0.0024
History of Oral Anti-coagulants *	0.1412	0.0001
History of Other Cardiovascular drugs *	0.0945	0.0002
Musculoskeletal Disease (from any data source)	-0.1212	<.0001
Hospitalization	0.4053	<.0001
Number of hospitalizations	-0.0572	0.0152
Polypharmacy	0.0686	0.0047
Psychological Disease (from any data source)	0.0241	0.3329
Respiratory Disease (from any data source)	0.1452	<.0001
Cancer (from hospital data)	0.1946	0.0015
Number of the other 9 Cardiovascular drugs	0.0343	0.0067

Males 85 and over

Variable	Coefficient	p-value
Intercept	-6.5943	<.0001
Number of Chronic Conditions (from any data source)=1	0.0148	0.7836
Number of Chronic Conditions (from any data source)=2	0.0567	0.3273
Number of Chronic Conditions (from any data source)=3	0.1133	0.0832
Number of Chronic Conditions (from any data source)=4	0.1108	0.1382
Number of Chronic Conditions (from any data source)=5	0.1027	0.2348
Number of Chronic Conditions (from any data source)=6 or more	-0.00058	0.9956
Number of Chronic Conditions (from hospital data)=1	0.2260	<.0001
Number of Chronic Conditions (from hospital data)=2	0.3490	<.0001
Number of Chronic Conditions (from hospital data)=3	0.3454	<.0001
Number of Chronic Conditions (from hospital data)=4 or more	0.2443	0.0024
Number of Chronic Conditions (from home health prescription)=1	0.5620	<.0001
Number of Chronic Conditions (from home health prescription)=2 or more	0.9112	<.0001
Age on 12/ 31/ 2012	0.0552	<.0001
Cardiovascular Disease (from home health prescription)	-0.1808	0.0009
Neurologic Diseases (from home health prescription)	-0.2313	0.0125
Cancer (chemo or radiation)	0.4807	0.0028
Aortic Stenosis-Stage 3	0.3770	0.0178
Coronary Artery Disease-Stage 2	0.1810	0.0345
Essential Hypertension-Stage 1	-0.1334	0.0049
Mitral Stenosis-Stage 3	0.7232	0.0092
Crohns Disease-Stage 2 or 3	2.0123	0.0823
Renal Failure-Stage 2 or 3	0.2166	<.0001
Neoplasm, Malignant Hematologic-Stage 2	0.9737	0.0279
Cerebrovascular Disease-Stage 2	-0.3039	0.0005

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1	0.4636	0.0033
Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1	1.0085	<.0001
Any Cancer - Stage 3 (from hospital data)	1.2151	<.0001
Any Cardiovascular - Stage 3 (from hospital data)	0.1683	0.0006
Any Eye Disease – All stages (from hospital data)	-0.3186	0.0024
Any Psychologic Disease - Stage 1 (from hospital data)	-0.3000	0.0483
Any Psychologic Disease - Stage 2 (from hospital data)	0.3406	0.0468
Cancer (from any data source)	0.0881	0.0255
Cardiovascular Disease (from drug prescriptions)	-0.0958	0.0203
Eye Disease (from drug prescriptions)	-0.1138	0.0019
Genitourinary Disease (from drug prescriptions)	0.3297	0.0004
Hepatobiliary Disease (from drug prescriptions)	1.1786	0.0031
Respiratory Disease (from drug prescriptions)	0.2047	<.0001
Day hospitalization	-0.2140	0.0003
Statins	-0.1823	<.0001
ACE/ARB	-0.1220	<.0001
Digitalis glycosides	0.1380	0.0011
Nitrates	0.1895	<.0001
Diuretics	0.1133	<.0001
Number of ER visits labeled 'Yellow'	-0.3232	<.0001
Total number of ER visits	0.5087	<.0001
Genitourinary Disease (from any data source)	0.1686	0.0037
Hematologic Disease (from any data source)	0.1699	0.0006
History of Endocrine Disease (from drug prescriptions) *	0.0877	0.0078
History of Psychological Disease (from drug prescriptions) *	0.1168	0.0009
History of Aortic Stenosis-Stage 3 *	0.3893	0.0002
History of Arrhythmias-Stage 2 *	0.1078	0.0013
History of Congestive Heart Failure-Stage 3 *	0.2010	0.0002
History of Coronary Artery Disease-Stage 1 *	0.1312	0.0002
History of Coronary Artery Disease-Stage 2 *	0.1103	0.0392
History of Coronary Artery Disease-Stage 3 *	0.1075	0.0476
History of Essential Hypertension-Stage 3 *	0.1204	0.0079
History of Pericarditis: Viral or Traumatic-Stage 2 or 3 *	-0.7456	0.0101
History of Diabetes Mellitus Type 1 or Type 2-Stage 1 *	0.1159	0.0141
History of Neoplasm, Malignant: Bladder, Urinary-Stage 1*	0.1311	0.0546
History of Renal Failure-Stage 2 or 3 *	0.1660	<.0001
History of Neoplasm, Malignant Hematologic-Stage 1 *	0.3476	0.0069
History of Neoplasm, Malignant Hematologic-Stage 2 *	-1.1984	0.0057
History of Cirrhosis of the Liver-Stage 2 or 3 *	0.2949	0.0083
History of Progressive Systemic Sclerosis-Stage 1 *	2.4748	0.0284
History of Cerebrovascular Disease-Stage 1 *	0.2552	<.0001
History of Cerebrovascular Disease-Stage 3 *	0.1474	0.0006
History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 1 *	0.2598	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

History of Dementia: Primary Degenerative (Alzheimer or Pick)-Stage 2 or 3 *	0.2971	0.0457
History of Potentially inappropriate prescribing - always to be avoided drugs *	-0.0705	0.0104
History of polypharmacy *	0.1276	0.0001
History of Chronic Obstructive Pulmonary Disease-Stage 1 or 2 *	0.2083	<.0001
History of Chronic Obstructive Pulmonary Disease-Stage 3 *	0.2686	0.0030
History of Neoplasm, Malignant: Lungs, Bronchi, or Mediastinum-Stage 1 *	0.4042	0.0142
History of Pneumonia: Bacterial-Stage 2 *	0.4137	0.0392
History of Other Cardiovascular drugs *	0.1277	0.0005
Immunologic Disease (from any data source)	0.9577	0.0461
Male Genital System (from any data source)	-0.1301	<.0001
Neurologic Disease (from any data source)	0.1639	<.0001
Polypharmacy	0.0778	0.0206
Psychological Disease (from any data source)	0.1247	0.0017
Endocrine Disease (from hospital data)	0.1413	0.0154
Respiratory Disease (from hospital data)	0.3324	<.0001

*History variables are calculated on previous 5 years of exposure data

For peer review only

Predicting Risk Using a Population Based Longitudinal Database
Louis et al
STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	See abstract section: Design.
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	The outcome measures used and what was found are summarized in the sections of the abstract: Main outcome measures and Results
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	See the Introduction section pages 4 and 5
Objectives	3	State specific objectives, including any prespecified hypotheses	Our objectives are described in the last two paragraphs of the Introduction
Methods			
Study design	4	Present key elements of study design early in the paper	See last paragraph of the Introduction and the Methods section
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	See Study Data and Study Population at beginning of Methods section
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	Our study includes 100% of the adult population See Study Data and Study Population at beginning of Methods section
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	See the Dependent variable and Independent variable sections in the Methods section
Data sources/ measurement	8*	For each variable of interest, give sources of data	See the Dependent variable

For Peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	and Independent variable sections in the Methods section. In addition we have included an Appendix with detailed mapping to independent variable.
Bias	9	Describe any efforts to address potential sources of bias	See Evaluation of the Models (page 10) in the Methods section
Study size	10	Explain how the study size was arrived at	Our study includes 100% of the adult population
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	See the Dependent variable and Independent variable sections in the Methods section. In addition we have included an Appendix with detailed mapping to independent variable.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	See the Modelling section and Evaluation of the Models section (pages 9 and 10) of the Methods
		(*b*) Describe any methods used to examine subgroups and interactions	See the Modelling section (pages 9 and 10) of the Methods
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	Our study includes 100% of the adult population
		(*e*) Describe any sensitivity analyses	See Evaluation of the Models last paragraph of Methods.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	See Page 11 Results section
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	See Table 1 and Results section (page 11-14)
		(b) Indicate number of participants with missing data for each variable of interest	

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	See Table 2 and figure and pages 14 and 15 of Results section
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	See pages 13
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	See Table 2 and page 13 and 14.
Discussion			
Key results	18	Summarise key results with reference to study objectives	See Discussion section page 16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	See Discussion section pages 17 and 18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	See Discussion section page 18
Generalisability	21	Discuss the generalisability (external validity) of the study results	See Discussion section pages 18 and 19
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	The study was funded by the Emilia-Romagna, region of Italy. See page 3.

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml