Claims-Based and Hybrid Measures of 30-Day Mortality Following Acute Ischemic Stroke Hospitalization Incorporating Risk Adjustment for Stroke Severity

Technical Report
(Version 1.0)

Submitted By

Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE)

Prepared For:

Centers for Medicare & Medicaid Services (CMS)

July 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CENTER FOR OUTCOMES RESEARCH AND EVALUATION (CORE) PROJECT TEAM

Jennifer Schwartz, PhD, MPH* – Lead
Yongfei Wang, MS* –Lead Analyst
Li Qin, MS, PhD – Supporting Analyst
Amena Keshawarz, MPH – Project Coordinator
Nicole Cormier, MPH – Research Associate
Hayley Dykhoff, BA – Research Assistant
Robert L. McNamara, MD, MHS* – Clinical Investigator
Karen B. Dorsey, MD, PhD* – Clinical Investigator
Megan Keenan, MPH – Project Manager
Lisa Suter, MD – Associate Director
Susannah Bernheim, MD, MHS – Project Director
Harlan M. Krumholz, MD, SM* -- Principal Investigator

*Yale School of Medicine

ACKNOWLEDGEMENTS

This work is a collaborative effort and the authors gratefully acknowledge and thank our many colleagues and collaborators for their thoughtful and instructive input.

Specifically, we would like to acknowledge the contribution of data and insights from our colleagues affiliated with the American Heart Association and American Stroke Association, and those affiliated with the Yale School of Medicine. Without their continued support, this project would not have been possible:

Gregg Fonarow, MD
Lee Schwamm, MD
Jason Sico, MD
Kevin Sheth, MD

We appreciate the ongoing contributions to this work from our clinical consultants and colleagues at CORE. These individuals include:

Kanchana Bhat, M.P.H.
Jacqueline Grady, M.S.
Lisa Suter, M.D.
Jessica Brewer, M.P.H.
Lori Schroeder, LL.M., J.D.

Finally, we would like to thank our Contracting Officer Representative at the Centers for Medicare & Medicaid Services, Dr. Lein Han, for her continued support of our work.

1. EXECUTIVE SUMMARY

The Centers for Medicare & Medicaid Services (CMS) publicly reports a 30-day hospital-level stroke mortality measure on *Hospital Compare* as part of the Inpatient Quality Reporting (IQR) program. CMS contracted with Yale New Haven Services Corporation, Center for Outcomes Research and Evaluation (CORE) to develop new stroke mortality measures that include an assessment of stroke severity in the risk adjustment models. This work was initiated in response to stakeholder feedback about the current measure and grows out of CMS' commitment to continually improve on quality measures and to seek opportunities to develop measures with clinical data. This new measure work became possible in part due to changes in clinical guidelines and hospital practices that allow for more standard collection of stroke severity. Based on a review of the literature, community comments, and current clinical guidelines for stroke care, we selected the National Institutes of Health Stroke Scale (NIHSS) as the stroke severity assessment to be incorporated into the measures.

CORE collaborated with the American Heart Association/American Stroke Association (AHA/ASA) to complete this work. Early in the project, we determined that a measure could be developed using NIHSS scores obtained from either Medicare administrative claims or from the electronic health record (EHR). An International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10)-based stroke severity score could be added to the claims-based model, whereas an EHR-based stroke severity score could be used to develop a measure that uses both claims and EHR data (hybrid measure). Because the most preferable form of the measures was unclear at the outset of development, this report describes the development of two types of 30-day hospital-level stroke mortality measures that contain the NIHSS: a claims-only measure and hybrid measures, which utilize both claims and clinical EHR data.

The cohorts and outcomes of the measures are aligned with CMS's original, publicly reported stroke mortality measure. Both new measures were developed using a linked dataset consisting of Medicare fee-for-service (FFS) claims and AHA/ASA Get With The Guidelines® (GWTG)-Stroke registry data. GWTG-Stroke registry data were used because at the time of measure development (2015), it was the largest database that included both the NIHSS and clinical EHR variables.

To build the risk adjustment model for the claims-only measure, the 41 claims-derived variables in the current publicly reported measure were considered as candidate variables in addition to the NIHSS. Variables that were predictive of mortality in the multivariate model were included in the final claims-only risk model. To build the risk-adjustment models for the hybrid measures, 14 clinical EHR variables were considered in addition to the 41 claims-derived variables and NIHSS. Similarly, variables that were predictive of mortality in the multivariate model were included in the final hybrid risk models. In an effort to streamline measure calculation and reduce dependence on claims data, the third measure is a hybrid measure that only utilizes data elements that could be extracted from an EHR for risk adjustment.

This report presents three risk-adjustment models for stroke mortality: The updated claims-only model includes 19 claims-derived variables and the NIHSS; one hybrid model includes 17 claims-derived variables, 3 clinical EHR variables, and the NIHSS; and the other hybrid model includes 8 clinical EHR variables and the NIHSS. All of the measures have modestly higher c-statistics and more parsimonious risk models than the current publicly reported stroke mortality measure. By developing measures using two different data sources, CMS has options with regards to their approach to the implementation of a stroke mortality measure that includes stroke severity.

2. INTRODUCTION

2.1 Overview

The Centers for Medicare & Medicaid Services (CMS) publicly reports a 30-day hospital-level stroke mortality measure on *Hospital Compare* as part of the Inpatient Quality Reporting (IQR) program. CMS contracted with Yale New Haven Services Corporation, Center for Outcomes Research and Evaluation (CORE) to develop new stroke mortality measures that include an assessment of stroke severity in the risk adjustment models. This work was initiated in response to community comments about the current measure, a commitment to seek opportunities to develop measures with richer clinical data, changes in clinical guidelines and hospital practices that allow for more standard collection of stroke severity, and in an effort to continually improve on quality measures and be responsive to stakeholder feedback.

CORE collaborated with the American Heart Association/American Stroke Association (AHA/ASA) to form a working group to determine how best to incorporate an assessment of a patient's stroke severity into hospital-level outcome measures. The AHA/ASA were chosen because of their position as one of the nation's preeminent non-profit public health organizations, their commitment to this area, and their possession of relevant data for use in the development of the measure. Based on a review of the literature and current clinical guidelines for stroke care, the National Institutes of Health Stroke Scale (NIHSS) was selected as the stroke severity assessment to be incorporated in the measure. At the outset of measure development, we assumed that NIHSS data would be obtained from electronic health records (EHRs) because it is collected with a standard definition and entered in structured fields, which indicates that it can be feasibly extracted. However, early in the development process, we learned that codes for the NIHSS will be added to International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) in October 2016. Therefore, this report describes the development of two types of hospital-level ischemic stroke mortality measures:

1. Updated Claims-Only Measure - The risk model for this measure utilizes both the first captured NIHSS and administrative claims data (demographics, comorbidities, and patient medical history).

2. Hybrid Measure- Hybrid measures are quality measures that utilize more than one source of data. This type of measure will rely on Medicare administrative claims and clinical EHR data. Since we planned to obtain the first captured NIHSS from the EHR, we also considered other data elements that are feasibly obtained from the EHR and taken upon a patient's arrival. We developed two hybrid risk-adjustment models:

    a) A risk model that utilizes the NIHSS, administrative claims data (demographics, comorbidities, and patient medical history), and clinical EHR data (demographics, laboratory results, and vital signs).

    b) A risk model that utilizes both the NIHSS and clinical EHR data (demographics, laboratory results, and vital signs) only.

The hybrid measure and both model options were developed for the Medicare population and rely on administrative claims data to derive the cohort and outcome, and are not intended to be fully specified electronic clinical quality measures.

These stroke mortality measures were developed using a linked dataset consisting of Medicare fee-for-service (FFS) administrative claims data and AHA/ASA Get With The Guidelines® (GWTG)-Stroke registry data abstracted from medical records. Data from this registry were used to develop the measures because at the time of measure development, the GWTG-Stroke registry was the largest database that included both the NIHSS and clinical EHR variables. However, the intent of these measures is to be implemented nationally using NIHSS collected through either claims or the EHRs.

This work also builds upon prior work performed by CORE to establish a set of core clinical data elements (CCDE) that could be feasibly extracted from EHRs and used to risk adjust outcome measures. In 2013, CMS contracted with CORE to identify a set of CCDE that are feasibly extracted from hospital EHRs and are related to patients' clinical status at the start of an inpatient encounter. CCDE capture the first set of vital signs (within 2 hours of the start of the episode of care), and the results of the first complete blood count and basic chemistry panel (within 24 hours). Preliminary work had established that the CCDE could be used to risk adjust measures of 30-day mortality across a variety of common and costly medical conditions.[1,2] Specifically, the development of the hybrid stroke mortality measure builds on this prior work and advances CMS in their goal of moving toward including EHR data in quality measurement programs where feasible, as we examined the utility of the CCDE as candidate clinical risk-adjustment variables to include in this measure.

In alignment with the current, publicly reported stroke mortality measure, the new measures estimate the hospital-level, risk-standardized mortality rate (RSMR) for patients discharged from the hospital with a principal discharge diagnosis of acute ischemic stroke. The outcome in each measure is all-cause 30-day mortality, defined as death from any cause within 30 days of the index admission date, whether in-hospital or not.

2.2 Importance of Stroke Mortality Measures

Stroke is the fourth most common cause of death in the United States, affecting approximately 795,000 people annually, and has a 30-day mortality rate that varies by age from 9% in patients 65 to 74 years of age, 13.1% in those 74 to 84 years of age, and 23% in those ≥85 years of age.[3-5] Stroke is also a leading cause of disability in the United States, which can lead to increased dependency on the health care system and higher subsequent health care related costs. Mortality following stroke is an important adverse outcome that can be measured reliably and objectively and that is influenced by the quality of care provided to patients during their initial hospitalization; therefore, mortality is an appropriate measure of quality of care following stroke hospitalization.[6,7] Specifically, post-stroke mortality rates have been shown to be influenced by critical aspects of care such as response to complications, speediness of delivery of care, organization of care, and appropriate imaging.[8-11] This work demonstrates the relationship between hospital organizational factors and performance on the stroke mortality measure and supports the ability of hospitals to impact these rates.

Measurement of patient mortality allows for a broad view of quality of care that encompasses more than what can be captured by individual process-of-care measures. The goal of outcome measurement is to identify institutions whose performance is better or worse than would be expected based on their patient case mix, by risk adjusting for patients' conditions at the time of hospital admission. The goal of reporting a stroke outcome measure is to improve patient outcomes by providing patients, physicians, and hospitals with information about hospital-level risk-standardized mortality rates (RSMRs) following hospitalization for acute ischemic stroke.

2.3 Rationale for Updated Claims-only and New Hybrid Stroke Mortality Measures

Clinicians and stakeholders, including the AHA/ASA and other professional organizations, continue to express preference for using patient-level clinical data for risk-adjustment over administrative claims data and to highlight the importance of including an assessment of stroke severity in risk-adjustment models of stroke mortality. The current publicly reported stroke mortality measure uses only administrative claims data for risk adjustment and does not include an assessment of stroke severity. Therefore, these new measures seek to satisfy stakeholders' and CMS's preferences by incorporating extractable clinical EHR data and an assessment of stroke severity into the risk adjustment models, while improving discrimination of the stroke mortality risk models.

Several studies have demonstrated that the initial stroke severity score is one of the strongest predictors of mortality in ischemic stroke patients.[12-14] The NIHSS, which was created in 1989 and is widely used in routine stroke care, is collected in the GWTG-Stroke registry, which has over 1,700 hospitals throughout the U.S.[15] The NIHSS is a 15-item neurologic examination stroke scale used to provide a quantitative measure of stroke related neurologic deficit by evaluating the effect of acute ischemic stroke on the levels of consciousness, language, neglect, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss. The NIHSS is designed to be a simple, valid, and reliable tool that can be administered at the bedside consistently by physicians, nurses, or therapists. The use of the NIHSS to measure stroke severity upon acute ischemic stroke patient presentation is Class I recommended in the AHA/ASA guidelines. We are now able to incorporate an assessment of stroke severity into these measures due to a recent increase in the collection and ability to obtain these assessments, which is likely due in part to the AHA/ASA guidelines that recommend using a stroke severity scale – specifically, the NIHSS – on all stroke patients.[16]

The current project is timely, as the NIHSS will soon be included in claims data for ICD-10, thereby providing several options with regard to incorporating this stroke severity assessment; the NIHSS can be extracted either from an EHR or claims data. Also, the recent proliferation of EHR systems, standardization of data extraction for quality reporting, and advancements in clinical practice to incorporate new clinical assessments have made it possible to integrate these data into outcome measures of hospital performance.

There are several potential benefits to incorporating clinical data from EHRs into hospital outcome measures. Utilizing the same variables to calculate hospital performance that are used to support clinical decision-making would be clinically sensible and cost-effective, as they reduce the burden of EHR data mapping and extraction required for quality reporting. In addition, medical record data captured in EHRs are recorded by clinicians who are interacting with the patient and who value the accuracy of the data to guide the care they provide. Therefore, many clinical data elements that are captured in real-time to support patient care are less susceptible to gaming, coding drift, and variations in billing practices compared with administrative data used for billing purposes. Lastly, the incorporation of clinical EHR data advances one of CMS's goals of shifting toward including EHR data in quality measurement reporting programs where feasible. Because the EHR data do not currently include follow-up data for capturing outcomes, it is not practical to create new measures without use of Medicare administrative claims data. Thus, in order to build the best hybrid measure, we considered approaches that link patient-level clinical data to Medicare claims data for risk adjustment, cohort derivation, and outcome determination.

In summary, we aimed to develop two types of measures – one that could be implemented using

claims data only and another that has two options for incorporating EHR and claims data. By updating the original stroke mortality measure and creating a new hybrid measure, CMS now has options with regard to their preferred and most feasible approach to implementation.

3. METHODOLOGY

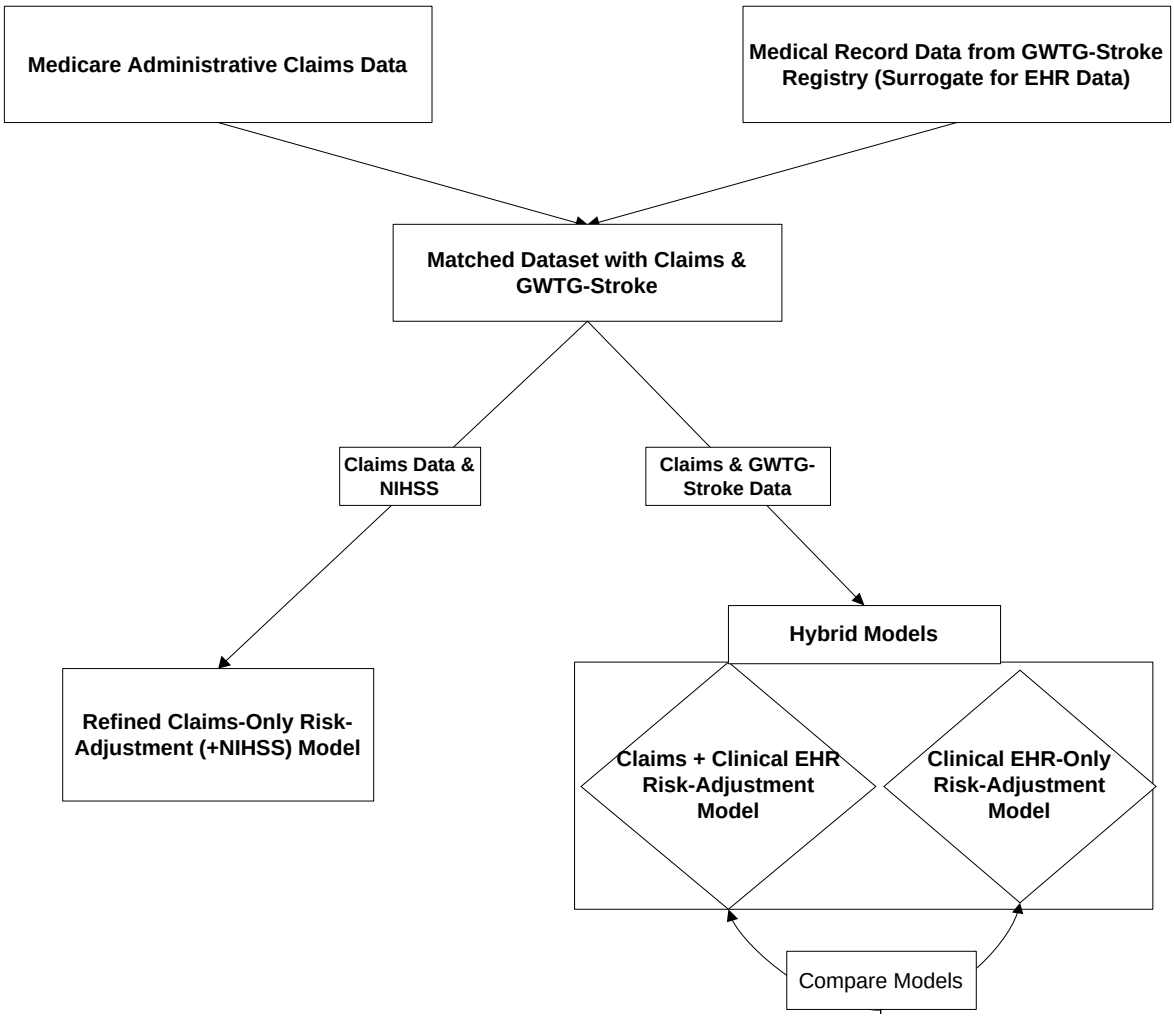
3.1 Overview

This section provides details about the development of the hospital risk-standardized stroke mortality measures, including the identification of a relevant data sources, the cohort definition, variable selection for the risk-adjustment models, and model testing. In developing the measures we followed the standards set forth in the development of prior outcome performance measures, specifically using guidance from the National Quality Forum, the CMS Measures Management System,[17] and the AHA's scientific statement, "Standards for Statistical Models Used for Public Reporting of Health Outcomes."[18]

3.2 Approach to Development

We developed three distinct stroke mortality risk-adjustment models using AHA/ASA's GWTG-Stroke registry data linked with Medicare administrative claims data; one of the models was developed to be implemented as a claims-only measure, and two of the models were designed to be implemented as hybrid measures (Figure 3.2.1). In the hybrid models, the registry data was used during development as a surrogate for EHR data, as there is currently no large national dataset that includes patient-level EHR data related to patients with acute ischemic stroke as well as a measure of stroke severity. We developed two hybrid risk models in order to compare and evaluate model performance, and provide CMS with risk adjustment model options. In the updated claims-only model, registry data was used solely as a source for NIHSS scores for the purposes of measure development. However, NIHSS scores will be included in ICD-10-CM coding system and obtainable from claims beginning October 2016. We reviewed and largely retained the cohort and outcome definitions of the currently publicly reported stroke mortality measure, focusing on updating and improving risk-adjustment strategies.

Figure 3.2.1 Approach to Developing Three Stroke Mortality Risk-Adjustment Models



3.2.1 CORE and AHA/ASA Working Group and Expert Input

Development of the stroke mortality measures involved input from a number of experts, including a working group convened by CORE and the AHA/ASA that consisted of clinical and methodological experts with extensive experience in both performance measure development and stroke. The group included stroke neurologists, members of the AHA/ASA, health sciences researchers, and other professionals with expertise in biostatistics, measure methodology, and quality improvement (Appendix C). The working group provided regular key input on all measure decisions, including cohort derivation, model development, and model testing. Working group meetings were typically held twice per month and addressed key issues to ensure the measures would be meaningful, useful, and well-designed.

3.3 Data Sources

For model development purposes only, we used two data sources: Medicare Administrative claims and the AHA/ASA GWTG-Stroke Registry. Both data sources were linked to create the dataset used for measure development. Registry data was used to obtain the NIHSS, which was included in both

measures, but the NIHSS will soon be available in claims and is available and feasibly extracted from most EHRs.

3.3.1 Medicare Administrative Datasets

Administrative claims data were obtained from Medicare Inpatient/Outpatient Claims Databases, Physician's Carrier Claims Database, as well as Medicare's Enrollment Database (EDB), containing Medicare beneficiary demographic (including age, gender, and birth date), benefit/coverage, and vital status information (such as whether the patient was dead or alive, and date of death).

3.3.2 AHA/ASA GWTG-Stroke Registry Data

Because there is currently no large national dataset that includes patient-level EHR data related to patients with acute ischemic stroke, we used data collected in the AHA/ASA's GWTG-Stroke registry for model development and validation. These registry data were used for model development as a proxy for clinical EHR data, but will not be used for implementation. Hospitals across the United States voluntarily participate in the registry, which includes information on stroke patients collected using the GWTG-Stroke Patient Management Tool™ (PMT). It includes patient characteristics such as age and sex; arrival and admission information; medical history such as atrial fibrillation/flutter, previous stroke/transient ischemic attack (TIA), previous myocardial infarction, diabetes mellitus, hypertension and heart failure; clinical diagnoses; medications prior to admission; measurements such as total cholesterol, triglycerides, high-density lipoprotein (HDL), low-density lipoprotein (LDL), blood glucose, serum creatinine, international normalized ratio (INR), heart rate, blood pressure and weight; and the NIHSS score, a 15-item neurological examination that is used to evaluate the effect of acute ischemic stroke on the levels of sensory loss, dysarthria, ataxia, motor strength, extraocular movement, visual field loss, neglect, language, and consciousness. In order to most accurately risk-adjust, the models include the first NIHSS score captured on patients.

A wide spectrum of hospitals across the country participates in the GWTG-Stroke registry. We compared the characteristics of hospitals that participated in GWTG-Stroke in 2013 with those of hospitals that did not using data from the American Hospital Association Survey. Compared with hospitals that did not participate in GWTG-Stroke, hospitals that did participate were larger (had a greater number of beds), more likely to be teaching hospitals, and less likely to be safety net hospitals. They were also more likely to be not-for-profit rather than government or for-profit hospitals, and to be located in metropolitan rather than rural areas (Table 3.3.1). Hospitals that participated in the GWTG-Stroke registry were not representative of all hospitals in the United States, but the diversity among them generates a valid dataset for measure development.

The AHA employs a number of strategies to ensure that data submitted to GWTG-Stroke are complete, consistent, and accurate. To optimize data quality, the GWTG-Stroke Program includes detailed training of site chart abstractors, standardized case definitions and coding instructions, predefined logic and range checks on data fields at data entry, audit trails, and regular data quality reports for all sites. Source documentation quality audits at the individual state and site levels are performed and have shown high data quality.

Table 3.3.1. Comparison of hospitals that participated and hospitals that did not participate in GWTG-Stroke registry among hospitals with at least one patient with ischemic stroke between July 2011 and June 2014, and hospitals in the 2013 American Heart Association (AHA) Hospital Survey Data.

Description	Hospitals in GWTG-Stroke (N=1,555) %	Hospitals not in GWTG-Stroke (N=2,845) %
Number of beds		
<100	15.8	66.4
100 to 300	47.4	26.7
>300	36.8	6.8
Mean (SD)	284.4 (220.3)	108.0 (148.2)
Ownership		
Government	11.3	28.5
Not-for-profit	73.1	54.4
For-profit	15.6	17.0
Region		
Associated area	0.5	1.5
New England	5.5	3.1
Middle Atlantic	14.2	5.5
South Atlantic	19.6	12.0
East North Central	16.5	15.0
East South Central	4.9	10.4
West North Central	8.0	17.9
West South Central	10.5	15.8
Mountain	6.4	9.1
Pacific	13.9	9.7
Teaching status		
Council of Teaching Hospitals	12.8	2.0
Other teaching	24.8	7.4
Non-teaching	62.4	90.6
Core-based statistical area**		
Division	24.0	8.4
Metro	60.1	32.7
Micro	10.4	23.0
Rural	5.5	36.0
Safety-Net Hospital*		
No	82.1	63.8
Yes	17.9	36.2

* Defined as government hospitals or non-government hospitals with high Medicaid caseload

**Core-based statistical areas are defined on the basis of the population contained within them:
Division: >2.5 million inhabitants
Metro: 50,000 – 2.5 million inhabitants
Micro: 10,000 – 50,000 inhabitants
Rural: <10,000 inhabitants

3.4 Cohort Derivation

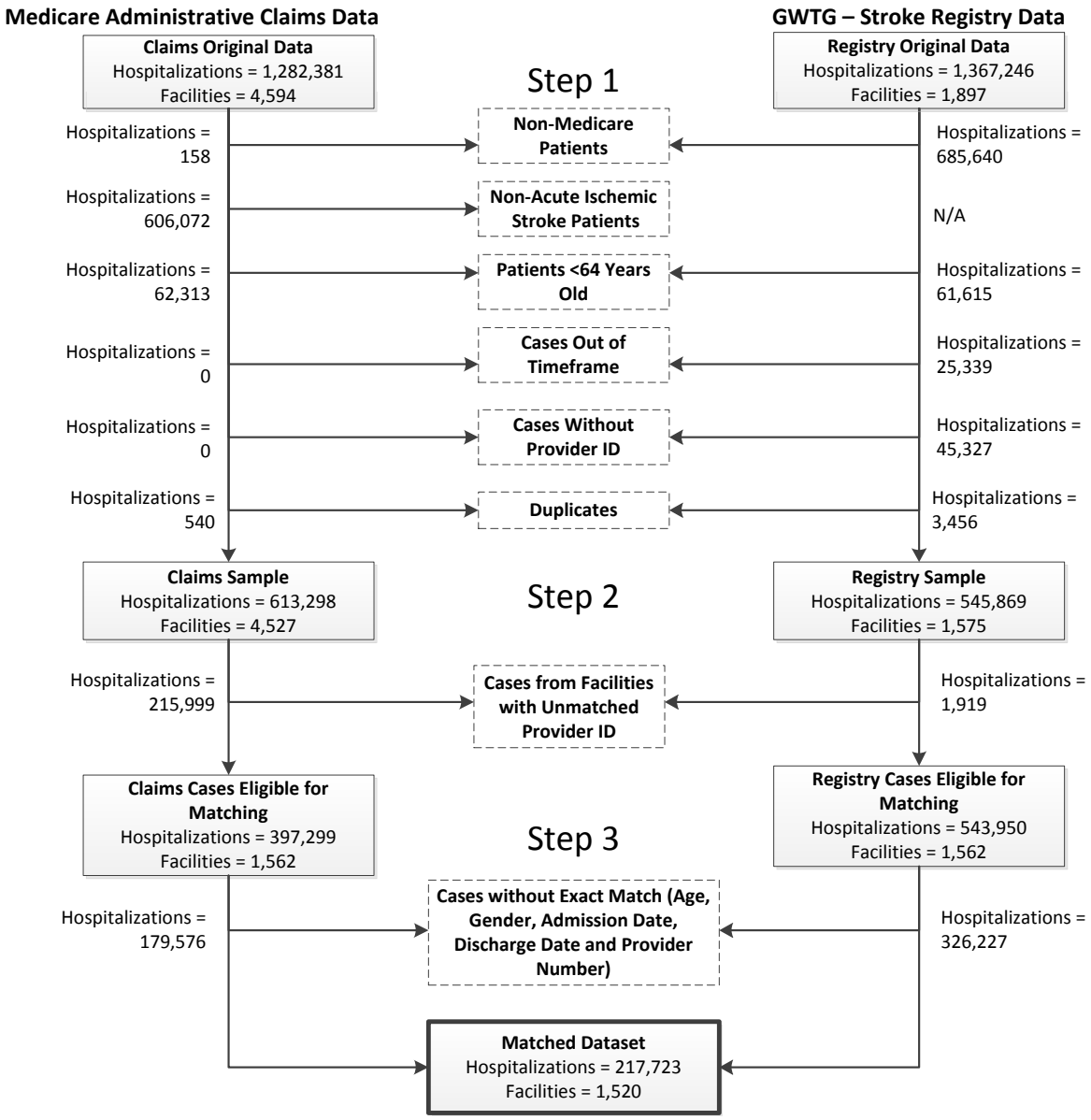
To build the dataset used in development of the models, we used discharges for ischemic stroke included in the GWTG-Stroke and Medicare claims datasets from July 1, 2011 through June 31, 2014.

Both claims and registry data had ischemic stroke patients. In the claims dataset, we identified discharges with ischemic stroke by International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) principal discharge diagnosis codes 433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434.01, 434.11, 434.91, and 436 (see Table A 1 in Appendix A). In the GWTG-Stroke dataset, ischemic stroke was identified clinically.

3.4.1 Deterministic Matching of GWTG-Stroke and Medicare Claims Datasets

In order to obtain a comparable cohort of patients in each dataset (GWTG-Stroke and Medicare claims data) in preparation for deterministic matching, we applied several exclusion criteria and then deterministically matched the remaining hospitalizations using a hospital ID number, patient age (within one year), sex, admission date, and discharge date as the linking fields. Admissions that did not match based on all five linking fields were excluded. Figure 3.4.1 depicts the steps followed to derive the matched Medicare-GWTG-Stroke cohort, followed by a detailed description of each step.

Figure 3.4.1. Deterministic matching to derive cohort for model development



Steps 1 and 2: Preparation of datasets for deterministic matching

In order to obtain a comparable cohort of patients within each dataset (GWTG-Stroke and Medicare claims data) in preparation for deterministic matching, we applied the following exclusion criteria to one or both datasets:

- **Non-Medicare patients (applied only to the GWTG-Stroke dataset)**

  Rationale: Non-Medicare patients in GWTG-Stroke would not be included in the Medicare dataset, and therefore could not be matched.

- **Patients without a principal discharge diagnosis of ischemic stroke**

  Rationale: Ischemic stroke is the condition targeted for measurement. The outcome of these measures is mortality from any cause within 30 days of the index admission date for patients hospitalized with ischemic stroke. In the claims dataset, we identified discharges with ischemic stroke by ICD-9-CM principal discharge diagnosis codes 433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434.01, 434.11, 434.91, and 436 (see Table A 1 in Appendix A). In the GWTG-Stroke dataset, ischemic stroke was identified clinically.

- **Age <64 years (Medicare claims and GWTG-Stroke data)**

  Rationale: Admissions for patients aged <64 years at the time of admission were excluded, as these patients may be in the Medicare dataset but are not targeted for measurement.

- **Cases out of timeframe**

  Rationale: Cases in the GWTG-Stroke dataset with admission date outside of the timeframe from July 1, 2011 through June 31, 2014 were excluded because these cases would not be included in the Medicare dataset, and therefore could not be matched.

- **Index admissions missing a provider ID**

  Rationale: Provider ID was necessary in order to match datasets.

- **Duplicate admissions within each separate dataset (Medicare claims and GWTG-Stroke data)**

  Rationale: Admissions for patients who have identical information in a single dataset indicated for age, sex, admission date, discharge date, and MPN are excluded to avoid making matching errors upon merging of the two datasets.

- **Cases from facilities with unmatched provider ID**

  Rationale: Cases from facilities with unmatched provider IDs are not included, as provider IDs must be matched in order to perform hospital-level analyses.

Step 3: Deterministic match of GWTG-Stroke and Medicare claims datasets

The remaining hospitalizations in both datasets were then matched using a hospital ID number, patient age (within one year), sex, admission date, and discharge date as the linking fields. Admissions that did not match based on all five linking fields were excluded.

Among admissions eligible for matching within the Medicare claims dataset, 55% were successfully matched to GWTG-Stroke data; Table 3.4.1 compares matched and unmatched admissions. Admissions in the claims dataset that matched to the GWTG-stroke data as compared with patients in the unmatched cohort were older; slightly less likely to be male; less likely to have cerebral hemorrhage,

previous ischemic or unspecified stroke, precerebral arterial occlusion and transient cerebral ischemia, cerebral atherosclerosis and aneurysm, and hemiplegia/hemiparesis; and more likely to have previous valvular and rheumatic heart disease, congenital cardiac/circulatory defects, and heart arrhythmias (all $p<0.05$). Possible explanations for mismatch include differences in selection criteria for the two databases; miscoding of principal discharge diagnoses in the Medicare data; failure to include an eligible patient in GWTG-Stroke; the use of sampling, as some GWTG-Stroke hospitals use sampling techniques consistent with The Joint Commission/CMS standards for sampling;[19] inaccuracies within the CMS or GWTG-Stroke data for linking fields (e.g., substituting age for date of birth); and other data entry errors.

Among admissions eligible for matching in the GWTG-Stroke registry, 40% were successfully matched to Medicare claims data. As shown in Table 3.4.2, the observed characteristics of admissions in GWTG-stroke that matched to claims admissions were similar to admissions that did not match, including similar age, cholesterol, and medical history. Possible explanations for the failure of 60% of the admissions to match include differences in selection criteria for the two databases, miscoding of principal discharge diagnoses in the Medicare data, admissions for patients ineligible for Medicare (e.g., non-U.S. citizens), admissions for patients in Medicare Advantage (not in Medicare FFS) or with non-governmental insurance, inclusion of non-ischemic stroke patients in the GWTG-Stroke dataset, inaccuracies within the CMS or GWTG-Stroke data for linking fields (e.g., substituting age for date of birth), and other data entry errors.

Table 3.4.1. Selected patient characteristics and outcomes in Medicare claims data for patients who matched and unmatched to GWTG-Stroke data

Description	Matched (N=217,723) %	Unmatched (N=179,576) %
Transfer from another ED	10.27	9.63
Demographic		
Age (continuous): Mean (SD)	79.47 (8.55)	78.98 (8.65)
Male	43.39	43.75
Cardiovascular/Cerebrovascular		
Congestive heart failure	23.88	23.69
Valvular and rheumatic heart disease	26.04	24.81
Congenital cardiac/circulatory defects	2.86	2.62
Hypertensive heart disease	5.01	5.11
Specified heart arrhythmias	30.70	29.44
Cerebral hemorrhage	2.25	2.41
Ischemic or unspecified stroke	23.40	27.79
Precerebral arterial occlusion and transient cerebral ischemia	22.52	24.93
Cerebral atherosclerosis and aneurysm	12.22	13.02
Hemiplegia/hemiparesis	6.27	7.04
Comorbidities		
History of infection	27.24	25.86

Description	Matched (N=217,723) %	Unmatched (N=179,576) %
Metastatic cancer and acute leukemia and other major cancers	3.93	3.98
Lymphatic, head and neck, brain, breast, colorectal and other major cancers	23.99	23.23
Protein-calorie malnutrition	6.54	6.83
Other significant endocrine and metabolic disorders	87.44	86.70
Other gastrointestinal disorders	48.92	49.35
Disorders of the vertebrae and spinal discs	19.48	19.47
Osteoarthritis of hip or knee	11.28	11.13
Other musculoskeletal and connective tissue disorders	68.01	67.66
Iron deficiency and other/unspecified anemia and blood disease	36.73	37.39
Dementia or senility	29.92	30.72
Major psychiatric disorders	9.86	10.40
Quadriplegia, other extensive paralysis	1.49	1.61
Multiple sclerosis	13.01	13.22
Seizure disorders and convulsions	7.61	8.29
Hypertension	92.05	92.07
Peripheral vascular disease	24.10	23.78
Chronic obstructive pulmonary disease	21.44	22.13
Pneumonia	15.22	15.71
Pleural effusion/pneumothorax	7.50	7.39
Other eye disorders	19.69	19.37
Other ear, nose, throat, and mouth disorders	27.40	27.61
Dialysis status	1.63	1.79
Renal failure	20.42	20.90
Urinary tract infection	20.42	20.75
Male genital disorders	14.44	14.25
Decubitus ulcer of skin	2.57	2.61
Chronic ulcer of skin, except decubitus	5.28	5.10
Other dermatological disorders	30.96	29.81
Outcomes		
In-hospital mortality	5.12	5.27
30-day mortality	14.13	14.58
90-day mortality	19.46	20.02

Table 3.4.2. Selected patient characteristics and outcomes in GWTG-Stroke Registry data for patients who matched and unmatched to CMS data

Description	Matched (N=217,723) %	Unmatched (N=326,227) %
Demographics		
Age (continuous): Mean (SD)	79.47 (8.55)	78.61 (8.24)
Female	56.61	55.81
Race - White	83.47	81.82
Race - Black	10.95	11.30
Race - Other	5.58	6.88
Medical history		
Atrial Fib/Flutter	26.11	21.53
Prosthetic Heart Valve	1.48	1.61
CAD/prior MI	29.96	29.23
Carotid Stenosis	4.29	6.69
Diabetes Mellitus	31.63	31.28
PVD	5.58	5.75
Hypertension	80.28	77.70
Smoker	10.01	8.76
Dyslipidemia	48.08	48.81
HF	11.23	9.74
Sickle Cell	0.02	0.02
Previous Stroke	27.06	24.82
Previous TIA	10.56	12.50
Drugs/Alcohol Abuse	1.46	1.29
Family History of Stroke	3.59	3.36
HRT	0.26	0.26
Migraine	0.64	0.71
Obesity/Overweight	6.55	7.09
Renal insufficiency - chronic	4.48	4.51
Sleep Apnea	0.29	0.32
Depression	1.07	1.20
Diagnosis and Evaluation		
Stroke symptoms resolved at time of presentation	7.37	16.14
First NIHSS score: Mean (SD)	7.58 (8.14)	6.08 (8.13)
Measurements: Mean (SD)		
Total Cholesterol	164.83 (52.36)	163.5 (255.48)
Triglycerides	123.87 (93.81)	123.73 (315.15)
HDL	45.91 (16.90)	46.78 (17.67)
LDL	95.25 (62.71)	92.98 (75.04)

Description	Matched (N=217,723) %	Unmatched (N=326,227) %
A1C	7.52 (24.53)	7.43 (23.53)
Blood glucose (mg/dL)	136.61 (60.94)	136.70 (65.36)
Serum Creatinine	1.58 (7.52)	1.58 (8.33)
Initial Platelet Count at Hospital Arrival	251.03 (80.35)	234.22 (108.33)
INR	1.18 (0.70)	1.26 (1.09)
Vital Signs - Heart Rate	79.62 (18.29)	79.07 (210.60)
Vital Signs - Blood Pressure Systolic	158.16 (69.17)	156.86 (30.79)
Vital Signs - Blood Pressure Diastolic	80.84 (17.89)	79.79 (18.06)
Height (cm)	167.14 (18.97)	166.83 (14.73)
Weight (kg)	76.31 (47.73)	76.48 (49.29)
Waist Circumference (cm)	102.48 (46.74)	99.21 (45.08)
BMI	27.72 (113.55)	27.69 (28.00)
In-Hospital Death	5.12	7.64

Inclusion and Exclusion Criteria Applied to Matched Dataset

After performing the deterministic match, as shown in Figure 3.4.1, we applied inclusion and exclusion criteria to derive the final cohort of patients for building the risk-adjustment models. These criteria are very similar to those in the currently publicly reported claims-based stroke mortality measure[20]. Figure 3.4.2 illustrates the steps followed to apply the inclusion and exclusion criteria to derive the final study cohort.

In addition to the principal discharge diagnosis of ischemic stroke and Medicare FFS enrollment criteria applied when matching the Medicare claims and GWTG-Stroke data, we also included index hospital admissions for patients who:

1. **Are aged 65 years or older**

   Rationale: Medicare patients younger than 65 usually qualify for the program due to severe disability. They are not included in the measure because they are considered to be too clinically distinct from Medicare patients 65 and over. The characteristics and outcomes of these patients may not be representative of the larger population of stroke patients.

2. **Were not transferred after being admitted to another acute care facility**

   Rationale: Death is attributed to the hospital where the patient was initially admitted. Transferred patients are still included in the measure cohort, but the "transfer-in" hospitalization is excluded as an index admission.

3. **Were enrolled in Part A and Part B Medicare for the 12 months prior to the date of admission, and enrolled in Part A during the index admission***

   Rationale: The 12-month prior enrollment criterion ensures that patients are Medicare FFS beneficiaries and that their comorbidities are captured from claims for risk adjustment. Medicare Part A is required during the index admission to ensure that no Medicare Advantage patients are included in the measures. ***This inclusion criterion could be removed for

implementation of a measure that uses only EHR data for risk adjustment.

The measure excludes index hospital admissions for patients who (Figure 3.4.2):

1. **Have inconsistent or unknown vital status or other unreliable data**

   Rationale: We do not include stays for patients where the age is greater than 115, where the gender is neither male nor female, where the admission date is after the date of death in the Medicare Enrollment Database, or where the date of death occurs before the date of discharge but the patient was discharged alive.

2. **Enrolled in the Medicare hospice program any time in the 12 months prior to the index admission, including the first day of the index admission**

   Rationale: These patients are likely continuing to seek comfort measures only; thus, mortality is not necessarily an adverse outcome or signal of poor quality care for these patients.

3. **Discharged against medical advice (AMA)**

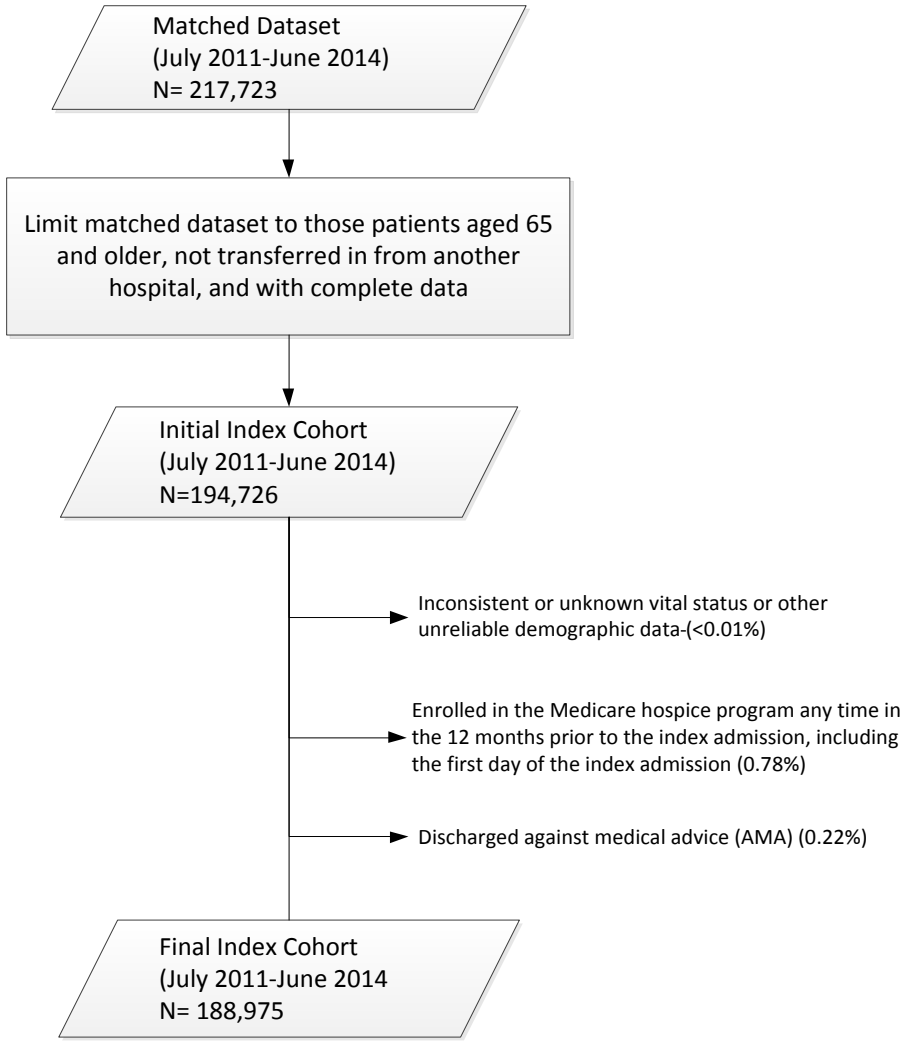
   Rationale: Providers did not have the opportunity to deliver full care and prepare the patient for discharge.

For patients with more than one admission for stroke in a given year between July of the current year to June of the following year, only one index admission is randomly selected for inclusion in the cohort.

As a part of claims data processing prior to the measure calculation, records are removed for non-short-term acute care facilities such as psychiatric facilities, rehabilitation facilities, or long-term care hospitals. Additional data cleaning steps include removing claims with stays longer than one year, claims with overlapping dates, and stays for patients not listed in the Medicare enrollment database as well as records for providers with invalid provider IDs.

Finally, for index admissions that occur during the transition between measure reporting periods, June and July of each year, the measures include admissions only if they were the first to occur in the 30 days prior to a patient's death; additional admissions in that 30-day period are excluded. This exclusion criterion is applied after one admission per patient per year is randomly selected to avoid assigning a single death to two admissions in two separate reporting periods. For example, consider a patient who is admitted on June 18, 2012, readmitted on July 2, 2012, and subsequently dies on July 15, 2012. If both admissions are randomly selected for inclusion (one for the July 2011-June 2012 time period and the other for the July 2012-June 2013 time period), the measure will exclude the July 2, 2012 admission to avoid assigning the death to two admissions.

Figure 3.4.2: Exclusions applied to the July 2011-June 2014 matched dataset



3.4.2 Transfers between Hospitals

The stroke mortality measures use the same methodology as the original stroke mortality measure to define transfers and to attribute mortality outcomes. For patients whose index admission includes one or more transfers between hospitals, the mortality outcome is attributed to the hospital where the patient was first admitted for stroke. For patients seen in the emergency department of one hospital and who are then admitted to another hospital, the measure assigns them to the admitting hospital.

3.4.3 Development and Validation Samples

In order to develop and test the stroke mortality measures, we randomly split the final index cohort (N=188,975), as shown in Figure 3.4.2, into two samples. The first sample – the development sample (N=94,466) – was used to develop the risk-adjusted models, and the second sample – the validation

sample (N=94,509) – was used to validate the models.

3.5 Outcome Assessment

The approach to assessment of the mortality outcome is identical to the original stroke mortality measure methodology. The outcome is 30-day all-cause mortality, defined as death from any cause within 30 days of the index admission date. We identify deaths for Medicare FFS patients in the Medicare Enrollment Database (EDB).

3.5.1 All-Cause Mortality

There are a number of reasons for counting all deaths in this measure. First, from a patient perspective, death from any cause is an adverse event. In addition, making inferences about quality issues and accountability based solely on the documented cause of death is difficult. For example, a patient with stroke who develops a hospital-acquired infection may ultimately die of sepsis and multi-organ failure. In this context, considering the patient's death to be unrelated to the care the patient received for stroke during the index admission would be inappropriate.

3.5.2 30-Day Time Period

The measures assess mortality within a 30-day period from the date of the index admission. This standard time period is necessary so that the outcome for each patient is measured uniformly. The measures use a 30-day time frame because outcomes occurring within 30 days of admission can be influenced by hospital care and the early transition to the outpatient setting. The use of the 30-day time frame is a clinically meaningful period for hospitals to collaborate with their communities in an effort to reduce mortality.[21]

3.6 Approach to Risk Adjustment

For the current project, we aimed to identify risk factors for the models that are clinically relevant, have strong relationships with the mortality outcome, and any clinical variables that can be feasibly extracted from the EHR (for the hybrid models). For the claims-only measure, risk adjustment variables were obtained from inpatient, outpatient, and physician carrier Medicare claims data extending 12 months prior to, and including, the index admission. For the hybrid model that includes clinical EHR data-only, risk–adjustment variables were obtained from the GWTG-Stroke registry (for development only). Risk-adjustment variables for the hybrid model with claims and clinical EHR data were obtained from both claims data and the GWTG-Stroke registry (for development only).

The measures align with other CMS hospital-level outcome measures, which seek to adjust for case mix differences among hospitals based on the clinical status of the patient at the time of the index admission. Accordingly, only comorbidities that convey information about the patient at the time of admission or in the 12 months prior, and not complications that arise during the course of the index hospitalization, are considered as risk-variables.

3.6.1 Candidate Risk-Adjustment Variables

We sought to develop three separate risk-adjustment models that included the NIHSS, other key variables that are clinically relevant that demonstrate a strong statistical association with 30-day mortality, and relevant clinical variables that can be feasibly extracted from the EHR (for the hybrid models). Specifically, we developed:

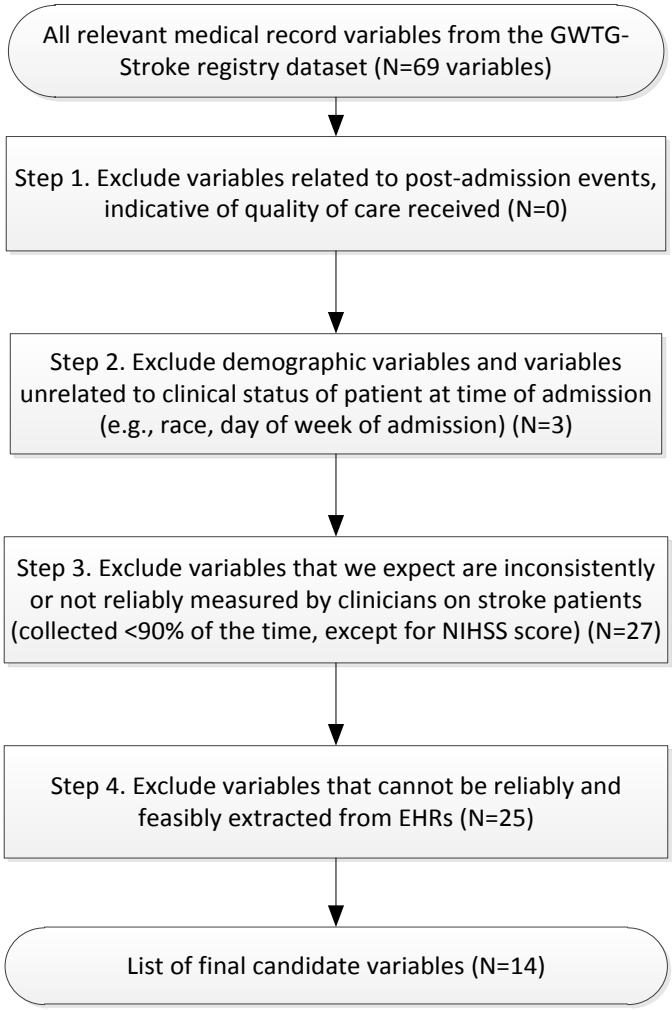
- A claims-only model that included the NIHSS and only variables available from claims data for risk adjustment, which would include patient medical history;
- One hybrid model that included the NIHSS and variables from both claims (patient medical history) and clinical EHR data (vitals and basic laboratory results) for risk adjustment; and
- One hybrid model that included only clinical EHR variables (vitals and basic laboratory results) for risk adjustment.

