THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

November 22, 2016

Josef Coresh, M.D., Ph.D.
ARIC Publications Committee
The Johns Hopkins University
Department of Epidemiology
615 North Wolfe Street; Room W6009
Baltimore, Maryland 21205

Re: ARIC Proposal 2881 "External Validation of the REGARDS Sepsis Risk Score"

Dear Dr. Coresh:

Thank you for your review of my application for use of ARIC data.

There was one comment from the publications committee:

> "Please clarify when the components of the sepsis risk score are measured/"

As modified on Page 4 of the proposal:

> "The components of the SRS and SSRS scores will be determined using the earliest available measurement. Some components were not measured at the beginning of ARIC; for example, urinary albumin and creatinine. Thus, we may shift the follow-up period to accommodate earliest available baseline data."

We look forward to your additional review.

Sincerely,

Henry E. Wang, MD, MS
Professor and Vice Chair for Research
Department of Emergency Medicine
University of Alabama at Birmingham
(205)-996-6526
hwang@uabmc.edu

Department of Emergency Medicine
266 N Jefferson Tower
619 So. 19th Street
Birmingham, AL 35249-7013
205.975.9358
Fax 205.975.4662

ARIC Manuscript Proposal #2881

PC Reviewed: 12/13/16 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: **External Validation of the REGARDS Sepsis Risk Score**

b. Abbreviated Title (Length 26 characters): **SRS Validation**

2. Writing Group:

Henry Wang, MD, MS (Department of Emergency Medicine, University of Alabama at Birmingham)
Sachin Yende, MD (Department of Critical Care Medicine, University of Pittsburgh)
John Donnelly, MPH (Department of Emergency Medicine, Department of Epidemiology, University of Alabama at Birmingham)
Gerardo Heiss, MD, PhD (Department of Epidemiology, University of North Carolina Chapel Hill)

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

First author: **Henry E. Wang, MD, MS**
Address: Department of Emergency Medicine
University of Alabama at Birmingham
OHB 251
Birmingham, AL 35249

Phone: 205-996-6526 Fax:
E-mail: hwang@uabmc.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: **Gerardo Heiss, PhD**
Address: University of North Carolina Chapel Hill
CVS Plaza
137 East Franklin Street, Suite 306
CB #8050
Chapel Hill, NC 27514

Phone: (919) 962-3253 Fax:
E-mail: gerardo_heiss@unc.edu

3. Timeline:

We expect to finish this project within 12 months of receipt of dataset.

4. Rationale:

Sepsis is the syndrome of microbial infection complicated by systemic inflammation. In the United States (US), sepsis places an immense burden on the healthcare system, resulting in 750,000 hospitalizations, 570,000 Emergency Department visits and 200,000 deaths annually.[1-3]

In a prior effort, using data from the Reasons for Geographic and Racial Differences in Stroke (REGARDS) cohort, we derived a Sepsis Risk Score (SRS) and a Severe Sepsis Risk Score (SSRS) characterizing community-dwelling individuals' long-term risks of sepsis and severe sepsis. (Wang, Critical Care Medicine 2016) The sepsis and severe sepsis events in REGARDS-Sepsis were based upon participant report of hospitalizations for a serious infection and confirmed by manual review of Emergency Department and hospital admission data.

The objective of the current proposal is to externally validate the SRS and SSRS using hospitalization event data from the Atherosclerosis Risk in Communities (ARIC) cohort.

This analysis will occur in parallel with two other analyses:

- Internal validation of the SRS and SSRS using linked REGARDS-Medicare claims data.
- External validation of the SRS and SSRS using data from the Cardiovascular Health Study (CHS) cohort.

We will compose a single manuscript summarizing the combined results of the three parallel validation efforts.

5. Main Hypothesis/Study Questions:

- Using data from the ARIC cohort, validate the ability of the REGARDS Sepsis Risk Score (SRS) and Severe Sepsis Risk Score (SSRS) to predict sepsis and severe sepsis hospitalizations.

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 the following variables (variables listed in order of REGARDS-SRS) (Potential ARIC variables listed in Appendix 1):

1) Chronic Lung Disease (defined by pulmonary function tests or medical history)
2) Age
3) Peripheral Artery Disease
4) Diabetes
5) Tobacco Use (Present smoker)

6) Race
7) Stroke
8) Atrial Fibrillation
9) Coronary Artery Disease
10) Obesity (Determined from BMI and waist circumference)
11) Hypertension
12) Deep Vein Thrombosis
13) Gender
14) hsCRP (calculated correction)
15) Cystatin-C
16) eGFR
17) Albumin-to-Creatinine Ratio

18) All ICD-9 discharge diagnoses for hospitalization events with discharge diagnosis for serious infection (Appendix 2) or sepsis [038-039.9, 020.0, 790.7, 117.9, 112.5, 112.81, 995.91, 995.92 and 785.52].

19) Dates of all serious infection hospitalization events (days from enrollment)

Population: The analysis will include all adult participants in the ARIC cohort with valid follow-up from ARIC enrollment (1989-90) through present.

Study Outcomes: The primary endpoints are hospitalization for sepsis and severe sepsis.

To identify sepsis hospitalizations, we will use the ICD-9 taxonomy of Martin, et al.[4] (Appendix 2) We will classify any hospitalization with a primary or secondary discharge diagnosis satisfying the Martin criteria as sepsis.

To identify severe sepsis hospitalizations, we will use the ICD-9 taxonomy of Angus, et al.[1] (Appendices 3 and 4) Rather than using sepsis-specific ICD-9 codes, the Angus system identifies severe sepsis as hospitalizations with discharge diagnoses for both a serious infection and organ dysfunction. We will also classify the presence of the codes 995.92 (Severe Sepsis) and 785.52 (Septic Shock) as severe sepsis.

We will identify sepsis events from a 10-year follow-up time window. We will shift the time window based upon the timing of biomarkers available in ARIC. For example, we may choose the 10-year period follow procurement of urinary albumin and creatinine.

Brief analysis plan: For each ARIC participant, we will calculate the REGARDS Sepsis Risk Score (SRS) and Severe Sepsis Risk Score (SSRS) and corresponding risk quintiles. The components of the SRS and SSRS scores will be determined using the earliest available measurement. Some components were not measured at the beginning of ARIC; for example, urinary albumin and creatinine. Thus, we may shift the follow-up period to accommodate earliest available baseline data.

We will fit a Cox regression model using time to first "Martin sepsis" event as the outcome. The primary exposure will be SRS point score or quintile of predicted risk. We will validate the SRS model by determining model discrimination (Harrell's C) and calibration (PI deciles). We will repeat the analysis using two forms of the SRS with and without biomarkers; we will calculate the Net Reclassification Improvement (NRI) for the SRS with biomarkers over the SRS without biomarkers. We will repeat the same process using time to first "Angus severe sepsis" event to validate the SSRS.

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? _ X ___ Yes _ ___ No
(This file ICTDER 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"? __X__ 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)?

There are no related papers.

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

11.b. If yes, is the proposal

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

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

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

12a. 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.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __X__ No.

REFERENCES

1. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Critical care medicine 2001;29:1303-10.

2. Wang HE, Shapiro NI, Angus DC, Yealy DM. National estimates of severe sepsis in United States emergency departments. Critical care medicine 2007;35:1928-36.

3. Dellinger RP, Carlet JM, Masur H, et al. Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. Intensive care medicine 2004;30:536-55.

4. Martin GS, Mannino DM, Eaton S, Moss M. The epidemiology of sepsis in the United States from 1979 through 2000. The New England journal of medicine 2003;348:1546-54.

APPENDIX 1

Potential ARIC Variables to be used in the Analysis

- Chronic Lung Disease – will be defined by pulmonary function tests or medical history

Pulmonary function data:

Visit	Version	Form	Variable name	Definition
Visit 1 (Pulmonary Function Test)	V1M4	PFTA	PFTA17	FEV(1) Over FEV(6)
	V1M4	PFTA	PFTA24	FVC (Liters)
	V1M4	PFTA	PFTA25	FEV(.5) (Liters)
	V1M4	PFTA	PFTA26	FEV(1) (Liters)
	V1M4	PFTA	PFTA27	FEV(3) (Liters)
	V1M4	PFTA	PFTA28	FEV(6) (Liters)
	V1M4	PFTA	PFTA29	FEV(3) Over FEV(6)
	V1M4	PFTA	PFTA30	FEV(.5) Over FVC
	V1M4	PFTA	PFTA31	FEV(1) Over FVC
	V1M4	PFTA	PFTA32	FEV(3) Over FVC
Visit 1 (Pulmonary Derived Variables)	V1M4	PULM	FVC01	FVC Predicted (liters)
	V1M4	PULM	FEV_501	FEV(.5) Predicted (liters)
	V1M4	PULM	FEV_101	FEV(1) Predicted (liters)
	V1M4	PULM	FEV_301	FEV(3) Predicted (liters)
	V1M4	PULM	FEV1FVC1	FEV(1)/FVC Predicted (%)
	V1M4	PULM	FEV3FVC1	FEV(3)/FVC Predicted (%)
Visit 2 (Pulmonary Function Test)	V1M4	PFTB	PFTB17	FEV(1) Over FEV(6)
	V1M4	PFTB	PFTB24	FVC (Liters)
	V1M4	PFTB	PFTB25	FEV(.5) (Liters)
	V1M4	PFTB	PFTB26	FEV(1) (Liters)
	V1M4	PFTB	PFTB27	FEV(3) (Liters)
	V1M4	PFTB	PFTB28	FEV(6) (Liters)
	V1M4	PFTB	PFTB29	FEV(3) Over FEV(6)
	V1M4	PFTB	PFTB30	FEV(.5) Over FVC
	V1M4	PFTB	PFTB31	FEV(1) Over FVC
	V1M4	PFTB	PFTB32	FEV(3) Over FVC
Visit 2 (Pulmonary Derived Variables)	V1M4	PULM21	FEF_122	FEV(1)/FVC Predicted (%)
	V1M4	PULM21	FEF_322	FEV(3)/FVC Predicted (%)
	V1M4	PULM21	FEF_522	FEV(.5) Predicted (Liters)
	V1M4	PULM21	FEV1FVC2	FEF(1)/FVC Predicted(%)
	V1M4	PULM21	FEV3FVC2	FEF(3)/FVC Predicted(%)
	V1M4	PULM21	FVC22	FVC Predicted (Liters)

Home interview/health history:

Visit	Version	Form	Variable name	Definition
Visit 1 (home interview)	V1M4	HOM	HOM10G	Chronic lung disease, such as chronic bronchitis, or emphysema
Visit 2 (health history)	V1M4	HHXB	HHXB05E	Chronic Lung Disease, Such As Chronic Bronchitis, Or Emphysema

- Age

Visit	Version	Form	Variable name	definition
Visit 1	V1m4	Derive10	V1age01	Age at visit 1
Visit 2	V1m4	Derive28	V2age22	Age at visit 2
Visit 3	V2M4	Derive35	V3age31	Age at visit 3
Visit 4	V4m3	derive44	V4age41	Age at visit 4
Visit 5	V1m1	Derive51_131218	V5age52	Corrected age at visit 4

- Peripheral Artery Disease

Visit	Version	Form	Variable name	definition
Visit 1	V1m4	Derive11	PAD01	PAD at visit 1, definition 1
	V1m4	Derive11	PAD02	PAD at visit 1, definition 2
Visit 3	V2m4	Derive37	PAD31	PAD at visit 3, definition 1
	V2m4	Derive37	PAD32	PAD at visit 3, definition 2
Visit 4	V4m3	derive46	PAD41	PAD at visit 4, definition 1
	V4m3	derive46	PAD42	PAD at visit 4, definition 2

- Diabetes

Visit	Version	Form	Variable name	definition
Visit 1	V1m4	Derive13	Diabts02	Diabetes at visit 1
	V1m4	Derive13	Diabts03	Diabetes with fasting glucose using cutoff of 126
Visit 2	V1m4	Derive28	Diabts22	Diabetes at visit 2
	V1m4	Derive28	Diabts23	Diabetes with fasting glucose using cutoff of 126
Visit 3	V2m4	Derive35	Diabts33	Diabetes at visit 3 using fasting glucose cutoff of 140
	V2m4	Derive35	Diabts33	Diabetes at visit 3 using fasting glucose cutoff of 126
Visit 4	V4m3	Derive44	Diabts41	Diabetes at visit 4 using fasting glucose cutoff of 140
	V4m3	Derive44	Diabts42	Diabetes at visit 4 using fasting glucose cutoff of 126
Visit	V1m1	Derive51_140409	Diabts53	Diabetes at visit 5 using fasting glucose

5				cutoff of 140
	V1m1	Derive51_140409	Diabts54	Diabetes at visit 5 using fasting glucose cutoff of 126
	V1m1	Derive51_140409	Diabts55	Diabetes at visit 5—lab and med only, using fasting glucose cutoff of 140
	V1m1	Derive51_140409	Diabts56	Diabetes at visit 5 using HbA1c cutoff of 6.5%
	V1m1	Derive51_140409	Diabts57	Diabetes at visit 5 (DM medication or DM reported on AFU)

- Tobacco Use (Present smoker)

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	Cigr01	Cigar smoking status
	V1m4	Derive10	Cigryr01	Cigar years of smoking
	V1m4	Derive10	Cigt01	Cigarette smoking status
	V1m4	Derive10	Cigtyr01	Cigarette years of smoking
	V1m4	Derive10	Cursmk01	Current smoking
	V1m4	Derive10	Evrsmk01	Ever smoked cigarette
	V1m4	Derive10	Forsmk01	Former cigarette smoking
Visit 2	V1m4	Derive2_10	Evrsmk21	Ever smoked cigarette
	V1m4	Derive2_10	Forsmk21	Former cigarette smoking
	V1m4	Derive2_10	Cigt21	Cigarette smoking status
	V1m4	Derive2_10	Cursmk21	Current smoking
Visit 3	V2m4	Derive37	Evrsmk31	Ever smoked cigarette
	V2m4	Derive37	Forsmk31	Former cigarette smoking
	V2m4	Derive37	Cigt31	Cigarette smoking status
	V2m4	Derive37	Cursmk31	Current smoking
Visit 4	V4m3	Derive44	Evrsmk41	Ever smoked cigarette
	V4m3	Derive44	Forsmk41	Former cigarette smoking
	V4m3	Derive44	Cigt41	Cigarette smoking status
	V4m3	Derive44	Cursmk41	Current smoking
Visit 5	V1m1	Derive51_140409	Evrsmk52	Ever smoked cigarette
	V1m1	Derive51_140409	Forsmk52	Former cigarette smoking
	V1m1	Derive51_140409	Cigt52	Cigarette smoking status
	V1m1	Derive51_140409	Cursmk52	Current smoking

