



Published in final edited form as:

J Am Geriatr Soc. 2016 July ; 64(7): 1439–1447. doi:10.1111/jgs.14228.

Predicting Two-Year Risk of Developing Pneumonia in Older Adults Without Dementia

Michael L. Jackson, PhD¹, Rod Walker, MS¹, Sei Lee, MD², Eric Larson, MD¹, and Sascha Dublin, MD, PhD^{1,3}

¹Group Health Research Institute, Seattle WA

²Division of Geriatrics, University of California, San Francisco CA

³Department of Epidemiology, University of Washington, Seattle WA

Abstract

Background—Pneumonia is a common cause of illness and death in older adults (65 years of age). Pneumonia prediction models could be used by clinicians in counseling patients and by policy makers and researchers for risk adjustment.

Objectives—To develop three prognostic indices, which vary in degree of detail required, for two-year pneumonia risk in older adults.

Setting—Community-dwelling enrollees in Group Health (GH), an integrated healthcare delivery system.

Participants—The study included 3,375 subjects enrolled in the Adult Changes in Thought study. Participants were 65 years of age, dementia-free, and enrolled in GH for at least two years prior to start of follow-up. Subjects were divided into development (n=2,250) and validation (n=1,125) cohorts.

Exposures—Questionnaire data and interviewer assessments on functional status, medical history, smoking and alcohol use, cognitive function, personal care, problem solving, physical measures including grip strength and gait speed, and administrative database information on comorbid illnesses, laboratory tests, and prescriptions dispensed.

Main outcome—Incident community-acquired pneumonia, defined presumptively from administrative data and validated by medical record review.

Results—Participants (59% female) contributed 12,998 visits at which risk factors were assessed; 642 pneumonia events were observed during follow-up. Age, sex, chronic obstructive pulmonary disease and congestive heart failure, body mass index, and use of inhaled or oral corticosteroids were key predictors in all prognostic indices. A risk score based on these seven

Corresponding author: Michael L. Jackson, 1730 Minor Ave, Suite 1600, Seattle WA, 98101 jackson.ml@ghc.org, phone: 206-287-2220.

Alternate: Sascha Dublin, 1730 Minor Ave, Suite 1600, Seattle WA, 98101, dublin.s@ghc.org, 206-287-2870

Author's Contributions: Study data were created by Michael L. Jackson, Eric Larson, and Rod Walker. Michael L. Jackson, Rod Walker, Sei Lee, and Sascha Dublin designed the study protocol, and Rod Walker analyzed the data. All authors critically reviewed the results. Michael L. Jackson drafted the manuscript; the remaining authors critically revised it.

variables, which are commonly available in electronic medical records, had equal or better performance (c-index, 0.69 in the validation cohort) than scores including more detailed data such as functional status.

Conclusion—Data commonly available in electronic medical records can stratify older adults into groups with varying subsequent two-year pneumonia risk.

Keywords

Pneumonia; Predictive Value of Tests; Aged; Aged, 80 and over

INTRODUCTION

Community-acquired pneumonia (CAP) is an important cause of morbidity and mortality among older adults (> 65 years of age). In the United States, the incidence CAP in older adults is 20 to 30 cases per 1,000 person-years.^{1–3} Approximately 40% of older adults with CAP are hospitalized,^{2,3} and older adults hospitalized for CAP have high rates of complications, including elevated risks of both short- and long-term mortality.^{4,5} Risk factors for CAP in older adults include increasing age, chronic lung disease, smoking, low body mass index (BMI), and functional status impairments.^{6–11}

Risk scores have been developed to predict the risk of complications or death among persons hospitalized for CAP.^{4,12–15} However, other than studies predicting the risk of pneumonia in persons hospitalized for stroke,^{16,17} little work has been done to predict risk of developing CAP in older adults. Such a prediction score could be used clinically, for example to encourage pneumococcal vaccination in high-risk patients, or to counsel patients contemplating behavioral change such as smoking cessation. Such prediction could also have research or policy uses, such as identifying patients at high risk for pneumonia for clinical trials, or adjusting for baseline pneumonia risk in observational studies or when calculating performance measures for healthcare systems.

We sought to develop risk scores that could predict the two-year risk of developing CAP in dementia-free older adults. We aimed to create three separate scores: a clinical score that would be sufficiently rapid and accurate for use in clinical practice; a research score involving more intensive data collection but potentially higher predictive accuracy; and a score that uses only data typically available in electronic medical record (EMR) systems that could be computed automatically.

METHODS

Study population

The study population was enrollees in the Adult Changes in Thought (ACT) study, which has been described in detail elsewhere.¹⁸ In brief, ACT is a prospective cohort study conducted among enrollees in Group Health Cooperative (GH), an integrated healthcare delivery system in Washington and Idaho States. ACT enrolled 3,392 older adults enrolled in GH living in the Seattle area between 1994–1996 and 2000–2003. Eligible participants were community-dwelling, cognitively intact, aged 65 years or older, and free from dementia.

Additional details on the study population are available in the Supplemental Methods section.

At enrollment, ACT participants underwent detailed evaluation of physical and cognitive function. Follow-up visits occurred at two-year intervals, with repeated evaluation of most of the characteristics measured at enrollment. Participants who scored less than 86 on the Cognitive Abilities Screening Instrument¹⁹ or who had symptoms or signs of dementia received a more detailed assessment. Participants judged by a multidisciplinary consensus committee to be dementia-free after this evaluation continued the biennial follow-up schedule. For the present study, we included all ACT enrollees and followed them from initial ACT enrollment until death, onset of dementia, disenrollment from ACT or GH, or end of the study period (December 31st, 2007).

Study outcome: incident CAP

We identified incident CAP in our study subjects using data and methods from a prior study of pneumonia among GH members, the Pneumonia Surveillance Study (PSS).²⁰ The PSS study first identified presumptive episodes of pneumonia among all GH members between 1998 and 2004, defined based on International Classification of Diseases, Version 9, Clinical Modification (ICD-9-CM) codes 480–487.0 or 507.0 assigned to outpatient and inpatient medical encounters. Presumptive pneumonia episodes were considered validated if manual review of chest radiograph reports within 30 days of first pneumonia diagnosis indicated the presence of an infiltrate not known to be chronic. For GH members hospitalized in non-GH hospitals (for whom radiograph reports may not have been available), presumptive pneumonia episodes were validated by reviewing hospital admission, consultation, and discharge summaries.

In the present study, we linked the ACT cohort to the PSS outcomes to identify validated CAP cases between 1998 and 2004. We repeated the PSS methods to identify incident pneumonias between 1994 and 1997 and between 2005 and 2007.

Data on potential pneumonia predictors

We used several data sources to create a broad list of variables that we judged could have prognostic value for CAP based on review of the literature (e.g.^{6,10,11}). First, we used variables available from the ACT enrollment and biennial visits. These visits included extensive in-person evaluations, with questionnaire data, standardized cognitive function measures, interviewer assessments of memory, orientation, personal care, and problem solving; and physical measures such as grip strength and gait speed.

We also collected data on study subjects from GH's electronic databases, which combine data from the GH Electronic Medical Record (EMR) system and claims for outside services. We used ICD-9-CM codes assigned to inpatient and outpatient encounters in the 24 months prior to each ACT study visit to define conditions such as serious cancer, malnutrition, and prior pneumonia. We used laboratory data on hematocrit and creatinine-based estimated glomerular filtration rate (EGFR) in the 24 months prior to each ACT study visit. We used prescriptions dispensed in the 12 months prior to each ACT visit to identify use of medications such as statins, angiotensin-converting enzyme (ACE) inhibitors, insulin, and

proton pump inhibitors. Finally, we used GH administrative records to identify frequency of outpatient and inpatient medical encounters. Notably, we did not include pneumococcal vaccination as a possible predictor. Pneumococcal polysaccharide vaccines such as PPV23 were recommended for use in older adults during our study period.²¹ These have been shown to protect against invasive pneumococcal disease but not against CAP.²² This study was conducted prior to the recommendations for use of the newer pneumococcal conjugate vaccines in older adults,²³ and so there was little use of these vaccines in our population.

Using the above sources, we assembled a list of 71 variables that could be predictive of CAP risk. The candidate predictor variables, along with a limited number of pre-specified possible interactions, were grouped into six domains for model development: demographic, comorbidity, functional status, behavioral, pharmacological, and detailed (eTable Supplement). The “detailed” domain includes potential predictors that are not typically available in the context of a clinic visit but that could be collected in a research setting, such as grip strength and gait speed.

We then characterized each of these variables based on completeness (i.e., extent of missing information) and on distribution of values (e.g., prevalence for binary variables, distributions for categorical or continuous variables) across all ACT study visits. Based on these univariate statistics, we discarded variables for which the distribution of values was too narrow for predictive discrimination (e.g., 99% of the population had the same values for a covariate). These decisions were made without assessing the association between the variables and pneumonia.

To account for missing data in our potential predictors, we used multiple imputation via chained equations²⁴ to generate 50 imputed datasets. Because our model building included step-down variable selection (described below), we erred on the conservative side in the number of imputations performed.

Statistical model

Potential predictors were assessed at the time of each ACT visit (enrollment or biennial). As our goal was to develop a model that predicts risk of CAP within two years, we treated each ACT visit as an observation and looked out up to two years from the visit for the earliest of CAP, subsequent biennial visit, or a subject’s end of follow-up (as described under the study population). In this context, death and disenrollment were considered competing risks for the CAP outcome, and the other endpoints were considered censoring events. We then estimated the 2-year cumulative incidence of CAP, adjusted for covariates, using the Fine and Gray model, which models cumulative incidence in the presence of competing risks.²⁵ Additional information on the model and on model development is available in the Supplemental Methods section.

Model development

To develop the prediction models, we first separated our study population of individuals (and their corresponding ACT visits) into a development cohort and a validation cohort, using an approximately 2/3–1/3 split, respectively. We then used a systematic two-step backwards selection process to determine which variables to include in the final clinical model. First,

we used backward selection separately within each of five variable domains (demographic, comorbidity, functional status, behavioral, pharmacological) to identify a smaller set of candidate variables with highest apparent prognostic ability, using a cut-off of $p < 0.2$. We then combined these identified variables from each of the five domains into a single model and repeated the backwards selection process to potentially further reduce this set of predictors, using a cut-off of $p < 0.05$. The variables remaining after this selection constituted our final clinical prediction model.

We developed the EMR model using the same process, but including only variables that would be expected to be available as coded values (i.e. not in free text) from an EMR system. Finally, we developed a research model that could include variables from the detailed variable domain. For the research model, we entered all the variables from the final clinical model with the detailed domain variables into a single model and repeated the backward selection process.

Model validation

We assessed performance of our prediction models in terms of calibration (a measure of the agreement between observed and predicted pneumonia risk) and discrimination (a measure of the models' ability to separate those who do vs. do not develop pneumonia).²⁶ For calibration, we grouped observations into deciles of predicted two-year pneumonia risk (as predicted by the final selected models), computed the observed risk in each group, and plotted observed versus average predicted risk within each group. This provides a visual assessment of calibration. In a perfectly calibrated model, all points would lie along the diagonal where observed and predicted CAP risk is equal. This process was done for each of the imputed datasets in the development cohort and separately, the validation cohort.

To assess discrimination of our prediction models we computed estimates of a concordance (*c*) index.²⁷ This *c*-index is directly analogous to the usual *c*-statistic reported for logistic regression models but accounts for the use of survival analysis with competing risks.²⁸ Higher values of the *c*-index indicate better discrimination. The *c*-index was estimated in the development cohort and separately in the validation cohort.

Finally we examined whether discrimination was notably impacted by a coarser representation of the models, in which each of the variables in the final risk models was assigned an integer point score based on the regression coefficients. This simplification would allow a risk score to be calculated easily for any given patient by summing the points for all the patient's risk factors. We plotted the observed risk for groups defined by these risk scores. All data management, statistical modeling, and descriptive summaries were conducted using a combination of statistical software including: SAS software, version 9.2 (SAS Institute, Inc., Cary, NC); R, version 2.15.3 (R Foundation for Statistical Computing, Vienna, Austria); and Stata 12.1 (Stat Corp., College Station, TX), along with the user-written module *stcompet*.²⁹

RESULTS

Characteristics of participants

The study population consisted of 3,375 subjects (17 ACT enrollees were excluded due to insufficient GH enrollment history), who contributed 12,998 ACT visits. During the follow-up period these older adults experienced 642 incident pneumonias, and 574 died (Table 1). The majority of participants (59%) were female. At ACT enrollment, 723 (21%) were 65–69 years of age, 1,785 (53%) were 70–79 years of age, 793 (23%) were 80–89 years of age, and 74 (2%) were 90 years of age or older. The study population was randomly divided into a development cohort (N=8,579 visits among 2,250 subjects) and a validation cohort (N = 4,419 visits among 1,125 subjects). The distributions of demographic characteristics, comorbidities, and other characteristics were similar between the two cohorts (Table 1, eTable in Supplement). Most of the variables were significantly associated with CAP risk in bivariate analyses; exceptions included measures of weekly exercise, self-reported diabetes or recent cancer diagnosis, and use of immunosuppressive medications, among others.

Multivariable models

After backward selection within each domain and then in the full model, the final clinical model included 12 predictors: age; sex; self-reported histories of chronic obstructive pulmonary disease (COPD), congestive heart failure (CHF), pneumonia; use of home oxygen; difficulties in dressing or walking half a mile; history of alcohol-related aggressive behavior; body mass index; and prescriptions for inhaled corticosteroids or for oral corticosteroids (Table 2). In the development cohort, the c-index for this model was 0.72 (95% CI 0.70–0.75) which indicates moderate discrimination. The high agreement between observed and predicted risk, evidenced by the light grey points aligning closely with the diagonal (top panel of Figure 1), indicates good calibration. In the validation cohort, the model showed slightly lower discrimination (c-index 0.70, 95% CI 0.66–0.73) than in the development cohort, and the calibration was worse, as evidenced by the deviation of the dark grey points from the diagonal (top panel of Figure 1). In particular, the model tended to underestimate the actual risk for groups with an average predicted risk in the 5–10% range in the validation cohort.

When we allowed the detailed research covariates to compete with the clinical model variables, none of the laboratory values, physical function measures, or healthcare utilization measures from the detailed domain were retained in the model. Only the covariate indicating prior pneumonia (based on administrative data codes in the prior two years) ultimately was added. This detailed model did not meaningfully differ from the clinical model in terms of calibration or discrimination (c-index in development cohort, 0.73, 95% CI 0.70–0.76; c-index in validation cohort, 0.71, 95% CI 0.67–0.74) and for simplicity is not presented further here.

When repeating the entire variable selection process restricting to variables that could easily be obtained from an EMR, the final EMR model included seven variables: age; sex; histories of COPD or CHF; body mass index; and prescriptions for inhaled corticosteroids or for oral corticosteroids (Table 2). The c-index for this model was similar in both the development

cohort (c-index 0.69, 95% CI 0.67–0.72) and the validation cohort (c-index 0.69, 95% CI 0.66–0.73). Further, the model showed similar calibration in both cohorts (bottom panel of Figure 1).

Risk scores

For simplification, each of the variables in the clinical and EMR models was assigned a point score based on the regression coefficients (Table 2). For the clinical model, the c-index in the development cohort based on the exact model coefficients (0.72, reported earlier) was virtually the same as the c-index based on the converted point-based risk score (0.72, 95% CI 0.69–0.75); the same was true in the validation cohort (point-based risk score c-index 0.69, 95% CI 0.66–0.73, compared to c-index of 0.70 reported earlier). Similar results were seen when comparing discrimination for the EMR model and its point-based analogue.

Finally, Figure 2 shows plots of the observed risk in groups classified based on their computed clinical risk score (top panel) and computed EMR risk score (bottom panel), both in the development cohort (light grey) and the validation cohort (dark grey). Both scores showed similar patterns in the development cohort, with two-year pneumonia risk of <5% for older adults with scores <4, pneumonia risk of approximately 5–17% in those with scores of 4–7, and pneumonia risk of approximately 25% for older adults with scores of 8 or higher. In the validation cohort, however, observed pneumonia risk tended to be similar for scores of 5 and greater predicted from the clinical risk score (dark grey in top panel), while observed risk increased approximately linearly with increasing predicted risk score from the EMR model (dark grey in bottom panel). In the validation cohort, 70% of subjects had clinical risk scores <4, 26% had scores 4–7, and 4% had scores of 8 or higher. For the EMR score, 79% had scores <4, 18% had scores 4–7, and 3% had scores of 8 or higher.

DISCUSSION

We set out to create three different pneumonia risk scores that could be used in different settings, depending on the expected time and data available for calculating the risk score. We had expected that a detailed risk score, which could include information on comorbidities, functional status, exercise, prescriptions, laboratory test results, and strength/gait tests, would show the greatest discrimination and calibration. Such a score could be used to identify patients at high risk of pneumonia in clinical trials. We likewise expected the clinical and EMR scores, designed to be quickly (clinical) or automatically (EMR) calculated in the context of a provider visit based on more limited data, to show acceptable calibration and discrimination, but not to perform as well as the research score.

