Title

Epidemiology and Long-term Clinical and Biologic Risk Factors for Pneumonia in Community-Dwelling Older Americans Analysis of Three Cohorts

Journal

CHEST Journal, 144(3)

ISSN

0012-3692

Authors

Yende, Sachin
Alvarez, Karina
Loehr, Laura
et al.

Publication Date

2013-09-01

DOI

10.1378/chest.12-2818

Peer reviewed

CHEST

Original Research

CHEST INFECTIONS

Epidemiology and Long-term Clinical and Biologic Risk Factors for Pneumonia in Community-Dwelling Older Americans

Analysis of Three Cohorts

Sachin Yende, MD; Karina Alvarez, MS; Laura Loehr, MD, PhD; Aaron R. Folsom, MD; Anne B. Newman, MD, MPH; Lisa A. Weissfeld, PhD; Richard G. Wunderink, MD, FCCP; Stephen B. Kritchevsky, PhD; Kenneth J. Mukamal, MD; Stephanie J. London, MD; Tamara B. Harris; MD; Doug C. Bauer, MD; and Derek C. Angus, MD, MPH, FCCP; for the Atherosclerosis Risk in Communities Study, the Cardiovascular Health Study, and the Health, Aging, and Body Composition Study

Background: **Preventing pneumonia requires better understanding of incidence, mortality, and long-term clinical and biologic risk factors, particularly in younger individuals.**
Methods: **This was a cohort study in three population-based cohorts of community-dwelling individuals. A derivation cohort (n = 16,260) was used to determine incidence and survival and develop a risk prediction model. The prediction model was validated in two cohorts (n = 8,495). The primary outcome was 10-year risk of pneumonia hospitalization.**
Results: **The crude and age-adjusted incidences of pneumonia were 6.71 and 9.43 cases/1,000 person-years (10-year risk was 6.15%). The 30-day and 1-year mortality were 16.5% and 31.5%. Although age was the most important risk factor (range of crude incidence rates, 1.69-39.13 cases/1,000 person-years for each 5-year increment from 45-85 years), 38% of pneumonia cases occurred in adults < 65 years of age. The 30-day and 1-year mortality were 12.5% and 25.7% in those < 65 years of age. Although most comorbidities were associated with higher risk of pneumonia, reduced lung function was the most important risk factor (relative risk = 6.61 for severe reduction based on FEV_1 by spirometry). A clinical risk prediction model based on age, smoking, and lung function predicted 10-year risk (area under curve [AUC] = 0.77 and Hosmer-Lemeshow [HL] C statistic = 0.12). Model discrimination and calibration were similar in the internal validation cohort (AUC = 0.77; HL C statistic, 0.65) but lower in the external validation cohort (AUC = 0.62; HL C statistic, 0.45). The model also calibrated well in blacks and younger adults. C-reactive protein and IL-6 were associated with higher pneumonia risk but did not improve model performance.**
Conclusions: **Pneumonia hospitalization is common and associated with high mortality, even in younger healthy adults. Long-term risk of pneumonia can be predicted in community-dwelling adults with a simple clinical risk prediction model.** ***CHEST 2013; 144(3):1008–1017***

Abbreviations: ARIC = Atherosclerosis Risk in Communities; AUC = area under curve; CAP = community-acquired pneumonia; CDC = Centers for Disease Control and Prevention; CHS = Cardiovascular Health Study; CRP = C-reactive protein; Health ABC = Health, Aging, and Body Composition; HL = Hosmer-Lemeshow; ICD-9 = *International Classification of Diseases, Ninth Revision*; IDI = integrated discrimination improvement; IRB = institutional review board; NRI = net reclassification improvement; RR = relative risk

Community-acquired pneumonia (CAP) is the most common infectious cause of hospitalization in the United States.[1] CAP is associated with high short- and long-term mortality and many morbid sequelae that may last for several years.[2-4] However, it is not known whether the long-term risk of pneumonia can be predicted in community-dwelling individuals, similar to cardiovascular disease and cancer.[5,6]

Individuals who are older than 65 years or have a chronic health condition are considered to be at high

risk for pneumonia.[7] These findings are based on prior studies that were conducted in small cohorts, which included few young individuals and had limited ethnic diversity and short follow-up,[8-12] or in administrative datasets, which lacked accurate information about clinical risk factors.[13] Whether the long-term risk of pneumonia is similar across all older adults and for different chronic health conditions is not known. Animal and small human studies suggest that circulating biomarker levels may be associated with higher risk of pneumonia,[14] but their role has not been assessed systematically in a large population-based study.

Understanding the role of individual clinical and biologic risk factors and predicting the long-term risk of pneumonia in an individual is important to target preventive strategies that have adverse effects and are expensive, such as statins.[15-23] Clinical trials to test these interventions will have to be conducted in high-risk individuals first, and, if effective, these interventions could be targeted to broader populations based on the risk-benefit profile. However, there is no tool to predict risk of pneumonia, similar to risk-prediction tools for cardiovascular disease and cancer.[5,6] Current guidelines for pneumococcal vaccination target a very broad population and cannot be used to test interventions that are expensive or have adverse events. We, therefore, sought to determine the incidence and understand the role of clinical and biologic factors that increase the 10-year risk of pneumonia hospitalizations. We developed and validated a risk prediction model in > 24,000 participants from three nationally representative cohorts that span middle and older age.

Manuscript received November 18, 2012; revision accepted April 15, 2013.

Affiliations: From The Clinical Research, Investigation, and Systems Modeling of Acute Illness (CRISMA) Center (Drs Yende, Weissfeld, and Angus and Ms Alvarez), the Department of Critical Care Medicine (Drs Yende and Angus), the Department of Biostatistics (Ms Alvarez and Dr Weissfeld), and the Department of Epidemiology (Dr Newman), University of Pittsburgh, Pittsburgh, PA; the Department of Epidemiology (Dr Loehr), UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC; the School of Public Health (Dr Folsom), University of Minnesota, Minneapolis, MN; the Division of Pulmonary and Critical Care Medicine (Dr Wunderink), Northwestern University Feinberg School of Medicine, Chicago, IL; The Sticht Center on Aging (Dr Kritchevsky), Wake Forest School of Medicine, Winston-Salem, NC; the Department of Medicine (Dr Mukamal), Beth Israel Deaconess Medical Center, Boston, MA; the Epidemiology Branch (Dr London), National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, NC; the Laboratory of Epidemiology, Demography, and Biometry (Dr Harris), National Institute on Aging, Bethesda, MD; and the Departments of Medicine and Epidemiology and Biostatistics (Dr Bauer), University of California, San Francisco, CA.