To select candidate variables from claims, we considered those 41 risk-adjustment variables in the currently publicly reported claims-based measure as candidate variables, plus the NIHSS (Table B 1). To select candidate clinical EHR data, the members of the working group reviewed the entire list of variables available in the GWTG-Stroke registry database. (The complete set of variables collected with the PMT can be found at: http://www.heart.org/idc/groups/heart-public/@wcm/@gwtg/documents/downloadable/ucm_474783.pdf). To identify clinically meaningful variables to review in the candidate variable selection process, we excluded variables not suitable for use in risk adjustment (e.g., patient name, physician name). We systematically applied a series of exclusion criteria to the remaining clinical variables (Figure 3.6.1) to obtain a list of candidate variables for building the hybrid models. Refer to Figure 3.6.1 for the clinical variable selection strategy, and Table B 2 in Appendix B for a list of variables and rationale for exclusion.

For the hybrid model based on both claims and clinical EHR data, we included as candidate variables those 41 claims-based variables plus 15 candidate clinical EHR variables identified after systematically applying the above referenced exclusion criteria, which included the NIHSS (Table B 3). For the hybrid measure based on clinical EHR data-only for risk adjustment, we only included as candidate variables the 15 clinical variables, which included the NIHSS (Table B 4).

Although EHRs will likely evolve to link across clinical episodes of care and to contain accurate historical patient data in standardized fields, given the EHR environment at the time of measure development, and the inability to reliably obtain data from the outpatient setting prior to admission, we only variables that would be available and consistently collected at admission were considered for inclusion in the hybrid models. Determination was based on prior discussions with the CCDE Technical Expert Panel (TEP) convened in 2013,[1] which examined multiple data elements to determine feasibility of capture. Data element feasibility will be reevaluated As EHRs advance.

Figure 3.6.1: Clinical EHR candidate variable selection process for hybrid models



Step 1: Exclusion of variables related to events, indicative of quality of care received

The purpose of risk adjustment is to take a patient's health status upon arrival at the hospital into consideration so we sought to exclude all variables that could occur after clinical care has been administered, such as complications. This did not result in the exclusion of any variables in this dataset.

Step 2: Exclusion of demographic variables and variables unrelated to the clinical status of the patient at the time of admission

Next, we excluded demographic variables to align with CMS policy regarding risk adjustment for sociodemographic variables at the time of measure development. This resulted in the exclusion of 2 demographic variables. An additional variable was excluded at this step due to its lack of relatedness to the clinical status of the patient at the time of admission, for a total of 3 excluded variables.

Step 3: Exclusion of variables expected to be inconsistently or not reliably measured by clinicians on stroke patients

We excluded clinical EHR variables that were collected on fewer than 90 percent of patients admitted to the hospital for ischemic stroke, such as height and waist circumference because patients are typically supine during the admission. Similarly, we excluded variables that are not reliably measured across physicians because they lack a standard definition, method of collection, and units of measurement, such as symptom duration. This resulted in the exclusion of an additional 27 variables.

Based on prior discussions with the CCDE Technical Expert Panel (TEP) in 2013 and examination of clinical data captured at that time, we determined that the EHR variable "transfer from" was not captured reliably or in a standard manner in EHRs (still obtained through claims).[1] Therefore, we excluded the EHR variable "mode of hospital arrival" from the final list of clinical EHR candidate variables at this step. Nonetheless, because the variable "mode of hospital arrival" was deemed important due to its association with mortality in stroke patients, and because it will likely be feasibly extracted from EHRs in the near future, we tested this variable in both hybrid models to test the potential incremental benefit of including it. See Appendix C for results of the two hybrid models that include the variable "mode of hospital arrival": Table C 1 has results from the hybrid model with claims and clinical EHR data, and Table C 2 has results from the hybrid model that used clinical EHR data only. These analyses were performed for information only.

Step 4: Exclusion of variables that cannot be feasibly extracted from EHRs

Finally, we excluded clinical EHR variables that cannot be feasibly extracted from an EHR. Feasible extraction relies on data to be captured in numerical, pseudo-numerical, or list format and entered in structured fields rather than free-text. This resulted in the exclusion of 25 variables.

After these four steps were applied, 14 clinical EHR variables remained as candidates for inclusion in the hybrid measures.

3.6.2 Approach to Defining Continuous Clinical EHR Variables

We examined distributions of the 14 continuous candidate clinical EHR variables. We identified variables outside of a plausible range based on clinical suggestions from working group members and CORE physicians, and deemed set these variables as missing. Next, we performed multiple imputation by creating five copies of the dataset, rounding variable values where appropriate (e.g., NISSS scores should be integer numbers). Associations between continuous clinical EHR variables and mortality were visually inspected with bivariate plots to assess predictive relationships. Then, to reduce the effect of spurious outliers, we Winsorized[22,23] variable values to specific ranges based on the examination of the distributions of the variables, the predictive relationship between the clinical EHR variable and mortality, clinical suggestions, and prior experience. See Appendix D for the approach to defining continuous clinical EHR variables. After Winsorization of the continuous variables, with the pre-selected candidate variables and the outcome of 30-day mortality, we performed a bootstrap simulation with 1,000 iterations by allowing patients to be selected repeatedly, a process which allowed us to select model variables to be included in the final model. Based on visual inspection of the continuous candidate clinical EHR variable distributions, we included both the linear and quadratic terms of the variable when appropriate for bootstrapping. Following bootstrap simulation, if the quadratic version of the variable reached the 90% selected rate, we included both the linear and quadratic versions in the model.

3.6.3 Handling of missing data for measure development

Because all three measures use patient information for risk-adjustment in the statistical models,

specifically with the NIHSS as an assessment of stroke severity, and hospitals do not always collect and record this key risk variable on all patients, we must address missing data. For the claims-only model, only the NIHSS has missing values. Because only final action claims are included in the dataset and records with unreliable data are excluded before CORE receives the data from CMS, all other variables in the claims dataset are complete. The missing NIHSS values were imputed using the standard statistical method of multiple implantation based on the claims data, and full conditional specification (FCS) with a multi-logit regression model was used for the imputation. For the hybrid models with clinical EHR data included for risk adjustment, many candidate clinical EHR variables (e.g. lab values) have occasional missing values, including the NIHSS. Therefore, FCS was used with a linear regression model for the lab values, and a multi-logit regression model for the NIHSS in the multiple imputation techniques were used to impute the missing values. Five copies of imputation datasets were produced for the analyses. The results based on these data were aggregated according to the standard statistical methods for presentation of the results and for the measure score calculation.

In multiple imputation, missing variable values are predicted using available related patient variables. The predicted values are substituted for the missing values, which results in a full dataset without any missing variables (the imputed dataset). By repeating this process multiple times, we get multiple imputed datasets upon which we conduct analyses and for which we obtain results. The results based on multiple datasets are combined to produce the overall final results. In general, multiple imputation is used to preserve the important characteristics of the underlying dataset and the inherent relationships among the variables in the dataset. This approach allows us to make use of all possible available information to generate a range of plausible values to use in place of the missing values. The multiple imputation represents a random sample of the missing values according to the association of the non-missing values of all the variables considered, and the resulting inferences of multiple imputation are statistically valid, which reflect uncertainty due to missing values. [24,25]

3.7 Model Specification and Validation

For model development we used logistic regression models, with outcome Y_i for the i^{th} patient equal to 1 if the patient died within 30 days of admission and 0 otherwise. We developed separate logistic regression models of mortality using the three separate risk-adjustment strategies and the original stroke mortality measure approach. To develop the model based on the claims data only, we used the previous claims model variables and NIHSS as the candidate predictors for the 30-day mortality and selected the best model using the logistic regression model with the stepwise selection method based on 1,000 bootstrapping samples for each copy of the multiple imputed (MI) data. Variable selection rate for all the variables selected into the best model was calculated for each copy of the MI data, and variables were included into the final model if the minimum variable selection rate among the 5 copies of MI was 90% or more. To develop the hybrid measures that used only the clinical EHR data or both the claims and clinical EHR data for risk adjustment, we used the same approach with different lists of variables as the candidate predictors.

After identifying the appropriate models using the logistic regression model above, we estimated the hospital-specific RSMRs using hierarchical generalized linear models (*Hierarchical model*) in each copy of the imputed data. This strategy accounts for within-hospital correlation of the observed outcome and accommodates the assumption that underlying differences in quality across hospitals lead to systematic differences in outcomes. We model the probability of mortality as a function of patient age and clinically relevant comorbidities with an intercept for the hospital-specific random effect.

We used the following strategy to calculate the hospital-specific RSMRs in each copy of the imputed data, which we calculated as the ratio of a hospital's "predicted" mortality to "expected" mortality multiplied by the national observed mortality rate. The expected mortality for each hospital was estimated using its patient mix and the average hospital-specific intercept (i.e., the average intercept among all hospitals in the sample). The predicted mortality for each hospital was estimated given the same patient mix but an estimated hospital-specific intercept. Operationally, the expected mortality for each hospital was obtained by summing the expected probabilities of mortality for all patients in the hospital. The expected probability of mortality for each patient was calculated via the hierarchical model, which applies the estimated regression coefficients to the observed patient characteristics and adds the average of the hospital-specific intercept. The predicted mortality for each hospital was calculated by summing the predicted probabilities for all patients in the hospital. The predicted probability for each patient was calculated through the hierarchical model, which applies the estimated regression coefficients to the patient characteristics observed and adds the hospital-specific intercept.

More specifically, we used a hierarchical logistic regression model to account for the natural clustering of observations within hospitals. The model employs a logit link function to link the risk factors to the outcome with a hospital-specific random effect:

Let Y_{ij} denote the outcome (equal to 1 if patient *i* dies within 30 days, zero otherwise) for patient i at hospital *j*; $\mathbf{Z_{ij}}$ denotes a set of risk factors. We assume the outcome is related linearly to the covariates via a logit function with dispersion:

$$\text{logit}(\text{Prob}(Y_{ij} = 1)) = \alpha_j + \boldsymbol{\beta}^{*}\mathbf{Z_{ij}} + \varepsilon_{ii} \qquad (1)$$

$$\alpha_j = \mu + \omega_j \; ; \; \omega_j \sim N(0,\tau^2)$$

where Y_{ij} denotes the outcome (equal to 1 if patient *i* dies within 30 days, zero otherwise) for patient i at hospital *j*; $\mathbf{Z_{ij}} = (Z_1, Z_2, \ldots Z_k)$ is a set of *k* patient-level covariates; α_j represents the hospital-specific intercept; μ is the adjusted average hospital intercept over all hospitals; τ^2 is the between-hospital variance component; and $\varepsilon \sim N(0,\sigma^2)$ captures any over- or under-dispersion. This model separates within-hospital variation from between-hospital variation. The hierarchical logistic regression model was estimated using the SAS software system (GLIMMIX procedure).

With the hospital specific RSMRs in all copies of the imputed data, we take the average of these RSMRs of each hospital to get the final hospital specific RSMR as the measure score.

3.7.1 Hospital Performance Assessment

We used the results of each hierarchical logistic regression model to calculate the predicted number of deaths and the expected number of deaths at each hospital. The predicted number of mortalities was calculated using the hierarchical logistic regression model, as the sum of the predicted probability of mortality for each patient, including the hospital-specific (random) effect. The expected number of mortalities for each hospital was similarly calculated as the sum of the predicted probability of mortality for each patient, ignoring the hospital specific (random) effect.

3.7.2 Model Performance Assessment

Assessment of the stroke mortality measures' performance included model calibration (to assess over-fitting), discrimination in terms of predictive ability (the range of observed mortality rates across

deciles of predicted rates), and distribution of model residuals. These analyses were done in the development and validation samples.

3.8 Measure Reliability

To determine the extent to which the assessments of a hospital using different but randomly selected subsets of patients produces similar measures of hospital performance, we calculated separately the RSMRs from each of the three models based on the development and validation cohorts. Thus, we obtain two RSMRs per hospital per model, using an entirely distinct set of patients from the same time period. To the extent that the calculated measures of these two subsets agree, we have evidence that the measure is assessing an attribute of the hospital, not of the patients. As a metric of agreement, we calculated the intra-class correlation as defined by ICC (2,1) by Shrout and Fleiss (1979).[26,27] For the hospital event rates based on patient binomial outcomes like mortality (Yes/No), an ICC value of 0-0.2 indicates poor agreement; 0.3-0.4 indicates fair agreement; 0.5-0.6 indicates moderate agreement; 0.7-0.8 indicates strong agreement; and >0.8 indicates almost perfect agreement.[26]

3.9 Comparison of Models

We developed two types of stroke mortality measures (a claims-only measure and a hybrid measure) which utilize three risk-adjustment models, and have shown a comparison of each of the three model's results, which all perform well, to inform CMS in their future decision-making. The working group also discussed implementation challenges, potential burden on hospitals, and responsiveness to stakeholder preference with CMS.

4. RESULTS

4.1 Cohort – All Models

The inclusion and exclusion criteria that were applied to the matched dataset are presented in Section 3.4; specifically, Figure 3.4.2 displays the percentage of patients meeting each exclusion criterion in the three-year dataset (July 2011 to June 2014). The final index cohort consisted of 188,975 hospital admissions at 1,473 hospitals; the development and validation samples consisted of 94,466 and 94,509 hospital admissions, respectively.

4.2 Outcome – All Models

4.2.1 Assessment of the 30-Day All-Cause Mortality Outcome

We created three separate risk-adjustment models, all of which assess 30-day all-cause mortality as the outcome. They were all developed in a 50% sample of the full 2011-2014 dataset and validated using the remaining 50% of the dataset. The crude mortality rates in the final index cohort, the development sample, and the validation sample were 14.43%, 14.28%, and 14.58%, respectively. Therefore, the development sample had a slightly lower rate than the validation sample.

4.2.2 Distribution of 30-Day Mortality Rate

The hospital unadjusted 30-day mortality rate in July 2011-June 2014 data for the final index cohort ranged from 0.00% to 100.00% across 1,511 hospitals with a median (interquartile range) of 14.40% (11.93%, 16.48%).

In the following sections, we provide separate results for each of these models.

4.3 Model with Claims-Only Risk Adjustment

Following the bootstrapping simulation method for variable selection, those candidate claims risk-adjustment variables that were included more than 90% of the time for all the copies of the imputed data were retained in the final model. The final model included 19 claims-based risk-adjustment variables and the NIHSS, listed in Table 4.3.1 below.

Table 4.3.1. Final claims-only stroke mortality model variables

Category	ICD-9/CC	Description
Demographic	N/A	Age-65 (continuous, per 5 years)
Arrival Information	N/A	Transfer from another ED
Evaluation	N/A	NIHSS score (continuous, per 5 units)
Cardiovascular/ Cerebrovascular	CC 80	Congestive heart failure
Cardiovascular/ Cerebrovascular	CC 87-88	Congenital cardiac/circulatory defects
Cardiovascular/ Cerebrovascular	CC 92	Specified heart arrhythmias
Cardiovascular/ Cerebrovascular	CC 98	Cerebral atherosclerosis and aneurysm
Comorbidities	CC 7-8	Metastatic cancer and acute leukemia and other major cancers

Category	ICD-9/CC	Description
Comorbidities	CC 21	Protein-calorie malnutrition
Comorbidities	CC 22-24	Other significant endocrine and metabolic disorders
Comorbidities	CC 36	Other gastrointestinal disorders
Comorbidities	CC 39	Disorders of the vertebrae and spinal discs
Comorbidities	CC 40	Osteoarthritis of hip or knee
Comorbidities	CC 43	Other musculoskeletal and connective tissue disorders
Comorbidities	CC 47	Iron deficiency and other/unspecified anemia and blood disease
Comorbidities	CC 49-50	Dementia or other specified brain disorders
Comorbidities	CC 72, 76	Multiple sclerosis
Comorbidities	CC 74	Seizure disorders and convulsions
Comorbidities	CC 111-113	Pneumonia
Comorbidities	CC 131	Renal failure

4.3.1 Logistic Regression

The final logistic regression model performed very well, with a mean C-statistic of 0.812 and an adjusted R-squared of 0.268. The variable descriptions, estimates, and standard errors for the logistic regression model using the final model variables are shown in Table 4.3.2 below.