- Race

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	racegrp	Race (from FTRA23)
	V1m4	Derive10	V1corra1	Corrected race group
Visit 2	V1m4	Derive2_10	racegrp	Race (from FTRA23)
Visit 3	V2m4	Derive37	racegrp	Race (from FTRA23)
Visit 4	V4m3	Derive44	racegrp	Race (from FTRA23)
Visit 5	V1m1	Derive51_140409	racegrp	Race (from FTRA23)

	V1m1	Derive51_140409	Racegrp51	Corrected race group

- Stroke

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Stroke01	Stroke01	Stroke
Visit 2	V1m4	Stroke2	Prvstr21	Prevalent stroke at visit 2
Visit 3	V2m4	Stroke32	Stroke31	Stroke
Visit 4	V4m3	Derive44	Prvstr41	Prevalent stroke at visit 4
	V4m3	Stroke41	Stroke41	Stroke
Visit 5	V1m1	Derive51_140611	Prvstr51	Prevalent stroke at by the end of visit 5
	V1m1	Derive51_140611	Prvstr52	Prevalent stroke at by the end of visit 5--unverified

- Atrial Fibrillation

Visit	Version	Form	Variable name	Definition
Visit 5	V1m1	Derive51_140813	Prvaf51	V5 atrial fibrillation/flutter before v5
	V1m1	Derive51_140813	Prvaf52	V5 atrial fibrillation/flutter by the end of v5

- Coronary Artery Disease

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive13	PRVCHD05	V1 Prevalent coronary heart disease
Visit 2	V1m4	Derive28	PRVCHD21	V2 Prevalent coronary heart disease
	V1m4	Derive28	PRVCHD22	V2 Prevalent CHD, unverified
	V1m4	Derive28	PRVCHD23	V2 Prevalent CHD, definition 3
Visit 3	V2m4	Derive35	PRVCHD31	V3 Prevalent coronary heart disease
	V2m4	Derive35	PRVCHD32	V3 Prevalent CHD, unverified
	V2m4	Derive35	PRVCHD33	V4 Prevalent CHD, definition 3
Visit 4	V4m3	Derive44	PRVCHD42	V4 Prevalent CHD, unverified
	V4m3	Derive44	PRVCHD43	V4 Prevalent CHD, definition 3
Visit 5	V1m1	Derive51_140813	PRVCHD51	V5 prevalent CHD before v5
	V1m1	Derive51_140813	PRVCHD53	V5 prevalent CHD by the end of v5
	V1m1	Derive51_140813	PRVCHD54	V5 Prevalent CHD, unverified

- Obesity - Determined from BMI and waist circumference

BMI

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	BMI01	V1 BMI in kg/m^2
Visit 2	V1m4	Derive2_10	BMI21	V2 BMI in kg/m^2
Visit 3	V2m4	Derive37	BMI32	V3 BMI in kg/m^2
Visit 4	V4m3	Derive46	BMI41	V4 BMI in kg/m^2
Visit 5	V1m1	Derive51_141112	BMI51	V5 BMI in kg/m^2

Waist circumference

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	ANTA	ANTA07A	Waist girth to the nearest cm
Visit 2	V1m4	ANTB	ANTB04A	Waist girth to the nearest cm
Visit 3	V2m4	ANTC04	ANTC3A	Waist girth to the nearest cm
Visit 4	V4m3	ANTD04	ANTD3A	Waist girth to the nearest cm
Visit 5	V1m1	ANT_rviid	ANT10a	Waist girth to the nearest cm

- Hypertension

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	HYPERT04	Hypertension, definition 4
	V1m4	Derive10	HYPERT05	Hypertension, definition 5
	V1m4	Derive10	HYPERT06	Hypertension, definition 6
Visit 2	V1m4	Derive28	HYPERT24	Hypertension, definition 4
	V1m4	Derive28	HYPERT25	Hypertension, definition 5
	V1m4	Derive28	HYPERT26	Hypertension, definition 6
Visit 3	V2m4	Derive35	HYPERT34	Hypertension, definition 4
	V2m4	Derive35	HYPERT35	Hypertension, definition 5
	V2m4	Derive35	HYPERT36	Hypertension, definition 6
Visit 4	V4m3	Derive44	HYPERT44	Hypertension, definition 4
	V4m3	Derive44	HYPERT45	Hypertension, definition 5
	V4m3	Derive44	HYPERT46	Hypertension, definition 6
Visit 5	V1m1	Derive51_131218	HYPERT54	Hypertension, definition 4
	V1m1	Derive51_131218	HYPERT55	Hypertension, definition 5
	V1m1	Derive51_131218	HYPERT56	Hypertension, definition 6

- Deep Vein Thrombosis

Visit	Version	Form	Variable name	Definition
Visit 4	V4m3	MHQA04	MHQA9A	Ever told deep vein thrombosis

- Gender

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive11	Gender	Sex
Visit 2	V1m4	Derive2_10	gender	sex
Visit 3	V2m4	Derive35	gender	sex
Visit 4	V4m3	Derive44	gender	sex
Visit 5	V1m1	Derive51_140409	gender	sex

- hsCRP

Visit	Version	Form	Variable name	Definition
Visit 5 (2)*	V1m1	V1_v5_analytes_140108	CRP_V2	V2 hsCRP (mg/L)
Visit 5 (4)*	V1m1	V1_v5_analytes_140108	CRP_V4	V4 hsCRP (mg/L)

Visit 5	V1m1	V1_v5_analytes_140108	LIP33	hsCRP (mg/L)

* Recorded in the dataset at visit 5

- Cystatin-C visit 4: Cystatin_C visit5/NCS: CYSC3

Visit	Version	Form	Variable name	Definition
Visit 5 (2)*	V1m1	V1_v5_analytes_140108	CYSC_V2	V2 Cystatin-C (mg/L)
Visit 5 (4)*	V1m1	V1_v5_analytes_140108	CYSC_V4	V4 calibrated Cystatin-C (mg/L)
Visit 5	V1m1	V1_v5_analytes_140108	CYSC_V5	V5 Cystatin-C (mg/L) reference

*Recorded in the dataset at visit 5

- eGFR – Determined from serum creatinine

Visit	Version	Form	Variable name	Definition
Visit 5 (1)*	V1m1	V1_v5_analytes_140108	EGFRSCR_V2	V1 eGFR- creatinine (ml/min/1.73m^2)
Visit 5 (2)*	V1m1	V1_v5_analytes_140108	EGFRSCR_V2	V2 eGFR- creatinine (ml/min/1.73m^2)
Visit 5 (4)*	V1m1	V1_v5_analytes_140108	EGFRSCR_V4	V4 eGFR- creatinine (ml/min/1.73m^2)
Visit 5	V1m1	V1_v5_analytes_140108	EGFRSCR_V5	V5 eGFR- creatinine (ml/min/1.73m^2)

*Recorded in the dataset at visit 5

- Albumin-to-Creatinine Ratio – determined from urinary albumin and creatinine

Urinary albumin

Visit	Version	Form	Variable name	Definition
Visit 5	V1m1	Chm_rviid	CHM33	urine albumin—UMALCR (mg g/CR)
	V1m1	Chm_rviid	CHM39	urine albumin—UMALI (mg/l)

Creatinine

Visit	Version	Form	Variable name	Definition
Visit 4	V4m3	Lipd04	LIPD6A	Creatinine
Visit 5	V1m1	Chm_rviid	CHM45	Urine Creatinine (mg/dl)

Albumin-to-Creatinine Ratio

Visit	Version	Form	Variable name	Definition
labdata	V1m1	Uc4507_as2002_02_finallurinary_p	newacrfinal	Albumin-to-Creatinine Ratio, mg/g

- All ICD-9 discharge diagnoses for hospitalization events with discharge diagnosis for serious infection (Appendix 2) or sepsis [038-039.9, 020.0, 790.7, 117.9, 112.5, 112.81, 995.91, 995.92 and 785.52].