Funding/Support: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by the National Heart, Lung, and Blood Institute (NHLBI) [Contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022]. This research was supported by NHLBI [Contracts HHSN268201200036C, HHSN268200800007C, N01-HC-55222, N01-HC-85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, H01-HC-85086 and Grant HL080295], with additional contribution from National Institute of Neurological Disorders and Stroke. Additional support was provided by the National Institute on Aging (NIA) [Grant AG-023629]. A full list of principal Cardiovascular Health Study investigators and institutions can be found at CHS-NHLBI.org. The Health, Aging, and Body Composition Study is supported by NIA [Contracts N01-AG-6-2101, N01-AG-6-2103, N01-AG-6-2106, R01-AG028050, and R01-NR012459]. This study was in part funded by National Institutes of Health [Grant K23GM083215 to Dr Yende] and Intramural Research Programs of the National Institute of Environmental Health Sciences and NIA.

Correspondence to: Sachin Yende, MD, University of Pittsburgh, 606D Scaife Hall, 3550 Terrace St, Pittsburgh, PA 15261; e-mail: yendes@upmc.edu

DOI: 10.1378/chest.12-2818

Materials and Methods

Subjects and Design

See e-Appendix 1 for details. We used a cohort design, using three population-based prospective studies.[24-26] We pooled data from the Atherosclerosis Risk in Communities (ARIC) (n = 15,792) and Cardiovascular Health Study (CHS) (n = 5,888) cohorts because these cohorts included community-dwelling participants without disability and had nonoverlapping age ranges (ARIC, 45-64 years; CHS, > 65 years), they were recruited during the same time period (ARIC, 1987-1989; CHS, 1989-1990) and from similar geographic areas, they used similar study procedures, and they had similar follow-up (> 10 years). We randomly split the pooled data into three quarters to determine the incidence, understand the role of individual risk factors, and integrate these risk factors to develop a risk prediction model (derivation cohort). The remaining one-quarter of the pooled data were used for internal validation of the model (internal validation cohort). For external validation, we used the Health, Aging, and Body Composition (Health ABC) (n = 3,075) study, a more contemporary cohort recruited in 1997-1998 with 10-year follow-up data.

The institutional review boards (IRBs) at each site approved all three studies (ARIC IRB approval #11-0734 at University of North Carolina, and CHS IRB approval #950706 and Health ABC IRB approval # 960212 at University of Pittsburgh). Participants gave written informed consent.

Outcomes

The primary outcome was hospitalization with pneumonia within a 10-year period. We used a combination of *International Classification of Diseases, Ninth Revision* (ICD-9) codes with chart review and physician adjudication. For ARIC and CHS, we used ICD-9 codes 480 to 487 in the first five discharge diagnoses fields to identify hospitalizations with pneumonia. We also identified hospitalizations for bacterial pneumonia by including hospitalizations with ICD-9 codes that are commonly used for bacterial pneumonia (481, 482, 485, and 486) and are listed only in the primary diagnosis field. In Health ABC, physicians adjudicated hospitalizations for pneumonia after chart review.

The use of 480 to 487 ICD-9 codes to identify pneumonia hospitalizations has been validated in prior studies,[13,27,28] and we also reviewed 158 charts to assess accuracy of these ICD-9 codes in our cohorts. For ICD-9 codes used for bacterial pneumonia, a

clinical diagnosis of CAP was recorded in 94% of cases, and a radiologic diagnosis was recorded in 96% of cases. For ICD-9 codes 480 to 487 in the first five discharge diagnoses fields, a clinical and radiologic diagnosis was recorded in 89% and 88% of cases. For cases in which a diagnosis of pneumonia was recorded in second through fifth diagnoses fields, pneumonia was the primary reason for hospitalization in 65% of cases.

Clinical Risk Factors

Clinical risk factors for pneumonia were chosen based on literature review and our prior work in CHS and Health ABC.[8-12,24] We assessed the role of the following clinical risk factors at baseline: demographic characteristics (age, sex, and race), BMI, health behaviors (smoking and alcohol abuse), lung function (FEV_1 based on spirometry), and chronic diseases (heart failure, coronary heart disease, diabetes, and chronic kidney disease).

Inflammatory Biomarkers

See e-Appendix 1 for details. We sought to determine whether specific inflammatory biomarkers increased risk and improved the performance of the clinical risk prediction model. Higher proinflammatory cytokines increase risk of pneumococcal pneumonia in animal studies.[14,24,29] We examined circulating C-reactive protein (CRP) level because it is an important opsonin and may protect against pneumococcal pneumonia.[30,31] We examined IL-6 level because it may be a marker for increased cellular senescence[32] and may increase adherence of pneumococci to cells in the lung.[14,29] CRP levels were available in all three cohorts at different time points, but IL-6 levels were available in Health ABC only.

Statistical Analyses

We first determined frequency, crude and age-adjusted incidence rates of pneumonia by age groups, and mortality after pneumonia. We conducted univariate comparisons to determine relative risk (RR) for each risk factor. We then developed a simple and parsimonious clinical risk prediction model to predict 10-year risk of pneumonia hospitalization in the derivation cohort using logistic regression. We assessed the discriminative ability of each variable by calculating the area under curve (AUC). We sequentially added variables to age, which had the highest AUC, calculated AUC and net reclassification improvement (NRI) for each combination, and chose the combination that had the highest AUC and NRI.[33] We assessed calibration of the model by comparing the predicted and observed risk and calculating the Hosmer-Lemeshow (HL) C statistic. We validated the clinical model in the internal and external validation cohorts.

We conducted four sensitivity analyses: (1) compared estimates for different sets of ICD-9 codes due to differing sensitivity and specificity to determine hospitalizations for pneumonia; (2) compared estimates from logistic regression (RR) and Cox proportional hazards (hazard ratio) models, because hazard ratios account for censoring due to mortality; (3) assessed model calibration in blacks and those <65 years of age to determine generalizability; and (4) evaluated ability to predict 5-year risk, since clinical trials are often performed over 5 years, and some variables may disproportionately predict more proximate episodes of pneumonia.