Table 4.3.2. Final model: logistic regression results for claims-only model (N=94,466 patients in development cohort)

Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-3.7613	0.0437	-86.11	-	-
Age-65 (continuous, per 5 years)	0.2899	0.0071	40.72	1.34	1.32, 1.36
Transfer from another ED	0.2993	0.0328	9.13	1.35	1.26, 1.44
NIHSS (continuous, per 5 units)	0.4637	0.0068	68.03	1.59	1.57, 1.61
Congestive heart failure	0.2115	0.0263	8.03	1.24	1.17, 1.3
Congenital cardiac/circulatory defects	-0.4088	0.0783	-5.22	0.66	0.57, 0.77
Specified heart arrhythmias	0.3011	0.0232	12.99	1.35	1.29, 1.41
Cerebral atherosclerosis and aneurysm	-0.2115	0.0340	-6.22	0.81	0.76, 0.87
Metastatic cancer and acute leukemia and other major cancers	1.0597	0.0471	22.51	2.89	2.63, 3.16
Protein-calorie malnutrition	0.4771	0.0349	13.66	1.61	1.5, 1.73
Other significant endocrine and metabolic disorders	-0.3601	0.0311	-11.59	0.70	0.66, 0.74
Other gastrointestinal disorders	-0.1038	0.0225	-4.61	0.90	0.86, 0.94
Disorders of the vertebrae and spinal discs	-0.1271	0.0289	-4.39	0.88	0.83, 0.93
Osteoarthritis of hip or knee	-0.1407	0.0347	-4.05	0.87	0.81, 0.93
Other musculoskeletal and connective tissue disorders	-0.1092	0.0246	-4.43	0.90	0.85, 0.94
Iron deficiency and other/unspecified anemia and blood disease	0.1767	0.0236	7.49	1.19	1.14, 1.25

Description	Estimate	Standard Error	T	OR	95% CI
Dementia or other specified brain disorders	0.2296	0.0230	9.97	1.26	1.2, 1.32
Multiple sclerosis	-0.1409	0.0337	-4.18	0.87	0.81, 0.93
Seizure disorders and convulsions	0.2077	0.0372	5.58	1.23	1.14, 1.32
Pneumonia	0.2670	0.0273	9.77	1.31	1.24, 1.38
Renal failure	0.1324	0.0263	5.03	1.14	1.08, 1.2

4.3.2 Hierarchical Logistic Regression Model

In the final hierarchical logistic regression model, the estimated mean between-hospital variance in the log-odds of mortality was 0.0420 (mean standard error=0.0072). This result implies that the odds of mortality for a high-mortality hospital (+1 standard deviation) were 1.51 times those for a low-mortality hospital (-1 standard deviation). Model variable descriptions, estimates, standard errors, and odds ratios are shown in Table 4.3.3 below. 1,473 hospitals with between-hospital variance=0.0420, standard error=0.0072.

Table 4.3.3. Final model: hierarchical logistic regression model results (N=94,466 patients in development cohort)

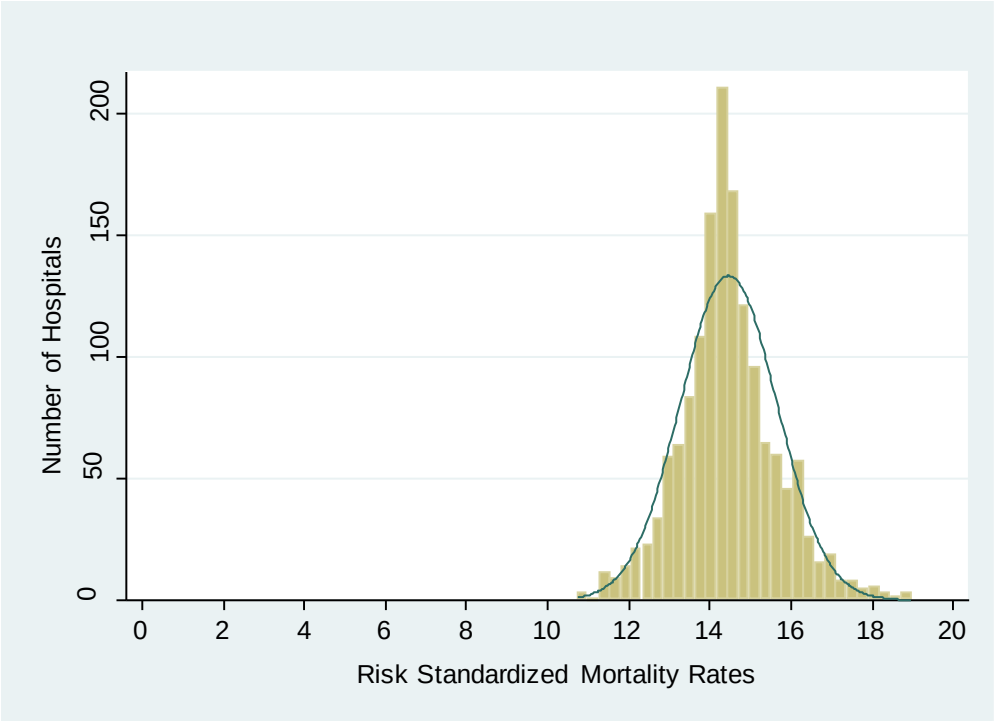
Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-3.7748	0.0445	-84.86	-	-
Age-65 (continuous, per 5 years)	0.2916	0.0072	40.72	1.34	1.32, 1.36
Transfer from another ED	0.2764	0.0343	8.07	1.32	1.23, 1.41
NIHSS (continuous, per 5 units)	0.4650	0.0069	68.03	1.59	1.57, 1.61
Congestive heart failure	0.2120	0.0265	8.01	1.24	1.17, 1.3
Congenital cardiac/circulatory defects	-0.4093	0.0784	-5.22	0.66	0.57, 0.77
Specified heart arrhythmias	0.2981	0.0233	12.81	1.35	1.29, 1.41
Cerebral atherosclerosis and aneurysm	-0.2114	0.0342	-6.18	0.81	0.76, 0.87
Metastatic cancer and acute leukemia and other major cancers	1.0686	0.0472	22.63	2.91	2.65, 3.19
Protein-calorie malnutrition	0.4813	0.0352	13.67	1.62	1.51, 1.73
Other significant endocrine and metabolic disorders	-0.3606	0.0312	-11.56	0.70	0.66, 0.74
Other gastrointestinal disorders	-0.1064	0.0226	-4.71	0.90	0.86, 0.94
Disorders of the vertebrae and spinal discs	-0.1262	0.0291	-4.34	0.88	0.83, 0.93
Osteoarthritis of hip or knee	-0.1390	0.0348	-4.00	0.87	0.81, 0.93
Other musculoskeletal and connective tissue disorders	-0.1083	0.0248	-4.37	0.90	0.85, 0.94
Iron deficiency and other/unspecified anemia and blood disease	0.1826	0.0237	7.70	1.20	1.15, 1.26

Description	Estimate	Standard Error	T	OR	95% CI
Dementia or other specified brain disorders	0.2343	0.0232	10.12	1.26	1.21, 1.32
Multiple sclerosis	-0.1419	0.0338	-4.20	0.87	0.81, 0.93
Seizure disorders and convulsions	0.2062	0.0374	5.51	1.23	1.14, 1.32
Pneumonia	0.2673	0.0274	9.75	1.31	1.24, 1.38
Renal failure	0.1310	0.0264	4.96	1.14	1.08, 1.20

4.3.3 Distribution of 30-Day Mortality Rate

After adjusting for patient characteristics and clustering within hospitals, RSMRs at the hospital level were more normally distributed, ranging from 11.75% to 18.98%. The median (interquartile range) RSMR was 14.48% (13.52%, 15.56%) (Figure 4.3.1).

Figure 4.3.1. Distribution of Hospital Risk-Standardized Mortality Rates (July 2011 – June 2014)



4.3.4 Validation of Claims-Only Model

We computed five summary statistics for assessing model performance: over-fitting indices, predictive ability, area under the receiver operating characteristic (ROC) curve, distribution of residuals, and model chi-square. Model performance was similar in each dataset, with strong model discrimination and fit. Predictive ability was also similar across datasets. The C-statistic (area under the ROC curve) was 0.81 and 0.82 for the development and validation datasets, respectively (Table 4.3.4).

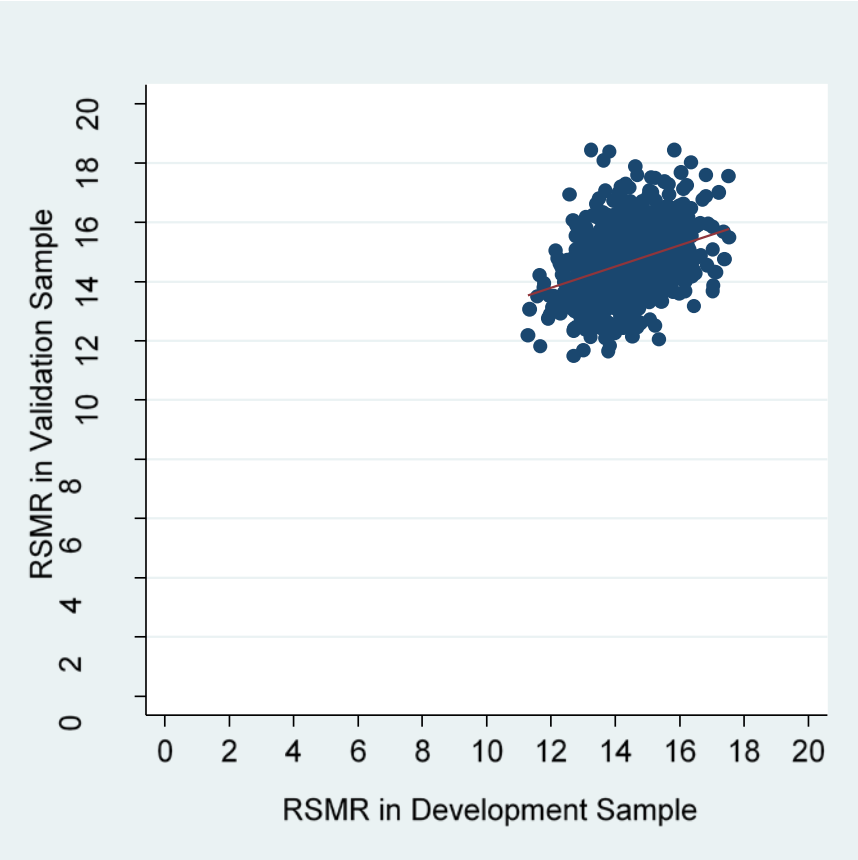
Table 4.3.4. Claims-only model performance results based on logistic regression model

Indices	Development Sample	Validation Sample
Number of Admissions	94,466	94,509
Mortality Rate	14.3	14.6
Calibration (γ0, γ1)	0.00, 1.00	0.00, 1.00
Discrimination: Adjusted R-square	0.2681	0.2764
Discrimination: C-statistic	0.81	0.82
Predictive Ability, % (lowest decile, highest decile)	1.3, 50.0	1.3, 51.6
Residuals Lack of Fit (Pearson Residual Fall %)	-	-
<-2	0.15	0.17
[-2, 0)	85.57	85.3
[0, 2)	8.63	8.99
[2+	5.65	5.58
Model χ2 (number of covariates)	12201.90 (20)	12584.62 (20)

4.3.5 Measure Testing - Reliability of Measure Results

When comparing the hospitals' RSMRs in the development and validation samples for the claims-only stroke mortality measure, hospital-level RSMRs were moderately correlated (correlation coefficient = 0.336), as shown in Figure 4.3.2. The reliability (ICC) for the full three years of data was 0.556, indicating moderate agreement between the development and validation samples.

Figure 4.3.2. Correlation of RSMRs in Development and Validation Samples for Claims-Only Stroke Mortality Measure for Hospitals with ≥ 12 Cases



4.4 Hybrid Model with Claims and Clinical EHR Data for Risk-Adjustment

Following the bootstrapping simulation method for variable selection, those candidate risk-adjustment variables that were included more than 90% of the time for all the copies of the imputed data were retained in the final model. The final model included 21 risk-adjustment variables (quadratic terms of three of the linear clinical EHR variables were retained in the model), listed in Table 4.4.1 below.

Table 4.4.1. Hybrid claims and clinical EHR data stroke mortality model final risk variables

Category	ICD-9/CC	Description
Demographics	--	Age (continuous, per 5 units)
Arrival Information	--	Transfer from another ED
Evaluation	--	NIHSS score (continuous, per 5 units)
Laboratory results	--	Blood glucose/10 (mg/dl): Linear
Laboratory results	--	Blood glucose/10 (mg/dl): Square
Vital signs	--	Heart rate: Linear
Vital signs	--	Heart rate: Square
Vital signs	--	DBP: Linear

Category	ICD-9/CC	Description
Vital signs	--	DBP: Square
Cardiovascular/ Cerebrovascular	CC 80	Congestive heart failure
Cardiovascular/ Cerebrovascular	CC 87-88	Congenital cardiac/circulatory defects
Cardiovascular/ Cerebrovascular	CC 92	Specified heart arrhythmias
Cardiovascular/ Cerebrovascular	CC 97	Precerebral arterial occlusion and transient cerebral ischemia
Cardiovascular/ Cerebrovascular	CC 98	Cerebral atherosclerosis and aneurysm
Comorbidities	CC 7-8	Metastatic cancer and acute leukemia and other major cancers
Comorbidities	CC 21	Protein-calorie malnutrition
Comorbidities	CC 22-24	Other significant endocrine and metabolic disorders
Comorbidities	CC 39	Disorders of the vertebrae and spinal discs
Comorbidities	CC 36	Other gastrointestinal disorders
Comorbidities	CC 43	Other musculoskeletal and connective tissue disorders
Comorbidities	CC 47	Iron deficiency and other/unspecified anemia and blood disease
Comorbidities	CC 49-50	Dementia or other specified brain disorders
Comorbidities	CC 111-113	Pneumonia
Comorbidities	CC 148	Decubitus ulcer of skin

4.4.1 Logistic Regression

The final logistic regression model performed very well, with a mean C-statistic of 0.8176 and an adjusted R-squared of 0.2794. The variable descriptions, estimates, and standard errors for the logistic regression model using the final model are shown in Table 4.4.2. For model variables that are included in both linear and square forms, odds ratios and 95% CIs are not displayed in the table. The odds ratio for a one unit increase in x is $\exp(\beta_1 + \beta_2(2x+1))$, where β_1 is the estimate of the linear term, and β_2 is the estimate of the square term. The formula still contains x, so it is not a constant across x, but a function of x.

Table 4.4.2. Final model: logistic regression results for claims and clinical EHR model (N=94,466 patients; development sample)

Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-6.8728	0.2621	-26.23	-	-
Age (continuous, per 5 units)	0.2947	0.0072	40.72	1.34	1.32, 1.36
Transfer from another ED	0.3829	0.0337	11.37	1.47	1.37, 1.57
NIHSS score (continuous, per 5 units)	0.4423	0.0065	68.03	1.56	1.54, 1.58
Blood glucose/10 (mg/dl): Linear	0.0649	0.0065	10.05	-	-
Blood glucose/10 (mg/dl): Square	-0.0011	0.0001	-7.32	-	-
Heart rate: Linear	-0.0109	0.0041	-2.66	-	-
Heart rate: Square	0.0001	0.0000	5.15	-	-
DBP: Linear	-0.0327	0.0034	-9.75	-	-
DBP: Square	0.0002	0.0000	10.23	-	-

Description	Estimate	Standard Error	T	OR	95% CI
Congestive heart failure	0.2491	0.0255	9.75	1.28	1.22, 1.35
Congenital cardiac/circulatory defects	-0.3890	0.0791	-4.92	0.68	0.58, 0.79
Specified heart arrhythmias	0.3357	0.0238	14.12	1.40	1.34, 1.47
Precerebral arterial occlusion and transient cerebral ischemia	-0.1449	0.0285	-5.09	0.87	0.82, 0.91
Cerebral atherosclerosis and aneurysm	-0.1733	0.0347	-4.99	0.84	0.79, 0.9
Metastatic cancer and acute leukemia and other major cancers	1.0599	0.0454	23.33	2.89	2.64, 3.15
Protein-calorie malnutrition	0.5536	0.0355	15.60	1.74	1.62, 1.86
Other significant endocrine and metabolic disorders	-0.3633	0.0313	-11.59	0.70	0.65, 0.74
Disorders of the vertebrae and spinal discs	-0.0985	0.0226	-4.35	0.91	0.87, 0.95
Other gastrointestinal disorders	-0.1588	0.0282	-5.63	0.85	0.81, 0.9
Other musculoskeletal and connective tissue disorders	-0.1197	0.0251	-4.77	0.89	0.84, 0.93
Iron deficiency and other/unspecified anemia and blood disease	0.2104	0.0228	9.23	1.23	1.18, 1.29
Dementia or other specified brain disorders	0.3221	0.0230	13.98	1.38	1.32, 1.44
Pneumonia	0.3059	0.0277	11.05	1.36	1.29, 1.43
Decubitus ulcer of skin	0.3131	0.0537	5.83	1.37	1.23, 1.52

4.4.2 Hierarchical Logistic Regression Model

In the final hierarchical logistic regression model, the estimated between-hospital variance in the log-odds of mortality was 0.0432 (standard error=0.0073). This result implies that the odds of mortality for a high-mortality hospital (+1 standard deviation) were 1.52 times those for a low-mortality hospital (-1 standard deviation). Model variable descriptions, estimates, standard errors, and odds ratios are shown in Table 4.4.3. For model variables that are included in both linear and square forms, odds ratios and 95% CIs are not displayed in the table. The odds ratio for a one unit increase in x is $\exp(\beta_1 + \beta_2(2x+1))$, where β_1 is the estimate of the linear term, and β_2 is the estimate of the square term. The formula still contains x, so it is not a constant across x, but a function of x.