Instead, we found that all three scores tended to include the same variables. Even when detailed data on laboratory results, strength/gait tests, and indicators of serious disease were allowed to compete in the detailed model, these variables did not meaningfully predict two-year pneumonia incidence beyond the more readily available clinical model variables. Because they included similar variables, all three scores showed similar discrimination and calibration in the validation cohort. In fact, the EMR risk score, which only included seven variables, appeared to be at least as well calibrated and discriminative as the clinical risk score, which included 12. Given the trivial differences in discrimination between the EMR

risk score and the detailed risk score, the seven-variable EMR risk score will likely be of greatest utility both for clinical practice and for research use. For research, this score could be used in observational studies for pneumonia prevention, treatment, or outcomes, which can use EMR data to control for baseline risk of pneumonia. It could also be used to identify high risk patients to recruit into clinical trials of potential interventions (such as recent trials of pneumococcal conjugate vaccine in older adults). The EMR score could also have clinical uses. A discussion of personal pneumonia risk could be used to encourage use of pneumococcal conjugate vaccines in patients that have not been vaccinated or only received polysaccharide vaccine. Pneumonia risk could help patients and their clinicians weigh the risks and benefits of interventions. For example, the benefits of corticosteroids for patients with osteoarthritis may be outweighed by the risk of CAP in patients at high risk for CAP. CAP risk could also be useful for advance care planning, in which older adults at high risk for CAP could be asked about the types of interventions they may or may not want (e.g. ventilator use, resuscitation). Since the EMR score relies solely on data available in medical records, EMR systems could automatically calculate the score for providers, reducing providers' burden of using the risk score.

The comparative success of the EMR risk score suggests that the majority of pneumonia risk in older adults might be driven by age, sex, chronic heart/lung disease, and BMI, which are known risk factors.⁷⁻⁹ Some risk factors such as alcohol abuse and functional limitations appeared in the clinical risk score but did not meaningfully improve pneumonia prediction relative to the EMR risk score. Other known risk factors, such as smoking and lack of physical activity,^{9,30} did not appear in any of the risk scores. This highlights an important feature of risk prediction, which is that strong etiologic risk factors (such as smoking) are not always predictors in multivariable models, particularly when the risk factors are mediated through intermediates (such as COPD) that are also included in the model.

One interesting finding was that, in a model controlling for factors such as age, comorbidity, and prescription medications, increasing BMI was associated with decreased risk of CAP. Our study was not aimed at testing the causal association between BMI and CAP risk, so these results should be interpreted with caution, as there may be confounders of the BMI/CAP association that were not included in our models. However, there is some evidence that higher BMI may be linked to a lower pneumonia risk. Studies not restricted to older adults have found that overweight individuals have a lower risk of CAP than underweight or normal weight individuals,³¹ with conflicting data on the risk in obese individuals.^{32,33}

The risk scores presented in this paper have several limitations. Our study population was restricted to community-dwelling older adults without dementia. The scores will likely not be applicable to older adults in nursing homes or other institutions, whose pneumonia risk factors may differ from community-dwelling persons.³⁴ The risk scores may also be less predictive in older adults with dementia. Second, while the risk scores have been internally validated through the use of development and validation cohorts, further work is needed to validate the risk scores in other populations. Finally, we recognize that our model selection process likely eliminated some covariates with actual prognostic ability or that we might have failed to consider some predictive interactions. Our goal, though, was to construct a

parsimonious tool that would not be too cumbersome to assess in clinical settings, even if achieving this simplicity meant sacrificing some model performance.

The seven-variable EMR risk score for pneumonia in older adults, based on readily available data, may be useful both for clinical practice and for public health research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Lisa Jackson, MD, (Group Health Research Institute) for the use of the pneumonia outcome data.

Funding: This project was funded by the Branta Foundation and by National Institute of Aging grant K23AG028954 (PI: Sascha Dublin). Data on pneumonia outcomes was funded by a cooperative agreement between the Centers for Disease Control and Prevention and the Association of Teachers of Preventive Medicine U50 CCU300860 TS-1244 (PI: Lisa Jackson). The Adult Changes in Thought study is funded by the National Institute on Aging grant U01AG006781 (PI: Eric Larson)

Conflict of Interest: Sascha Dublin received a Merck/American Geriatrics Society New Investigator Award. Rod Walker has received funding as a biostatistician from a research grant awarded to Group Health Research Institute from Pfizer.

REFERENCES

1. Almirall J, Bolibar I, Vidal J, et al. Epidemiology of community-acquired pneumonia in adults: A population-based study. *Eur Respir J*. 2000; 15:757–763. [PubMed: 10780770]
2. Jackson ML, Neuzil KM, Thompson WW, et al. The burden of community-acquired pneumonia in seniors: Results of a population-based study. *Clin Infect Dis*. 2004; 39:1642–1650. [PubMed: 15578365]
3. Jokinen C, Heiskanen L, Juvonen H, et al. Incidence of community-acquired pneumonia in the population of four municipalities in eastern Finland. *Am J Epidemiol*. 1993; 137:977–988. [PubMed: 8317455]
4. Akram AR, Chalmers JD, Taylor JK, et al. An evaluation of clinical stability criteria to predict hospital course in community-acquired pneumonia. *Clin Microbiol Infect*. 2013; 19:1174–1180. [PubMed: 23438068]
5. Koivula I, Sten M, Makela PH. Prognosis after community-acquired pneumonia in the elderly: A population-based 12-year follow-up study. *Arch Intern Med*. 1999; 159:1550–1555. [PubMed: 10421277]
6. Jackson ML, Nelson JC, Jackson LA. Risk factors for community-acquired pneumonia in immunocompetent seniors. *J Am Geriatr Soc*. 2009; 57:882–888. [PubMed: 19453307]
7. Almirall J, Bolibar I, Balanzo X, et al. Risk factors for community-acquired pneumonia in adults: A population-based case-control study. *Eur Respir J*. 1999 Feb; 13(2):349–355. 1999. [PubMed: 10065680]
8. Koivula I, Sten M, Makela PH. Risk factors for pneumonia in the elderly. *Am J Med*. 1994; 96:313–320. [PubMed: 8166149]
9. LaCroix AZ, Lipson S, Miles TP, et al. Prospective study of pneumonia hospitalizations and mortality of U.S. older people: The role of chronic conditions, health behaviors, and nutritional status. *Public Health Rep*. 1989; 104:350–360. [PubMed: 2502806]
10. Juthani-Mehta M, De Rekeneire N, Allore H, et al. Modifiable risk factors for pneumonia requiring hospitalization of community-dwelling older adults: The Health, Aging, and Body Composition Study. *J Am Geriatr Soc*. 2013; 61:1111–1118. [PubMed: 23772872]