We compared the clinical usefulness of the risk prediction model and the Centers for Disease Control and Prevention (CDC) guidelines for pneumococcal vaccination by comparing the proportion of the population who would have to be treated and the proportion of pneumonia cases that would be identified using different thresholds of the clinical risk prediction model and for the CDC guidelines.[7,34] Finally, we assessed incremental benefit of adding biomarkers to the clinical risk prediction model by estimating improvement in AUC and calculating NRI and integrated discrimination improvement (IDI).[33]

Results

Baseline characteristics of the three individual cohorts and the derivation (n = 16,260), internal validation (n = 5,420), and external validation (n = 3,075) cohorts are described in Table 1 and e-Appendix 1. The derivation cohort included 24% black participants, and the age distribution was similar to the age distribution of the US population ≥ 45 years old in 1990, with slightly fewer older individuals (e-Table 1). Thus, the derivation cohort was fairly representative of community-dwelling individuals from the US population. Most clinical characteristics were virtually identical in the derivation and internal validation cohorts. The prevalence of chronic kidney disease, as evidenced by glomerular filtration rate <60 mL/min/1.73 m^2, was lower in the derivation and internal validation cohorts compared with the external validation cohorts (7.3% vs 13.31%).

Incidence and Mortality

Of the 16,260 participants, 1,000 participants (6.15%) were hospitalized with pneumonia at least once over 10 years. Of these, most were hospitalized for bacterial pneumonia (n = 660). The median time to hospitalization was 5.95 years (interquartile range = 3.48-8.19 years). The crude and age-adjusted incidences were 6.71 and 9.43 cases/1,000 person-years. The 30-day and 1-year mortality after pneumonia were 16.5% and 31.5%. Even subjects hospitalized with bacterial pneumonia, for whom an ICD-9 code for pneumonia was recorded as the primary diagnosis code, had high 30-day case fatality and 1-year mortality (12.12% and 26.67%).

Risk Factors

Univariate analyses with RR, AUC, and missing data for individual risk factors are shown in Table 2. Age was an important risk factor, and the average age of developing pneumonia was 75 years (SD = 10 years). Age was associated with increased risk (Fig 1). Although most pneumonia cases (62%) occurred among those >65 years of age, 38% occurred among those 45 to 65 years old (Fig 1). The 30-day and 1-year mortality for those 45 to 65 years old were 12.5% and 25.7%. The 1-year mortality after pneumonia hospitalization remained high among younger participants, when we excluded those with a chronic health condition (14.29%) and additionally excluded those who ever smoked (12.82%).

Severe reduction in lung function, as evidenced by percent predicted $FEV_1 < 50\%$, increased the risk of CAP 6.61-fold (95% CI = 5.25-8.32) compared with

Table 1—*Clinical Characteristics of Various Cohorts at Enrollment*

Variables	ARIC (n = 15,792)	CHS (n = 5,888)	Health ABC (n = 3,075)	Combined ARIC and CHS: Derivation Cohort (n = 16,260)	Combined ARIC and CHS: Internal Validation (n = 5,420)
Demographics					
Age at enrollment, mean (SD)	54.16 (5.76)	72.35 (5.42)	73.62 (2.87)	59.23 (10.06)	59.24 (10.19)
Age at hospitalization for pneumonia, mean (SD), y	70.12 (6.35)	82.44 (6.00)	78.97 (3.74)	75.08 (10.36)	74.97 (10.43)
Sex, female (%)	8,710 (55.15)	3,393 (57.63)	1,584 (51.51)	9,088 (55.89)	3,015 (55.63)
Race, black (%)	4,266 (27.01)	924 (15.69)	1,281 (41.66)	3,896 (24.01)	1,294 (23.94)
Common risk factors for pneumonia at enrollment					
Smoking					
Current	4,132 (26.16)	700 (11.90)	318 (10.36)	3,577 (22.02)	1,255 (23.18)
Former	5,972 (32.15)	2,444 (41.55)	1,404 (45.73)	5,622 (34.61)	1,894 (34.98)
Never	6,572 (41.66)	2,738 (46.55)	1,348 (43.91)	7,045 (43.37)	2,265 (41.84)
Pack-y of smoking, mean (SD, median)	16.06 (21.84)	18.43 (27.17)	19.06 (28.00)	16.48 (23.04)	17.13 (23.98)
BMI, kg/m^2	27.71 (5.37)	26.88 (4.77)	27.39 (4.82)	27.40 (5.21)	27.35 (5.19)
Heart failure (%)	752 (4.85)	275 (4.67)	95 (3.11)	766 (4.77)	261 (4.89)
Coronary heart disease (%)	766 (4.96)	1,154 (19.60)	625 (20.81)	1,418 (8.86)	502 (9.43)
Diabetes (%)	1,870 (11.95)	1,493 (28.89)	468 (15.28)	1,109 (7.11)	384 (7.37)
Chronic kidney disease[a] (%)	598 (5.11)	764 (13.24)	683 (22.21)	1,169 (7.30)	193 (13.31)
Lung function using FEV_1					
% predicted FEV_1, mean (SD)	93.27 (17.42)	90.00 (22.26)	94.08 (21.89)	92.53 (18.81)	92.07 (18.98)
% predicted FEV_1					
< 50%	307 (1.96)	322 (5.80)	93 (3.25)	476 (2.99)	153 (2.88)
50%-80%	2,697 (17.22)	1,252 (22.54)	595 (20.78)	2,907 (18.28)	1,042 (19.61)
> 80%	12,654 (80.81)	3,980 (71.66)	2,175 (75.97)	12,516 (78.72)	4,118 (77.51)

All variables measured at baseline (study entry). ARIC = Atherosclerosis Risk in Communities; CHS = Cardiovascular Health Study; GFR = glomerular filtration rate; Health ABC = Health, Aging, and Body Composition.

[a]Based on estimated GFR using Modification of Diet in Renal Disease study equation. Chronic kidney disease defined as GFR < 60 mL/min/1.73 m^2.

those with $FEV_1 \geq 80\%$. A modest increase in risk was observed for those with mild or moderate reduction in lung function (percent predicted FEV_1, 50%-80%; RR = 2.38; 95% CI = 2.05-2.76) and other chronic health conditions. Smoking was also associated with a higher risk (RR and 95% CI were 1.61 [1.37-1.90] and 1.40 [1.21-1.63]) for current smokers and past smokers compared with never smokers).