Table 4.4.3. Final model: hierarchical logistic regression model results (N=94,466 patients; development sample)

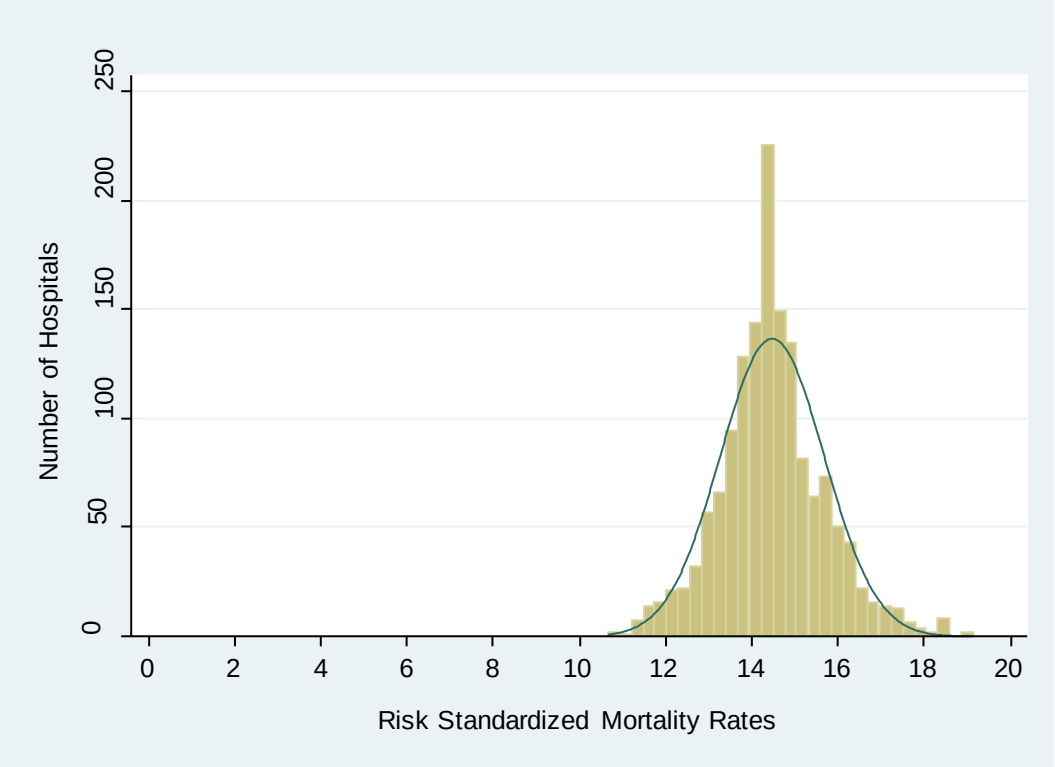
Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-6.9170	0.2616	-26.44	-	-
Age (continuous, per 5 units)	0.2965	0.0073	40.72	1.35	1.33, 1.36
Transfer from another ED	0.3634	0.0353	10.29	1.44	1.34, 1.54
NIHSS score (continuous, per 5 units)	0.4438	0.0066	68.03	1.56	1.54, 1.58
Blood glucose/10 (mg/dl): Linear	0.0652	0.0064	10.19	-	-
Blood glucose/10 (mg/dl): Square	-0.0011	0.0001	-7.36	-	-
Heart rate: Linear	-0.0111	0.0041	-2.71	-	-
Heart rate: Square	0.0001	0.0000	5.18	-	-

Description	Estimate	Standard Error	T	OR	95% CI
DBP: Linear	-0.0325	0.0034	-9.62	-	-
DBP: Square	0.0002	0.0000	10.12	-	-
Congestive heart failure	0.2499	0.0257	9.73	1.28	1.22, 1.35
Congenital cardiac/circulatory defects	-0.3880	0.0792	-4.90	0.68	0.58, 0.79
Specified heart arrhythmias	0.3323	0.0239	13.92	1.39	1.33, 1.46
Precerebral arterial occlusion and transient cerebral ischemia	-0.1426	0.0286	-4.98	0.87	0.82, 0.92
Cerebral atherosclerosis and aneurysm	-0.1735	0.0349	-4.98	0.84	0.79, 0.9
Metastatic cancer and acute leukemia and other major cancers	1.0682	0.0457	23.40	2.91	2.66, 3.18
Protein-calorie malnutrition	0.5589	0.0358	15.60	1.75	1.63, 1.88
Other significant endocrine and metabolic disorders	-0.3643	0.0315	-11.56	0.69	0.65, 0.74
Disorders of the vertebrae and spinal discs	-0.1011	0.0227	-4.45	0.90	0.86, 0.95
Other gastrointestinal disorders	-0.1584	0.0283	-5.60	0.85	0.81, 0.9
Other musculoskeletal and connective tissue disorders	-0.1189	0.0252	-4.71	0.89	0.85, 0.93
Iron deficiency and other/unspecified anemia and blood disease	0.2156	0.0230	9.39	1.24	1.19, 1.3
Dementia or other specified brain disorders	0.3271	0.0231	14.15	1.39	1.33, 1.45
Pneumonia	0.3060	0.0279	10.97	1.36	1.29, 1.43
Decubitus ulcer of skin	0.3204	0.0539	5.94	1.38	1.24, 1.53

4.4.3 Distribution of 30-Day Mortality Rate

After adjusting for patient characteristics and clustering within hospitals, RSMRs at the hospital level were found to be more normally distributed, ranging from 10.67% to 19.15%. The median (interquartile range) RSMR was 14.50% (13.50% to 15.61%) (Figure 4.4.1).

Figure 4.4.1. Distribution of Hospital Risk-standardized Mortality Rates (July 2011-June 2014) for the hybrid model with claims and clinical EHR risk-adjusment variables



4.4.4 Validation of Hybrid Model with Claims and Clinical EHR Data

We computed five summary statistics for assessing model performance: over-fitting indices, predictive ability, area under the receiver operating characteristic (ROC) curve, distribution of residuals, and model chi-square. Model performance was similar in each dataset, with strong model discrimination and fit. Predictive ability was also similar across datasets. The C-statistic (area under the ROC curve) was 0.82 for both datasets (Table 4.4.4).

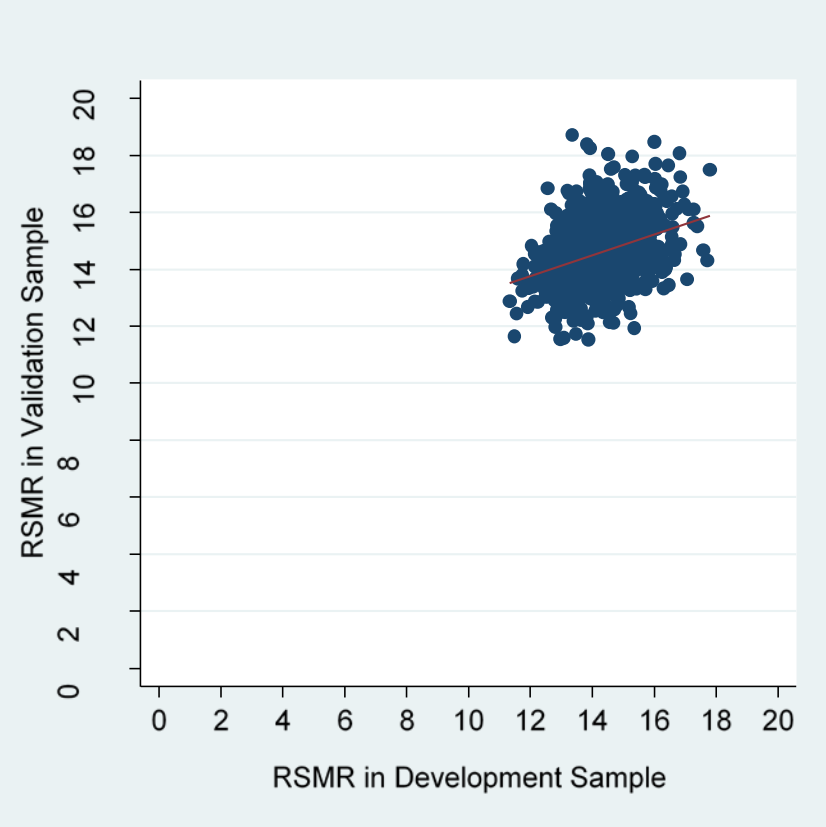
Table 4.4.4. Model performance: results based on the logistic regression model

Indices	Development Sample	Validation Sample
Number of Admissions	94,466	94,509
Mortality Rate	14.3	14.6
Calibration (γ0, γ1)	0.00, 1.00	0.00, 1.00
Discrimination: Adjusted R-square	0.2794	0.2851
Discrimination: C-statistic	0.82	0.82
Predictive Ability, % (lowest decile, highest decile)	1.3, 51.1	1.1, 51.9
Residuals Lack of Fit (Pearson Residual Fall %)	-	-
<-2	0.15	0.18
[-2, 0)	85.57	85.25
[0, 2)	8.74	9.08
[2+	5.54	5.49
Model χ2 (number of covariates)	12443.61 (24)	12754.41 (24)

4.4.5 Measure Testing - Reliability of Measure Results

When comparing the hospitals' RSMRs in the development and validation samples for the hybrid (claims and clinical EHR data risk-adjustment) stroke mortality measure among hospitals with at least 12 cases, hospital-level RSMRs were moderately correlated (correlation coefficient = 0.338), as shown in Figure 4.4.2. The reliability (ICC) for the full three years of data was 0.561, indicating moderate agreement between the development and validation samples.

Figure 4.4.2. Correlation of RSMRs in Development and Validation Samples for the hybrid (claims and clinical EHR risk-adjustment) stroke mortality model for Hospitals with ≥12 Cases



4.5 Hybrid Model with Clinical EHR Data-Only For Risk-Adjustment

Following the bootstrapping simulation method for variable selection, those candidate risk-adjustment variables that were included more than 90% of the time for all the copies of the imputed data were retained in the final model. The final model included 9 risk-adjustment variables, seven of which are CCDE (quadratic terms of five of the linear clinical EHR variables were retained in the model). With the exception of the NIHSS and international normalized ratio (INR), all of these variables are included in the CCDE (Table 4.5.1.).

Table 4.5.1. Clinical EHR data stroke mortality model final risk variables

Category	Description
Demographic	Age
Demographic	Male
Evaluation	NIHSS score
Laboratory results	Blood glucose/10 (mg/dL): Linear
Laboratory results	Blood glucose/10 (mg/dL): Square
Laboratory results	INR: Linear
Vital signs	Heart Rate: Linear
Vital signs	Heart Rate: Square
Vital signs	Weight: Linear
Vital signs	Weight: Square
Vital signs	SBP: Linear
Vital signs	SBP: Square
Vital signs	DBP: Linear
Vital signs	DBP: Square

4.5.1 Logistic Regression

The final logistic regression model performed very well, with a mean C-statistic of 0.7939 and an adjusted R-squared of 0.2442. The variable descriptions, estimates, and standard errors for the logistic regression model using the final model are shown in Table 4.5.2. For model variables that are included in both linear and square forms, odds ratios and 95% CIs are not displayed in the table. The odds ratio for a one unit increase in x is $\exp(\beta_1 + \beta_2(2x+1))$, where β_1 is the estimate of the linear term, and β_2 is the estimate of the square term. The formula still contains x, so it is not a constant across x, but a function of x.

Table 4.5.2. Final model: logistic regression results for clinical EHR data-only model (N=94,466 patients; development sample)

Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-4.5520	0.3604	-12.63	-	-
Age (continuous, per 5 units)	0.3113	0.0075	40.72	1.37	1.35, 1.39
Male	0.1867	0.0252	7.40	1.21	1.15, 1.27
NIHSS score (continuous, per 5 units)	0.4685	0.0064	68.03	1.60	1.58, 1.62
Blood glucose/10 (mg/dL): Linear	0.0658	0.0065	10.05	-	-
Blood glucose/10 (mg/dL): Square	-0.0011	0.0001	-7.46	-	-
INR	0.1745	0.0251	6.96	1.19	1.13, 1.25
Heart Rate: Linear	-0.0101	0.0040	-2.55	-	-
Heart Rate: Square	0.0001	0.0000	5.20	-	-
Weight: Linear	-0.0253	0.0034	-7.45	-	-
Weight: Square	0.0001	0.0000	6.20	-	-
SBP: Linear	-0.0233	0.0027	-8.58	-	-
SBP: Square	0.0001	0.0000	7.88	-	-
DBP: Linear	-0.0265	0.0034	-7.69	-	-

Description	Estimate	Standard Error	T	OR	95% CI
DBP: Square	0.0002	0.0000	8.32	*	*

4.5.2 Hierarchical Logistic Regression Model

In the final hierarchical logistic regression model, the estimated between-hospital variance in the log-odds of mortality was 0.0433 (standard error=0.0071). This result implies that the odds of mortality for a high-mortality hospital (+1 standard deviation) were 1.52 times those for a low-mortality hospital (-1 standard deviation). Model variable descriptions, estimates, standard errors, and odds ratios are shown in Table 4.5.3. For model variables that are included in both linear and square forms, odds ratios and 95% CIs are not displayed in the table. The odds ratio for a one unit increase in x is $\exp(\beta_1 + \beta_2(2x+1))$, where β_1 is the estimate of the linear term, and β_2 is the estimate of the square term. The formula still contains x, so it is not a constant across x, but a function of x.

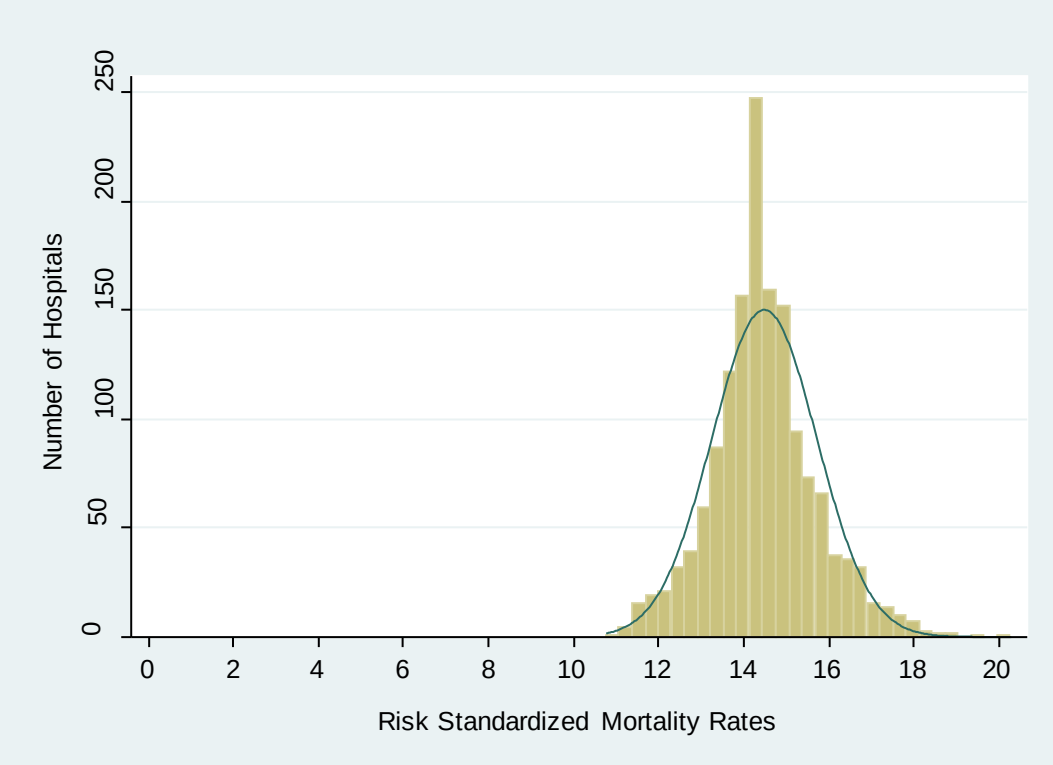
Table 4.5.3. Final model: hierarchical logistic regression model results (N=94,466 patients; development sample)

Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-4.5656	0.3592	-12.71	-	-
Age (continuous, per 5 units)	0.3137	0.0075	40.72	1.37	1.35, 1.39
Male	0.1879	0.0253	7.42	1.21	1.15, 1.27
NIHSS score (continuous, per 5 units)	0.4694	0.0066	68.03	1.60	1.58, 1.62
Blood glucose/10 (mg/dL): Linear	0.0660	0.0065	10.18	-	-
Blood glucose/10 (mg/dL): Square	-0.0011	0.0001	-7.49	-	-
INR: Linear	0.1739	0.0251	6.93	1.19	1.13, 1.25
Heart Rate: Linear	-0.0102	0.0040	-2.58	-	-
Heart Rate: Square	0.0001	0.0000	5.23	-	-
Weight: Linear	-0.0255	0.0034	-7.52	-	-
Weight: Square	0.0001	0.0000	6.20	-	-
SBP: Linear	-0.0236	0.0027	-8.64	-	-
SBP: Square	0.0001	0.0000	7.95	-	-
DBP: Linear	-0.0264	0.0035	-7.63	-	-
DBP: Square	0.0002	0.0000	8.25	-	-

4.5.3 Distribution of 30-day Mortality Rate

After adjusting for patient characteristics and clustering within hospitals, RSMRs at the hospital level were found to be more normally distributed, ranging from 10.77% to 20.26%. The median (interquartile range) RSMR was 14.46% (13.54%, 15.64%) (Figure 4.5.1).