Visit	Version	Form	Variable name	Definition
Cohort surveillence	V9m1	C10celb1	CELB10A, CELB10B, CELB10C, CELB10D, CELB10E, CELB10F, CELB10G, CELB10H, CELB10I, CELB10J, CELB10K, CELB10L, CELB10M, CELB10N, CELB10O, CELB10P, CELB10Q, CELB10R, CELB10S, CELB10T, CELB10U, CELB10V, CELB10W, CELB10X, CELB10Y, CELB10Z,	ICD-9 code

- Dates of all serious infection hospitalization events (days from enrollment)

Visit	Version	Form	Variable name	Definition
Cohort surveillence	V9m1	C10celb1	CELB04D	Day of discharge or death
Cohort surveillence	V9m1	C10celb1	CELB04M	Month of discharge or death
Cohort surveillence	V9m1	C10celb1	CELB04Y	Year of discharge or death

APPENDIX 2

ICD-9 discharge diagnoses for sepsis, adopted from Martin, et al.[4]

038-0.38.9	Septicemia
020.0	Septicemic
790.7	Bacteremia
117.9	Disseminated fungal infection
112.5	Disseminated candida infection
112.81	Disseminated fungal endocarditis
995.91	Sepsis
995.92	Severe sepsis
785.52	Septic shock

APPENDIX 3

Angus, et al. ICD-9 codes for a serious infection.[1] Includes all subgroup under each major code.

Infection Cat	ICD-9 Code	Description
Parasitic	001	Cholera
	002	Typhoid/paratyphoid fever
	003	Other salmonella infection
	004	Shigellosis
	005	Other food poisoning
	008	Intestinal infections due to Escherichia coli
	008.1	Intestinal infections due to Arizona group of paracolon bacillus
	008.2	Intestinal infections due to Aerobacter aerogenes
	008.3	Intestinal infections due to Proteus (mirabilis morganii)
	008.4	Intestinal infections due to unspecified bacteria
	008.5	Bacterial enteritis, unspecified
	009	Ill-defined intestinal infection
	013	CNS tuberculosis
	018	Miliary tuberculosis
	020	Plague
	021	Tularemia
	022	Anthrax
	023	Brucellosis
	024	Glanders
	025	Melioidosis
	026	Rat-bite fever
	027	Other bacterial zoonoses
	032	Diphtheria
	033	Whooping cough
	034	Streptococcal throat/scarlet fever
	035	Erysipelas
	036	Meningococcal infection
	037	Tetanus
	038	Septicemia
	039	Actinomycotic infections
	040	Other bacterial diseases
	041	Bacterial infection in other diseases not specified
	098	Gonococcal infections
	100	Leptospirosis
	101	Vincent's angina
	112	Candidiasis, of mouth
	112.4	Candidiasis, of lung
	112.5	Candidiasis, disseminated
	112.8	Candidiasis, of other specified sites
	114	Coccidioidomycosis

	115	Histoplasmosis
	116	Blastomycotic infection
	117	Other mycoses
	118	Opportunistic mycoses
Nervous	320	Bacterial meningitis
	321	Cryptococcal meningitis
	321.1	Meningitis in other fungal diseases
	324	CNS abcess
	325	Phlebitis of intracranial sinus
	360	Purulent endophthalmitis
	376	Acute inflammation of orbit
	380.14	Malignant otitis externa
	383	Acute mastoiditis
Circulatory	420.99	Acute pericarditis due to other specified organisms
	421	Acute or subacute endocarditis
Respiratory	461	Acute sinusitis
	462	Acute pharyngitis
	463	Acute tonsillitis
	464	Acute laryngitis/tracheitis
	465	Acute upper respiratory infection of multiple sites/not otherwise specified
	475	Peritonsillar abscess
	481	Pneumococcal pneumonia
	482	Other bacterial pneumonia
	485	Bronchopneumonia with organism not otherwise specified
	486	Pneumonia, organism not otherwise specified
	491.21	Acute exacerbation of obstructive chronic bronchitis
	494	Bronchiectasis
	510	Empyema
	513	Abscess of lung and mediastinum
Digestive	522.5	Periapical abscess without sinus
	522.7	Periapical abscess with sinus
	526.4	Inflammatory conditions of the jaw
	527.3	Abscess of the salivary glands
	528.3	Cellulitis and abscess of oral soft tissue
	540	Acute appendicitis
	541	Appendicitis not otherwise specified
	542	Other appendicitis
	562.01	Diverticulitis of the small intestine without hemorrhage
	562.03	Diverticulitis of the small intestine with hemorrhage
	562.11	Diverticulitis of colon without hemorrhage
	562.13	Diverticulitis of colon with hemorrhage

	566	Abscess of the anal and rectal regions
	567	Peritonitis
	569.5	Intestinal abscess
	569.61	Infection of colostomy or enterostomy
	569.83	Perforation of intestine
	572	Abscess of liver
	572.1	Portal pyemia
	575	Acute cholecystitis
Genitourinary	590	Kidney infection
	599	Urinary tract infection not otherwise specified
	601	Prostatic inflammation
	604	Orchitis and epididymitis
	614	Female pelvic inflammation disease
	615	Uterine inflammatory disease
	616.3	Abscess of Bartholin's gland
	616.4	Other abscess of vulva
Pregnancy infection	634	Spontaneous abortion, complicated by genital tract and pelvic
	635	Legally induced abortion, complicated by genital tract and pelvic infection
	636	Illegally induced abortion, complicated by genital tract and pelvic infection
	637	Unspecified abortion, complicated by genital tract and pelvic infection
	638	Failed attempted abortion, complicated by genital tract and pelvic infection
	639	Complications following abortion and ectopic and molar pregnancies
	646.6	Infections of genitourinary tract in pregnancy
	658.4	Infection of amniotic cavity
	670	Major puerperal infection
	675.1	Abscess of breast
Skin	681	Cellulitis, finger/toe
	682	Other cellulitis or abscess
	683	Acute lymphadenitis
	685	Pilonidal cyst, with abscess
	686	Other local skin infection
Musculoskeletal	711	Pyogenic arthritis
	728.86	Necrotizing fasciitis
	730	Osteomyelitis
Other	790.7	Bacteremia

958.3	Posttraumatic wound infection, not elsewhere classified
996.6	Infection or inflammation of device/graft
998.5	Postoperative infection
999.3	Infectious complication of medical care not otherwise classified
995.91	Sepsis

APPENDIX 4

Angus, et al. ICD-9 codes for organ dysfunction.[1]

Organ System	ICD-9 Code	Description
Cardiovascular	458	Orthostatic hypotension
	458.8	Other specified hypotension
	458.9	Hypotension, unspecified
	785.5	Shock without mention of trauma
Hematologic	286.6	Defibrination syndrome
	286.9	Other and unspecified coagulation defects
	287.4	Secondary thrombocytopenia
	287.5	Thombocytopenia, unspecified
Hepatic	570	Acute and subacute necrosis of liver
	573.4	Hepatic infarction
Neurologic	293	Transient organic psychosis
	348.1	Anoxic brain damage
	348.3	Encephalopathy
Renal	584	Acute renal failure
Respiratory	518.8	Respiratory failure
	786.03	Apnea
	799.1	Respiratory arrest

Angus, et al. ICD-9 codes for explicitly coded severe sepsis.[1]

ICD-9-CM Code	ICD-9-CM Code Description
995.92	Severe Sepsis
785.52	Septic Shock

ARIC Manuscript Proposal #2881

PC Reviewed: 12/13/16 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: **External Validation of the REGARDS Sepsis Risk Score**

b. Abbreviated Title (Length 26 characters): **SRS Validation**

2. Writing Group:

Henry Wang, MD, MS (Department of Emergency Medicine, University of Alabama at Birmingham)
Sachin Yende, MD (Department of Critical Care Medicine, University of Pittsburgh)
John Donnelly, MPH (Department of Emergency Medicine, Department of Epidemiology, University of Alabama at Birmingham)
Gerardo Heiss, MD, PhD (Department of Epidemiology, University of North Carolina Chapel Hill)

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

First author: **Henry E. Wang, MD, MS**
Address: Department of Emergency Medicine
University of Alabama at Birmingham
OHB 251
Birmingham, AL 35249

Phone: 205-996-6526 Fax:
E-mail: hwang@uabmc.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: **Gerardo Heiss, PhD**
Address: University of North Carolina Chapel Hill
CVS Plaza
137 East Franklin Street, Suite 306
CB #8050
Chapel Hill, NC 27514

Phone: (919) 962-3253 Fax:
E-mail: gerardo_heiss@unc.edu

3. Timeline:

We expect to finish this project within 12 months of receipt of dataset.

4. Rationale:

Sepsis is the syndrome of microbial infection complicated by systemic inflammation. In the United States (US), sepsis places an immense burden on the healthcare system, resulting in 750,000 hospitalizations, 570,000 Emergency Department visits and 200,000 deaths annually.[1-3]

In a prior effort, using data from the Reasons for Geographic and Racial Differences in Stroke (REGARDS) cohort, we derived a Sepsis Risk Score (SRS) and a Severe Sepsis Risk Score (SSRS) characterizing community-dwelling individuals' long-term risks of sepsis and severe sepsis. (Wang, Critical Care Medicine 2016) The sepsis and severe sepsis events in REGARDS-Sepsis were based upon participant report of hospitalizations for a serious infection and confirmed by manual review of Emergency Department and hospital admission data.

The objective of the current proposal is to externally validate the SRS and SSRS using hospitalization event data from the Atherosclerosis Risk in Communities (ARIC) cohort.

This analysis will occur in parallel with two other analyses:

- Internal validation of the SRS and SSRS using linked REGARDS-Medicare claims data.
- External validation of the SRS and SSRS using data from the Cardiovascular Health Study (CHS) cohort.

We will compose a single manuscript summarizing the combined results of the three parallel validation efforts.

5. Main Hypothesis/Study Questions:

- Using data from the ARIC cohort, validate the ability of the REGARDS Sepsis Risk Score (SRS) and Severe Sepsis Risk Score (SSRS) to predict sepsis and severe sepsis hospitalizations.

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 the following variables (variables listed in order of REGARDS-SRS) (Potential ARIC variables listed in Appendix 1):

1) Chronic Lung Disease (defined by pulmonary function tests or medical history)
2) Age
3) Peripheral Artery Disease
4) Diabetes
5) Tobacco Use (Present smoker)

6) Race
7) Stroke
8) Atrial Fibrillation
9) Coronary Artery Disease
10) Obesity (Determined from BMI and waist circumference)
11) Hypertension
12) Deep Vein Thrombosis
13) Gender
14) hsCRP (calculated correction)
15) Cystatin-C
16) eGFR
17) Albumin-to-Creatinine Ratio

18) All ICD-9 discharge diagnoses for hospitalization events with discharge diagnosis for serious infection (Appendix 2) or sepsis [038-039.9, 020.0, 790.7, 117.9, 112.5, 112.81, 995.91, 995.92 and 785.52].

19) Dates of all serious infection hospitalization events (days from enrollment)

Population: The analysis will include all adult participants in the ARIC cohort with valid follow-up from ARIC enrollment (1989-90) through present.

Study Outcomes: The primary endpoints are hospitalization for sepsis and severe sepsis.

To identify sepsis hospitalizations, we will use the ICD-9 taxonomy of Martin, et al.[4] (Appendix 2) We will classify any hospitalization with a primary or secondary discharge diagnosis satisfying the Martin criteria as sepsis.

To identify severe sepsis hospitalizations, we will use the ICD-9 taxonomy of Angus, et al.[1] (Appendices 3 and 4) Rather than using sepsis-specific ICD-9 codes, the Angus system identifies severe sepsis as hospitalizations with discharge diagnoses for both a serious infection and organ dysfunction. We will also classify the presence of the codes 995.92 (Severe Sepsis) and 785.52 (Septic Shock) as severe sepsis.

We will identify sepsis events from a 10-year follow-up time window. We will shift the time window based upon the timing of biomarkers available in ARIC. For example, we may choose the 10-year period follow procurement of urinary albumin and creatinine.

Brief analysis plan: For each ARIC participant, we will calculate the REGARDS Sepsis Risk Score (SRS) and Severe Sepsis Risk Score (SSRS) and corresponding risk quintiles. The components of the SRS and SSRS scores will be determined using the earliest available measurement. Some components were not measured at the beginning of ARIC; for example, urinary albumin and creatinine. Thus, we may shift the follow-up period to accommodate earliest available baseline data.

We will fit a Cox regression model using time to first "Martin sepsis" event as the outcome. The primary exposure will be SRS point score or quintile of predicted risk. We will validate the SRS model by determining model discrimination (Harrell's C) and calibration (PI deciles). We will repeat the analysis using two forms of the SRS with and without biomarkers; we will calculate the Net Reclassification Improvement (NRI) for the SRS with biomarkers over the SRS without biomarkers. We will repeat the same process using time to first "Angus severe sepsis" event to validate the SSRS.

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? _ X ___ Yes _ ___ No
(This file ICTDER 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"? __X__ 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)?

There are no related papers.

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

11.b. If yes, is the proposal

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

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

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

12a. 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.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __X__ No.

REFERENCES

1. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Critical care medicine 2001;29:1303-10.

2. Wang HE, Shapiro NI, Angus DC, Yealy DM. National estimates of severe sepsis in United States emergency departments. Critical care medicine 2007;35:1928-36.

3. Dellinger RP, Carlet JM, Masur H, et al. Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. Intensive care medicine 2004;30:536-55.

4. Martin GS, Mannino DM, Eaton S, Moss M. The epidemiology of sepsis in the United States from 1979 through 2000. The New England journal of medicine 2003;348:1546-54.