Development and Validation of Clinical Risk Prediction Model

Age had the highest AUC (AUC = 0.74); the AUC for reduced lung function was 0.65, whereas other chronic diseases, race, sex, and smoking had AUC values near 0.5 (Table 2). The sequential addition of FEV_1 and smoking to age had the highest improvement in NRI (Table 3), and small improvement in AUC and NRI occurred when additional variables were added. Thus, the final clinical risk prediction model included age, % predicted FEV_1, and smoking (Table 4). Interactions of all age groups with smoking and FEV_1 were assessed and were significant for those 70 to 85 years of age (Table 4). The risk prediction model had moderate discrimination and excellent calibration (AUC = 0.77; 95% CI, 0.76-0.79; HL C statistic, 0.12) (Fig 2). Model discrimination was similar in the internal validation cohort (AUC = 0.77; HL C statistic, 0.65). Model discrimination was lower in external validation cohort (AUC = 0.62), but calibration was excellent (HL C statistic, 0.45) (Fig 2).

Sensitivity Analyses

The estimates remained unchanged using the more specific definition for hospitalization for bacterial pneumonia and using Cox proportional hazards model to predict risk of hospitalization with pneumonia (e-Table 2). Model calibration was good among black participants (HL C statistics in derivation and internal and external validation cohorts = 0.51, 0.38, and 0.78), among younger participants (HL C statistics in derivation and internal validation cohorts = 0.13 and 0.99), and the clinical prediction model calibrated well to predict 5-year risk of pneumonia (HL C statistic in derivation cohort = 0.26).

Inflammatory Biomarkers

See e-Appendix 1 for details. In all three cohorts, baseline CRP was associated with higher risk of pneumonia hospitalization (ORs with 95% CIs per SD

Table 2—*Univariate Analyses, AUC, and Missing Data for Risk Factors for Pneumonia Hospitalization in the Derivation Cohort*

Variable	Hospitalized for Pneumonia (n = 1,000)	Never Hospitalized for Pneumonia[a] (n = 15,260)	RR (95% CI)	AUC	Missing, %
Demographics					
Age at enrollment in the study, mean (SD), y	67.92 (10.65)	58.67 (9.75)	…[b]	0.74	0
Sex, female (%)	496 (49.60)	8,592 (56.30)	1.31 (1.15-1.49)	0.53	0
Race, black (%)	202 (20.24)	3,694 (24.26)	0.85 (0.73-0.99)	0.51	0.2
Alcohol abuse (%)	90 (9.08)	1422(9.38)	0.97 (0.77-1.21)	0.50	0.6
Smoking, current/former/never (%)	27.50/38.00/34.50	21.66/34.39/43.95	1.61 (1.37-1.90)[c] 1.40 (1.21-1.63)[d]	0.55	0.1
Pack-y of smoking, mean (SD)	24.71 (29.12)	15.85 (22.52)	1.01 (1.01-1.02)	0.58	2.1
BMI, kg/m^2 (< 18.5 and > 25 vs 18.5-25)	26.55 (5.08)	27.46 (5.21)	1.96 (1.29-2.97) 0.84 (0.74-0.97)	0.54	3.4
Heart failure (%)	91 (9.17)	675 (4.48)	2.15 (1.71-2.70)	0.52	1.2
Coronary heart disease (%)	202 (20.38)	1216 (8.09)	2.90 (2.46-3.42)	0.56	1.5
Diabetes (%)	620 (67.17)	4,668 (31.82)	2.03 (1.76-2.35)	0.53	4.1
Chronic kidney disease[e] (%)	145 (14.83)	1,024 (6.81)	2.38 (1.97-2.87)	0.54	1.5
Lung function					
% predicted FEV_1, mean (SD)	81.49 (23.86)	93.22 (18.24)	…[b]	0.65	2.2
FEV1, % predicted < 50, 50-80 vs > 80	11.73/30.28/58.00	2.45/17.53/80.02	6.61 (5.25-8.32) 2.38 (2.05-2.76)	0.62	…

AUC = area under curve; RR = relative risk. See Table 1 legend for expansion of other abbreviations.

[a]*P* values for all comparisons were < .0001, except for race, where *P* = .04.

[b]Incidence rates for different age groups described in text.

[c]RR for current vs never smokers.

[d]RR for former vs never smokers.

[e]Based on estimated GFR using Modification of Diet in Renal Disease Study equation. Chronic kidney disease defined as GFR < 60 mL/min/1.73 m^2.

increase in CRP for ARIC, CHS, and Health ABC were 1.21 [1.15-1.28], 1.08 [1.02-1.15], and 1.12 [1.02-1.21], respectively, and *P* values were < .0001, < .0001, and .04, respectively). However, the addition of CRP to the clinical model resulted in a small improvement in model discrimination. For example, in ARIC, the AUC improved from 0.67 to 0.69, 5.7% of cases were reclassified (*P* = .02) (e-Table 3), and IDI was 0.5% (*P* < .0001). In CHS, there was no improvement in the AUC, no improvement in NRI, and very small improvement in IDI (0.1%, *P* = .03) (e-Table 4). In Health ABC, the addition of CRP to the clinical risk

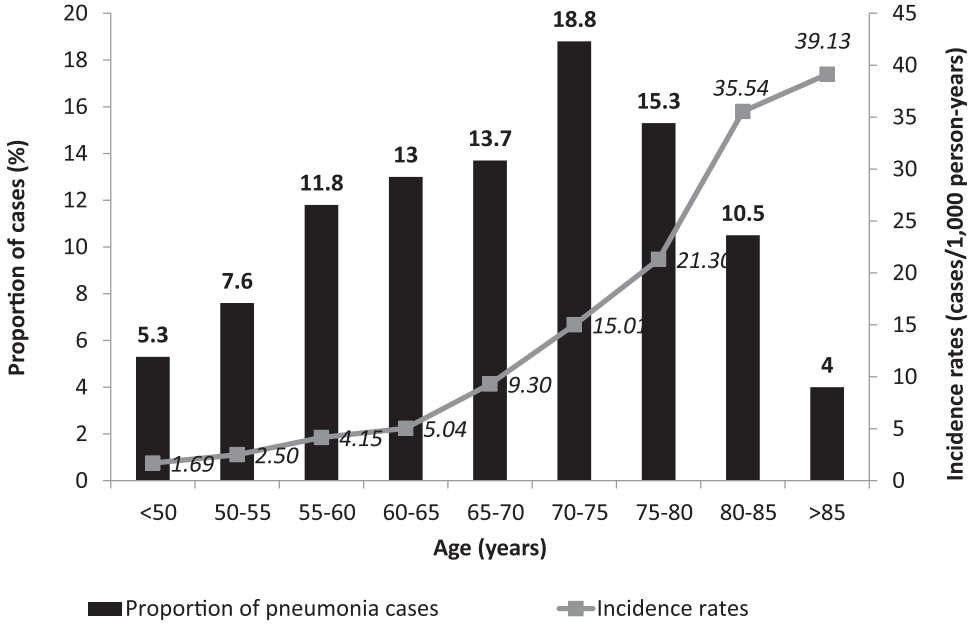

FIGURE 1. Proportion of pneumonia cases and crude incidence rates across different age groups.