Figure 4.5.1. Distribution of Hospital Risk-standardized Mortality Rates (July 2011-June 2014) for the Clinical EHR Data-Only Model



4.5.4 Validation of the Clinical EHR Data-only Model

Model performance was similar in each dataset, with strong model discrimination and fit. Predictive ability was also similar across datasets. The C-statistic (area under the ROC curve) was 0.79 and 0.80 for the development and validation samples, respectively (Table 4.5.4).

Table 4.5.4. Model performance: results based on the logistic regression model

Indices	Development Sample	Validation Sample
Number of Admissions	94,466	94,509
Mortality Rate	14.3	14.6
Calibration (γ0, γ1)	0.00, 1.00	0.00, 1.00
Discrimination: Adjusted R-square	0.2442	0.2523
Discrimination: C-statistic	0.79	0.80
Predictive Ability, % (lowest decile, highest decile)	2.2, 49.5	1.9, 50.9
Residuals Lack of Fit (Pearson Residual Fall %)	-	-
<-2	0.09	0.11
[-2, 0)	85.62	85.31
[0, 2)	8.21	8.54
[2+	6.07	6.03
Model χ2 (number of covariates)	11420.75 (14)	11826.5 (14)

4.5.5 Measure Testing - Reliability of Measure Results

When comparing the hospitals' RSMRs in the development and validation samples for the clinical EHR data-only stroke mortality measure among hospitals with at least 12 cases, hospital-level RSMRs were moderately correlated (correlation coefficient = 0.347), as shown in Figure 4.5.2. The reliability (ICC) for the full three years of data was 0.579, indicating moderate agreement between the development and validation samples.

Figure 4.5.2. Correlation of RSMRs in Development and Validation Samples for the Clinical EHR Data-Only Stroke Mortality Model for Hospitals with ≥12 Cases

4.6 Comparison of Model Results

We developed three stroke mortality risk-adjustment models, one claims-only model, and two hybrid models (claims and clinical EHR data; and clinical EHR data-only) in order to evaluate and compare model performance and their responsiveness to stakeholder concerns. The two models that include clinical EHR data were developed in order to provide CMS with options for approaches to risk adjustment of measures that include EHR data.

As shown in Table 4.6.1, the model with the highest c-statistic and the highest r-squared value was the hybrid model that combined claims and clinical EHR data in the risk-adjustment, with a c-statistic of 0.8176 and an r-squared value of 0.2794. The most parsimonious model was the hybrid model that used clinical EHR data-only in the risk-adjustment, with 9 variables.

Table 4.6.1. Comparison of three hybrid models that include the NIHSS: Claims-only risk-adjustment model, claims and clinical EHR risk-adjustment model, and clinical EHR-only risk-adjustment model

Domain of comparison		Claims only risk adjustment model	Claims + clinical EHR risk adjustment model	Clinical EHR only risk adjustment model
Model performance	C-statistic	0.81	0.82	0.79
	Adjusted R-squared	0.2681	0.2794	0.2442
	# Covariates	20	21	9

5. SUMMARY

This technical report describes the methodology used to develop measures of 30-day, hospital-level stroke mortality that include stroke severity in their risk models. This work is aligned with clinical guidelines to collect the NIHSS on patients admitted to the hospital with ischemic stroke. It is also responsive to stakeholder preference to include a stroke severity score in the risk models to improve their predictive ability and face validity. Given the two different data collection pathways for the NIHSS and rapidly expanding capabilities of EHR systems, we developed three risk models to give CMS different options for implementation of a stroke mortality measure.

The updated claims-only measure is designed to be implemented using comorbidity variables and the NIHSS, once available, from administrative claims. In addition to a modestly higher c-statistic, the updated claims-only measure has a more parsimonious risk model than the publicly reported stroke mortality measure. Calculation of this measure would require three years of NIHSS claims data. We anticipate that the NIHSS will be assigned ICD-10 codes in the fall of 2016 and data collection could begin from that point forward. If there is a preference for this version of the measure, CMS will work collaboratively with stakeholders to determine the best approach for data submission.

Alternatively, hybrid measures that use both claims and EHR data could be implemented using one of two approaches to risk adjustment. One approach utilizes the NIHSS and additional clinical variables from the EHR, as well as comorbidity variables from Medicare claims. The other takes a more parsimonious approach and uses the NIHSS and clinical variables from the EHR but omits the Medicare claims variables. Calculation of both versions of the hybrid measure will require collection of three years of prospective data from hospital EHRs. Because the hybrid stroke mortality measure uses data from both claims and EHR sources, implementation will require several additional steps. For example, all clinical EHR data elements included in the measure must be expressed in electronic clinical quality measure standards. If there is a preference for one of the hybrid versions of the measure, CMS will work collaboratively with stakeholders to develop the data submission standards.

By developing two types of measures and three different risk adjustment approaches, each of which has a model performance in line with those of other CMS publicly reported outcome measures, CMS and stakeholders have options regarding their preferred and most feasible approach to implementation of a stroke mortality measure that includes an assessment of stroke severity.

6. GLOSSARY OF TERMS

- *Administrative claims data:* An electronic environment in which hospitals capture data to submit claims to insurance providers for payment. These databases allow providers to complete the Universal Bill required to submit Medicare claims and contain patient data, such as dates of birth, names, national and unique medical record identification numbers, dates of admission, dates of discharge, principal discharge diagnoses, and all hospital charges that might be included in a bill for care provided.
- *Case Mix:* The particular illness severity and age characteristics of patients with index admissions at a given hospital.
- *Cohort:* The index admissions used to calculate the measure after inclusion and exclusion criteria have been applied.
- *Complications:* Medical conditions that are acquired during the index admission and might be a consequence of care rendered during hospitalization.
- *Comorbidities:* Medical conditions that the patient had in addition to his/her primary reason for admission to the hospital.
- *Core Clinical Data Elements (CCDE)*: A standardized set of clinical data that are consistently obtained on adult hospital inpatients that could be feasibly extracted from electronic health records, to be used in risk-adjustment for hospital quality outcome measures.
- *Electronic health records (EHR)*: A record in digital format that allows for systematic collection of electronic health information about individual patients or populations. It theoretically allows for sharing of information across different health care settings.
- *Expected mortality:* The number of deaths expected based on average hospital performance with a given hospital's case mix.
- *Feasibility*: Data elements that are consistently captured in current clinical practice, captured with a standard definition, and entered in structured fields across individuals as well as EHR and hospital systems.
- *Hierarchical model:* A widely accepted statistical method that enables fair evaluation of relative hospital performance by accounting for patient risk factors, as well as the number of patients a hospital treats. This statistical model accounts for the structure of the data (patients clustered within hospitals) and calculates (1) how much variation in hospital mortality rates overall is accounted for by patients' individual risk factors (such as age and other medical conditions); and (2) how much variation is accounted for by hospital contribution to mortality risk.
- *Hospital-specific intercept:* A measure of the hospital quality of care calculated based on the hospital's actual mortality rate relative to hospitals with similar patients, considering how many patients it served, its patients' risk factors, and how many died or were readmitted. The hospital-specific effect will be negative for a better-than-average hospital, positive for a worse-than-average hospital, and close to zero for an average hospital. The hospital-specific effect is used in the numerator to calculate "predicted" mortality.
- *Hybrid measure:* Quality measure that utilizes more than one source of data, such as patient clinical data captured in the EHR and CMS administrative claims data.
- *Index admission:* Any admission included in the measure calculation as the initial admission for an episode of care to which the outcome is attributed.

- *Medicare fee-for-service (FFS):* Original Medicare plan in which providers receive a fee or payment for each individual service provided directly from Medicare. All services rendered are unbundled and paid for separately. Only beneficiaries in Medicare FFS, not in managed care (Medicare Advantage), are included in the measure.
- *NIHSS:* National Institutes of Health Stroke Scale. This is a 15-item neurologic examination used to evaluate the effect of acute cerebral infarction on the levels of consciousness, language, neglect, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss.
- *Outcome:* The result of a broad set of healthcare activities that affect patients' well-being. For this mortality measure, the outcome is mortality within 30 days of discharge.
- *Predicted mortality:* The number of deaths within 30 days predicted based on the hospital's performance with its observed case mix.
- *Risk-adjustment:* Patient demographics and comorbidities used to standardize rates for differences in case mix across hospitals.

7. REFERENCES

1. Dorsey K WY, Zhang W, et al. 2013 Core Clinical Data Elements Technical Report (Version 1.1). 2015.
2. McNamara R BS, Mody P, et al. Hybrid 30-day Risk-standardized Acute Myocardial Infarction (AMI) Mortality Measure with Electronic Health Record (EHR)-Extracted Risk Factors Technical Report (Version 1.1), . 2013.
3. Go AS, Mozaffarian D, Roger VL, et al. Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation.* Jan 21 2014;129(3):e28-e292.
4. Kochanek KD MS, Xu JQ, Arias E. Mortality in the United States, 2013. *NCHS data brief, no 178.* 2014.
5. Casper ML NI, Croft JB, Nilasena DS. *Atlas of Stroke Hospitalizations Among Medicare Beneficiaries.* 2008.
6. Weir NU, Sandercock PA, Lewis SC, Signorini DF, Warlow CP. Variations between countries in outcome after stroke in the International Stroke Trial (IST). *Stroke.* Jun 2001;32(6):1370-1377.
7. DesHarnais SI, Chesney JD, Wroblewski RT, Fleming ST, McMahon LF, Jr. The Risk-Adjusted Mortality Index. A new measure of hospital performance. *Med Care.* Dec 1988;26(12):1129-1148.
8. Hong KS, Kang DW, Koo JS, et al. Impact of neurological and medical complications on 3-month outcomes in acute ischaemic stroke. *European journal of neurology : the official journal of the European Federation of Neurological Societies.* Dec 2008;15(12):1324-1331.
9. Lingsma HF, Dippel DW, Hoeks SE, et al. Variation between hospitals in patient outcome after stroke is only partly explained by differences in quality of care: results from the Netherlands Stroke Survey.[Reprint in Ned Tijdschr Geneeskd. 2008 Sep 27;152(39):2126-32; PMID: 18856030]. *Journal of Neurology, Neurosurgery & Psychiatry.* 2008;79(8):888-894.
10. Reeves MJ, Smith E, Fonarow G, Hernandez A, Pan W, Schwamm LH. Off-hour admission and in-hospital stroke case fatality in the get with the guidelines-stroke program. *Stroke.* Feb 2009;40(2):569-576.
11. Smith MA, Liou JI, Frytak JR, Finch MD. 30-day survival and rehospitalization for stroke patients according to physician specialty. *Cerebrovascular diseases (Basel, Switzerland).* 2006;22(1):21-26.
12. Fonarow GC, Saver JL, Smith EE, et al. Relationship of national institutes of health stroke scale to 30-day mortality in medicare beneficiaries with acute ischemic stroke. *J Am Heart Assoc.* Feb 2012;1(1):42-50.
13. Nedeltchev K, Renz N, Karameshev A, et al. Predictors of early mortality after acute ischaemic stroke. *Swiss Medical Weekly.* 2010;140(17-18):254-259.
14. Smith EE, Shobha N, Dai D, et al. Risk score for in-hospital ischemic stroke mortality derived and validated within the Get With the Guidelines-Stroke Program. *Circulation.* Oct 12 2010;122(15):1496-1504.
15. Fonarow GC, Alberts MJ, Broderick JP, et al. Stroke outcomes measures must be appropriately risk adjusted to ensure quality care of patients: a presidential advisory from the American Heart Association/American Stroke Association. *Stroke.* May 2014;45(5):1589-1601.
16. Jauch EC, Saver JL, Adams HP, Jr., et al. Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* Mar 2013;44(3):870-947.
17. Health Services Advisory Group I. A Blueprint for the CMS Measures Management System. 2014.

18. Krumholz HM, Brindis RG, Brush JE, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. *Circulation.* Jan 24 2006;113(3):456-462.
19. Commission TJ. Specifications Manual for Joint Commission Natinoal Quality Measures (v2013A1): Population and Sampling Specifications. 2013.
20. Bernheim S WC, Want Y, et al. . *Hospital 30-Day Mortality Following Acute Ischemic Stroke Hospitalization Measure Methodology Report.* 2010.
21. Drye EE, Normand SL, Wang Y, et al. Comparison of hospital risk-standardized mortality rates calculated by using in-hospital and 30-day models: an observational study with implications for hospital profiling. *Ann Intern Med.* Jan 3 2012;156(1 Pt 1):19-26.
22. HP A. Estimation of radioimmunoassay data using robust nonlinear regression methods. *Proceddings of the 10th Symposium on Computational Statistics.* 1992:367-372.
23. Dixon WJ YK. Trimming and winsorization: A review. *sStatistische Hefte.* 1974;15(2-3):157-170.
24. Rubin DB. Multiple Imputation for Nonresponse in Surveys. 2008.
25. He R, Belin T. Multiple imputation for high-dimensional mixed incomplete continuous and binary data. *Stat Med.* Jun 15 2014;33(13):2251-2262.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* Mar 1977;33(1):159-174.
27. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin.* Mar 1979;86(2):420-428.

8. APPENDICES

Appendix A. Cohort Definition for All Three Models

Table A 1. ICD-9-CM codes for stroke cohort

ICD-9-CM Codes	Description
433.01	Occlusion and stenosis of basilar artery with cerebral infarction
433.11	Occlusion and stenosis of carotid artery with cerebral infarction
433.21	Occlusion and stenosis of vertebral artery with cerebral infarction
433.31	Occlusion and stenosis of multiple and bilateral precerebral arteries with cerebral infarction
433.81	Occlusion and stenosis of other specified precerebral artery with cerebral infarction
433.91	Occlusion and stenosis of unspecified precerebral artery with cerebral infarction
434.01	Cerebral thrombosis with cerebral infarction
434.11	Cerebral embolism with cerebral infarction
434.91	Cerebral artery occlusion, unspecified with cerebral infarction
436	Acute, but ill-defined, cerebrovascular disease

Outcome

1. **30-day time frame**
   Rationale: Outcomes occurring within 30 days of discharge can be influenced by hospital care and the early transition to outpatient settings. The use of the 30-day time frame is a clinically meaningful period for hospitals to collaborate with their communities to reduce mortality.
2. **All-cause mortality**
   Rationale: From a patient perspective, death from any cause is an adverse event.

