ARIC Manuscript Proposal # 1837

PC Reviewed: 8/9/11	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Clinical Risk Factors and Biomarkers to Predict Risk of Hospitalization With Pneumonia: Analyses of Three Multicenter Cohorts

b. Abbreviated Title (Length 26 characters): Predicting risk of pneumonia

2. Writing Group:

Writing group members: Sachin Yende, Laura Loehr, Aaron Folsom, and Stephanie London.

Authors from other cohorts (CHS and ARIC) include: Karina Alvarez, Anne Newman, Lisa A. Weissfeld, Richard Wunderink, Stephen B, Kritchevsky, Kenneth Mukumal, Stephanie J. London, Tamara B. Harris, Derek C. Angus.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _SY____ **[please confirm with your initials electronically or in writing]**

First author: **Sachin Yende, MD MS**
Address: 606D Scaife Hall, 3550 Terrace Street,
University of Pittsburgh, Pittsburgh, PA 15261.
Phone: 412 647 8467 Fax: 412 647 8060
E-mail: yendes@upmc.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: **Laura Loehr**
Address: Department of Epidemiology,
UNC Gillings School of Global Public Health,
University of North Carolina at Chapel Hill
Phone: 919-619-5023 Fax:
E-mail: lloehr@email.unc.edu

3. Timeline: September/October 2011

4. Rationale: The Centers for Disease Control and Prevention (CDC) guidelines for pneumococcal vaccine are used to target prevention strategies for pneumonia. Pharmacologic interventions, such as statins, have shown promise to prevent

pneumonia,[1-5] but more precise prediction models are needed to verify efficacy and subsequently implement these therapies because they are expensive and have adverse effects. This manuscript proposal will develop and validate a clinical risk prediction model for pneumonia hospitalization in population-based cohorts of community-dwelling individuals. We will also test the hypothesis that biomarkers (C-reactive protein [CRP]) improve performance of clinical risk prediction model.
We will combine participants in ARIC and Cardiovascular Health Study (CHS, n=5,888) to develop the clinical risk prediction model because these cohorts included individuals with non-overlapping age ranges (ARIC: 45-64 years and CHS: >65 years); they were recruited during the same time period (ARIC: 1987-89 and CHS: 1989-90) and from similar geographic areas; they used similar study procedures, and had similar follow-up (over 10 years). We will randomly split the data into three-quarters for model development (derivation cohort) and one-quarter for internal validation (internal validation cohort). We will also perform external validation in the Health, Aging, and Body Composition (Health ABC, n=3075) cohort, a more contemporary cohort recruited in 1997-98 with 10-year follow-up data. Biomarker analyses will be performed in each cohort separately and will include a subset of ARIC that had CRP measured.

5. Main Hypothesis/Study Questions:
Aim #1: To develop and validate a clinical risk prediction model to predict hospitalization with pneumonia in community-dwelling individuals.
Aim #2: To test the hypothesis that biomarkers (CRP and interleukin-6) improve performance of the clinical risk prediction model. For ARIC, analyses will be restricted to CRP alone.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).
We will use an inception cohort design to develop and validate the risk prediction model. The primary outcome measure will be hospitalization with pneumonia within a 10 year period. For ARIC and CHS, pneumonia will be identified by International Classification of Diseases, ninth edition (ICD-9) codes. We will define hospitalization with pneumonia based on presence of ICD-9 codes 481-487 in the first five discharge diagnoses fields. 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 discharge diagnosis field.

We have combined 15,792 ARIC participants and 5,888 CHS participants. We randomly split the combined cohort into a derivation cohort (three quarters of the combined cohort, n=16,260) and internal validation (one-quarter of the combined cohort, n=5,420). Of the 16,260 participants in the derivation cohort, we have identified 1,000 hospitalizations with pneumonia.

Covariates for model development: Demographics (age, sex, and race), health behaviors (smoking – never, previous, current), body mass index, and chronic health conditions (lung function [FEV_1], coronary heart disease, heart failure, diabetes, and chronic kidney disease [GFR based on creatinine]).

Analyses: We will develop a parsimonious risk prediction model in the derivation cohort based on area under curve (AUC) and net reclassification improvement (NRI). Model performance will be assessed in the validation cohorts by comparing AUC and model calibration (Hosmer Lemeshow statistics). Details of statistical methods are provided in the Appendix.

For biomarker analyses, we will measure CRP from year 9 samples in ARIC participants. Therefore, models for ARIC participants that assessed the role of CRP excluded pneumonia events during the first 9 years and the 10-year risk of hospitalization with pneumonia from year 9 onwards was estimated. Of the 15,792 ARIC participants, CRP levels were measured in 11,334 participants. Of these, 334 participants had CRP >20 mg/L and were excluded because such high levels are usually associated with acute infections.