APPENDIX 1

Potential ARIC Variables to be used in the Analysis

- Chronic Lung Disease – will be defined by pulmonary function tests or medical history

Pulmonary function data:

Visit	Version	Form	Variable name	Definition
Visit 1 (Pulmonary Function Test)	V1M4	PFTA	PFTA17	FEV(1) Over FEV(6)
	V1M4	PFTA	PFTA24	FVC (Liters)
	V1M4	PFTA	PFTA25	FEV(.5) (Liters)
	V1M4	PFTA	PFTA26	FEV(1) (Liters)
	V1M4	PFTA	PFTA27	FEV(3) (Liters)
	V1M4	PFTA	PFTA28	FEV(6) (Liters)
	V1M4	PFTA	PFTA29	FEV(3) Over FEV(6)
	V1M4	PFTA	PFTA30	FEV(.5) Over FVC
	V1M4	PFTA	PFTA31	FEV(1) Over FVC
	V1M4	PFTA	PFTA32	FEV(3) Over FVC
Visit 1 (Pulmonary Derived Variables)	V1M4	PULM	FVC01	FVC Predicted (liters)
	V1M4	PULM	FEV_501	FEV(.5) Predicted (liters)
	V1M4	PULM	FEV_101	FEV(1) Predicted (liters)
	V1M4	PULM	FEV_301	FEV(3) Predicted (liters)
	V1M4	PULM	FEV1FVC1	FEV(1)/FVC Predicted (%)
	V1M4	PULM	FEV3FVC1	FEV(3)/FVC Predicted (%)
Visit 2 (Pulmonary Function Test)	V1M4	PFTB	PFTB17	FEV(1) Over FEV(6)
	V1M4	PFTB	PFTB24	FVC (Liters)
	V1M4	PFTB	PFTB25	FEV(.5) (Liters)
	V1M4	PFTB	PFTB26	FEV(1) (Liters)
	V1M4	PFTB	PFTB27	FEV(3) (Liters)
	V1M4	PFTB	PFTB28	FEV(6) (Liters)
	V1M4	PFTB	PFTB29	FEV(3) Over FEV(6)
	V1M4	PFTB	PFTB30	FEV(.5) Over FVC
	V1M4	PFTB	PFTB31	FEV(1) Over FVC
	V1M4	PFTB	PFTB32	FEV(3) Over FVC
Visit 2 (Pulmonary Derived Variables)	V1M4	PULM21	FEF_122	FEV(1)/FVC Predicted (%)
	V1M4	PULM21	FEF_322	FEV(3)/FVC Predicted (%)
	V1M4	PULM21	FEF_522	FEV(.5) Predicted (Liters)
	V1M4	PULM21	FEV1FVC2	FEF(1)/FVC Predicted(%)
	V1M4	PULM21	FEV3FVC2	FEF(3)/FVC Predicted(%)
	V1M4	PULM21	FVC22	FVC Predicted (Liters)

Home interview/health history:

Visit	Version	Form	Variable name	Definition
Visit 1 (home interview)	V1M4	HOM	HOM10G	Chronic lung disease, such as chronic bronchitis, or emphysema
Visit 2 (health history)	V1M4	HHXB	HHXB05E	Chronic Lung Disease, Such As Chronic Bronchitis, Or Emphysema

- Age

Visit	Version	Form	Variable name	definition
Visit 1	V1m4	Derive10	V1age01	Age at visit 1
Visit 2	V1m4	Derive28	V2age22	Age at visit 2
Visit 3	V2M4	Derive35	V3age31	Age at visit 3
Visit 4	V4m3	derive44	V4age41	Age at visit 4
Visit 5	V1m1	Derive51_131218	V5age52	Corrected age at visit 4

- Peripheral Artery Disease

Visit	Version	Form	Variable name	definition
Visit 1	V1m4	Derive11	PAD01	PAD at visit 1, definition 1
	V1m4	Derive11	PAD02	PAD at visit 1, definition 2
Visit 3	V2m4	Derive37	PAD31	PAD at visit 3, definition 1
	V2m4	Derive37	PAD32	PAD at visit 3, definition 2
Visit 4	V4m3	derive46	PAD41	PAD at visit 4, definition 1
	V4m3	derive46	PAD42	PAD at visit 4, definition 2

- Diabetes

Visit	Version	Form	Variable name	definition
Visit 1	V1m4	Derive13	Diabts02	Diabetes at visit 1
	V1m4	Derive13	Diabts03	Diabetes with fasting glucose using cutoff of 126
Visit 2	V1m4	Derive28	Diabts22	Diabetes at visit 2
	V1m4	Derive28	Diabts23	Diabetes with fasting glucose using cutoff of 126
Visit 3	V2m4	Derive35	Diabts33	Diabetes at visit 3 using fasting glucose cutoff of 140
	V2m4	Derive35	Diabts33	Diabetes at visit 3 using fasting glucose cutoff of 126
Visit 4	V4m3	Derive44	Diabts41	Diabetes at visit 4 using fasting glucose cutoff of 140
	V4m3	Derive44	Diabts42	Diabetes at visit 4 using fasting glucose cutoff of 126
Visit	V1m1	Derive51_140409	Diabts53	Diabetes at visit 5 using fasting glucose

5				cutoff of 140
	V1m1	Derive51_140409	Diabts54	Diabetes at visit 5 using fasting glucose cutoff of 126
	V1m1	Derive51_140409	Diabts55	Diabetes at visit 5—lab and med only, using fasting glucose cutoff of 140
	V1m1	Derive51_140409	Diabts56	Diabetes at visit 5 using HbA1c cutoff of 6.5%
	V1m1	Derive51_140409	Diabts57	Diabetes at visit 5 (DM medication or DM reported on AFU)

- Tobacco Use (Present smoker)

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	Cigr01	Cigar smoking status
	V1m4	Derive10	Cigryr01	Cigar years of smoking
	V1m4	Derive10	Cigt01	Cigarette smoking status
	V1m4	Derive10	Cigtyr01	Cigarette years of smoking
	V1m4	Derive10	Cursmk01	Current smoking
	V1m4	Derive10	Evrsmk01	Ever smoked cigarette
	V1m4	Derive10	Forsmk01	Former cigarette smoking
Visit 2	V1m4	Derive2_10	Evrsmk21	Ever smoked cigarette
	V1m4	Derive2_10	Forsmk21	Former cigarette smoking
	V1m4	Derive2_10	Cigt21	Cigarette smoking status
	V1m4	Derive2_10	Cursmk21	Current smoking
Visit 3	V2m4	Derive37	Evrsmk31	Ever smoked cigarette
	V2m4	Derive37	Forsmk31	Former cigarette smoking
	V2m4	Derive37	Cigt31	Cigarette smoking status
	V2m4	Derive37	Cursmk31	Current smoking
Visit 4	V4m3	Derive44	Evrsmk41	Ever smoked cigarette
	V4m3	Derive44	Forsmk41	Former cigarette smoking
	V4m3	Derive44	Cigt41	Cigarette smoking status
	V4m3	Derive44	Cursmk41	Current smoking
Visit 5	V1m1	Derive51_140409	Evrsmk52	Ever smoked cigarette
	V1m1	Derive51_140409	Forsmk52	Former cigarette smoking
	V1m1	Derive51_140409	Cigt52	Cigarette smoking status
	V1m1	Derive51_140409	Cursmk52	Current smoking

- Race

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	racegrp	Race (from FTRA23)
	V1m4	Derive10	V1corra1	Corrected race group
Visit 2	V1m4	Derive2_10	racegrp	Race (from FTRA23)
Visit 3	V2m4	Derive37	racegrp	Race (from FTRA23)
Visit 4	V4m3	Derive44	racegrp	Race (from FTRA23)
Visit 5	V1m1	Derive51_140409	racegrp	Race (from FTRA23)