11. Yende S, Alvarez K, Loehr L, et al. Epidemiology and long-term clinical and biologic risk factors for pneumonia in community-dwelling older Americans: Analysis of three cohorts. *Chest*. 2013; 144:1008–1017. [PubMed: 23744106]
12. Mather JF, Fortunato GJ, Ash JL, et al. Prediction of Pneumonia 30-day readmissions: A single-center attempt to increase model performance. *Respir Care*. 2014; 59:199–208. [PubMed: 23942750]
13. Musonda P, Sankaran P, Subramanian DN, et al. Prediction of mortality in community-acquired pneumonia in hospitalized patients. *Am J Med Sci*. 2011; 342:489–493. [PubMed: 21876427]
14. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med*. 1997; 336:243–250. [PubMed: 8995086]
15. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: An international derivation and validation study. *Thorax*. 2003; 58:377–382. [PubMed: 12728155]
16. Ji R, Shen H, Pan Y, et al. Novel risk score to predict pneumonia after acute ischemic stroke. *Stroke*. 2013; 44:1303–1309. [PubMed: 23482598]
17. Harms H, Grittner U, Droge H, et al. Predicting post-stroke pneumonia: The PANTHERIS score. *Acta Neurol Scand*. 2013; 128:178–184. [PubMed: 23461541]
18. Kukull WA, Higdon R, Bowen JD, et al. Dementia and Alzheimer disease incidence: A prospective cohort study. *Arch Neurol*. 2002; 59:1737–1746. [PubMed: 12433261]
19. Teng EL, Hasegawa K, Homma A, et al. The Cognitive Abilities Screening Instrument (CASI): A practical test for cross-cultural epidemiological studies of dementia. *Int Psychogeriatr*. Spring. 1994; 6:45–58. discussion 62.
20. Nelson JC, Jackson M, Yu O, et al. Impact of the introduction of pneumococcal conjugate vaccine on rates of community acquired pneumonia in children and adults. *Vaccine*. 2008; 26:4947–4954. [PubMed: 18662735]
21. Update: pneumococcal polysaccharide vaccine usage--United States. *MMWR Morb Mortal Wkly Rep*. 1984; 33:273–276. 281. [PubMed: 6425629]
22. Jackson LA, Neuzil KM, Yu O, et al. Effectiveness of pneumococcal polysaccharide vaccine in older adults. *New Engl J Med*. 2003; 348:1747–1755. [PubMed: 12724480]
23. Tomczyk S, Bennett NM, Stoecker C, et al. Use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine among adults aged ≥ 65 years: Recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Morb Mortal Wkly Rep*. 2014; 63:822–825. [PubMed: 25233284]
24. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007; 16:219–242. [PubMed: 17621469]
25. Fine JP, Gray RJ. A Proportional Hazards Model for the subdivision of a competing risk. *J Am Stat Assoc*. 1999; 94:496–509.
26. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology*. 2010; 21:128–138. [PubMed: 20010215]
27. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15:361–387. [PubMed: 8668867]
28. Wolbers M, Koller MT, Wittman JC, et al. Prognostic models with competing risks: methods and application to coronary risk prediction. *Epidemiology*. 2009; 20:555–561. [PubMed: 19367167]
29. Coviello V, Boggess M. Cumulative incidence estimation in the presence of competing risks. *Stata J*. 2004; 4:103–112.
30. Baik I, Curhan GC, Rimm EB, et al. A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women. *Arch Intern Med*. 2000; 160:3082–3088. [PubMed: 11074737]
31. Phung DT, Wang Z, Rutherford S, et al. Body mass index and risk of pneumonia: A systematic review and meta-analysis. *Obesity Rev*. 2013; 14:839–857.

32. Blumentals WA, Nevitt A, Peng MM, et al. Body mass index and the incidence of influenza-associated pneumonia in a UK primary care cohort. *Influenza Other Respir Viruses*. 2012; 6:28–36. [PubMed: 21668664]
33. Kornum JB, Norgaard M, Dethlefsen C, et al. Obesity and risk of subsequent hospitalisation with pneumonia. *Eur Respir J*. 2010; 36:1330–1336. [PubMed: 20351023]
34. Vergis EN, Brennen C, Wagener M, et al. Pneumonia in long-term care: A prospective case-control study of risk factors and impact on survival. *Arch Intern Med*. 2001; 161:2378–2381. [PubMed: 11606155]