7.a. Will the data be used for non-CVD analysis in this manuscript? __x__ Yes ____ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? ____ Yes ____ No - Not applicable
(This file ICTDER03 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ____ Yes __x__ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"?
__ __ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

___x___ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new

proposal or collaboration)? Proposals reviewed at http://www.cscc.unc.edu/aric/published_public.php?sort=au_last Authors from related pulmonary manuscripts have been included here

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? _x___ Yes ____ No

11.b. If yes, is the proposal

___ A. primarily the result of an ancillary study (list number* _____)

_x__ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* 2010.22_____ _____ _____)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

References

1. van de Garde EM, Hak E, Souverein PC, Hoes AW, van den Bosch JM, Leufkens HG. Statin therapy and reduced risk of pneumonia in patients with diabetes. Thorax 2006.

2. 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;27:325-32.

3. Etminan M, Zhang B, Fitzgerald M, Brophy JM. Do angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers decrease the risk of hospitalization secondary to community-acquired pneumonia? A nested case-control study. Pharmacotherapy 2006;26:479-82.

4. Takahashi T, Morimoto S, Okaishi K, et al. Reduction of pneumonia risk by an angiotensin I-converting enzyme inhibitor in elderly Japanese inpatients according to insertion/deletion polymorphism of the angiotensin I-converting enzyme gene. Am J Hypertens 2005;18:1353-9.

5. van de Garde EM, Souverein PC, van den Bosch JM, Deneer VH, Leufkens HG. Angiotensin-converting enzyme inhibitor use and pneumonia risk in a general population. Eur Respir J 2006;27:1217-22.

Appendix

Pooled data from ARIC and CHS will be randomly split into three-quarter (derivation cohort) and one-quarter (internal validation cohort) samples using PROC SURVERYSELECT in SAS. We will conduct univariate comparisons in the derivation cohort to identify clinical variables and biomarkers associated with increased risk of pneumonia hospitalization. We will categorize continuous variables (lung function, glomerular filtration rate [GFR], and BMI). Cutpoints for categories of percentage predicted FEV1 were >80%, 50-80%, and < 50% based on the American Thoracic Society and GOLD criteria; for GFR were >60 and <60 mL/min/1.7 m^2 based on the Modification of Diet in Renal Disease Study Group recommendation; and for BMI we examined several cutoffs similar to the CDC criteria, modeling risk for underweight, normal weight, overweight, and obese separately.

We will develop a simple and parsimonious clinical risk prediction model to predict 10-year risk of pneumonia using logistic regression in the derivation cohort. To construct the clinical risk prediction model, we will assess the discriminative ability of each variable by calculating the area under curve (AUC). Age had the highest AUC. We will then sequentially add variables to age, calculate AUC and net reclassification improvement (NRI) for each combination, and chose the combination that had the highest AUC and NRI. We will assess calibration of the model by comparing the predicted and observed risk and calculating the Hosmer-Lemeshow (HL) statistic. We will divide the predicted risk into 4 categories: very low risk (<2.5%), low risk (2.5-5%), intermediate risk (5-15%), and high risk (>15%), similar to the Framingham risk prediction model. We will validate the clinical risk prediction model in the internal and external validation cohorts by estimating AUC, the observed and predicted risk, and calculating the HL statistic.

We will conduct several sensitivity analyses. First, we recognize limitations of ICD-9 codes to identify CAP hospitalizations. ICD-9 codes for bacterial pneumonia had higher specificity during chart review and included only those cases where ICD-9 codes for bacterial pneumonia were listed in the primary discharge diagnosis field. Therefore, we will compare the estimates for each risk factor in the clinical risk prediction model developed above and the same model to predict risk of bacterial pneumonia. Second, we will compare odds ratios for each risk factor using logistic regression and Cox proportional hazards model because logistic regression models do not account for differences in time to occurrence of pneumonia and censoring due to deaths or dropouts prior to occurrence of CAP. Third, to determine generalizability of the model we compared observed and predicted risk and calculated the HL statistic for blacks and all participants younger than 65 years. Fourth, we assessed model performance to predict 5-year risk since clinical trials are often performed over 5 years.

We will compare the clinical utility of the clinical risk prediction model and the CDC criteria for pneumococcal vaccination. We will compare 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 clinical risk prediction model and the CDC criteria.

We will assess incremental benefit of adding biomarkers to the clinical risk prediction model by estimating improvement in AUC and calculating NRI and integrated discrimination improvement (IDI), as described by Pencina et al. Analytic methods for each assay differed across cohorts, and therefore, we analyzed biomarkers for each cohort separately.