Table 3—*AUC and NRI for Different Combination of Variables to Develop the Risk Prediction Model*

Variable	AUC	NRI[a]
Age	0.74	…
With % predicted FEV_1	0.76	15.01
With smoking	0.76	12.40
With heart failure	0.75	4.60
With diabetes	0.74	3.10
With coronary heart disease	0.74	5.70
With sex	0.74	2.20
With race	0.74	< 0.01
With BMI	0.74	1.50
With chronic kidney disease	0.74	0
Age and % predicted FEV_1	0.76	…
With smoking	0.77	6.14
With heart failure	0.77	3.70
With diabetes	0.77	2.0
With coronary heart disease	0.77	0.90
With sex	0.77	1.50
With race	0.76	1.90
With BMI	0.77	0.09
With chronic kidney disease	0.77	−0.01
Age and smoking	0.76	…
With % predicted FEV_1	0.77	9.90
With heart failure	0.76	1.40
With diabetes	0.77	0.97
With coronary heart disease	0.76	0.67
With sex	0.76	−0.85
With race	0.76	−0.76
With BMI	0.76	−0.16
With chronic kidney disease	0.76	−0.007
Age, % predicted FEV_1, and smoking	0.77	…
With heart failure	0.77	2.60
With diabetes	0.78	2.30
With coronary heart disease	0.77	2.0

NRI = net reclassification improvement. See Table 1 legend for expansion of other abbreviation.

[a]NRI is calculated as described by Pencina et al.[33] Briefly, it is calculated as follows: NRI = [P(up|D = 1) − P(down|D = 1)] − [P(up|D = 0) − P(down|D = 0)], where Pup and Pdown are the probabilities of participants reclassified to a higher and lower risk category and D = 0 and D = 1 indicate groups in which pneumonia did and did not occur.

Table 4—*Risk Prediction Model for Pneumonia Hospitalization*

	Hospitalization for Pneumonia	
Variable	RR (95% CI)	*P* Value
Age, y		
< 50	…	…
50-55	1.43 (0.99-2.05)	.05
55-60	2.20 (1.57-3.08)	< .0001
60-65	2.66 (1.91-3.72)	< .0001
65-70	4.81 (3.43-6.75)	< .0001
70-75	9.93 (7.00-14.08)	< .0001
75-80	14.72 (10.17-21.31)	< .0001
80-85	27.87 (18.67-41.58)	< .0001
> 85	16.13 (9.73-26.76)	< .0001
Lung function		
Predicted FEV_1 > 80%	…	…
Predicted FEV_1 50%-80%	2.14 (1.75-2.62)	< .0001
Predicted FEV_1 < 50%	5.73 (4.22-7.79)	< .0001
Smoking		
Never smoker	…	…
Past smoker	1.26 (1.07-1.49)	.005
Current smoker	2.06 (1.70-2.50)	< .0001

Logit(p) = −4.5885 + 0.7597*(if $PPFEV_1$: 50%-80%) + 1.7462*(if $PPFEV_1$: < 50%) + 0.3552 (if Age: 50-55 y) + 0.7885 (if Age: 55-60 y) + 0.9792 (if Age: 60-65 y) + 1.5714 (if Age: 65-70 y) + 2.2951 (if Age: 70-75 y) + 2.6892 (if Age: 75-80 y) + 3.3274 (if Age: 80-85 y) + 2.7808 (if Age: > 85 y) + 0.2338 (if Previous Smoker) + 0.7240 (if Current Smoker) + −0.2977*(if $PPFEV_1$: 50%-80% and Age: 70-75 y) + −0.8383*(if $PPFEV_1$: < 50% and Age: 70-75 y) + −0.3415*(if $PPFEV_1$: 50%-80% and Age: 75-80 y) + −1.1867*(if $PPFEV_1$ < 50% and Age: 75-80 y) + −0.9957*(if $PPFEV_1$: 50%-80% and Age: 80-85 y) + −1.3284*(if $PPFEV_1$ < 50% and Age: 80-85 y). Predicted risk of community-acquired pneumonia = 1/1 + e-logit(p). $PPFEV_1$ = percent predicted FEV_1. See Table 2 legend for expansion of other abbreviation.

prediction model did not improve AUC (e-Table 5). Similarly, IL-6 was also associated with a higher risk of pneumonia hospitalization (OR per SD increase in IL-6 was 1.16 [1.06-1.28], $P = .002$), but did not improve performance of the clinical risk prediction model (e-Appendix 1, e-Table 6).

Clinical Application of Prediction Model

We developed a simplified decision rule to categorize participants into different risk categories based on hierarchical application of the identified risk factors (e-Appendix 1, e-Fig 1). We also determined the clinical application of the model in the external validation cohort by estimating the number of pneumonia cases identified using different risk cutoffs (Fig 3). Most subjects had a predicted risk > 5%, and since all participants in the external validation cohort were > 65 years of age, the CDC guidelines for pneumococcal vaccination would require targeting all subjects for prevention. Targeting those with predicted risk > 5%, > 10%, and > 15% predicted risk would target 90.4%, 39.1%, and 14.4% of the subjects in this cohort and identify 88.9%, 55.1%, and 26.8% of pneumonia cases, respectively. To illustrate a strategy of targeting a more select population, such as those with risks of ≥ 10% in the clinical risk prediction model, would require treating approximately one-third of the population and identify one-half of pneumonia cases. Targeting only high-risk participants (≥ 15% risk) or 14% of the population would identify one-fourth of pneumonia cases, a subgroup that could be targeted initially to conduct randomized clinical trials to test pharmacologic interventions cost-effectively.