Appendix B. Candidate Variables

Table B 1. Candidate variables for the updated claims-only risk adjustment model

Variable	Description
n/a	Age minus 65 (years above 65, continuous)
n/a	Male
n/a	NIHSS (continuous)
n/a	Transfer from ED
ICD-9 codes V45.82, 00.66, 36.06, 36.07	History of Percutaneous Transluminal Coronary Angioplasty (PTCA)
ICD-9 codes V45.81, 36.10–36.16	History of Coronary Artery Bypass Graft (CABG)
CC 80	Congestive heart failure
CC 81	Acute myocardial infarction
CC 82	Other acute/subacute forms of ischemic heart disease
ICD-9 codes 410.00-410.12	Anterior myocardial infarction
ICD-9 codes 410.20-410.62	Other location of myocardial infarction
CC 83-84	Coronary atherosclerosis or angina
CC 79	Cardio-respiratory failure or shock
CC 86	Valvular or rheumatic heart disease
CC 89, 91	Hypertension
CC 95-96	Stroke
CC 97-99, 103	Cerebrovascular disease
CC 131	Renal failure
CC 108	Chronic obstructive pulmonary disease (COPD)
CC 111-113	Pneumonia
CC 15-20, 120	Diabetes mellitus (DM) or DM complications except proliferative retinopathy
CC 21	Protein-calorie malnutrition
CC 49-50	Dementia or other specified brain disorders
CC 67-69, 100-102, 177-178	Hemiplegia, paraplegia, paralysis, functional disability
CC 104-105	Vascular disease and complications
CC 7-8	Metastatic cancer, acute leukemia and other severe cancers
CC 154-156, 158-162	Trauma in last year
CC 54-56	Major psychiatric disorders
CC 25-27	Chronic liver disease

Table B 2. Selection process: candidate clinical variables from GWTG-Stroke registry (to be extracted from EHRs)

Category	Element	Step 1: Exclude variables related to post admission events	Step 2: Exclude variables unrelated to clinical status of patient at time of admission	Step 3: Exclude variables inconsistently or not reliably collected on stroke patients by clinicians (<90% capture rate)	Step 4: Exclude variables that may not be reliably & feasibly extracted from EHRs	Final candidate variables
Demographics	Age					Age
	Gender					Gender
	Race		X			
	Ethnicity		X			
History	Atrial Fibrillation/Prior MI				X	
	CAD/AMI				X	
	Carotid Stenosis				X	
	PVD				X	
	Heart failure				X	
	Stroke				X	
	TIA				X	
	Obesity/overweight				X	
	Hypertension				X	
	Diabetes				X	
	Dyslipidemia				X	
	Renal insufficiency				X	
	Smoking				X	
	Drug/alcohol abuse				X	
	Sleep apnea				X	
	Depression				X	
	HRT				X	
	Sickle cell				X	
Dx and Evaluation	Symptom Duration			X		

Category	Element	Step 1: Exclude variables related to post admission events	Step 2: Exclude variables unrelated to clinical status of patient at time of admission	Step 3: Exclude variables inconsistently or not reliably collected on stroke patients by clinicians (<90% capture rate)	Step 4: Exclude variables that may not be reliably & feasibly extracted from EHRs	Final candidate variables
	Had symptoms resolved at time of presentation				X	
	NIH Stroke Scale score					NIH Stroke Scale score
	Weakness/Paresis				X	
	Altered Level of Consciousness				X	
	Aphasia/Language Disturbance				X	
	Other neurological signs/symptoms				X	
	No neurological signs/symptoms				X	
	Ambulatory status on admission				X	
Medications	Antiplatelet			X		
	Anticoagulant			X		
	Aspirin			X		
	Aspirin/dipyridamole (Aggrenox)			X		
	Clopidogrel			X		
	Ticagrelor			X		
	Ticlopidine			X		
	Unfractionated heparin IV			X		
	Full dose LMW heparin			X		
	Warfarin			X		
	Dabigatran			X		

Category	Element	Step 1: Exclude variables related to post admission events	Step 2: Exclude variables unrelated to clinical status of patient at time of admission	Step 3: Exclude variables inconsistently or not reliably collected on stroke patients by clinicians (<90% capture rate)	Step 4: Exclude variables that may not be reliably & feasibly extracted from EHRs	Final candidate variables
	Argatroban			X		
	Desirudin			X		
	Fondaparinux			X		
	Rivaroxaban			X		
	Apixaban			X		
	Lepirudin			X		
	Antihypertensive			X		
	Cholesterol-reducer			X		
	Diabetic medication			X		
	Antidepressant medication			X		
Measurements	Total Cholesterol					Total Cholesterol
	HDL					HDL
	LDL					LDL
	Triglycerides					Triglycerides
	Blood glucose					Blood glucose
	A1C			X		
	INR					INR
	Serum creatinine					Serum creatinine
	Heart rate					Heart rate
	SBP					SBP
	DBP					DBP
	Height			X		
	Weight					Weight
	Waist Circumference			X		
	BMI			X		
Arrival and Admission	Mode of hospital arrival			X		

Category	Element	Step 1: Exclude variables related to post admission events	Step 2: Exclude variables unrelated to clinical status of patient at time of admission	Step 3: Exclude variables inconsistently or not reliably collected on stroke patients by clinicians (<90% capture rate)	Step 4: Exclude variables that may not be reliably & feasibly extracted from EHRs	Final candidate variables
Information						
	Arrival date		X			

Table B 3. Candidate variables for the claims and clinical EHR data risk adjustment model

Variable	Description
n/a	Age minus 65 (years above 65, continuous)
n/a	Male
n/a	NIHSS (continuous)
n/a	Transfer from ED
n/a	Total Cholesterol
n/a	HDL
n/a	LDL
n/a	Triglycerides
n/a	Blood glucose
n/a	INR
n/a	Serum creatinine
n/a	Heart rate
n/a	SBP
n/a	DBP
n/a	Weight
ICD-9 codes V45.82, 00.66, 36.06, 36.07	History of Percutaneous Transluminal Coronary Angioplasty (PTCA)
ICD-9 codes V45.81, 36.10–36.16	History of Coronary Artery Bypass Graft (CABG)
CC 80	Congestive heart failure
CC 81	Acute myocardial infarction
CC 82	Other acute/subacute forms of ischemic heart disease
ICD-9 codes 410.00-410.12	Anterior myocardial infarction
ICD-9 codes 410.20-410.62	Other location of myocardial infarction
CC 83-84	Coronary atherosclerosis or angina
CC 79	Cardio-respiratory failure or shock
CC 86	Valvular or rheumatic heart disease
CC 89, 91	Hypertension
CC 95-96	Stroke
CC 97-99, 103	Cerebrovascular disease
CC 131	Renal failure
CC 108	Chronic obstructive pulmonary disease (COPD)
CC 111-113	Pneumonia
CC 15-20, 120	Diabetes mellitus (DM) or DM complications except proliferative retinopathy
CC 21	Protein-calorie malnutrition
CC 49-50	Dementia or other specified brain disorders
CC 67-69, 100-102, 177-178	Hemiplegia, paraplegia, paralysis, functional disability
CC 104-105	Vascular disease and complications
CC 7-8	Metastatic cancer, acute leukemia and other severe cancers

Variable	Description
CC 154-156, 158-162	Trauma in last year
CC 54-56	Major psychiatric disorders
CC 25-27	Chronic liver disease

Table B 4. Candidate variables for the clinical EHR data-only risk adjustment model

Description
Age (continuous)
Gender
NIHSS score (continuous)
Total Cholesterol
HDL
LDL
Triglycerides
Blood glucose
INR
Serum creatinine
Heart rate
SBP
DBP
Weight

Appendix C. Results of Hybrid Models with "Mode of Arrival" Data Element Added

The variable "mode of hospital arrival" has previously been shown to be associated with mortality in stroke patients; however, as described in Section 3.6.1 above, this variable was excluded from the list of candidate variables because it is inconsistently or not reliably collected on stroke patients by clinicians. Because this variable was available in the GWTG-Stroke registry data and is an important predictor of mortality after stroke, we assessed the incremental benefit of including this variable in both the claims and clinical EHR model (Table C 1), and the clinical EHR data-only model. The c-statistics for these models were 0.8258 and 0.8070, respectively. The adjusted r-squared values for these models were 0.2964 and 0.2676, respectively.

Table C 1. Logistic regression results for hybrid model with claims and clinical EHR risk adjustment with the addition of "mode of hospital arrival"

Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-6.3275	0.2631	-24.05	-	-
Age (continuous, per 5 units)	0.1434	0.0661	2.17	1.15	1.01, 1.31
Transfer from another ED	0.2763	0.0072	40.72	1.32	1.3, 1.34
Mode of Hospital Arrival: Unknown	-0.1143	0.0465	-2.46	0.89	0.81, 0.98
Mode of Hospital Arrival: Private transport/taxi/other from home/scene	-1.0508	0.0358	-29.32	0.35	0.33, 0.38
Mode of Hospital Arrival: Transfer from other hospital	0.0786	0.0655	1.20	1.08	0.95, 1.23
NIHSS score (continuous, per 5 units)	0.4116	0.0067	68.03	1.51	1.49, 1.53
Blood glucose/10 (mg/dl): Linear	0.0615	0.0063	9.79	-	-
Blood glucose/10 (mg/dl): Square	-0.0010	0.0001	-6.99	-	-
Heart rate: Linear	-0.0105	0.0041	-2.59	-	-
Heart rate: Square	0.0001	0.0000	4.95	-	-
DBP: Linear	-0.0304	0.0034	-9.01	-	-
DBP: Square	0.0002	0.0000	9.39	-	-
Congestive heart failure	0.2302	0.0257	8.96	1.26	1.2, 1.32
Congenital cardiac/circulatory defects	-0.4019	0.0790	-5.09	0.67	0.57, 0.78
Specified heart arrhythmias	0.3044	0.0237	12.84	1.36	1.29, 1.42
Precerebral arterial occlusion and transient cerebral ischemia	-0.1294	0.0283	-4.57	0.88	0.83, 0.93
Cerebral atherosclerosis and aneurysm	-0.1633	0.0347	-4.71	0.85	0.79, 0.91
Metastatic cancer and acute leukemia and other major cancers	1.1022	0.0458	24.05	3.01	2.75, 3.29
Protein-calorie malnutrition	0.5305	0.0353	15.02	1.70	1.59, 1.82
Other significant endocrine and metabolic disorders	-0.3516	0.0314	-11.20	0.70	0.66, 0.75
Disorders of the vertebrae and spinal discs	-0.1018	0.0227	-4.48	0.90	0.86, 0.94
Other gastrointestinal disorders	-0.1535	0.0284	-5.41	0.86	0.81, 0.91
Other musculoskeletal and connective tissue disorders	-0.1229	0.0253	-4.86	0.88	0.84, 0.93
Iron deficiency and other/unspecified anemia and blood disease	0.1920	0.0230	8.36	1.21	1.16, 1.27
Dementia or other specified brain disorders	0.2984	0.0230	12.96	1.35	1.29, 1.41

Description	Estimate	Standard Error	T	OR	95% CI
Pneumonia	0.2848	0.0278	10.26	1.33	1.26, 1.4
Decubitus ulcer of skin	0.2823	0.0532	5.30	1.33	1.19, 1.47

Please note that for model variables that are included in both linear and square forms, odds ratios and 95% CIs are not displayed in the table. The odds ratio for a one unit increase in x is $\exp(\beta_1 + \beta_2(2x+1))$, where β_1 is the estimate of the linear term, and β_2 is the estimate of the square term. The formula still contains x, so it is not a constant across x, but a function of x.

Table C 2. Logistic regression results for clinical EHR data-only model with the addition of "mode of hospital arrival"

Description	Estimate	Standard Error	T	OR	95% CI
Intercept	-3.9128	0.3551	-11.02	-	-
Age (continuous, per 5 units)	0.2925	0.0075	40.72	1.34	1.32, 1.36
Male	0.2083	0.0253	8.24	1.23	1.17, 1.29
Mode of Hospital Arrival: Unknown	-0.1395	0.0447	-3.12	0.87	0.8, 0.95
Mode of Hospital Arrival: Private transport/taxi/other from home/scene	-1.1479	0.0356	-32.27	0.32	0.3, 0.34
Mode of Hospital Arrival: Transfer from other hospital	0.1474	0.0326	4.52	1.16	1.09, 1.24
NIHSS score (continuous, per 5 units)	0.4298	0.0066	68.03	1.54	1.52, 1.56
Blood glucose/10 (mg/dL): Linear	0.0625	0.0064	9.83	-	-
Blood glucose/10 (mg/dL): Square	-0.0010	0.0001	-7.11	-	-
INR	0.1731	0.0257	6.74	1.19	1.13, 1.25
Heart Rate: Linear	-0.0098	0.0039	-2.48	-	-
Heart Rate: Square	0.0001	0.0000	5.00	-	-
Weight: Linear	-0.0247	0.0027	-9.20	-	-
Weight: Square	0.0001	0.0000	8.52	-	-
SBP: Linear	-0.0235	0.0035	-6.76	-	-
SBP: Square	0.0001	0.0000	7.32	-	-
DBP: Linear	-0.0262	0.0035	-7.41	-	-
DBP: Square	0.0001	0.0000	6.01	-	-

Please note that for model variables that are included in both linear and square forms, odds ratios and 95% CIs are not displayed in the table. The odds ratio for a one unit increase in x is $\exp(\beta_1 + \beta_2(2x+1))$, where β_1 is the estimate of the linear term, and β_2 is the estimate of the square term. The formula still contains x, so it is not a constant across x, but a function of x.

Appendix D. Approach to Defining Continuous Clinical EHR Variables

Table D 1. Approach to defining and including continuous clinical EHR variables in risk models

Age	**Plausible range**	65 – 115 years
	Winsorization Decision	No Winsorization on age
	Form(s) Included in Model(s)	Linear
National Institutes of Health Stroke Severity (NIHSS) Scale	**Plausible range**	0 – 42
	Winsorization Decision	No Winsorization on NIHSS
	Form(s) Included in Model(s)	Linear
Total Cholesterol	**Plausible range**	50 – 800 mg/dL
	Winsorization Decision	Winsorize the upper limit to 326 mg/dL
	Form(s) Included in Model(s)	Linear and quadratic
Triglycerides	**Plausible range**	30 – 2000 mg/dL
	Winsorization Decision	Winsorize the upper limit to 416 mg/dL
	Form(s) Included in Model(s)	Linear
High Density Lipoprotein	**Plausible range**	10 – 130 mg/dL
	Winsorization Decision	Winsorize the upper limit to 113 mg/dL
	Form(s) Included in Model(s)	Linear, quadratic, and cubic
Low Density Lipoprotein	**Plausible range**	25 – 400 mg/dL
	Winsorization Decision	Winsorize the upper limit to 235 mg/dL
	Form(s) Included in Model(s)	Linear
Blood Glucose	**Plausible range**	10 – 1000 mg/dL
	Winsorization Decision	Winsorize the upper limit to 600 mg/dL
	Form(s) Included in Model(s)	Linear and quadratic
Serum Creatinine	**Plausible range**	0 – 15 mg/dL
	Winsorization Decision	Winsorize the upper limit to 4.0 mg/dL
	Form(s) Included in Model(s)	Linear and quadratic
International Normalized Ratio	**Plausible range**	0 – 8
	Winsorization Decision	Winsorize the upper limit to 5.5
	Form(s) Included in Model(s)	Linear
Heart Rate	**Plausible range**	30 – 250 bpm
	Winsorization Decision	Winsorize the upper limit to 164 bpm
	Form(s) Included in Model(s)	Linear and quadratic
Systolic Blood Pressure	**Plausible range**	60 – 280 mmHg
	Winsorization Decision	Winsorize the upper limit to 261 mmHg
	Form(s) Included in Model(s)	Linear and quadratic
Diastolic Blood Pressure	**Plausible range**	20 – 200 mmHg
	Winsorization Decision	Winsorize the upper limit to 153 mmHg
	Form(s) Included in Model(s)	Linear and quadratic
Weight	**Plausible range**	30 – 400 kg
	Winsorization Decision	Winsorize the upper limit to 161 kg
	Form(s) Included in Model(s)	Linear and quadratic

Appendix E. Working Group Members

Table E 1. List of measure development working group members and affiliations

Name	Organization/Affiliation
Gregg Fonarow, MD	• Professor of Medicine, David Geffen School of Medicine at UCLA • Director, Ahmanson-UCLA Cardiomyopathy Center • Co-Director, UCLA Preventative Cardiology Program • Associate Chief, UCLA Division of Cardiology • The Eliot Corday Chair in Cardiovascular Medicine and Science
Lee Schwamm, MD	• Professor of Neurology, Harvard Medical School • Vice Chairman, Department of Neurology, Massachusetts General Hospital • Director, TeleStroke & Acute Stroke Services
Kevin Sheth, MD	• Associate Professor of Neurology and of Neurosurgery • Division Chief, Neurocritical Care and Emergency Neurology • Director, Neuroscience ICU • Chief, Clinical Research, Department of Neurology
Jason Sico, MD	• Assistant Professor of Neurology • Assistant Professor Internal Medicine (Section of General Medicine) • Director, Stroke Care VA Connecticut Healthcare System