	V1m1	Derive51_140409	Racegrp51	Corrected race group

- Stroke

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Stroke01	Stroke01	Stroke
Visit 2	V1m4	Stroke2	Prvstr21	Prevalent stroke at visit 2
Visit 3	V2m4	Stroke32	Stroke31	Stroke
Visit 4	V4m3	Derive44	Prvstr41	Prevalent stroke at visit 4
	V4m3	Stroke41	Stroke41	Stroke
Visit 5	V1m1	Derive51_140611	Prvstr51	Prevalent stroke at by the end of visit 5
	V1m1	Derive51_140611	Prvstr52	Prevalent stroke at by the end of visit 5--unverified

- Atrial Fibrillation

Visit	Version	Form	Variable name	Definition
Visit 5	V1m1	Derive51_140813	Prvaf51	V5 atrial fibrillation/flutter before v5
	V1m1	Derive51_140813	Prvaf52	V5 atrial fibrillation/flutter by the end of v5

- Coronary Artery Disease

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive13	PRVCHD05	V1 Prevalent coronary heart disease
Visit 2	V1m4	Derive28	PRVCHD21	V2 Prevalent coronary heart disease
	V1m4	Derive28	PRVCHD22	V2 Prevalent CHD, unverified
	V1m4	Derive28	PRVCHD23	V2 Prevalent CHD, definition 3
Visit 3	V2m4	Derive35	PRVCHD31	V3 Prevalent coronary heart disease
	V2m4	Derive35	PRVCHD32	V3 Prevalent CHD, unverified
	V2m4	Derive35	PRVCHD33	V4 Prevalent CHD, definition 3
Visit 4	V4m3	Derive44	PRVCHD42	V4 Prevalent CHD, unverified
	V4m3	Derive44	PRVCHD43	V4 Prevalent CHD, definition 3
Visit 5	V1m1	Derive51_140813	PRVCHD51	V5 prevalent CHD before v5
	V1m1	Derive51_140813	PRVCHD53	V5 prevalent CHD by the end of v5
	V1m1	Derive51_140813	PRVCHD54	V5 Prevalent CHD, unverified

- Obesity - Determined from BMI and waist circumference

BMI

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	BMI01	V1 BMI in kg/m^2
Visit 2	V1m4	Derive2_10	BMI21	V2 BMI in kg/m^2
Visit 3	V2m4	Derive37	BMI32	V3 BMI in kg/m^2
Visit 4	V4m3	Derive46	BMI41	V4 BMI in kg/m^2
Visit 5	V1m1	Derive51_141112	BMI51	V5 BMI in kg/m^2

Waist circumference

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	ANTA	ANTA07A	Waist girth to the nearest cm
Visit 2	V1m4	ANTB	ANTB04A	Waist girth to the nearest cm
Visit 3	V2m4	ANTC04	ANTC3A	Waist girth to the nearest cm
Visit 4	V4m3	ANTD04	ANTD3A	Waist girth to the nearest cm
Visit 5	V1m1	ANT_rviid	ANT10a	Waist girth to the nearest cm

- Hypertension

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive10	HYPERT04	Hypertension, definition 4
	V1m4	Derive10	HYPERT05	Hypertension, definition 5
	V1m4	Derive10	HYPERT06	Hypertension, definition 6
Visit 2	V1m4	Derive28	HYPERT24	Hypertension, definition 4
	V1m4	Derive28	HYPERT25	Hypertension, definition 5
	V1m4	Derive28	HYPERT26	Hypertension, definition 6
Visit 3	V2m4	Derive35	HYPERT34	Hypertension, definition 4
	V2m4	Derive35	HYPERT35	Hypertension, definition 5
	V2m4	Derive35	HYPERT36	Hypertension, definition 6
Visit 4	V4m3	Derive44	HYPERT44	Hypertension, definition 4
	V4m3	Derive44	HYPERT45	Hypertension, definition 5
	V4m3	Derive44	HYPERT46	Hypertension, definition 6
Visit 5	V1m1	Derive51_131218	HYPERT54	Hypertension, definition 4
	V1m1	Derive51_131218	HYPERT55	Hypertension, definition 5
	V1m1	Derive51_131218	HYPERT56	Hypertension, definition 6

- Deep Vein Thrombosis

Visit	Version	Form	Variable name	Definition
Visit 4	V4m3	MHQA04	MHQA9A	Ever told deep vein thrombosis

- Gender

Visit	Version	Form	Variable name	Definition
Visit 1	V1m4	Derive11	Gender	Sex
Visit 2	V1m4	Derive2_10	gender	sex
Visit 3	V2m4	Derive35	gender	sex
Visit 4	V4m3	Derive44	gender	sex
Visit 5	V1m1	Derive51_140409	gender	sex

- hsCRP

Visit	Version	Form	Variable name	Definition
Visit 5 (2)*	V1m1	V1_v5_analytes_140108	CRP_V2	V2 hsCRP (mg/L)
Visit 5 (4)*	V1m1	V1_v5_analytes_140108	CRP_V4	V4 hsCRP (mg/L)

Visit 5	V1m1	V1_v5_analytes_140108	LIP33	hsCRP (mg/L)

* Recorded in the dataset at visit 5

- Cystatin-C visit 4: Cystatin_C visit5/NCS: CYSC3

Visit	Version	Form	Variable name	Definition
Visit 5 (2)*	V1m1	V1_v5_analytes_140108	CYSC_V2	V2 Cystatin-C (mg/L)
Visit 5 (4)*	V1m1	V1_v5_analytes_140108	CYSC_V4	V4 calibrated Cystatin-C (mg/L)
Visit 5	V1m1	V1_v5_analytes_140108	CYSC_V5	V5 Cystatin-C (mg/L) reference

*Recorded in the dataset at visit 5

- eGFR – Determined from serum creatinine

Visit	Version	Form	Variable name	Definition
Visit 5 (1)*	V1m1	V1_v5_analytes_140108	EGFRSCR_V2	V1 eGFR- creatinine (ml/min/1.73m^2)
Visit 5 (2)*	V1m1	V1_v5_analytes_140108	EGFRSCR_V2	V2 eGFR- creatinine (ml/min/1.73m^2)
Visit 5 (4)*	V1m1	V1_v5_analytes_140108	EGFRSCR_V4	V4 eGFR- creatinine (ml/min/1.73m^2)
Visit 5	V1m1	V1_v5_analytes_140108	EGFRSCR_V5	V5 eGFR- creatinine (ml/min/1.73m^2)

*Recorded in the dataset at visit 5

- Albumin-to-Creatinine Ratio – determined from urinary albumin and creatinine

Urinary albumin

Visit	Version	Form	Variable name	Definition
Visit 5	V1m1	Chm_rviid	CHM33	urine albumin—UMALCR (mg g/CR)
	V1m1	Chm_rviid	CHM39	urine albumin—UMALI (mg/l)

Creatinine

Visit	Version	Form	Variable name	Definition
Visit 4	V4m3	Lipd04	LIPD6A	Creatinine
Visit 5	V1m1	Chm_rviid	CHM45	Urine Creatinine (mg/dl)

Albumin-to-Creatinine Ratio

Visit	Version	Form	Variable name	Definition
labdata	V1m1	Uc4507_as2002_02_finallurinary_p	newacrfinal	Albumin-to-Creatinine Ratio, mg/g

- All ICD-9 discharge diagnoses for hospitalization events with discharge diagnosis for serious infection (Appendix 2) or sepsis [038-039.9, 020.0, 790.7, 117.9, 112.5, 112.81, 995.91, 995.92 and 785.52].