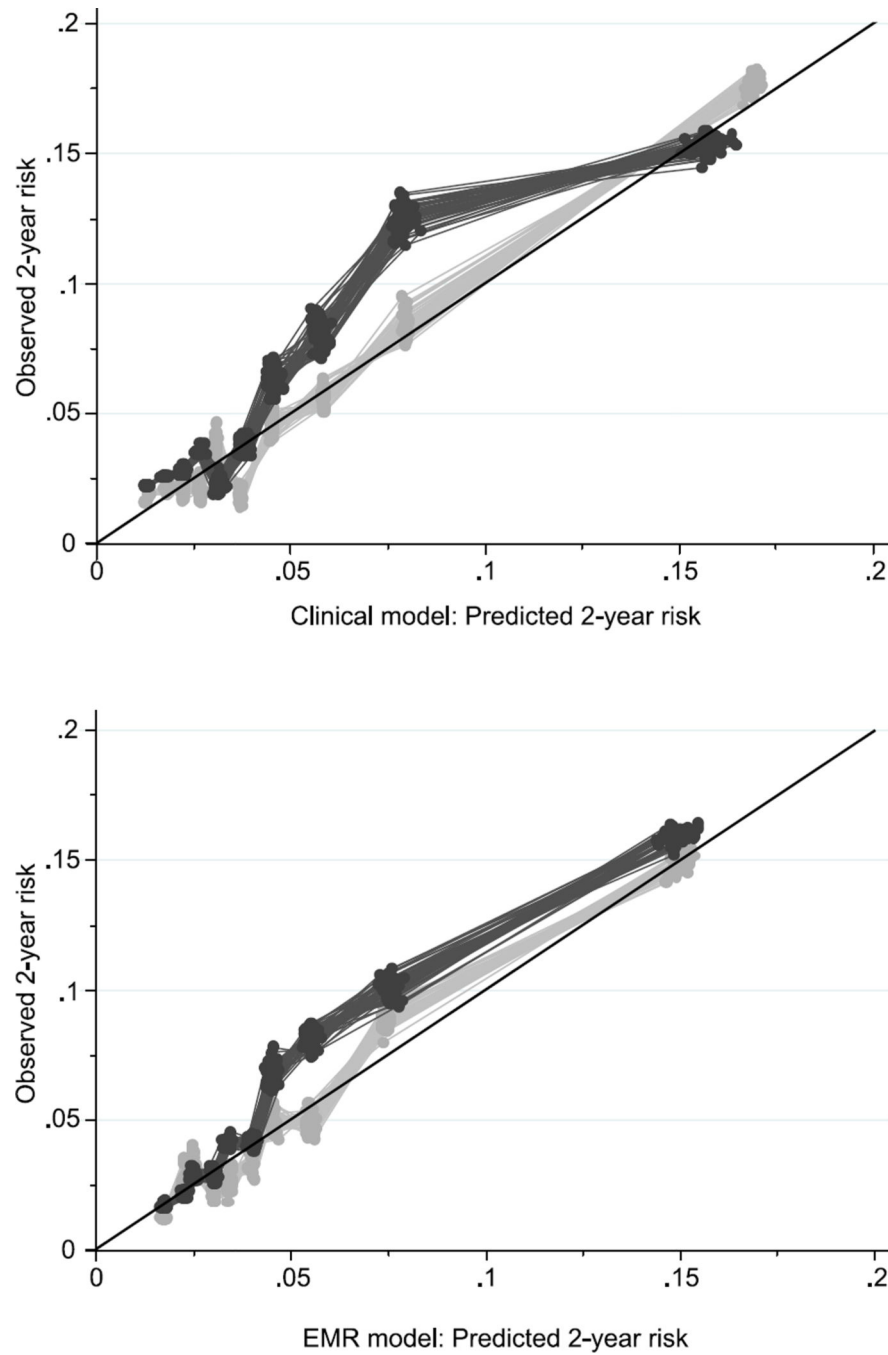


Figure 1. Predicted vs. observed two-year pneumonia risk in dementia-free older adults from two prediction models, across 50 imputed datasets. Light grey dots, pneumonia risk within the development cohort; dark grey dots, pneumonia risk within the validation cohort. (Top panel) Clinical model; (Bottom panel) Electronic Medical Record (EMR) model. Diagonal lines indicate perfect agreement between observed and predicted risk.