Similar analyses to test targeted strategies were conducted in the derivation cohort (e-Fig 2A). Finally, the clinical risk prediction model was also particularly likely to identify fatal pneumonia cases, as evidenced by higher 30-day and 1-year mortality for participants with higher predicted risk (e-Fig 2B).

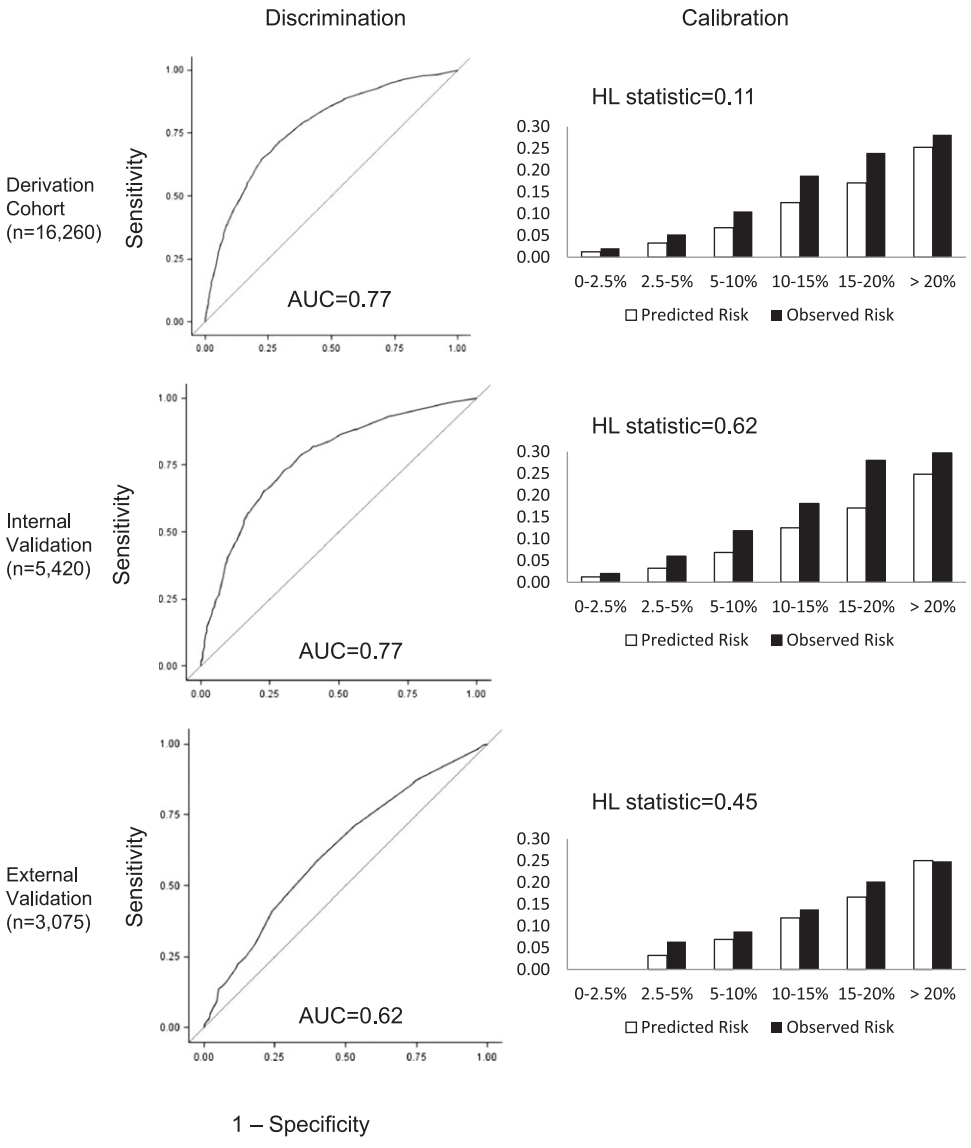

FIGURE 2. Model discrimination (using AUC) and calibration (using HL C statistic) for the pneumonia clinical risk prediction model in the derivation, internal validation, and external validation cohorts. AUC = area under curve; HL = Hosmer-Lemeshow.

DISCUSSION

Our results showed that hospitalization with pneumonia is common, even in younger adults. More than one-third of all pneumonia cases occurred in those younger than 65 years, and the short- and long-term mortality remained high in these individuals. Long-term risk of pneumonia hospitalization can be predicted in community-dwelling well-functioning adults using a simple clinical risk prediction model based on age, smoking status, and lung function. Circulating CRP and IL-6 levels do not materially improve discrimination.

Perhaps strikingly, the incidence of pneumonia requiring hospitalization was similar or higher than the incidence of acute myocardial infarction or stroke in the general population.[35,36] In contrast to the incidence of acute myocardial infarction (2.74 cases/1,000 person-years), the incidence of pneumonia hospitalization was 2.5-fold higher (6.71 cases/1,000 person-years), and these estimates are similar to results of prior studies using administrative data.[13,27] Compared with cardiovascular disease, the short- and long-term mortality after pneumonia was high. Even in younger adults, the 1-year mortality after pneumonia was 12.82% among those without a major chronic health condition and those who never smoked, a subgroup in whom mortality can be attributed to pneumonia rather than underlying chronic disease. Furthermore, the higher mortality was also observed in cases in which pneumonia was the primary reason for hospitalization and unlikely to be explained by inclusion of subjects with hospital-acquired pneumonia. Thus, pneumonia is a major public

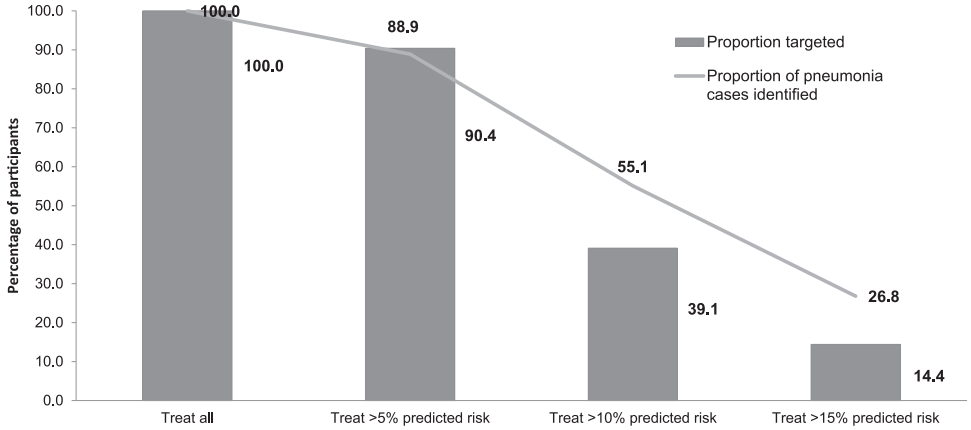

FIGURE 3. Proportion of pneumonia cases and the proportion of population that would have to be targeted using different thresholds of the clinical risk prediction model in the external validation cohort. To illustrate a strategy of targeting a more select population, such as those with risks of >10%, would require treating approximately one-third of the population and identify one-half of pneumonia cases. Targeting only high-risk participants (>15% risk) or 14% of the population would identify one-fourth of pneumonia cases, a subgroup that could be targeted initially to conduct randomized clinical trials to test pharmacologic interventions cost-effectively.

health problem, and designing new preventive strategies to reduce risk of pneumonia is important.

We developed and validated a simple risk prediction model that could be used to identify different risk groups and target expensive preventive strategies that have adverse events. In contrast to the CDC guidelines for pneumococcal vaccine, which target a broad population based on age and presence of any chronic disease, our risk prediction model offers a better opportunity to evaluate the balance between the threshold of risk chosen and the number of individuals at risk (e-Fig 1). For example, one would have to recruit 3,500 high-risk subjects and follow them over 5 years to test a preventive intervention that would reduce risk of pneumonia by 25%. If this intervention is effective, then our prediction model could be used to target this intervention, especially in younger adults.

Our study has several strengths. First, CHS and ARIC cohorts included probability-based samples of community-dwelling well-functioning individuals from seven US states. The distribution of participants was fairly similar to the US population, and one-fourth of participants were black. Second, we had objective assessment of the burden and severity of chronic diseases rather than relying on self-report. Third, we assessed long-term risk of pneumonia over 10 years, much longer than most studies have done. Finally, we examined the role of biomarkers to predict risk of pneumonia. Although associated with higher risk, inflammatory biomarkers did not improve performance of our clinical risk prediction model.

Our study has limitations. First, CAP is a clinical syndrome diagnosed based on clinical symptoms and radiographic findings, and microbiologic confirmation is usually available in fewer than one-third of cases in routine clinical practice.[37] We used ICD-9 codes shown to have good sensitivity and specificity in prior studies and validated these codes in our study.[27,28] We used two criteria, a sensitive criterion for CAP based on ICD-9 codes in first five discharge diagnoses fields and a more specific criterion using ICD-9 codes for bacterial pneumonia in the first discharge diagnosis field. About two-thirds of all hospitalizations with pneumonia met the latter criteria, and in two-thirds of the remainder, pneumonia was the primary reason for hospitalization. Thus, few pneumonia cases in our derivation cohort are likely to be hospital acquired. The relative risks for the risk prediction model were similar using both criteria, suggesting that misclassification of cases is unlikely to affect our results. Furthermore, the model was validated in Health ABC, in which physicians conducted chart review to confirm all cases of hospitalized pneumonia.

Second, we focused on hospitalization for pneumonia in this study and did not include nonhospitalized cases. We did so in part because chest radiographs are not routinely obtained in outpatient settings, and it is difficult to distinguish between bronchitis and pneumonia. Outpatient pneumonia cases also have lower mortality and may, therefore, be less important to prevent. However, due to exclusion of these cases we may have overestimated the effect of age, as younger individuals with pneumonia are less likely to be hospitalized.

Third, we assessed the role of baseline characteristics to develop the risk prediction model. Future studies should include repeated measures of potential predictors to determine whether these risk factors may be associated with pneumonia susceptibility in shorter time frames. Finally, the three cohorts had

very few Hispanics, and whether our risk prediction model can be generalized to this ethnic group remains unknown.

In conclusion, pneumonia is common across a wide spectrum of age and associated with high short- and long-term mortality even in younger adults. Using the key risk factors of age, smoking, and lung function, the risk of pneumonia hospitalization can be predicted using a simple clinical risk prediction model. In the future, this model can be used to design clinical trials to test new preventive strategies for pneumonia and target proven prevention efforts to those at highest risk of this condition.

Acknowledgments

Author contributions: Dr Yende had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Dr Yende: contributed to study conception and design, obtaining funding, analysis and interpretation, statistical analysis, and drafting of the manuscript.
Ms Alvarez: contributed to statistical analysis and critical revisions and approval of the final manuscript.
Dr Loehr: contributed to obtaining funding, acquisition of data, and critical revisions and approval of the final manuscript.
Dr Folsom: contributed to obtaining funding, acquisition of data, and critical revisions and approval of the final manuscript.
Dr Newman: contributed to study conception and design, obtaining funding, acquisition of data, and critical revisions and approval of the final manuscript.
Dr Weissfeld: contributed to study conception and design, analysis and interpretation, statistical analysis, and critical revisions and approval of the final manuscript.
Dr Wunderink: contributed to study conception and design, and critical revisions and approval of the final manuscript.
Dr Kritchevsky: contributed to obtaining funding, acquisition of data, and critical revisions and approval of the final manuscript.
Dr Mukamal: contributed to critical revisions and approval of the final manuscript.
Dr London: contributed to critical revisions and approval of the final manuscript.
Dr Harris: contributed to obtaining funding, acquisition of data, and critical revisions and approval of the final manuscript.
Dr Bauer: contributed to obtaining funding, acquisition of data, and critical revisions and approval of the final manuscript.
Dr Angus: contributed to study conception and design, obtaining funding, analysis and interpretation, and drafting of the manuscript.
Financial/nonfinancial disclosures: The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.
Role of sponsors: The sponsors had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.
Additional information: The e-Appendix, e-Figures, and e-Tables can be found in the "Supplemental Materials" area of the online article.

References

1. Simonsen L, Conn LA, Pinner RW, Teutsch SM. Trends in infectious disease hospitalizations in the United States, 1980-1994. *Arch Intern Med*. 1998;158(17):1923-1928.
2. Yende S, Angus DC, Ali IS, et al. Influence of comorbid conditions on long-term mortality after pneumonia in older people. *J Am Geriatr Soc*. 2007;55(4):518-525.
3. Herridge MS, Cheung AM, Tansey CM, et al; Canadian Critical Care Trials Group. One-year outcomes in survivors of the acute respiratory distress syndrome. *N Engl J Med*. 2003;348(8):683-693.
4. Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA*. 2010;304(16):1787-1794.
5. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989;81(24):1879-1886.
6. Armstrong K, Eisen A, Weber B. Assessing the risk of breast cancer. *N Engl J Med*. 2000;342(8):564-571.
7. Centers for Disease Control and Prevention. Vaccines and preventable diseases: pneumococcal vaccination. Centers for Disease Control and Prevention website. http://www.cdc.gov/vaccines/vpd-vac/pneumo/default.htm. Accessed February 2012.
8. 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(9):977-988.
9. LaCroix AZ, Lipson S, Miles TP, White L. 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(4):350-360.
10. Baik I, Curhan GC, Rimm EB, Bendich A, Willett WC, Fawzi WW. A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women. *Arch Intern Med*. 2000;160(20):3082-3088.
11. Koivula I, Sten M, Mäkelä PH. Risk factors for pneumonia in the elderly. *Am J Med*. 1994;96(4):313-320.
12. O'Meara ES, White M, Siscovick DS, Lyles MF, Kuller LH. Hospitalization for pneumonia in the Cardiovascular Health Study: incidence, mortality, and influence on longer-term survival. *J Am Geriatr Soc*. 2005;53(7):1108-1116.
13. Fry AM, Shay DK, Holman RC, Curns AT, Anderson LJ. Trends in hospitalizations for pneumonia among persons aged 65 years or older in the United States, 1988-2002. *JAMA*. 2005;294(21):2712-2719.
14. Cundell DR, Gerard NP, Gerard C, Idanpaan-Heikkila I, Tuomanen EI. Streptococcus pneumoniae anchor to activated human cells by the receptor for platelet-activating factor. *Nature*. 1995;377(6548):435-438.
15. Mancini GB, Etminan M, Zhang B, Levesque LE, FitzGerald JM, Brophy JM. Reduction of morbidity and mortality by statins, angiotensin-converting enzyme inhibitors, and angiotensin receptor blockers in patients with chronic obstructive pulmonary disease. *J Am Coll Cardiol*. 2006;47(12):2554-2560.
16. Goldfine AB. Statins: is it really time to reassess benefits and risks? *N Engl J Med*. 2012;366(19):1752-1755.
17. Schlienger RG, Fedson DS, Jick SS, Jick H, Meier CR. Statins and the risk of pneumonia: a population-based, nested case-control study. *Pharmacotherapy*. 2007;27(3):325-332.
18. Almog Y, Shefer A, Novack V, et al. Prior statin therapy is associated with a decreased rate of severe sepsis. *Circulation*. 2004;110(7):880-885.
19. Gupta R, Plantinga LC, Fink NE, et al. Statin use and sepsis events [corrected] in patients with chronic kidney disease. *JAMA*. 2007;297(13):1455-1464.
20. Redberg RF, Katz MH. Healthy men should not take statins. *JAMA*. 2012;307(14):1491-1492.
21. Tleyjeh IM, Kashour T, Hakim FA, et al. Statins for the prevention and treatment of infections: a systematic review and meta-analysis. *Arch Intern Med*. 2009;169(18):1658-1667.
22. Preiss D, Seshasai SR, Welsh P, et al. Risk of incident diabetes with intensive-dose compared with moderate-dose statin therapy: a meta-analysis. *JAMA*. 2011;305(24):2556-2564.

23. van de Garde EM, Hak E, Souverein PC, Hoes AW, van den Bosch JM, Leufkens HG. Statin treatment and reduced risk of pneumonia in patients with diabetes. *Thorax*. 2006;61(11): 957-961.
24. Yende S, Tuomanen EI, Wunderink RG, et al. Preinfection systemic inflammatory markers and risk of hospitalization due to pneumonia. *Am J Respir Crit Care Med*. 2005;172(11): 1440-1446.
25. Folsom AR, Chambless LE, Ballantyne CM, et al. An assessment of incremental coronary risk prediction using C-reactive protein and other novel risk markers: the atherosclerosis risk in communities study. *Arch Intern Med*. 2006;166(13): 1368-1373.
26. Ives DG, Fitzpatrick AL, Bild DE, et al. Surveillance and ascertainment of cardiovascular events. The Cardiovascular Health Study. *Ann Epidemiol*. 1995;5(4):278-285.
27. Lindenauer PK, Lagu T, Shieh MS, Pekow PS, Rothberg MB. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003-2009. *JAMA*. 2012;307(13):1405-1413.
28. Guevara RE, Butler JC, Marston BJ, Plouffe JF, File TM Jr, Breiman RF. Accuracy of ICD-9-CM codes in detecting community-acquired pneumococcal pneumonia for incidence and vaccine efficacy studies. *Am J Epidemiol*. 1999;149(3): 282-289.
29. Tuomanen EI, Austrian R, Masure HR. Pathogenesis of pneumococcal infection. *N Engl J Med*. 1995;332(19):1280-1284.
30. Szalai AJ, Briles DE, Volanakis JE. Role of complement in C-reactive-protein-mediated protection of mice from Streptococcus pneumoniae. *Infect Immun*. 1996;64(11): 4850-4853.
31. Horowitz J, Volanakis JE, Briles DE. Blood clearance of Streptococcus pneumoniae by C-reactive protein. *J Immunol*. 1987;138(8):2598-2603.
32. Freund A, Orjalo AV, Desprez PY, Campisi J. Inflammatory networks during cellular senescence: causes and consequences. *Trends Mol Med*. 2010;16(5):238-246.
33. 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. *Stat Med*. 2008;27(2):157-172.
34. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. *Ann Intern Med*. 2008;149(10):751-760.
35. Centers for Disease Control and Prevention (CDC). Prevalence of stroke—United States, 2005. *MMWR Morb Mortal Wkly Rep*. 2007;56(19):469-474.
36. Yeh RW, Sidney S, Chandra M, Sorel M, Selby JV, Go AS. Population trends in the incidence and outcomes of acute myocardial infarction. *N Engl J Med*. 2010;362(23):2155-2165.
37. Metersky ML, Ma A, Bratzler DW, Houck PM. Predicting bacteremia in patients with community-acquired pneumonia. *Am J Respir Crit Care Med*. 2004;169(3):342-347.