Visit	Version	Form	Variable name	Definition
Cohort surveillence	V9m1	C10celb1	CELB10A, CELB10B, CELB10C, CELB10D, CELB10E, CELB10F, CELB10G, CELB10H, CELB10I, CELB10J, CELB10K, CELB10L, CELB10M, CELB10N, CELB10O, CELB10P, CELB10Q, CELB10R, CELB10S, CELB10T, CELB10U, CELB10V, CELB10W, CELB10X, CELB10Y, CELB10Z,	ICD-9 code

- Dates of all serious infection hospitalization events (days from enrollment)

Visit	Version	Form	Variable name	Definition
Cohort surveillence	V9m1	C10celb1	CELB04D	Day of discharge or death
Cohort surveillence	V9m1	C10celb1	CELB04M	Month of discharge or death
Cohort surveillence	V9m1	C10celb1	CELB04Y	Year of discharge or death

APPENDIX 2

ICD-9 discharge diagnoses for sepsis, adopted from Martin, et al.[4]

038-0.38.9	Septicemia
020.0	Septicemic
790.7	Bacteremia
117.9	Disseminated fungal infection
112.5	Disseminated candida infection
112.81	Disseminated fungal endocarditis
995.91	Sepsis
995.92	Severe sepsis
785.52	Septic shock

APPENDIX 3

Angus, et al. ICD-9 codes for a serious infection.[1] Includes all subgroup under each major code.

Infection Cat	ICD-9 Code	Description
Parasitic	001	Cholera
	002	Typhoid/paratyphoid fever
	003	Other salmonella infection
	004	Shigellosis
	005	Other food poisoning
	008	Intestinal infections due to Escherichia coli
	008.1	Intestinal infections due to Arizona group of paracolon bacillus
	008.2	Intestinal infections due to Aerobacter aerogenes
	008.3	Intestinal infections due to Proteus (mirabilis morganii)
	008.4	Intestinal infections due to unspecified bacteria
	008.5	Bacterial enteritis, unspecified
	009	Ill-defined intestinal infection
	013	CNS tuberculosis
	018	Miliary tuberculosis
	020	Plague
	021	Tularemia
	022	Anthrax
	023	Brucellosis
	024	Glanders
	025	Melioidosis
	026	Rat-bite fever
	027	Other bacterial zoonoses
	032	Diphtheria
	033	Whooping cough
	034	Streptococcal throat/scarlet fever
	035	Erysipelas
	036	Meningococcal infection
	037	Tetanus
	038	Septicemia
	039	Actinomycotic infections
	040	Other bacterial diseases
	041	Bacterial infection in other diseases not specified
	098	Gonococcal infections
	100	Leptospirosis
	101	Vincent's angina
	112	Candidiasis, of mouth
	112.4	Candidiasis, of lung
	112.5	Candidiasis, disseminated
	112.8	Candidiasis, of other specified sites
	114	Coccidioidomycosis

	115	Histoplasmosis
	116	Blastomycotic infection
	117	Other mycoses
	118	Opportunistic mycoses
Nervous	320	Bacterial meningitis
	321	Cryptococcal meningitis
	321.1	Meningitis in other fungal diseases
	324	CNS abcess
	325	Phlebitis of intracranial sinus
	360	Purulent endophthalmitis
	376	Acute inflammation of orbit
	380.14	Malignant otitis externa
	383	Acute mastoiditis
Circulatory	420.99	Acute pericarditis due to other specified organisms
	421	Acute or subacute endocarditis
Respiratory	461	Acute sinusitis
	462	Acute pharyngitis
	463	Acute tonsillitis
	464	Acute laryngitis/tracheitis
	465	Acute upper respiratory infection of multiple sites/not otherwise specified
	475	Peritonsillar abscess
	481	Pneumococcal pneumonia
	482	Other bacterial pneumonia
	485	Bronchopneumonia with organism not otherwise specified
	486	Pneumonia, organism not otherwise specified
	491.21	Acute exacerbation of obstructive chronic bronchitis
	494	Bronchiectasis
	510	Empyema
	513	Abscess of lung and mediastinum
Digestive	522.5	Periapical abscess without sinus
	522.7	Periapical abscess with sinus
	526.4	Inflammatory conditions of the jaw
	527.3	Abscess of the salivary glands
	528.3	Cellulitis and abscess of oral soft tissue
	540	Acute appendicitis
	541	Appendicitis not otherwise specified
	542	Other appendicitis
	562.01	Diverticulitis of the small intestine without hemorrhage
	562.03	Diverticulitis of the small intestine with hemorrhage
	562.11	Diverticulitis of colon without hemorrhage
	562.13	Diverticulitis of colon with hemorrhage

	566	Abscess of the anal and rectal regions
	567	Peritonitis
	569.5	Intestinal abscess
	569.61	Infection of colostomy or enterostomy
	569.83	Perforation of intestine
	572	Abscess of liver
	572.1	Portal pyemia
	575	Acute cholecystitis
Genitourinary	590	Kidney infection
	599	Urinary tract infection not otherwise specified
	601	Prostatic inflammation
	604	Orchitis and epididymitis
	614	Female pelvic inflammation disease
	615	Uterine inflammatory disease
	616.3	Abscess of Bartholin's gland
	616.4	Other abscess of vulva
Pregnancy infection	634	Spontaneous abortion, complicated by genital tract and pelvic
	635	Legally induced abortion, complicated by genital tract and pelvic infection
	636	Illegally induced abortion, complicated by genital tract and pelvic infection
	637	Unspecified abortion, complicated by genital tract and pelvic infection
	638	Failed attempted abortion, complicated by genital tract and pelvic infection
	639	Complications following abortion and ectopic and molar pregnancies
	646.6	Infections of genitourinary tract in pregnancy
	658.4	Infection of amniotic cavity
	670	Major puerperal infection
	675.1	Abscess of breast
Skin	681	Cellulitis, finger/toe
	682	Other cellulitis or abscess
	683	Acute lymphadenitis
	685	Pilonidal cyst, with abscess
	686	Other local skin infection
Musculoskeletal	711	Pyogenic arthritis
	728.86	Necrotizing fasciitis
	730	Osteomyelitis
Other	790.7	Bacteremia

958.3	Posttraumatic wound infection, not elsewhere classified
996.6	Infection or inflammation of device/graft
998.5	Postoperative infection
999.3	Infectious complication of medical care not otherwise classified
995.91	Sepsis

APPENDIX 4

Angus, et al. ICD-9 codes for organ dysfunction.[1]

Organ System	ICD-9 Code	Description
Cardiovascular	458	Orthostatic hypotension
	458.8	Other specified hypotension
	458.9	Hypotension, unspecified
	785.5	Shock without mention of trauma
Hematologic	286.6	Defibrination syndrome
	286.9	Other and unspecified coagulation defects
	287.4	Secondary thrombocytopenia
	287.5	Thombocytopenia, unspecified
Hepatic	570	Acute and subacute necrosis of liver
	573.4	Hepatic infarction
Neurologic	293	Transient organic psychosis
	348.1	Anoxic brain damage
	348.3	Encephalopathy
Renal	584	Acute renal failure
Respiratory	518.8	Respiratory failure
	786.03	Apnea
	799.1	Respiratory arrest

Angus, et al. ICD-9 codes for explicitly coded severe sepsis.[1]

ICD-9-CM Code	ICD-9-CM Code Description
995.92	Severe Sepsis
785.52	Septic Shock