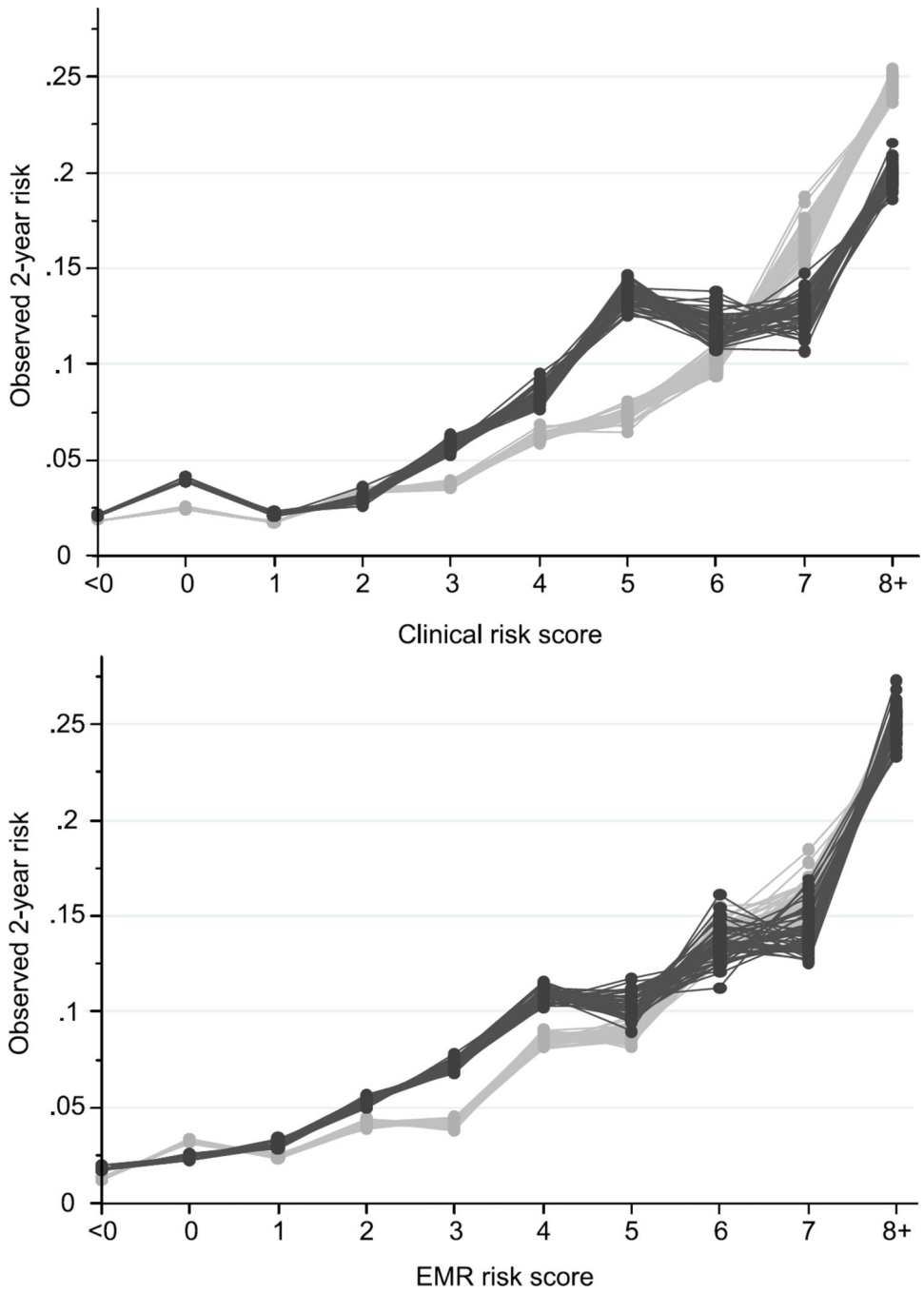


Figure 2. Pneumonia risk score vs. observed pneumonia risk in dementia-free seniors from two prediction models, across 50 imputed datasets. Light grey dots, pneumonia risk within the development cohort; dark grey dots, pneumonia risk within the validation cohort. (Top panel) Clinical risk score; (Bottom panel) Electronic Medical Record (EMR) risk score

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Distribution of selected covariates and outcomes in study population, development cohort, and validation cohort. Distribution in the study population is among all participants at the time of Adult Changes in Thought (ACT) enrollment. Distribution in the development and validation cohorts is among all visits over the course of the study.

	Study population	Development Cohort	Validation Cohort	Association with pneumonia risk in development cohort	
				Hazard ratio* (95% CI)	p-value
Participants	3,375	2,250	1,125		
Visits among the participants	12,998	8,579	4,419		
Outcomes	N participants (%)	N visits (%)	N visits (%)		
Pneumonia events	642 (19)	401 (5)	241 (5)		
Deaths	574 (17)	399 (5)	175 (4)		
Disenrollment from GH or ACT	423 (13)	278 (3)	145 (3)		
Demographic domain					
Age					<0.001
65–69	723 (21)	649 (8)	349 (8)	REF	
70–74	990 (29)	2000 (23)	1030 (23)	0.94 (0.58, 1.53)	
75–79	795 (24)	2507 (29)	1275 (29)	1.26 (0.80, 1.98)	
80–84	546 (16)	1996 (23)	1036 (23)	1.47 (0.92, 2.34)	
85–89	247 (7)	1043 (12)	494 (11)	2.67 (1.67, 4.27)	
90+	74 (2)	384 (4)	235 (5)	3.19 (1.88, 5.43)	
Male	1385 (41)	3455 (40)	1758 (40)	1.49 (1.19, 1.86)	<0.001
Education					0.810
Less than high school	431 (13)	968 (11)	502 (11)	1.20 (0.81, 1.78)	
Complete high school	880 (26)	2159 (25)	1181 (27)	REF	
Some college	903 (27)	2284 (27)	1169 (26)	1.00 (0.73, 1.38)	
Completed college	1161 (34)	3168 (37)	1567 (35)	1.03 (0.77, 1.38)	
Comorbidity domain					

	Study population	Development Cohort	Validation Cohort	Association with pneumonia risk in development cohort	
				Hazard ratio* (95% CI)	p-value
COPD	355 (11)	1385 (16)	628 (14)	2.48 (1.96, 3.12)	<0.001
CHF	147 (4)	591 (7)	311 (7)	2.50 (1.86, 3.37)	<0.001
History of pneumonia	959 (28)	3241 (38)	1644 (37)	2.55 (2.07, 3.13)	<0.001
Home oxygen use in prior 24 months	43 (1)	92 (1)	49 (1)	5.18 (3.27, 8.21)	<0.001
Any forearm, hip, or spine fracture in prior 24 months	54 (2)	156 (2)	88 (2)	2.02 (1.19, 3.43)	0.009
Functional domain					
Any difficulty with:					
Getting to or using a toilet	67 (2)	321 (4)	158 (4)	1.95 (1.32, 2.90)	0.001
Dressing yourself	170 (5)	622 (7)	306 (7)	2.27 (1.73, 3.00)	<0.001
Level of difficulty with walking a half- mile					<0.001
None	2502 (74)	5667 (67)	2912 (67)	REF	
Some	390 (12)	1120 (13)	580 (13)	1.56 (1.16, 2.09)	
Much / Cannot do	470 (14)	1685 (20)	876 (20)	2.31 (1.82, 2.92)	
Missing	13 (0)	107 (1)	51 (1)		
Use of rehabilitation hospital	44 (1)	258 (3)	137 (3)	2.48 (1.66, 3.72)	<0.001
Vision problems	2736 (81)	7059 (82)	3729 (84)	1.12 (0.83, 1.51)	0.448
Behavioral domain					
Smoking status					0.003
Never smoker	1595 (47)	4024 (47)	2172 (50)	REF	
Former smoker	1573 (47)	4105 (48)	2020 (46)	1.49 (1.18, 1.88)	
Current smoker	205 (6)	388 (5)	195 (4)	1.21 (0.73, 1.98)	
History of alcohol-related aggressive behavior [†]	130 (4)	349 (4)	155 (4)	1.97 (1.30, 2.99)	0.001
Weekly exercise					
Aerobics					0.853
0 days	2979 (88)	7433 (87)	3878 (89)	REF	

	Study population	Development Cohort	Validation Cohort	Association with pneumonia risk in development cohort	
				Hazard ratio* (95% CI)	p-value
1-3 days	230 (7)	674 (8)	339 (8)	1.03 (0.67, 1.58)	
4-7 days	166 (5)	396 (5)	162 (4)	0.87 (0.53, 1.43)	
Walking					0.118
0 days	1279 (38)	4266 (50)	2223 (51)	REF	
1-3 days	840 (25)	1720 (20)	858 (20)	0.80 (0.61, 1.07)	
4-7 days	1256 (37)	2518 (30)	1300 (30)	0.79 (0.61, 1.02)	
Body mass index					<0.001
<18.5	38 (1)	105 (1)	48 (1)	2.08 (1.16, 3.72)	
18.5 to <25	1100 (33)	2782 (34)	1403 (34)	REF	
25 to <30	1349 (41)	3317 (41)	1769 (42)	0.76 (0.59, 0.97)	
30 to <35	621 (19)	1434 (18)	757 (18)	0.76 (0.55, 1.04)	
35	209 (6)	501 (6)	199 (5)	0.48 (0.27, 0.86)	
Pharmacy domain (use in prior 12 months)					
Bronchodilator	208 (6)	567 (7)	361 (8)	3.60 (2.76, 4.70)	<0.001
Inhaled corticosteroid	156 (5)	419 (5)	249 (6)	3.57 (2.63, 4.86)	<0.001
Oral corticosteroid	158 (5)	363 (4)	244 (6)	3.35 (2.43, 4.63)	<0.001
Statin	271 (8)	1381 (16)	666 (15)	1.40 (1.07, 1.82)	0.013
Angiotensin-converting enzyme inhibitor	443 (13)	1661 (19)	832 (19)	1.24 (0.96, 1.59)	0.098
Immunosuppressive	35 (1)	98 (1)	53 (1)	1.56 (0.69, 3.50)	0.286
Antiglaucoma	266 (8)	824 (10)	383 (9)	1.10 (0.78, 1.54)	0.589
Detailed domain					
Balance tests					0.001
Unable to complete	802 (25)	1944 (27)	987 (26)	REF	
Completed	2360 (75)	5370 (73)	2779 (74)	0.67 (0.54, 0.85)	
Timed chair stand test (time to complete five chair stands)					<0.001

	Study population	Development Cohort	Validation Cohort	Association with pneumonia risk in development cohort	
				Hazard ratio* (95% CI)	p-value
Unable to complete	323 (10)	1340 (17)	661 (16)	REF	
Completed in 20+ seconds	404 (12)	977 (12)	491 (12)	0.64 (0.46, 0.91)	
Completed in <20 seconds	2577 (78)	5780 (71)	3015 (72)	0.49 (0.38, 0.62)	
Administrative data code (in prior 24 months) for :					
Pneumonia	206 (6)	741 (9)	411 (9)	3.53 (2.75, 4.53)	<0.001
Malnutrition	79 (2)	332 (4)	192 (4)	2.28 (1.54, 3.37)	<0.001
Serious cancer	86 (3)	303 (4)	167 (4)	1.80 (1.19, 2.72)	0.005
Hematocrit level within prior 24 months:					0.044
<34	123 (4)	328 (5)	177 (5)	1.69 (1.09, 2.62)	
34-49 (males), 34-47 (females)	2688 (95)	6527 (94)	3393 (94)	REF	
>49 (males), >47 (females)	25 (1)	85 (1)	44 (1)	0.27 (0.04, 1.99)	

* Estimate from the Fine and Gray model, also known as a proportional sub-distribution hazards model (see Supplemental Methods)

[†] Participants were asked, "Have you had problems with aggressive behavior (e.g. fighting) while under the influence of alcohol?"

Table 2

Model coefficients and risk scores

Covariates	Clinical model		EMR model	
	Hazard ratio* (95% CI)	Points	Hazard ratio* (95% CI)	Points
Age				
65–69	REF	0	REF	0
70–74	0.87 (0.53, 1.42)		0.91 (0.56, 1.49)	
75–79	1.09 (0.69, 1.72)		1.18 (0.74, 1.86)	
80–84	1.16 (0.73, 1.87)	1	1.31 (0.82, 2.11)	3
85–89	1.78 (1.11, 2.87)	2	2.19 (1.35, 3.54)	
90+	1.77 (1.01, 3.10)		2.43 (1.40, 4.20)	
Gender				
Male	1.75 (1.40, 2.19)	2	1.73 (1.38, 2.17)	2
Comorbidity				
COPD	1.59 (1.23, 2.06)	2	2.03 (1.57, 2.62)	3
CHF	1.61 (1.17, 2.19)	2	1.91 (1.40, 2.60)	2
Pneumonia	1.94 (1.56, 2.41)	2	NI [†]	NI
Home oxygen use in prior 24 months	1.68 (1.02, 2.76)	2	NI	NI
Functional				
Any difficulty dressing yourself	1.42 (1.04, 1.93)	1	NA [‡]	NA
Difficulty walking a half- mile				
Some difficulty	1.33 (0.98, 1.80)	1	NA	NA
Much difficulty/cannot do	1.44 (1.08, 1.91)		NA	NA
Behavioral				
History of alcohol- related aggressive behavior	1.69 (1.12, 2.54)	2	NA	NA
Body mass index				
<18.5	1.39 (0.76, 2.54)	1	1.41 (0.77, 2.60)	1
18.5 to <25	REF	0	REF	0
25 to <30	0.75 (0.58, 0.95)	–1	0.77 (0.60, 0.98)	–1
30 to <35	0.69 (0.49, 0.97)		0.78 (0.56, 1.09)	
35	0.45 (0.25, 0.81)	–3	0.54 (0.30, 0.96)	–2
Pharmacy				
Inhaled corticosteroid	1.49 (1.05, 2.12)	1	1.75 (1.21, 2.51)	2
Oral corticosteroid	1.87 (1.32, 2.66)	2	2.22 (1.56, 3.15)	3
C-index (95% CI) Development cohort	0.72 (0.70, 0.75)	0.72 (0.69, 0.75)	0.69 (0.67, 0.72)	0.69 (0.66, 0.72)
C-index (95% CI) Validation cohort	0.70 (0.66, 0.73)	0.69 (0.66, 0.73)	0.69 (0.66, 0.73)	0.69 (0.66, 0.73)

* Estimate from the Fine and Gray model, also known as a proportional sub-distribution hazards model (see Supplemental Methods)

[†]NI = Not included in final model

[‡]NA = Not applicable (i.e. variable not expected to be available in EMR)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript