HHS Public Access

Author manuscript

Pharmacoepidemiol Drug Saf. Author manuscript; available in PMC 2020 May 01.

Published in final edited form as:

Pharmacoepidemiol Drug Saf. 2019 May ; 28(5): 584–592. doi:10.1002/pds.4769.

Development and validation of a five-year mortality prediction model using regularized regression and Medicare data

Jennifer L. Lund, PhD[1,2], **Tzy-Mey Kuo, PhD**[2], **M. Alan Brookhart, PhD**[1,2], **Anne-Marie Meyer, PhD**[1,3], **Alexandra F. Dalton, PhD**[4], **Christine E. Kistler, MD, MASc**[2,5], **Stephanie B. Wheeler, PhD**[2,6], and **Carmen L. Lewis, MD, MPH**[4]

[1)] Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA [2)] Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA [3)] IQVIA, St. Prex, Switzerland [4)] Department of Medicine, University of Colorado, Denver, CO, USA [5)] Department of Family Medicine, University of North Carolina, Chapel Hill, NC, USA [6)] Department of Health Policy and Management, University of North Carolina, Chapel Hill, NC, USA

Abstract

Purpose: De-implementation of low-value services among patients with limited life expectancy is challenging. Robust mortality prediction models using routinely collected healthcare data can enhance healthcare stakeholders' ability to identify populations with limited life expectancy. We developed and validated a claims-based prediction model for five-year mortality using regularized regression methods.

Methods: Medicare beneficiaries age ≥66 with an office visit and at least 12-months of pre-visit continuous Medicare A/B enrollment were identified in 2008. Five-year mortality was assessed through 2013. Secondary outcomes included 30-, 90-, 180-day and one-year mortality. Claims-based predictors, including comorbidities and indicators of disability, frailty, and functional impairment, were selected using regularized logistic regression, applying the least absolute shrinkage and selection operator (LASSO) in a random 80% training sample. Model performance was assessed and compared with the Gagne comorbidity score in the 20% validation sample.

Results: Overall, 183,204 (24%) individuals died. In addition to demographics, 161 indicators of comorbidity and function were included in the final model. In the validation sample, the c-statistic was 0.825 (0.823, 0.828). Median predicted probability of five-year mortality was 14%; almost 4% of the cohort had a predicted probability >80%. Compared to the Gagne score, the LASSO model led to improved five-year mortality classification (net reclassification index=9.9%; integrated discrimination index=5.2%).

Conclusions: Our claims-based model predicting five-year mortality showed excellent discrimination and calibration, similar to the Gagne score model, but resulted in improved

Corresponding Author: Jennifer L. Lund, Department of Epidemiology, 2102-D McGavran-Greenberg Hall, Gillings School of Global Public Health, CB#7435, University of North Carolina, Chapel Hill, NC 27599-7435, Jennifer.Lund@unc.edu, Phone: (919) 966-7440, Fax: (919) 966-2089.

Conflict of interest:
The authors have no conflicts of interest to declare.

mortality classification. Regularized regression is a feasible approach for developing prediction tools that could enhance healthcare research and evaluation of care quality.

Keywords

comorbidity; frailty; machine learning; Medicare; prediction modeling

Introduction

Efficient use of limited healthcare resources is key to controlling healthcare costs. Healthcare cost control includes de-implementation of unnecessary services that do not enhance patient outcomes.[1] One approach to achieve efficient healthcare utilization is to decrease the use of low-value services, such as those identified by Choosing Wisely.[2] Another approach is to decrease overuse of high-value services among patients who are unlikely to benefit. For example, individuals with limited life expectancy may not live long enough to benefit from an intervention and may experience net-harm.[3–6] Services for which the benefits accrue over long periods of time (e.g., colonoscopy,[7–9] cancer screening,[10–12] statin therapy for primary prevention,[13] cancer treatment[14–17]) are typically considered low-value in people with limited life expectancy.

Addressing overuse of low-value healthcare services in people with limited life expectancy is challenging. There is significant uncertainty when estimating life expectancies and the net-benefit of services for individual patients. In clinical practice, physicians support the exclusion of frail older adults from colorectal cancer screening and do not want to screen someone who would not be able to tolerate treatment, but also do not want to miss a treatable cancer.[18,19] Researchers and payers are also interested in quantifying and evaluating utilization of low-value services and their impacts on population-level costs and health outcomes.[20,21] Therefore, standardized, quantitative, robust risk prediction tools are needed to assist a variety of healthcare stakeholders in the identification and potential de-implementation of low-value services in subpopulations with limited life expectancy.

Several clinical mortality prediction models have been developed that require active collection of information about clinical (e.g., comorbid diseases) and functional (e.g., activities of daily living) health directly from patients or providers.[22–24] Given the busy clinical environment, adoption of these models has been low.[25] An alternative approach for risk prediction focuses on the use of existing data sources such as electronic health records (EHRs) or administrative claims data, which draw upon passive data generation processes. Several claims-based mortality prediction models exist and have demonstrated good predictive performance;[26] however, these models include only comorbid conditions as potential predictors. Thus, more refined claims-based mortality prediction models that include additional dimensions of health including disability, frailty, and functional impairment may enhance predictive performance.

The purpose of this study was to develop and internally validate a claims-based risk prediction model for five-year mortality among North Carolina Medicare beneficiaries. The long-term goal of this work is to improve the performance of risk prediction models to better inform and guide research, quality improvement, and de-implementation efforts. We used

regularized regression to build a five-year mortality prediction model in Medicare claims, incorporating indicators of comorbidity, disability, frailty, and functional impairment. We also sought to compare the performance of this model with an existing claims-based comorbidity score[24] shown to improve upon two standard comorbidity measures.[27,28]

Methods

Data Source

This study utilized Medicare enrollment and claims data from North Carolina from 2007–2013. Medicare enrollment files contain information about individuals' demographics, eligibility, coverage, and vital status (through December 31, 2013). Medicare Part A (hospital insurance) and Part B (medical insurance) claims are recorded for the purposes of reimbursement and contain information about diagnoses and procedures performed during a healthcare encounter.

Study Population

We identified a cohort of adults age 66 years or older who had an office visit in 2008 (index date), allowing for a full five years of mortality follow-up (through 2013) and at least 12-months of continuous Medicare Parts A/B enrollment prior to their index date. This cohort was anchored to an office visit to ensure the capture of a minimum amount of healthcare information. The 12-month period prior to the index date was used to capture all potential claims-based predictor variables. Individuals with a hospice claim during the 12-month baseline period or the index month were excluded, as these individuals had a determination of limited life expectancy.

Measures

Outcome Variables—The primary outcome was five-year all-cause mortality, defined using vital status ascertained from Medicare enrollment files as of December 31, 2013. Secondary outcomes included 30-, 90-, and 180-day and 1-year all-cause mortality.

Potential Predictor Variables—Potential predictor variables included demographics and benefit eligibility, comorbidities, or indicators of frailty, disability, and functional impairment as described below.

Demographics and beneficiary information—Age at the index date, sex, race (black, white, other), dual-eligibility (ever enrolled in NC Medicaid in the 12-months prior to the index date), and benefit eligibility due to end-stage renal disease were assessed from enrollment files. Age was modeled continuously, including a squared and cubic term.

Comorbidities—We considered the 17 comorbid conditions included in the Romano adaptation of the Charlson comorbidity index,[27,29,30] and the 30 conditions from the Elixhauser comorbidity classification system.[28] When overlap between the two comorbidity groupings occurred, the more inclusive definition was used, as in Gagne et al.[24] In total, 37 comorbid conditions were included. Comorbid conditions were defined by observing at least one International Classification of Diseases, Clinical Modification, 9th Edition (ICD-9) code

on any claim occurring within the 13 months prior to the index date (including the index month).

Indicators of frailty, disability, or functional impairment—Since comorbid conditions capture only one dimension of health in older adults,[31] we also considered Medicare claims-based predictors of frailty, disability, and functional impairment.[32–34] Faurot and colleagues identified 41 potential indicators of frailty, based on theory, to develop a claims-based model of frailty.[32] Davidoff and colleagues identified 112 potential indicators to develop a claims-based model of poor performance or disability status.[33] Finally, Chrischilles and colleagues identified 24 function-related indicators related to poor prognosis in older adults that were distinct from comorbid conditions included in the common indices.[34] We assessed overlap between these potential predictors and merged indicators that were largely overlapping, resulting in 151 unique indicators. Each of these indicators was defined as present or absent, requiring at least 1 ICD-9 diagnosis or procedure code or CPT/HCPCS code within the 12 months prior to the index date. These codes could appear on any inpatient, outpatient, skilled nursing, home health, or durable medical equipment claim, except indicators from Davidoff, which were derived only from physician, and durable medical equipment claims.

We further assessed overlap between the frailty, disability, and functional impairment indicators and the 37 comorbidities above, resulting in a total of 176 predictors. Including demographic and beneficiary characteristics above, 184 unique candidate predictors were considered for inclusion in the model building phase.

Statistical analysis

The total eligible cohort was randomly split into training (80%) and validation (20%) samples. Descriptive statistics for all predictors were reported by cohort. The training sample was used for model selection, and the validation sample was used to determine model performance and avoid overfitting.

Model selection—To build the predictive model, we used regularized regression based on the least absolute shrinkage and selection operator (LASSO).[35] This approach performs both variable selection and regularization and avoids issues of multicollinearity and overfitting. LASSO regression specifically penalizes parameter estimates generated using L1 penalization that has the effect of shrinking the estimates towards zero, introducing some bias in order to decrease prediction variance. This penalization results in a reduction of the mean squared prediction error. We applied the LASSO algorithm (using the glmnet R package)[36] to a logistic regression model predicting five-year mortality and used ten-fold cross-validation to select the optimal value for regularization. Model intercept and parameter estimates from the final selected model are reported in the Appendix.

Model validation—We applied the LASSO model selected in the training cohort to the validation cohort. Predicted probabilities of five-year mortality were generated for each individual. The distribution of predicted probabilities was described in the validation sample. We also report the area under the receiver operating characteristic (ROC) curve or the c-

statistic and 95% confidence intervals (CIs). Calibration of the final model was visually assessed by comparing the observed and predicted probabilities of five-year mortality.

Comparison with the Gagne score model—We further compared the performance of the LASSO model with the Gagne score (including age and sex).[24] The Gagne score was selected as a comparison model, as it has been shown to be superior to both the Charlson comorbidity index[27,30] and the Elixhauser comorbidity classification system[28] when predicting one-year mortality. We plotted the ROC curves and reported c-statistics for each model when predicting 30-, 90-, and 180-day and one-year all-cause mortality. For reference, we also reported the ROC curve and c-statistics for a base model including age and sex alone. To highlight the clinical impact of these c-statistics, we also report the number needed to treat (NNT) using established thresholds.[37]

Because even small changes in the c-statistic of a new model can result in benefits in terms of both confounding control[38] and improved classification,[39,40] we also computed reclassification statistics comparing the LASSO model with the Gagne score. We used two metrics, the net reclassification index (NRI) and the integrated discrimination improvement (IDI), to determine classification improvement attributable to the LASSO model.[41] The NRI indicates the proportion of patients correctly reclassified by a new model compared with an existing or standard model, while the IDI indicates the change in difference in average predicted probabilities between those who died and those who did not in a new and existing model. For our analysis, positive numbers for both the NRI and the IDI indicate that the LASSO model performed better than the Gagne model in discriminating five-year mortality. These measures (NRI and IDI) provide complimentary information to the c-statistic, indicating the marginal strength of the new model to the existing model and the discrimination slope, respectively.[41,42] Calibration of the Gagne score was also plotted.

In sensitivity analyses, we re-estimated the weights for the Gagne score in the training sample data and used this re-estimated model for comparison.

All statistical analyses were conducted using SAS version 9.4 (SAS Institute, Inc., Cary, NC) and R (R Core Team, 2016).[43] This study was approved by the Institutional Review Board at the University of North Carolina at Chapel Hill.

Results

The study population consisted of 773,748 North Carolina Medicare beneficiaries age 66 years or older who met all inclusion criteria (see Figure 1, Appendix). Characteristics of the training and validation samples were nearly identical (Table 1) with an average age at office visit of 75 years, 60% female, predominantly white, and 14% with dual Medicare/Medicaid eligibility. The median Gagne combined comorbidity score was 0. In total, 24% of the training (n=146,701) and validation (n=36,503) samples died within the five-year study period.

Model development and validation

The prevalence of all 184 potential predictor variables are presented along with the parameter estimates for the 166 predictor variables selected via LASSO for the final model (see Appendix, Table 1) based on the training cohort. In addition to older age and male sex, the most influential variables positively associated with five-year mortality included: metastatic cancer (OR=4.16), end stage renal disease beneficiary status (OR=4.11), Parkinson's disease (OR=1.95), dementia/delirium (OR=1.71), heart failure (OR=1.59), liver disease (OR=1.55), and home oxygen (OR=1.54). The most influential variables inversely associated with five-year mortality included: major breast surgery (e.g., mastectomy, OR=0.52), orthopedic surgery (e.g., knee replacement, OR=0.55), and lithotripsy (a surgical procedure to destroy kidney stones, OR=0.67).

The c-statistic for the LASSO model estimated using the validation sample was 0.825 (95% CI: 0.823, 0.828). Using the parameter estimates from the LASSO model, we computed predicted probabilities of five-year mortality for everyone in the validation sample; the distribution of these probabilities is presented in Figure 1 (median=0.14; interquartile range: 0.08–0.32). More than 36% of the study population had a predicted probability of dying within five years of <10%, whereas more than 4% had a predicted probability of dying within five years of >80%. The proportion of individuals who died within five years closely corresponded to the proportion predicted to die within five-years based on the LASSO model (Figure 1). Predictions among the highest-risk individuals were slightly overestimated.

Comparison with the Gagne score and base model

ROC curves for predicted five-year mortality demonstrate that compared to the base (c-statistic=0.722) and Gagne (c-statistic=0.797) models, the LASSO model resulted in improved five-year mortality prediction (Figure 2) and at each of the secondary time-points (Table 2). The Gagne model also showed good calibration, but with slight overestimation among the highest risk individuals (Appendix, Figures 2A– B). Overall, the c-statistics for both the LASSO model and the Gagne score model translate into a number needed to treat of <2, indicating both models have excellent discriminative performance, according to established thresholds.[37]

For reclassification analyses, we first selected cut-points based on previously used methods[24] selecting the average predicted probability of five-year mortality in: (1) those who died and (2) those who survived using the Gagne score model, which were 0.42 and 0.18, respectively. Using these cut-points, we cross-classified people according to the Gagne model and the LASSO model, and further stratified by vital status at five-years (Table 3).

Among those who died within five years, 2543 individuals (7.0%) had improvement in their classification using the LASSO model compared to the Gagne model. Among those who survived the five-year period, 3488 individuals (2.9%) had improvement in their classification using the LASSO model compared with the Gagne model. Taken together, the LASSO model resulted in an NRI for five-year mortality of 9.9% reclassified, which translates into 6031 individuals or 4.0% of the validation cohort. For the integrated

discrimination index (IDI), or the discrimination slope, the average predicted probability of dying among those who died was higher for the LASSO model (45.3%) compared with the Gagne model (41.3%). Similarly, the predicted probability of dying among those who did not die was lower for the LASSO model (16.9%) compared with the Gagne model (18.1%). This resulted in an overall IDI of 5.2%, further confirming improvement in classification with the LASSO model. Sensitivity analyses using the re-calibrated Gagne score generated similar results with an NRI of 9.0% and IDI of 4.7% (Appendix, Table 3).

A reference list of codes for the final predictors is available in the Appendix (see Table 4). Additional documentation and SAS and R code for the final predictive model is available upon request.

Discussion

Using regularized regression methods, we developed and validated a claims-based model to predict five-year mortality among older Medicare beneficiaries with excellent discrimination and calibration. Compared to the Gagne score, our Medicare claims-based prediction model had similar discriminative performance, as measured by the c-statistic; however, it also resulted in improved classification of mortality, which can be important for further separating high- and low-risk patients.[39]

Regularized regression can be useful for generating accurate predictions in big data settings, where the number of potential predictors is large. These methods avoid overfitting, which can hamper performance when applied to external data.[35,44] Given the increasing use of Medicare claims data to study the effects of medical interventions on health outcomes in older adults, healthcare stakeholders can leverage this model to better control for confounding by an individuals' predicted risk of death, but also understand whether intervention effects are modified by this underlying mortality risk.

Healthcare payers, including the Center for Medicare and Medicaid Services (CMS), could use our model to develop or augment existing quality metrics benchmarking the use of low-value services. Our model identified almost 5% of the Medicare sample with an 80% chance or higher of dying within five years. This population would be unlikely to benefit from certain services that require extended (e.g., > five years) life expectancy.[45] Thus, health plans that report high utilization rates for these services in subgroups with high predicted-mortality may be flagged for targeted quality improvement. On the other hand, more than 36% of the study population had a very low predicted probability of dying within five years (i.e., <10%). These people are likely to be healthy with substantial life expectancies and would thus benefit from interventions requiring substantial lag-time.

Our model may also be useful for characterizing the population at high-risk of receiving low-value care. Currently, there is no uniform approach for defining low-value care, which has impeded de-implementation efforts. For example, in colon cancer screening, low-value care has been defined as screening among individuals with advanced cancer,[46] those in the top tertile of comorbidities among veterans,[47] and average-risk individuals undergoing too-frequent repeat colonoscopies.[8] Our model showed that metastatic cancer and end-stage

renal disease were the strongest drivers of five-year mortality; populations with these conditions may be particularly at-risk of receiving potentially low-value care.

The performance of our claims-based model opens the door to future opportunities to improve clinical care using EHR data. We and others have proposed individualized decision-making in clinical practice for older adults to maximize benefits and minimize harms from healthcare interventions.[3,48,49] Unfortunately, the uptake of these prognostic models[22,23,50] and decision support tools has been limited.[51] Another approach to de-implementing low-value care would be excluding patients who are unlikely to live long enough to benefit from interventions, providing an opportunity to systematically address net-harm for patients with limited life expectancy. This approach is consistent with a threshold approach to clinical decision-making, where clinicians no longer pursue interventions likely to cause net harm.[18,19,52–54] Further work extending these methods to EHR data, determining relevant threshold cut-points,[55] and integrating informatics tools within the EHR system directly may provide opportunities to enhance targeted de-implementation efforts.

Limitations to our study should be noted. Medicare claims data are generated for reimbursement purposes, and lack clinical detail which limits their utility to inform patient-level decision-making. However, claims are often used by healthcare stakeholders to evaluate overall quality and benefits and harms of services among beneficiaries. Thus, this claims-based model provides an additional tool to augment and develop new quality measures to monitor the use of potentially low-values services in populations with limited predicted life expectancy. Although automated methods including regularized regression can generate prediction models with excellent discrimination and calibration, this performance often comes at a cost.[44] Our five-year mortality prediction model included >160 variables, rendering routine calculation of individuals' predicted risk by a physician infeasible. Instead, embedding such a model on the backend of an EHR system and pre-populating data prior to a patient visit is likely the best way to encourage uptake. Furthermore, the reclassification statistics in this analysis used empirically-defined cut-points rather than those based on clinical relevance, which may influence their ultimate utility for decision-making. Finally, our study is limited to North Carolina residents insured by Medicare who had an office visit. Future efforts will evaluate model performance in a nationwide Medicare sample using more contemporary data to assure broad generalizability.

Conclusion

We demonstrated improvement in the predictive performance of a new Medicare claims-based model over an existing model when predicting five-year mortality among beneficiaries ≥66 years old. This model also improved classification of death within five years, an important step toward developing de-implementation strategies for low-value services among people with limited life expectancy. In the future, CMS could use this five-year mortality prediction model to inform quality measures identifying potential overuse of specific health services and to guide quality improvement projects. However, we also anticipate that extensions of this method and our approach can be developed for use in clinical care settings using EHR data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This work was support by a grant from the National Cancer Institute (R21 CA191454) and UNC Oncology Clinical Translational Research Training Program (NCI, K12CA120780).

References

1. Qaseem A, Alguire P, Dallas P, et al. Appropriate use of screening and diagnostic tests to foster high-value, cost-conscious care. Annals of internal medicine. 2012;156(2):147–149. [PubMed: 22250146]
2. ABIM Foundation. Choosing Wisely. 2017; http://www.choosingwisely.org/. Accessed August 3, 2017.
3. Walter LC, Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. JAMA : the journal of the American Medical Association. 2001;285(21):2750–2756. [PubMed: 11386931]
4. US Preventive Services Task Force. Screening for colorectal cancer: US preventive services task force recommendation statement. JAMA : the journal of the American Medical Association. 2016;315(23):2564–2575. [PubMed: 27304597]
5. United States Preventive Services Task Force. Screening for colorectal cancer. 2008; http://www.uspreventiveservicestaskforce.org/uspstf/uspscolo.htm. Accessed August 21, 2012.
6. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. JAMA : the journal of the American Medical Association. 2012;307(2):182–192. [PubMed: 22235089]
7. Kruse GR, Khan SM, Zaslavsky AM, Ayanian JZ, Sequist TD. Overuse of Colonoscopy for Colorectal Cancer Screening and Surveillance. Journal of general internal medicine. 2014.
8. Goodwin JS, Singh A, Reddy N, Riall TS, Kuo YF. Overuse of screening colonoscopy in the Medicare population. Archives of internal medicine. 2011;171(15):1335–1343. [PubMed: 21555653]
9. Tran A, Man Ngor E, Wu BU. Surveillance colonoscopy in elderly patients: A retrospective cohort study. JAMA Internal Medicine. 2014;174(10):1675–1682. [PubMed: 25111954]
10. Walter LC, Bertenthal D, Lindquist K, Konety BR. PSA screening among elderly men with limited life expectancies. JAMA : the journal of the American Medical Association. 2006;296(19):2336–2342. [PubMed: 17105796]
11. Freedman RA, Keating NL, Pace LE, Lii J, McCarthy EP, Schonberg MA. Use of Surveillance Mammography Among Older Breast Cancer Survivors by Life Expectancy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(27):3123–3130. [PubMed: 28749724]
12. Royce TJ, Hendrix LH, Stokes WA, Allen IM, Chen RC. Cancer screening rates in individuals with different life expectancies. JAMA Intern Med. 2014;174(10):1558–1565. [PubMed: 25133746]
13. Kutner JS, Blatchford PJ, Taylor DH Jr, et al. Safety and benefit of discontinuing statin therapy in the setting of advanced, life-limiting illness: A randomized clinical trial. JAMA Internal Medicine. 2015;175(5):691–700. [PubMed: 25798575]
14. Peyrade F, Gastaud L, Re D, Pacquelet-Cheli S, Thyss A. Treatment decisions for elderly patients with haematological malignancies: a dilemma. The Lancet Oncology. 2012;13(8):e344–352. [PubMed: 22846839]
15. Bian SX, Hoffman KE. Management of Prostate Cancer in Elderly Men. Seminars in Radiation Oncology. 2013;23(3):198–205. [PubMed: 23763886]
16. Delpierre C, Lamy S, Kelly-Irving M, et al. Life expectancy estimates as a key factor in over-treatment: the case of prostate cancer. Cancer epidemiology. 2013;37(4):462–468. [PubMed: 23623489]

17. Daskivich TJ, Tan H-J, Litwin MS, Hu JC. Life Expectancy and Variation in Treatment for Early Stage Kidney Cancer. The Journal of Urology. 2016;196(3):672–677. [PubMed: 27012644]
18. Lewis CL, Moore CG, Golin CE, Griffith J, Tytell-Brenner A, Pignone MP. Resident physicians' life expectancy estimates and colon cancer screening recommendations in elderly patients. Medical decision making : an international journal of the Society for Medical Decision Making. 2008;28(2):254–261. [PubMed: 18349429]
19. Dalton AF, Golin CE, Esserman D, Pignone MP, Pathman DE, Lewis CL. Relationship between Physicians' Uncertainty about Clinical Assessments and Patient-Centered Recommendations for Colorectal Cancer Screening in the Elderly. Medical decision making : an international journal of the Society for Medical Decision Making. 2015;35(4):458–466. [PubMed: 25712448]
20. Schwartz AL, Landon BE, Elshaug AG, Chernew ME, McWilliams JM. Measuring low-value care in Medicare. JAMA Intern Med. 2014;174(7):1067–1076. [PubMed: 24819824]
21. Barnett ML, Linder JA, Clark CR, Sommers BD. Low-Value Medical Services in the Safety-Net Population. JAMA Intern Med. 2017;177(6):829–837. [PubMed: 28395014]
22. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. JAMA : the journal of the American Medical Association. 2006;295(7):801–808. [PubMed: 16478903]
23. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. External validation of an index to predict up to 9-year mortality of community-dwelling adults aged 65 and older. Journal of the American Geriatrics Society. 2011;59(8):1444–1451. [PubMed: 21797837]
24. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. Journal of clinical epidemiology. 2011;64(7):749–759. [PubMed: 21208778]
25. Vogenberg FR. Predictive and Prognostic Models: Implications for Healthcare Decision-Making in a Modern Recession. American Health & Drug Benefits. 2009;2(6):218–222. [PubMed: 25126292]
26. Sharabiani MT, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. Medical care. 2012;50(12):1109–1118. [PubMed: 22929993]
27. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373–383. [PubMed: 3558716]
28. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. 1998;36(1):8–27. [PubMed: 9431328]
29. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. Journal of clinical epidemiology. 2000;53(12):1258–1267. [PubMed: 11146273]
30. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. Journal of clinical epidemiology. 1993;46(10):1075–1079; discussion 1081–1090. [PubMed: 8410092]
31. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. The journals of gerontology Series A, Biological sciences and medical sciences. 2004;59(3):255–263.
32. Faurot KR, Jonsson Funk M, Pate V, et al. Using claims data to predict dependency in activities of daily living as a proxy for frailty. Pharmacoepidemiology and drug safety. 2015;24(1):59–66. [PubMed: 25335470]
33. Davidoff AJ, Zuckerman IH, Pandya N, et al. A novel approach to improve health status measurement in observational claims-based studies of cancer treatment and outcomes. Journal of geriatric oncology. 2013;4(2):157–165. [PubMed: 23795223]
34. Chrischilles E, Schneider K, Wilwert J, et al. Beyond comorbidity: expanding the definition and measurement of complexity among older adults using administrative claims data. Medical care. 2014;52 Suppl 3:S75–84. [PubMed: 24561763]
35. Tibshirani R Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society Series B (Methodological). 1996;58(1):267–288.

36. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of statistical software. 2010;33(1):1–22. [PubMed: 20808728]
37. Cohen J Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988. In.
38. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol. 2001;154(9):854–864. [PubMed: 11682368]
39. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology (Cambridge, Mass). 2010;21(1):128–138.
40. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007;115(7):928–935. [PubMed: 17309939]
41. Pencina MJ, D'Agostino RB Sr., D'Agostino RB Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine. 2008;27(2):157–172; discussion 207–112. [PubMed: 17569110]
42. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. Am J Epidemiol. 2012;176(6):473–481. [PubMed: 22875755]
43. R Core Team. R: A language and environment for statistical computing. 2016; https://www.R-project.org/.
44. Obermeyer Z, Emanuel EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. New England Journal of Medicine. 2016;375(13):1216–1219. [PubMed: 27682033]
45. Braithwaite RS, Fiellin D, Justice AC. The payoff time: a flexible framework to help clinicians decide when patients with comorbid disease are not likely to benefit from practice guidelines. Medical care. 2009;47(6):610–617. [PubMed: 19433991]
46. Sima CS, Panageas KS, Schrag D. Cancer screening among patients with advanced cancer. JAMA : the journal of the American Medical Association. 2010;304(14):1584–1591. [PubMed: 20940384]
47. Walter LC, Lindquist K, Nugent S, et al. Impact of age and comorbidity on colorectal cancer screening among older veterans. Annals of internal medicine. 2009;150(7):465–473. [PubMed: 19349631]
48. Breslau ES, Gorin SS, Edwards HM, Schonberg MA, Saiontz N, Walter LC. An Individualized Approach to Cancer Screening Decisions in Older Adults: A Multilevel Framework. Journal of general internal medicine. 2016;31(5):539–547. [PubMed: 26941042]
49. Lewis CL, Golin CE, DeLeon C, et al. A targeted decision aid for the elderly to decide whether to undergo colorectal cancer screening: development and results of an uncontrolled trial. BMC medical informatics and decision making. 2010;10:54. [PubMed: 20849625]
50. Lee S ePrognosis - Estimating Prognosis for Elders. 2015; http://eprognosis.ucsf.edu/default.php.
51. Elwyn G, Scholl I, Tietbohl C, et al. "Many miles to go ...": a systematic review of the implementation of patient decision support interventions into routine clinical practice. BMC medical informatics and decision making. 2013;13 Suppl 2:S14. [PubMed: 24625083]
52. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. The New England journal of medicine. 1980;302(20):1109–1117. [PubMed: 7366635]
53. Lewis CL, Esserman D, DeLeon C, Pignone MP, Pathman DE, Golin C. Physician decision making for colorectal cancer screening in the elderly. Journal of general internal medicine. 2013;28(9):1202–1207. [PubMed: 23539281]
54. Lewis CL, Griffith J, Pignone MP, Golin C. Physicians' decisions about continuing or stopping colon cancer screening in the elderly: a qualitative study. Journal of general internal medicine. 2009;24(7):816–821. [PubMed: 19437080]
55. Eddy DM. Clinical decision making : from theory to practice : a collection of essays from the Journal of the American Medical Association. Sudbury; Boston, MA: Jones and Bartlett Publishers; 1996.

Key points:

1) Regularized regression methods can be used to improve the prediction of mortality among older adults, 2) Medicare claims data contain information about comorbidities and other indicators of disability, frailty, and functional impairment that be used to develop prediction models, and 3) A new Medicare claims-based five-year mortality prediction model, selected using regularized logistic regression, demonstrated improved classification of mortality at five years compared with a widely used claims-based comorbidity score, the Gagne combined index.

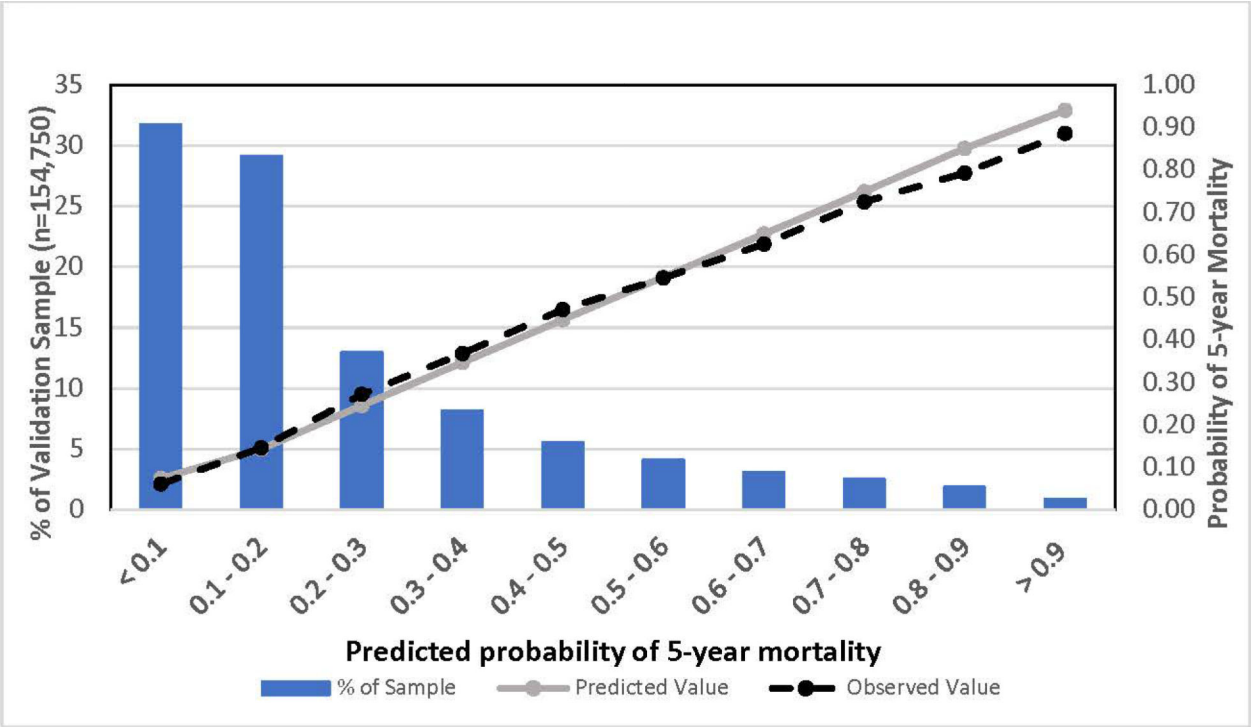

Figure 1. Observed and predicted 5-year mortality in validation sample (n=154,750).

The left vertical axis reports the proportion of the validation sample included in each of the categories of predicted probability of 5-year mortality (horizontal axis). The right vertical axis reports both the observed and predicted probabilities of 5-year mortality.

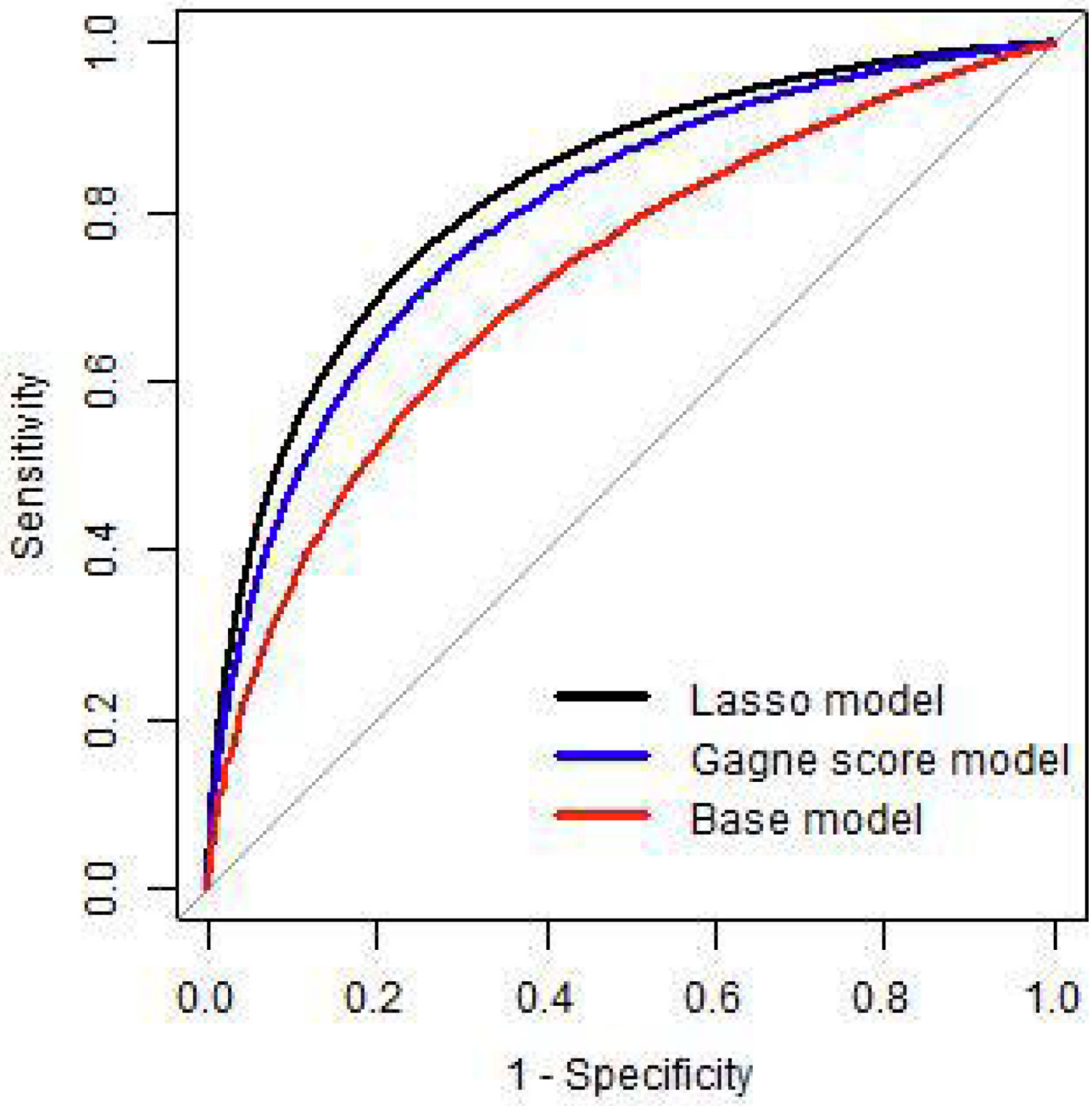

Figure 2. Receiver operating curves comparing three predictive models of 5-year mortality in the validation sample (n=154,750).

All models are adjusted for age (using polynomials) and sex. C-statistics for the age-sex model (red), Gagne model (blue), and LASSO model (black) were 0.722, 0.797, and 0.825, respectively.

Table 1.

Characteristics of the training (n=618,998) and validation (n=155,750) cohorts

	Training (N= 618,998)		Validation (N= 154,750)	
	n	%	n	%
Died within 5 years	146,701	24	36,503	24
Age, years (mean, SD)				
Age Categories	75.26	7.24	75.26	7.23
66–69	167,007	27	41,886	27
70–74	154,148	25	38,224	25
75–79	126,396	20	31,814	21
80–84	94,101	15	23,514	15
85+	77,346	13	19,312	12
Sex				
Male	249,390	40	62,100	40
Female	369,608	60	92,650	60
Race				
White	526,374	85	131,804	85
Black	82,382	13	20,416	13
Other	10,242	2	2,530	2
Medicare/Medicaid dual eligibility[a]	86,672	14	21,556	14
Gagne combined score[a] (median IQR)	0	(0, 2.0)	0	(0, 2.0)
Healthcare utilization[a]				
Nursing home visit	30,103	5	7,506	5
Hospital visit	99,967	16	24,867	16
Emergency room visit	139,291	23	34,910	23
Home visit	1,428	0.2	339	0.2

Abbreviations: SD=standard deviation; IQR=interquartile range

[a]Measured in the 12-months prior to the month of the qualifying office visit.

Table 2.

Comparison of c-statistics[a] for model discrimination in predicting mortality at varying timepoints in the validation sample (n=155,412)

All-Cause Mortality	Number of deaths (cumulative incidence)	Base model[b]	Gagne score model[b]	LASSO model
30-day mortality	506 (0.3%)	0.713 (0.690, 0.737)	0.849 (0.832, 0.866)	0.867 (0.852, 0.882)
90-day mortality	1,708 (1.1%)	0.709 (0.696, 0.722)	0.844 (0.835, 0.854)	0.864 (0.856, 0.873)
180-day mortality	3,478 (2.3%)	0.702 (0.693, 0.711)	0.827 (0.820, 0.834)	0.850 (0.844, 0.857)
One-year mortality	7,155 (4.6%)	0.699 (0.693, 0.706)	0.813 (0.808, 0.818)	0.840 (0.836, 0.845)
5-year mortality	36,503 (23.6%)	0.722 (0.720, 0.726)	0.797 (0.795, 0.800)	0.825 (0.823, 0.828)

Abbreviations: LASSO=least absolute shrinkage and selection operator

[a] C-statistics calculated using validation dataset (20% of the entire cohort). The data in parentheses are 95% confidence intervals.

[b] Age is included as polynomials.

Table 3.

Reclassification table indicating improvement[a] in five-year mortality prediction for the LASSO model compared with the Gagne model in the validation sample (n=155,412)

	Beneficiaries who died within 5-years				
		LASSO Model			
Gagne Model	**Predicted 5-year mortality**	<18%	18%−42%	>=42%	Total
	<18%	5758	2281	288	8327
	18%−42%	1794	6523	3819	12136
	>=42%	120	1931	13989	16040
	Total	7672	10735	18096	36503
	Beneficiaries who survived >5-years				
		LASSO Model			
Gagne Model	**Predicted 5-year mortality**	<18%	18%−42%	>=42%	Total
	<18%	72567	6814	259	79640
	18%−42%	9870	15599	3031	28500
	>=42%	386	3336	6385	10107
	Total	82823	25749	9675	118247
	All beneficiaries				
		LASSO Model			
Gagne Model	**Predicted 5-year mortality**	<18%	18%−42%	>=42%	Total
	<18%	78325	9095	547	87967
	18%−42%	11664	22122	6850	40636
	>=42%	506	5267	20374	26147
	Total	90495	36484	27771	154750

[a]For people who died (top section of table), improvement in classification was defined as: the number of people who moved upwards (i.e., to a higher predicted probability group) from the Gagne model to the LASSO model (grey cells) subtracted from the number of people who moved downwards (i.e., to a lower predicted probability group) from Gagne model to LASSO model (peach cells) divided by all patients who died [((2281+288+3819)−(1794+120+1931))/36503 = 2543/36503 = 7.0%]. For people who survived (top section of table), improvement in classification was defined as: the number of people who moved downwards (i.e., to a lower predicted probability group) from Gagne model to LASSO model (peach cells) subtracted from the number of people who moved upwards (i.e., to a higher predicted probability group) from Gagne model to LASSO model (grey cells) [((9870+386+3336)−(6814+259+3031))/118247 = 3488/118247 = 2.9%]. The net reclassification index (NRI) is the net improvement in classification of people who died and survived, which was 7.0%+2.9%=9.9%.