

**2013 Core Clinical Data Elements  
Technical Report  
(Version 1.1)**

**Submitted By**

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

**Prepared For:**

Centers for Medicare & Medicaid Services (CMS)

**February 2015**

## Table of Contents

1.	EXECUTIVE SUMMARY .....	8
1.1	Overview .....	8
1.2	Objectives.....	8
1.3	Methodology.....	8
1.4	Results.....	9
2.	INTRODUCTION .....	10
2.1	Conceptual Rationale .....	10
2.2	Report Overview .....	10
2.3	Report Update.....	11
3.	IDENTIFICATION OF FEASIBLE CLINICAL DATA CATEGORIES AND SUBCATEGORIES .....	12
3.1	Approach.....	12
3.2	Methods .....	12
3.2.1.	Electronic Health Information Framework .....	12
3.2.2.	Establishing Criteria to Identify High-Quality and High-Value Data .....	13
3.2.3.	Application of Criteria to Data Categories and Subcategories.....	15
3.3	Results.....	16
3.3.1.	Survey Results .....	16
3.3.2.	Discussion of Survey Results .....	18
3.3.3	Final Subcategories Selected for Further Feasibility Testing .....	21
4.	ASSESSMENT AND SELECTION OF DATA ELEMENTS FOR TESTING .....	24
4.1	Approach.....	24
4.2	Methodology.....	25
4.2.1.	Literature Review to Identify Common Predictors of Mortality across Conditions....	25
4.2.2.	Feasibility Testing in a Clinical Dataset .....	25
4.3	Results.....	30
4.3.1.	Consistency and Timing of First-Captured Value .....	30
4.3.2.	Variation across Hospitals and Hospital Units .....	33
4.4	Summary .....	37
5.	PRELIMINARY TESTING OF RISK-ADJUSTED 30-DAY MORTALITY MODELS.....	39
5.1	Approach.....	39
5.2	Methods.....	39
5.2.1.	Outcome.....	39
5.2.2.	Data Source and Cohort Derivation .....	39
5.2.3.	Preliminary Model Development.....	40

5.3 Results .....	41
5.4 Summary .....	42
6. CORE CLINICAL DATA ELEMENTS AND POTENTIAL STRATEGIES FOR INTEGRATION INTO CMS PROGRAMS .....	44
6.1 Summary .....	44
6.2 Rationale for a Clinical Dataset .....	45
7. GLOSSARY OF TERMS .....	47
REFERENCES .....	49
Appendix A .....	50
Appendix B .....	61

## LIST OF TABLES

Table 4.1: Readily Extractable Vital Signs from KPNC Databases .....	26
Table 4.2: Readily Extractable Laboratory Results from KPNC Databases.....	26
Table 4.3: Admissions by Specialty Cohort .....	28
Table 4.4: Population Statistics by Condition .....	29
Table 4.5: Proportion of Episodes with Captured Vital Signs at Various Time points .....	30
Table 4.6: Proportion of Admissions with at least 90% Captured Vital Signs at 2 Hours for the Common Medical Conditions .....	31
Table 4.7: Proportion of Admissions with Laboratory Results at Various Time Points .....	32
Table 4.8: Proportion of Admissions with at least 90% Capture of Laboratory Result at 24 hours for the Common Medical Conditions.....	32
Table 4.9: First Hospital Entry Locations by Condition Cohort .....	33
Table 4.10: Range of Values for Vital Signs in Full Cohort .....	35
Table 4.11: Range of Values for Laboratory Results in Full Cohort .....	36
Table 5.1: Odds Ratios and 95% Confidence Intervals for Data Elements Tested in Each Condition-Specific Risk-Adjustment Model .....	41
Table 6.1: Proposed Clinical Data Elements for Risk-Adjusted Mortality Measures .....	45
Table A.1: Evaluating Electronic Health Record (EHR) Data Elements for use in Hospital Quality Measures: Technical Expert Panel Members.....	53
Table A.2: Strong agreement on infeasibility ( $\geq 70\%$ ) .....	54
Table A.3: Complete agreement on infeasibility (100%) .....	54
Table B.1: Data Elements for Patient Characteristics in Mortality Indices .....	63
Table B.2: Data Elements for Physical Examination in Mortality Indices .....	63
Table B.3: Data Elements for Laboratory Results in Mortality Indices .....	64
Table B.4: Hospital System Characteristics .....	65
Table B.5: Identification of Hospital Entry Time .....	65
Table B.6: Capture and Extraction Capabilities of Core Clinical Data Elements .....	65
Table B.7: Additional Considerations for Laboratory Test Results .....	66

## LIST OF FIGURES

Figure 2.1: Project Overview.....	11
Figure 3.1: Identification of Potentially Feasible Clinical Data .....	23
Figure 4.1: Project Overview.....	24
Figure 4.2: Cohort Derivation for Feasibility Testing .....	28
Figure 4.3: Proportion of Admissions with First Hemoglobin Result by First Hospital Entry Location.....	34
Figure 4.4: Proportion of Admissions with First Sodium Result by First Hospital Entry Location .....	34
Figure 4.5: Range of Values for Hemoglobin by First Hospital Entry Location .....	36
Figure 4.6: Assessment and Selection of Clinical Data Elements for Testing .....	38
Figure 5.1: Identification of Core Clinical Data Elements .....	43
Figure 6.1: Project Overview.....	44
 Figure A.1: Example of QDM Logic Structure.....	 50
Figure B.1: Flow diagram of literature search initial results to articles that were full-text reviewed.....	63

## **YNHHSC/CORE PROJECT TEAM**

Karen Dorsey, M.D., Ph.D. – Project Lead  
Yongfei Wang, M.S. – Lead Project Analyst  
Weiwei Zhang, M.P.H. - Supporting Project Analyst  
Megan Keenan, M.P.H. – Project Coordinator  
Amena Keshawarz, M.P.H. – Research Assistant II  
Susannah Bernheim, M.D., M.H.S. – Project Director  
Harlan M. Krumholz, M.D., S.M.\* - Principal Investigator

### **Prepared with Contributions from the Hybrid Outcome Measures Reevaluation Team:**

Karen B. Dorsey, MD, PhD\* – Lead  
Yongfei Wang, MS – Lead Analyst  
Zhenqiu Lin PhD – Lead Analyst  
Megan Keenan, MPH – Project Manager  
Nicole Cormier, MPH – Research Associate  
Mallory Perez, BSPH – Research Assistant  
Weiwei Zhang, MPH – Supporting Analyst  
Susannah Bernheim, MD, MHS – Project Director  
Harlan M. Krumholz, M.D., S.M.\* – 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 the full description of the methodology used for electronic health data mapping and extraction from our colleagues affiliated with Kaiser Permanente of Northern California. We are also grateful for the ongoing support from the team:

Gabriel Escobar, M.D.  
Marla Gardner, B.A.  
John Greene, M.A.

We would also like to acknowledge support from individuals at the Office of the National Coordinator for Health Information Technology (ONC) who provided their expertise regarding eMeasures:

Kevin Larsen, M.D.  
Jesse James, M.D., M.B.A.  
Julia Skapik, M.D., M.P.H.

We would like to thank Triple Frog LLC for designing our online survey and report graphics.

We would also like to thank members of the Technical Expert Panel for their valuable contributions to this work. The full list of members is located in [Appendix A2](#).

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

Kanchana Bhat, M.P.H.	Jessica Brewer, M.P.H.
Nihar Desai, M.D., Ph.D.	Elizabeth Drye, M.D., M.S.
Elizabeth Eddy, B.A.	Jacqueline Grady, M.S.
Nancy Kim, M.D., Ph.D.	Alex Liu, B.S.
Robert McNamara, M.D., M.H.S.	Jeremy Michel, M.D.
Julia Montague, M.P.H.	Joseph Ross, M.D., M.H.S.
Lori Schroeder, LL.M., J.D.	Lisa Suter, M.D.

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

## 1. EXECUTIVE SUMMARY

### 1.1 Overview

A standardized set of clinical data that are consistently obtained on hospital inpatients will support critical advances in quality measurement, research, and health care surveillance. We designed and applied an approach to identify a set of core clinical data elements that are consistently available for adult hospitalized patients and that could be feasibly extracted from [electronic health records \(EHR\)](#). Our particular focus was on data elements that reflect patients' clinical status upon admission to the hospital and could be used in risk-adjustment for hospital quality outcome measures. However, the approach and resultant dataset will have broader applications.

Currently, hospital outcome measures that are publicly reported by the Centers for Medicare & Medicaid Services (CMS) on [Hospital Compare](#) are calculated using administrative claims data. These measures have been shown to provide valid information about hospital performance. However, the clinical community continues to express a preference for the use of clinical data to assess hospital performance. The recent and widespread transition from paper to electronic medical records among healthcare providers creates an opportunity to collect and integrate clinical information into hospital quality measurement without the substantial burden of manual chart abstraction. The purpose of this project is to demonstrate that clinical data can be used to risk-adjust outcome measures across a variety of medical conditions by using clinical information that is consistently captured for most adult patients during acute care hospital admissions and can be feasibly extracted from EHRs. We assessed the use of electronic data for risk adjustment as this is an area of growing interest and opportunity.

### 1.2 Objectives

In this project, we establish a set of criteria and processes to identify clinical data that are available for most adult hospitalized patients and provide accurate information relevant to the assessment of hospital performance. Specifically, we identify a set of data elements that are consistently captured in medical records under current clinical practice and that can be extracted from electronic or paper medical records. To demonstrate the utility of this set of clinical data elements, we examined their use as risk adjusters in models of 30-day mortality in adult patients hospitalized for a variety of common medical conditions.

These core clinical data elements will provide measure developers with a standard set of reliable data that can be used as a starting place when building risk-adjustment models for quality measures using clinical data. Furthermore, this dataset will provide an additional motivation for the standardization of capture, storage, and extraction of important clinical information. The benefits of standardizing common, discrete data elements extend beyond outcome measurement to other applications, like EHR interoperability, real-time clinical decision support, research, and public health surveillance.

### 1.3 Methodology

In order to identify the core clinical data elements for risk adjustment, we first conducted a qualitative assessment of the reliable capture, [accuracy](#), and extractability of clinical data, such as vital signs, laboratory test results, documentation of medical encounters, and medications. We established a set of criteria to assess the [consistency of data capture](#), relevance to hospital quality measures, and extractability from health records. We then convened a Technical Expert Panel (TEP) to apply these criteria to categories and subcategories of clinical data based on the Quality Data Model (QDM). Data



categories and subcategories were rated independently by TEP members. The ratings were tallied and TEP members identified a list of data subcategories that were potentially feasible for use in quality measurement.

Next, we directly examined the [feasibility](#) of clinical data from the TEP-identified data subcategories using a large multi-site database. We used data from Kaiser Permanente of Northern California (KPNC), an integrated healthcare delivery system serving over 3.3 million members and utilizing a fully automated EHR in all of its 21 acute care hospitals. We examined all admissions to KPNC acute care hospitals between 2010 and 2011. All clinical data were extracted from KPNC's EHR, an Epic system, and clinical laboratory database. We analyzed clinical data elements to determine the format of the data, the consistency and timing of capture, and the accuracy of the data elements. We examined the data elements across conditions, hospitals, and points of hospital entry.

To verify that the findings from our analysis of the KPNC database were generalizable to other hospitals and [electronic health systems](#), CORE partnered with Premier Inc., a collaborative healthcare alliance of approximately 2,900 U.S. community hospitals focused on measuring and improving their members' quality outcomes and safely reducing healthcare costs. We administered a survey to four of their member hospital systems that used a variety of EHR systems to confirm the availability of the clinical data elements.

Finally, we created statistical models to determine whether the identified and feasible clinical data elements could be used to risk adjust for 30-day mortality in cohorts of adult patients hospitalized with specific, common, and costly medical conditions.

## **1.4 Results**

We identified a standard set of clinical data elements that are captured for most adult hospitalized patients and that can be readily extracted from most currently operating EHRs. We established that this list of core clinical data elements can be used to risk adjust measures of 30-day mortality across a variety of common and costly medical conditions. These data elements can provide a foundation of reliable and accurate clinical information that can be extracted and potentially used for a variety of important applications, such as [eMeasure](#) development, health surveillance efforts, research, and clinical decision support.

We have also developed a rigorous, systematic, and replicable process for identifying and selecting high-value and high-quality data elements that are accurate, reliable, and relevant for use in hospital quality measures. In recent years, there has been a proliferation of electronic health systems designed to capture clinical data and integrate it with administrative data within health systems. This is due in large part to national investments in these systems to improve healthcare access, delivery, and quality. As a result, EHR utilization and functionality are on a dynamic and rapidly moving trajectory. The process described in this report can be reapplied in the future to evaluate the quality and value of new data elements to accommodate this evolution in EHR use and capability as well as innovations in clinical care.

## 2. INTRODUCTION

### 2.1 Conceptual Rationale

National efforts to leverage health technology and maximize the quality of health and health care have led to increasing interest in the use of clinical data in outcome measurement. Harmonization of clinical data for healthcare quality surveillance and improvement as well as public health efforts, clinical research, and electronic health system interoperability has the potential to benefit patients and lead to substantial reduction in healthcare costs.

In this project, we sought to identify data elements that are already routinely captured for most patients admitted to the hospital for common medical conditions under current clinical practice. These data elements should be consistently captured with a single, consensus definition across institutions and clinicians. Additionally, these data elements should be recorded in a structured format and extractable from medical records. We anticipated that many data elements meeting these criteria could be used effectively in hospital quality measures for risk adjustment.

### 2.2 Report Overview

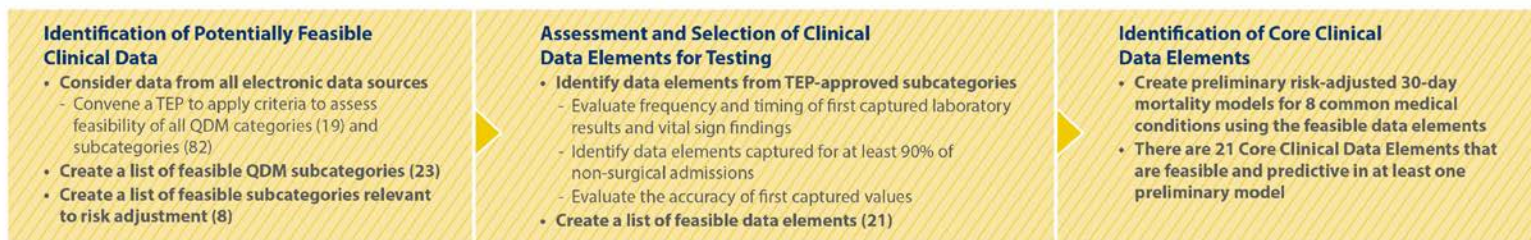
This technical report describes each phase of the development of a list of core clinical data elements as well as our recommendations for future steps toward implementation of that list in the following sections:

**Identification of Potentially Feasible Clinical Data** ([Section 3](#)): This section describes the qualitative evaluation of a comprehensive clinical data framework to identify categories and subcategories of feasibly extracted data. Specifically, we convened a Technical Expert Panel (TEP) and asked them to apply a set of data feasibility criteria to categories of clinical data elements in this framework. For this assessment, data feasibility was defined as whether data are consistently captured in current clinical practice, captured with a standard definition, and entered in structured fields. Experts identified those categories of data most likely to be feasible, accurate, and relevant to quality measurement.

**Assessment and Selection of Clinical Data Elements for Testing** ([Section 4](#)): This section describes the quantitative analysis of specific data elements in the clinical categories identified as potentially feasible by the TEP. We identified data elements shown to predict mortality based on a systematic review of the literature. Those data elements along with other currently [mapped](#) data elements were extracted from the Kaiser Permanente Northern California (KPNC) database for all adult patients admitted to their network of hospitals between 2010 and 2011. We examined the data format, the consistency and timing of capture, and the accuracy of these extracted data elements across different hospitals, across principal discharge diagnoses, and for various locations of patients' entry at each hospital. We identified a list of data elements that were consistently captured close to patients' first arrival at the hospital and that were accurate for potential use in quality measurement.

**Identification of Core Clinical Data Elements** ([Section 5](#)): This section describes statistical modeling of those data elements from the KPNC dataset that were shown to be feasible and accurate for risk adjustment in measures of 30-day mortality after admission for eight common medical conditions. We identified a list of data elements that were statistically significant predictors of the outcome for at least one of the eight conditions. Refer to [Figure 2.1](#) for an overview of these three sections.

**Figure 2.1: Project Overview**



Finally, [Section 6](#) describes the final core clinical data elements. We also describe several potential strategies to incentivize the collection of these data elements and to integrate them into publicly reported outcome measures.

## 2.3 Report Update

Please note that this report has been modified from its original version for posting with the Hospital Inpatient Prospective Payment Systems 2016 Proposed Rule.

We identified a standard set of core clinical data elements that are captured during routine clinical practice on most adult hospitalized patients and can be readily extracted from most currently operating EHRs. We established that this list of 21 core clinical data elements can be used to risk adjust measures of 30-day mortality across a variety of common and costly medical conditions.

The hospital 30-day risk-standardized acute myocardial infarction (AMI) mortality eMeasure (NQF #2473) (now referred to as a hybrid measure) originally identified several of the core clinical data elements for inclusion in the risk-adjustment model. The measure includes the following risk variables: age, heart rate, systolic blood pressure, and creatinine. It also includes one AMI-specific data element, the laboratory value for troponin ratio (initial troponin value / troponin upper range limit for hospital). For further details, please see the “Hybrid 30-day Risk-standardized Acute Myocardial Infarction Mortality Measure with Electronic Health Record Extracted Risk Factors (Version 1.1)” posted along with this report.

The hybrid hospital-wide 30-day readmission measure was developed to examine the use of the core clinical data elements in a broader cohort of hospitalized medical and surgical patients as well as to examine its utility in predicting hospital readmission. The measure is a composite of five models that group similar conditions and procedures. The following core clinical data elements are predictive in at least one of those models: age, heart rate, respiratory rate, temperature, systolic blood pressure, oxygen saturation, weight, hematocrit, white blood cell count, sodium, potassium, bicarbonate, creatinine and glucose. For further details, please see the “Hybrid Hospital-Wide Readmission Measure with Electronic Health Record Extracted Risk Factors (Version 1.1)” posted along with this report.

### 3. IDENTIFICATION OF FEASIBLE CLINICAL DATA CATEGORIES AND SUBCATEGORIES

#### 3.1 Approach

The first phase in identifying the core clinical data elements was to examine the full set of potential data categories available in electronic health systems and to select those that contained data elements that are consistently captured and extractable. This section describes our approach to identifying feasible data categories using a qualitative evaluation of the capture, accuracy, and extractability of various types of clinical data for use in hospital quality measures. This approach involved identifying an information model that organizes clinical data elements into comprehensive and conceptually coherent categories, and developing a [set of criteria](#) to evaluate those categories with respect to feasibility and relevance to hospital outcome measures. We convened a TEP and asked experts to assess the feasibility of each category of clinical data by applying the set of criteria. Their assessment provided a narrow and manageable list of categories containing high quality clinical data elements that could then be tested for availability and accuracy ([Section 4](#)) and tested in risk-adjustment models ([Section 5](#)).

We focused on the assessment of electronic sources of clinical data in order to align with national efforts to incentivize the adoption of these systems and to inform the development of eMeasures, or measures that utilize clinical data from electronic systems. However, the data assessed in this report are also available and extractable from paper medical records. Due to national efforts, most healthcare providers have transitioned or are in the process of transitioning from paper to electronic capture and storage of their patients' health information. The storage of health information using electronic systems greatly reduces the cost and effort required to extract and report clinical information compared with paper records. This creates an opportunity to increase integration of clinical data into quality measures. The electronic health systems referred to in this section and throughout this report include integrated and non-integrated [patient management databases](#), administrative or billing databases, clinical laboratory databases, and clinical electronic health record (EHR) databases ([Figure 3.1](#)). These systems contain important information collected during hospital admissions about patients' clinical status, the quality of care they receive, and their health outcomes.

Currently, most eMeasure development processes identify variables that perform well in measures before they consider data feasibility. As a result, many eMeasures cannot be implemented because they contain variables that may not be consistently captured or feasibly extracted. No standard procedure currently exists to evaluate the feasibility of clinical data for use in national quality measures. In this section, we present a replicable process for assessing the availability of clinical data captured in electronic health systems for use in quality measures. We describe the selection of the data information model, data feasibility criteria, and the results of the TEP's evaluation and discussion of categories of clinical data.

#### 3.2 Methods

##### 3.2.1. *Electronic Health Information Framework*

We first sought to identify a structured information model that describes all clinical data that are likely to be routinely captured and stored by hospitals. This model would provide a framework for identifying data that are consistently available for adult patients and that can be readily mapped and extracted from electronic databases for use in hospital outcome measures. After discussions with experts, we chose the Quality Data Model (QDM) for this project because its logic structure was

designed specifically to support quality measurement. For an example of the QDM logic structure, refer to [Figure A.1](#) in [Appendix A](#).

The QDM is a comprehensive data information model created by the National Quality Forum (NQF) to support eMeasure development. The QDM was first released publicly in 2010. Comments and suggestions for improvement were solicited from EHR software industry experts, eMeasure developers, healthcare quality researchers, EHR users, and experts from the regulatory and policy communities. These comments and suggestions were incorporated into several subsequent versions of the model. The version used was released in [December 2012](#) and provided a complete listing of 19 mutually exclusive data categories. Each category has several subcategories, referred to as data types in the QDM, to further organize data elements. For example, the category *laboratory test* has 6 subcategories, *test recommended*, *test order*, *test performed*, *test result*, *test intolerance*, and *test adverse event* ([Figure A.1](#)). The QDM is now part of the CMS Measure Authoring Tool (MAT) to support eMeasure development and [eSpecification](#).

The QDM has several features that make it well-suited for this project including:

- An intuitive logic structure to organize data elements into clinical categories and subcategories.
- Linkage of data elements with [metadata](#) required for analytic examination including attributes related to the timing, method, personnel, and other data capture information.
- A development process that includes open discussion and feedback from users and stakeholders.
- Common use by the Department of Health and Human Services-sponsored health information technology (HIT) efforts.
- Support for eSpecification of measures derived from electronic health systems.

The QDM has several important limitations with respect to developing hospital outcome measures. For example, the QDM does not provide a logic structure for elements commonly used for cohort selection and measure exclusions, such as care directives and discharge disposition. The QDM also does not provide a logic structure for concepts traditionally captured in hospitals' patient management systems, such as the exact timing of the start and end of an inpatient encounter or the details of patient transfers between hospital units. For example, a patient may initially be seen in a hospital's emergency department (ED) for pneumonia for several hours before they are admitted to the hospital. The QDM does not provide a clear data element for linking these encounters within the healthcare setting. Furthermore, there is no clear way to identify admissions for ongoing treatment of a single medical condition when patients are transferred between different facilities. However, these limitations are not unique to the QDM, so it was selected as the most appropriate existing EHR data information model for our evaluation.

### *3.2.2. Establishing Criteria to Identify High-Quality and High-Value Data*

Currently, there is no standard set of criteria for selecting feasible electronic data for use in hospital quality measures. In order to address the lack of feasibility standards and the need for standardized health data for quality measurement, CORE developed a set of feasibility criteria.

CORE previously developed the hybrid hospital 30-day risk-standardized acute myocardial infarction (AMI) mortality measure. The primary objective of this previous project was to develop a hybrid outcome measure that could be implemented without changing clinical workflows or requiring

adaptation of EHRs. “Hybrid” outcome measures are quality measures that utilize more than one source of data. During the development of this measure and through consultation with EHR experts, CORE developed three criteria to ensure all data elements used in the measure would meet this standard. These criteria required data elements to be:

1. Consistently obtained in the target population based on current clinical practice.
2. Captured with a standard definition and recorded in a standard format.
3. Entered in structured fields that are feasibly retrieved from current EHR systems.

The first criterion ensures that the measure will not rely on adoption of new clinical practices, such as requiring medical staff to routinely collect a laboratory test they might not otherwise order. The second criterion confirms that data elements used in the measure have the same meaning across providers. The third aligns with the intention to build a measure that could be feasibly implemented in current EHRs.

For this project, we sought to identify a broader set of routinely captured data elements that can be used to risk adjust 30-day mortality measures across a variety of common medical conditions, but we similarly aimed to identify data elements that could be captured without disruption of clinical workflow or adaptation of current EHRs. We therefore used the CORE criteria as a starting point for this project. We then re-engaged in discussion with the Office of the National Coordinator for HIT (ONC), other health information technology (HIT) experts, and quality measure development experts. We also examined [NQF’s eMeasure Feasibility Assessment Report](#) to ensure that our criteria were consistent with opinions of their expert panel. The NQF report emphasized four key aspects of feasibility. First, data should be structured or easily converted to a structured and interpretable format. Second, data should be accurate. Third, data should be easily associated with a standard set of codes to ensure consistent extraction across EHR environments. Finally, data should not require changes to current clinical practice or workflows.

We used all of this information to create a modified set of feasibility criteria divided into two groups: 1) data capture criteria related to how, with what consistency, and in what format data are entered into electronic health systems; and 2) data extraction criteria related to the ease of consistently extracting data from electronic health system databases. These criteria are listed below:

#### **Data Capture Criteria**

- Obtained consistently under current practice  
Routinely collected for patients admitted to the hospital under current clinical practice and EHR workflows
- Captured with a standard definition  
Consistent conceptual understanding, method of collection, and units of measurement
- Entered in a structured field  
Captured in numerical, pseudo-numerical, or list format

#### **Data Extraction Criteria**

- Encoded consistently  
Can be linked to a standard and uniform coding structure such as ICD-9 or



## Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT)

- Extractable from the EHR  
Can be readily and consistently identified and exported from current EHR databases
- Exported with metadata  
Additional information such as time stamps and reference values that are needed for interpretation are consistently available

Our criteria and review process are meant to evaluate current [data availability](#) and extractability. However, the evaluation process can be reapplied to clinical data elements in the future as electronic health system use and functionality changes over time.

### 3.2.3. *Application of Criteria to Data Categories and Subcategories*

In alignment with the CMS Measures Management System (MMS), CORE convened a TEP to complete an electronic health system data feasibility survey. The TEP members represented EHR vendors, HIT companies, hospital systems, quality improvement organizations, and specialty societies. CORE designed this survey to elicit the expert input of TEP members' about whether data in the QDM meet the data capture and data extraction criteria for adult hospitalized patients. The survey asked TEP members to apply both sets of criteria at the QDM [subcategory](#) level. For example, in the *laboratory test* category, TEP members applied the criteria to each of the six subcategories. Soliciting the TEP's expert rating at the subcategory level allowed for the collection of more granular information than we would collect if rating at the category level. Surveying at this level was also less burdensome than surveying at the individual data element level which would require analysis of thousands of data elements.

We designed the survey using a slightly modified version of the December 2012 version of the QDM. Specifically, we divided the physical examination category into three categories for more specific input: *vital signs*, *neurological assessment*, and *other*.

The final data feasibility survey contained 21 data categories and 85 subcategories ([Appendix A1](#)). TEP members were first asked to consider whether at least one data element contained in each subcategory currently meets all three [data capture criteria](#) for adult hospitalized patients. These criteria were grouped together because data elements that do not meet any one of these criteria are not feasible for use in quality measurement. If TEP members indicated that a subcategory currently meets the data capture criteria, they were asked to indicate all [data extraction criteria](#) that are met by at least one data element within that subcategory. Conversely, if no data elements met all of the data capture criteria, TEP members were asked to consider the future feasibility of that subcategory by answering whether elements would likely meet the data capture criteria in 2014, when Stage 2 of [Meaningful Use](#) will be implemented; 2016, when Stage 3 of Meaningful Use will be implemented; 2018, well into Meaningful Use implementation; or never.

After receiving an orientation to this project and to the survey, TEP members completed the survey independently and had the opportunity to provide additional comments and rationales for each response. TEP members also had the opportunity to skip questions that they did not feel they had the expertise to answer. Due to the length of the survey, we placed seven categories at the end of the survey that the CORE team and consultants identified as unlikely to meet the data capture

criteria and not high priority for risk adjustment. The TEP was given the option of assessing these categories only if they thought they met the feasibility criteria. These categories were *care goal*, *care experience*, *communication*, *intervention*, *risk category assessment*, *symptom*, and *system characteristics*.

### 3.3 Results

#### 3.3.1. Survey Results

CORE solicited potential TEP members through a posting on the CMS [Quality Measures Public Comment Page](#). The final list of TEP members and their affiliations is given in [Appendix A2](#). Of the 18 experts selected to serve on this TEP, 16 completed the survey in the time provided. Survey results were transmitted electronically and tabulated.

Based on the responses to the data capture questions, categories and subcategories were classified based on the level of agreement in scoring. Agreement was only considered for subcategories that had at least nine responses, or at least half of TEP members responding. The agreement classification groups are listed below:

- Complete Agreement on Feasibility: 100% of respondents agreed that at least 1 data element within the subcategory met all 3 data capture criteria ([Table 3.1](#))
- Strong Agreement on Feasibility: 70% or greater of respondents agreed that at least 1 data element in the subcategory met all 3 data capture criteria ([Table 3.2](#))
- No Agreement: Fewer than 70% of respondents agreed on data capture criteria ([Table 3.3](#))
- Strong Agreement on Infeasibility: 70% or greater of respondents agreed that no data elements within the subcategory met all 3 data capture criteria ([Table A2](#))
- Complete Agreement on Infeasibility: 100% of respondents agreed that no data elements within the subcategory met all 3 data capture criteria ([Table A3](#))

Patient characteristics that are collected upon admission to the hospital, such as date of birth, insurance payer, and sex, were unanimously scored as meeting the data capture criteria. Most TEP members also agreed that these data elements meet the data extraction criteria.

The TEP also consistently scored the subcategory *order* as meeting the data capture criteria across categories, with the exception of the *physical examination* categories. Among the data extraction criteria, TEP members consistently responded that orders would be easily extracted and would contain the necessary metadata, but they highlighted that they may not be consistently encoded. One TEP member explained that orders are coded for short-term use by each organization. Since there are no incentives or guidelines for standardization, few hospitals have spent their resources encoding orders.

The TEP consistently scored the subcategories of *recommended*, *intolerance*, *adverse event*, and *allergy* as not meeting the data capture criteria across categories. Most TEP members agreed that these subcategories of data are not consistently obtained and are typically captured as unstructured data. Agreement was not reached on the future feasibility of these data, which may be due to the fact that TEP members' opinions on future feasibility is largely dependent on each respondent's confidence in the progression of [natural language processing](#).



**Table 3.1: Complete Agreement on Feasibility (100%)**

Category	Subcategory	Met Data Capture Criteria (yes/no)
2. Individual Characteristics	2.1 Birth Date	16/0
2. Individual Characteristics	2.4 Payer	16/0
2. Individual Characteristics	2.5 Sex	16/0

**Table 3.2: Strong Agreement on Feasibility (≥70%)**

Category	Subcategory	Met Data Capture Criteria (yes/no)
1. Encounter	1.2 Order	12/4
1. Encounter	1.3 Active	12/3
1. Encounter	1.4 Performed	14/2
2. Individual Characteristics	2.2 Expired	13/2
2. Individual Characteristics	2.6 Ethnicity	12/2
2. Individual Characteristics	2.7 Race	12/2
3. Transfer	3.1 To (another facility)	12/2
8. Laboratory Test	8.2 Order	13/3
8. Laboratory Test	8.3 Performed	11/4
8. Laboratory Test	8.4 Result	15/1
9. Diagnostic Study	9.2 Order	14/2
9. Diagnostic Study	9.3 Performed	11/3
11. Procedure	11.2 Order	12/3
12. Device	12.2 Order	11/2
13. Medication	13.1 Order	15/1
13. Medication	13.3 Administered	12/4
13. Medication	13.8 Discharge	11/4
14. Substance	14.1 Order	12/4

**Table 3.3: No Agreement**

Category	Subcategory	Met Data Capture Criteria (yes/no)
2. Individual Characteristics	2.8 Patient Characteristic: Other	9/6
3. Transfer	3.2 From (another facility)	4/8
4. Physical Examination: Vital Signs	4.2 Order	8/6
4. Physical Examination: Vital Signs	4.3 Finding	11/5
5. Physical Examination: Neurological Status	5.2 Order	10/6
5. Physical Examination: Neurological Status	5.3 Finding	9/6
6. Physical Examination: Other	6.3 Finding	4/8
7. Functional Status	7.3 Result	7/8
9. Diagnostic Study	9.4 Result	7/8
10. Condition/Diagnosis/Problem	10.1 Active	11/5
10. Condition/Diagnosis/Problem	10.3 Resolved	5/11
11. Procedure	11.3 Performed	11/5
11. Procedure	11.4 Result	7/8
12. Device	12.3 Applied	9/4
13. Medication	13.2 Dispensed	9/5
13. Medication	13.4 Administered	12/4
13. Medication	13.5 Allergy	10/6
13. Medication	13.6 Intolerance	7/9
14. Substance	14.3 Administered	5/11
14. Substance	14.5 Allergy	6/10
15. System Characteristics	15.1 System Characteristics	0/7

Category	Subcategory	Met Data Capture Criteria (yes/no)
16. Symptom	16.1 Active	5/7
16. Symptom	16.2 Resolved	4/8
16. Symptom	16.3 Assessed	4/7
17. Risk Category Assessment	17.1 Risk Category Assessment	4/6
18. Care Goal	18.1 Care Goal	3/6
19. Intervention	19.2 Order	6/4
19. Intervention	19.3 Performed	5/5
19. Intervention	19.4 Result	5/5
20. Communication	20.2 From Provider to Patient	4/6

### 3.3.2. Discussion of Survey Results

A second call with the TEP members was convened to discuss the results of the survey and to gain clarity on scoring decisions for categories and subcategories of importance in quality measurement. A summary of this discussion can be found in [Appendix A4](#).

This discussion had three goals. The first goal was to achieve consensus, where possible, about which subcategories of data met the three data capture criteria. The second goal was to gain further input on ease of extraction based on the data extraction criteria. This part of the discussion addressed pragmatic issues related to mapping and extracting data for reporting. This focused on how consistently data elements are associated with standard codes and how consistently data are linked with the metadata needed for data analysis, such as time and date stamps. Each of these data extraction issues will be important in addressing the potential pathways to implementation of hybrid outcome measures. The third goal was to discuss the future feasibility of data relevant to hospital outcomes.\*

The survey, discussion, and consensus-building approach were designed as a replicable and transparent process for obtaining stakeholder input on the feasibility of electronic health data. The data subcategories selected for discussion were identified by the CORE team as those most relevant to risk adjustment of 30-day mortality for adult hospitalized patients. The selection was based on consultation with experts, results of the TEP survey, and a literature review to identify variables commonly included in mortality prediction scales or indices, as discussed further in [Section 4.2.1](#).

The TEP discussion of each category and subcategory is summarized below. For more details, please refer to [Appendix A4](#).

- i. *Encounter*
  - a. *Performed*

For the purposes of the TEP discussion, the subcategory *encounter performed* was defined as a hospital admission. The majority of TEP members agreed that the appearance of a bill in a hospital's [administrative database](#) to indicate a completed encounter is consistently captured and extractable from all hospitals' electronic administrative or billing databases. Several members of the TEP noted that time and date stamps are used to bill for hospitalizations, and hospitals are required to file the UB-04 (Universal Bill 2004) Medicare claims forms for payment. Therefore, time and date stamps are two of the most reliable electronic data elements

---

\* Subcategories that were consistently identified as not meeting the data capture criteria (70% agreement) were excluded from this discussion.

associated with hospitalization. The TEP reached consensus that administrative and billing databases remain the most feasible and most accurate source of information about the occurrence of a hospital admission.

The TEP members acknowledged that the time and date stamps associated with administrative data elements may not accurately reflect the time a patient spent at the facility receiving care. Patients often enter the hospital through different departments and receive treatment prior to official hospital admission. TEP members also noted that the exact start and end times of a hospital admission indicating when care is actually initiated and when care ends lack a standard definition, and they are most often captured in electronic patient management systems where the method of capture and the exact data elements used varies across hospitals, departments within a single hospital, and software systems.

*b. Transfer to and transfer from*

The TEP members reached broad agreement that data elements related to patient transfers to another facility would meet the data capture and extraction criteria. The current payment system incentivizes the capture of information related to patient discharge and discharge disposition. However, they noted the lack of standardized coding systems for identifying the type of facility a patient transfers to upon hospital discharge.

Information related to patients transferred from another facility is not consistently captured or entered into structured fields because no incentives currently exist for hospitals to systematically record that a patient arrival or admission is due to a transfer from an outside facility. TEP members reached consensus that the subcategory *transfer from* is not feasible for quality measurement at this time.

*ii. Physical Examination: Vital Signs Finding*

The TEP agreed that several data elements in the subcategory *vital signs finding* currently meet the data capture and extraction criteria when captured at specific time points during a hospital admission. For example, vital signs measurements performed at triage in the ED, upon admission to an intensive care unit (ICU), or upon admission to the inpatient floor are consistently captured using a standard definition and recorded in structured fields.

However, survey scores for *vital sign finding* overall were just below the 70% threshold for agreement. During the discussion, TEP members clarified that many instances of *vital sign finding* failed to meet the standard definition criterion because of the different modes of data measurement that can be used in various clinical situations and the failure to consistently capture metadata describing the details of measurement. They provided heart rate as an example of a vital sign that can be collected as a matter of routine surveillance, or in response to a patient's complaint or clinician's concern. Heart rate values may be collected at the bedside when a patient is at rest or during a cardiac stress test. The TEP members agreed that most EHRs currently lack the capacity to capture these descriptors of how, why, and under what clinical circumstances these data are measured. The TEP members noted that although the current coding system captures the most general concept of a vital sign finding, there is no structure within that coding system for these nuances of measurement.

The TEP reached agreement that the subcategory *vital sign finding* contains feasible data elements for quality measurement if time stamps are present and a temporal relationship with

the start of the admission can be determined.

*iii. Physical Examination: Neurological Assessment*

The TEP agreed that neurological assessments do not currently meet the data capture criteria and are not currently feasible due to the lack of an assessment of neurological status that is captured for all or most adult inpatients. The TEP noted that certain standard assessments, such as the Glasgow Coma Score (GCS), are routinely captured for trauma and ICU patients and are feasible among these specific patient populations. In addition, the TEP members remarked that non-standardized neurological assessments performed by clinicians are typically recorded in text and would require natural language processing for analysis. They did not recommend the use of clinical data elements in the category *physical examination: neurological assessments* for this project.

*iv. Laboratory Test Result*

The TEP members universally agreed that data elements in the subcategory *laboratory test result* currently meet the data capture criteria and can be extracted for use in quality measures. The TEP members also discussed the feasibility of laboratory test reference values to denote thresholds for normal and abnormal results. They noted that these data are typically captured as numerical data in clinical laboratory databases but are nearly universally exported as text or string data into EHRs or other software. This remains a barrier to using threshold or reference values to interpret laboratory test results and to standardizing results across hospitals and hospital systems. The TEP noted that these barriers apply only to tests that have thresholds or normal ranges that vary across laboratories, specific methods of testing, or clinical populations.

*v. Diagnostic Study Order and Diagnostic Study Result*

Most TEP members agreed that data elements in the subcategory *diagnostic study order* meet the data capture criteria. However, these orders are not likely to be associated with a standard set of codes and may not be as informative as data elements in the subcategory *diagnostic study results*, which are most commonly captured as text or string data and are not currently feasible. They did not recommend the use of clinical diagnostic study results for use in this project.

*vi. Condition/Diagnosis/Problem*

The TEP discussion of this category focused on the concepts of principal discharge diagnosis, primary admitting diagnosis, and secondary or comorbid conditions. The discussion addressed both administrative and clinical sources of electronic data.

Most TEP members agreed that the primary admitting diagnosis and secondary or comorbid conditions that are captured by clinicians in the clinical EHR are not currently feasible as they are not consistently captured, often lack a standard definition, and are often entered as string or text data in notes rather than as structured or encoded data on problem lists. In addition, only those comorbid diagnoses thought to be important for care are documented.

TEP members also noted that a clinician's problem list is focused on their estimation of the most important problem at the moment of care and documentation. This may not consistently identify the principal discharge diagnosis because these opinions often change over the course of an admission and from one clinician to the next.

The TEP agreed that conditions are feasible and encoded with standard ICD-9 codes on the administrative problem list used for billing. The conditions are assigned retrospectively based on abstracted data and on payment incentives. The TEP concluded that administrative or billing databases remain the most reliable and appropriate source of data related to conditions for use in quality measurement at this time.

#### *vii. Medication Order, Medication Administered, and Medication Discharge*

During the discussion, TEP members suggested that medications may not be consistently encoded using standard RxNorm codes by hospitals but could easily be mapped according to RxNorm standards. They noted that this would be fairly straightforward for data elements in the *medication order* subcategory.

However, the TEP raised concerns about data elements in the subcategories *medication administered* and *medication discharge*. Due to the complexity of metadata associated with administering medications over the course of a hospital admission, such as dosing schedules, routes of delivery, and changes during the course of care, standardizing these data for use in quality measurement would be difficult.

Similarly, several TEP members noted that discharge medications are often located in the component of the electronic health systems designated for prescriptions. However, prescriptions provided at discharge are often an incomplete list of all medications patients are instructed to take at home. They noted that medication reconciliation standards might improve the quality of these lists in the future. Despite these concerns, the TEP agreed that data elements in the subcategories *medication ordered*, *administered*, and *discharge* are feasible and could be considered for use in this project.

### *3.3.3 Final Subcategories Selected for Further Feasibility Testing*

For this project, we sought to assess the feasibility of data subcategories that contain data elements that may be particularly useful as risk adjusters for 30-day mortality measures. Specifically, we sought data that are captured during most adult hospital admissions regardless of their principal discharge diagnosis. For example, the TEP members agreed that GCS is routinely captured with a standard definition in a structured field for patients treated for trauma or for those admitted to the ICU. However, these data only meet the feasibility criteria for the smaller population of trauma and ICU patients.

In addition, some data subcategories were deemed feasible but were not carried forward for further feasibility testing due to the availability of a more appropriate source of information for risk adjustment. For example, the survey results indicated that data elements in the subcategories *laboratory test order* and *performed* are consistently captured with a standard definition and in a structured format. However, the *laboratory test result*, if available, is preferable for risk adjustment. A result indicates that a test was both ordered and performed; therefore, the *laboratory test result* subcategory was chosen for inclusion over the subcategories *laboratory test ordered* and *performed*.

Several additional subcategories that had 70% TEP agreement on feasibility were not carried forward for the next phase of feasibility testing because they were not appropriate or relevant for risk adjustment. For example, the subcategories *ethnicity* and *race* were not considered for feasibility testing because risk models do not include variables that may mask disparities in care based on race or socioeconomic status. The subcategory *transfer to* was also not considered for this

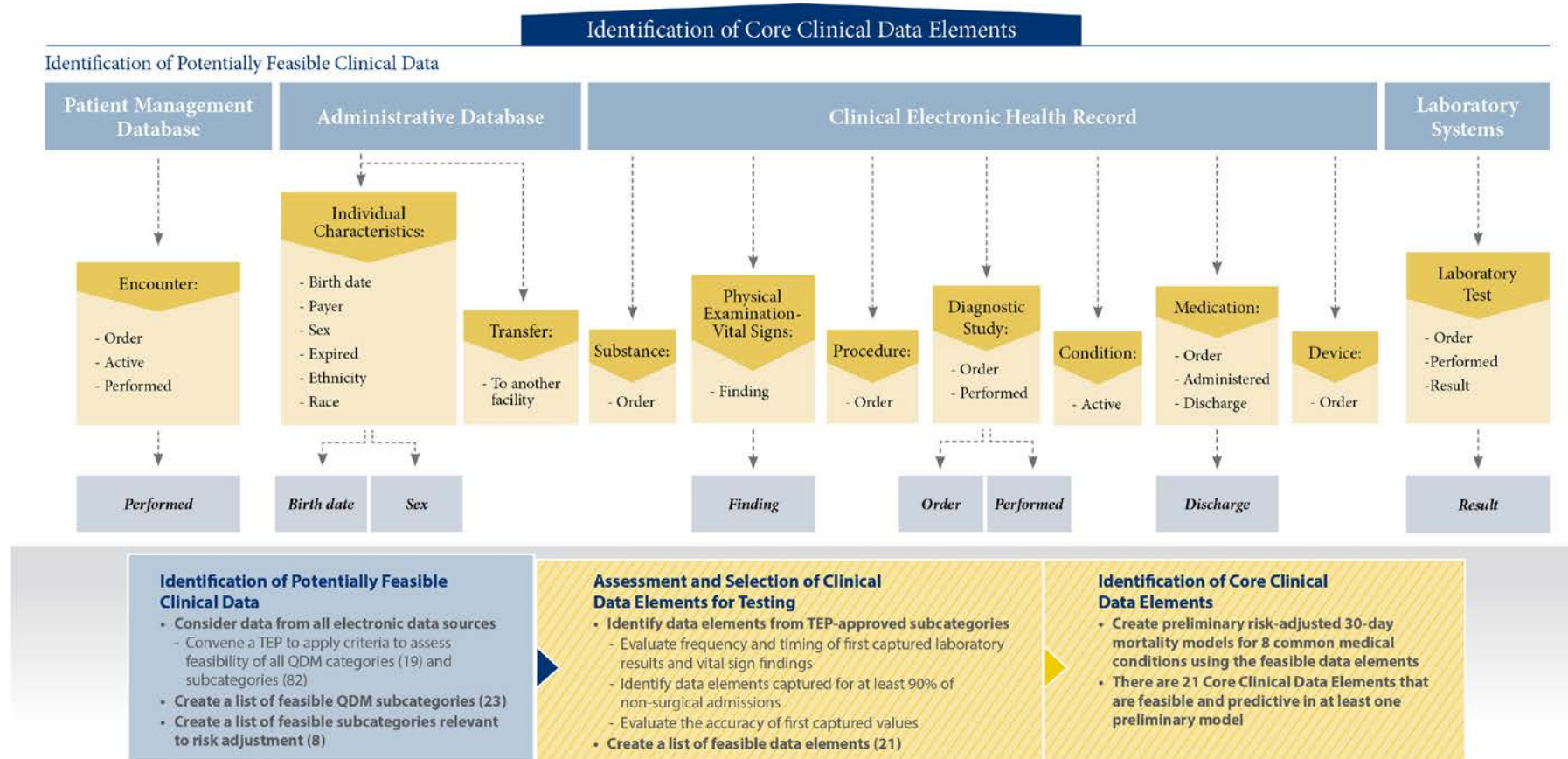
project, as this is not a data element used in hospital mortality measures for cohort selection or risk adjustment. The subcategories *medication order* and *administered* within the *medication* category were not carried forward as these were deemed to be inappropriate for risk adjustment by the CORE team because they are components of treatment and do not reflect a patient's clinical status before treatment is administered.

Based on these considerations and the recommendations of the TEP, the following EHR data subcategories were identified as feasible and as candidates for inclusion in the final clinical dataset:

- *Encounter Performed*
- *Birth date*
- *Sex*
- *Vital Signs Finding*
- *Diagnostic Study Order*
- *Diagnostic Study Performed*
- *Medication Discharge*
- *Laboratory Test Result*

For more details on the identification of the potentially feasible clinical data, please refer to [Figure 3.1](#).

**Figure 3.1: Identification of Potentially Feasible Clinical Data**



## 4. ASSESSMENT AND SELECTION OF DATA ELEMENTS FOR TESTING

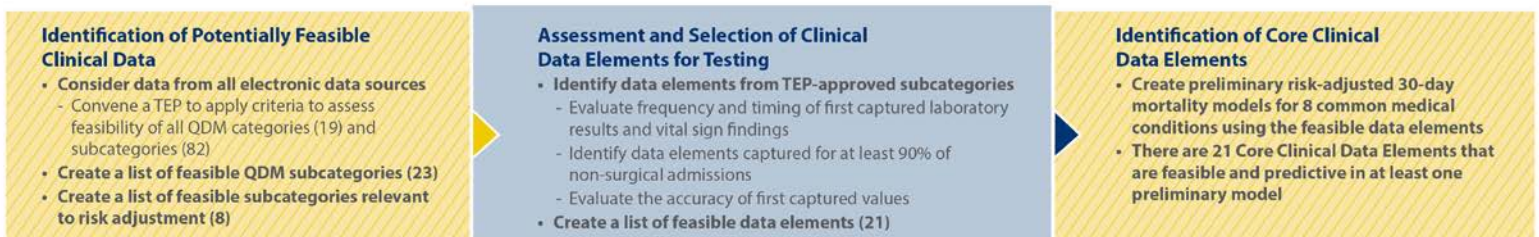
### 4.1 Approach

This section describes a series of analyses performed to identify a final list of core clinical data elements for testing in risk-adjusted mortality measures. We aimed to identify data elements from the subcategories identified as feasible by the TEP that are consistently captured and relevant to 30-day mortality across a broad variety of medical conditions following hospital admission. A two-part approach was used to ensure that feasible, accurate, and relevant data elements were identified. The first was a systematic literature review to identify clinical data elements commonly included in validated mortality indices for adult hospitalized patients. The second was a direct analysis of clinical data elements from the EHR database of the 21 KPNC member hospitals to assess the consistency and timing of data capture and data accuracy. We also conducted a survey of several health systems that are members of Premier Inc., a collaborative healthcare alliance focused on measuring and improving their members' quality outcomes and safely reducing costs. The purpose of the survey was to verify that the data elements identified through analysis of KPNC data are also consistently captured and extractable by other providers using alternate EHR systems.

We selected data elements for feasibility testing that were included among the data subcategories recommended by the TEP and present in validated mortality prediction indices identified through a systematic review of the literature. All data elements selected were found to be extractable from the KPNC EHR. A direct analysis of the selected clinical data elements was performed to verify that they are consistently captured during all hospital admissions for adult patients as well as for specific common medical conditions. The analysis included evaluation of the data element format, consistency and timing of data capture, and data accuracy.

This issue of timing of capture is important for risk adjustment because patients' clinical status must be assessed before the effects of treatment are realized. For this reason, the time and date stamps that indicate when a patient first arrived at the hospital for care, as well as the time and date stamps associated with capture of each data element, are required to be reported along with the data element values for inclusion in our final dataset. Using the time stamps, the CORE team of clinical and performance measurement experts examined the frequency of capture for each data element during the first 24 hours of hospitalization and determined the most appropriate window for inclusion of [first captured](#) values in risk adjustment models. These analyses were performed for all admissions as well as for each of the 21 hospitals, for each [hospital entry location](#) or area within each hospital where care was initiated, and for several common discharge diagnoses. The final list of first captured data elements that were shown to be consistently available and accurate were considered for testing in statistical models ([Figure 4.1](#)).

Figure 4.1: Project Overview





## 4.2 Methodology

### 4.2.1. Literature Review to Identify Common Predictors of Mortality across Conditions

A number of indices have been developed to predict mortality in adult hospitalized patients. These indices have been used for clinical research and to support clinical care of patient populations treated in a variety of settings including EDs, ICUs, surgical units, and general and specialty medical care inpatient floors. To ensure that all relevant data elements were included in the feasibility analysis, we performed a systematic review of the literature describing validated mortality prediction indices for adult hospitalized patient populations. We focused on identifying indices that were not specific to any single condition as we sought to identify clinical data elements that predict 30-day mortality following hospitalization for a broad spectrum of medical diagnoses.

The review yielded 15 indices that met our criteria. Refer to [Appendix B1](#) for more detailed methodology and the complete list of mortality prediction indices. Source documents describing the derivation and validation of the 15 scales were examined to identify data elements included in each index. Variables in these indices were sorted into the data categories and subcategories: *patient characteristics* ([Table B.1](#)), *physical examination* ([Table B.2](#)), and *laboratory results* ([Table B.3](#)). We found a great deal of overlap of the clinical data elements included among these indices and within the categories of clinical data identified as feasible by the TEP, such as age, sex, common laboratory test results, and vital sign findings. Although several mortality indices include information regarding conditions and procedures, we do not list them in this report because we focused on identifying clinical data elements that can be used as risk adjusters in place of the administrative data and the TEP did not recommend identifying conditions or procedures from clinical records at this time. Five of the 15 indices included patients' medications. However, no two scales included the same class of medications and no single index included more than two classes of medications. Due to this lack of consistency, medications were not included in the list of possible data elements for risk adjustment identified by the literature review.

The *patient characteristic* data element most commonly included across mortality prediction scales was age. The most common *physical examination* data elements were vital sign findings for heart rate and blood pressure. The most common *laboratory test result* data elements were creatinine, blood urea nitrogen, bilirubin, white blood cell count, and platelets. All of the laboratory tests and vital sign findings from the scales were included in our feasibility analyses. Although the GCS is commonly included in mortality prediction indices, this data element was not included in our analyses. This was due to the TEP's consensus that the data are not consistently captured across most adult inpatients and are not consistently captured as [structured data](#).

### 4.2.2. Feasibility Testing in a Clinical Dataset

#### *Data Extraction*

KPNC is an integrated health care delivery system that serves over 3.3 million members and utilizes a fully automated EHR in all of its 21 acute care hospitals. KPNC has worked to develop an extensive clinical risk-adjustment methodology for internal benchmarking and quality assurance and is in the process of developing the capability to use these clinical data in real time for clinical decision support and quality measurement. Their work has required mapping specific clinical data elements within their databases, extracting data, and validating their source and accuracy. We used only those data that had already been mapped and extracted by the KPNC team.

All KPNC hospitals use an integrated Epic system to capture and store patient management, administrative, and clinical data in their outpatient and inpatient healthcare settings. We partnered with KPNC to identify data elements in the feasible QDM categories that may be commonly captured and important for risk adjustment. For this work, we asked KPNC to provide a cohort of all adult patients with a hospital admission discharged between January 2010 and December 2011 from any of their member hospitals. In this cohort, all data elements are linked to a single hospital admission using a unique encounter identification number. Any individual patient may have one or more admissions in the database. We asked the KPNC team to extract all possible data elements for each hospital admission in the categories *individual characteristics*, *encounter* (start and stop times), *physical examination: vital signs*, and *laboratory results*.

The list of vital signs and laboratory results that were mapped and extractable are listed in the [Table 4.1](#) and [Table 4.2](#).

**Table 4.1: Readily Extractable Vital Signs from KPNC Databases**

Vital Sign	Form
Diastolic blood pressure	Numeric
Systolic blood pressure	Numeric
Pulse pressure	Numeric
Height	Numeric
Heart rate	Numeric
Respiratory rate	Numeric
Oxygen saturation	Numeric
Temperature	Numeric
Urine output	Numeric
Weight	Numeric
Oxygen flow	Numeric
Fraction inspired oxygen (FiO2)	Numeric
Oxygen mode: High flow nasal cannula	Text
Oxygen mode: Mask	Text
Oxygen mode: Nasal cannula	Text
Oxygen mode: Nasal continuous airway pressure	Text
Oxygen mode: Tracheal collar	Text
Oxygen mode: Ventilator	Text

**Table 4.2: Readily Extractable Laboratory Results from KPNC Databases**

Lab Type	Form
Albumin	Numeric
Anion gap	Numeric
Blood culture	Numeric
Bicarbonate	Numeric
Total serum bilirubin	Numeric
Blood urea nitrogen (BUN)	Numeric
Clostridium difficile (C. difficile)	Numeric
Creatinine	Numeric
Cerebrospinal fluid (CSF) culture	Numeric
Glucose	Numeric
Hematocrit	Numeric

Lab Type	Form
International normalized ratio (INR)	Numeric
Lactate	Numeric
Arterial partial pressure of carbon dioxide (PaCO2)	Numeric
Arterial partial pressure of oxygen (PaO2)	Numeric
Arterial pH	Numeric
Venous pH	Numeric
Platelet	Numeric
Sodium	Numeric
Troponin	Numeric
White blood cell (WBC) count	Numeric

All of the data elements from subcategories approved by the TEP and listed in mortality predictions scales were mapped and extracted from the KPNC dataset. Extractable data elements that had not been previously identified by the literature review included measures of oxygen flow, blood cultures, cerebrospinal fluid cultures, and lactate and bicarbonate labs, among others. The three potentially feasible and relevant subcategories identified by the TEP that did not have readily mapped data in the KPNC database were *diagnostic order* and *performed* as well as *medication discharge*. However, diagnostic studies and medications were seldom included in the mortality indices. For this reason we did not pursue further assessment of data elements in these subcategories.

Finally, clinicians were consulted to ensure that all common vital signs and laboratory results that would normally be collected on adults during a hospital admission were identified by the literature review and extracted into the KPNC dataset. There were a few labs that clinicians agreed would normally be ordered as part of a panel but that were not included in the validated mortality prediction indices, possibly due to collinear relationships among related data elements. However, we wanted to extract and examine all data elements from the same panel of tests. For example, aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase, and total protein may typically be ordered with albumin and bilirubin as part of a liver function panel. We asked KPNC to extract results from the entire panel for testing.

For each data element, analyses were conducted to evaluate the frequency of capture and timing relative to the [time of arrival](#) at the hospital for care as well as the accuracy of the values for each data element. We also examined the frequency of capture and timing of each data element for individual KPNC member hospitals, hospital entry location, and common conditions among hospitalized adults. Mortality models were constructed for these same conditions to test the data elements after the feasibility analysis ([Section 5](#)).

#### *Cohort derivation*

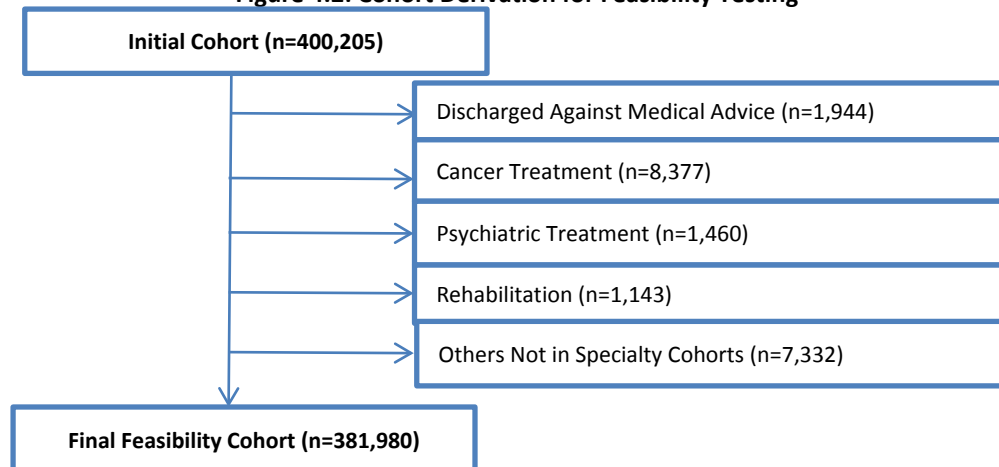
Admissions were divided into 5 [specialty cohorts](#) based on those utilized for the publicly reported [30-Day Hospital-Wide All Cause Readmission Measure](#) using the principal discharge diagnosis ICD-9 codes in the KPNC dataset ([Table 4.3](#)). Agency for Healthcare Research & Quality (AHRQ) [Clinical Classification Software \(CCS\)](#) categories were used to select ICD-9 principal discharge diagnosis codes associated with each of five specialty cohorts. These cohorts combine AHRQ CCS conditions that are typically cared for by the same team of clinicians. This methodology was employed for this work to evaluate data capture and accuracy across a representative and diverse set of common

medical conditions. Admissions that did not fall into any of these specialty cohorts, including cancer treatment, rehabilitation, or psychiatric therapy, were excluded. Additionally, patients discharged against medical advice were excluded because they do not receive the full medical treatment recommended for their conditions which could include the assessment of vital signs and laboratory tests. Details of the cohort derivation are shown in [Figure 4.2](#).

**Table 4.3: Admissions by Specialty Cohort**

Specialty Cohort	Number of Admissions
Cardiorespiratory	28,076
Cardiovascular	26,538
Medicine	129,308
Neurology	14,549
Surgery/gynecology	183,509

**Figure 4.2: Cohort Derivation for Feasibility Testing**



The final cohort contained 381,980 admissions. Patients in this cohort had a mean age of 58 with a standard deviation of 21 years. The cohort was 62.6% female.

We identified cohorts that could be used to assess data capture and then also used to build condition-specific mortality models. We wanted these cohorts to be common and broadly representative of a range of medical conditions, so we identified the most common medical conditions<sup>†</sup> in each of the specialty cohorts, with the exception of surgery/gynecology, which is procedure- and not condition-based ([Table 4.4](#)). The two most common conditions in the cardiorespiratory specialty cohort were congestive heart failure (CHF) and pneumonia (pneum). The two most common conditions with the cardiovascular cohort were atherosclerosis (athero) and other heart disease and cardiac dysrhythmias (arrhyth). We also included AMI from this group because it is a common and costly condition for which there is a publically reported mortality measure. The two most common medical conditions were septicemia (sepsis) and diabetes mellitus with complications (DM). The two most common neurological conditions were acute cerebrovascular disease (stroke) and other nervous system disorders. We eliminated the second condition, other nervous system disorders, as it is non-specific and had many fewer admissions

<sup>†</sup> Medical conditions are defined here by the Agency for Healthcare Research & Quality (AHRQ) Clinical Classification Software (CCS), which groups related ICD-9 codes.

compared with the other included conditions.

**Table 4.4: Population Statistics by Condition**

Cohort	# of Episodes (%)	Age Mean (SD)	Female %
All	381,980 (100.0)	58.3 (± 20.8)	239,248 (62.6)
Congestive heart failure (CCS 108)	10,494 (2.7)	76.0 (± 13.5)	5,242 (50.0)
Pneumonia (CCS 122)	6,676 (1.7)	71.9 (± 16.7)	3,438 (51.5)
AMI (CCS 100)	6,193 (1.6)	69.4 (± 14.1)	2,357 (38.1)
Coronary atherosclerosis (CCS 101)	6,632 (1.7)	66.4 (± 11.8)	1,995 (30.1)
Cardiac dysrhythmias (CCS 106)	6,427 (1.7)	71.0 (± 14.0)	3,259 (50.7)
Septicemia (except during labor) (CCS 2)	30,127 (7.9)	70.8 (± 16.5)	15,611 (51.8)
Diabetes mellitus with complications (CCS 50)	6,098 (1.6)	57.9 (± 17.9)	2,771 (45.4)
Acute cerebrovascular disease (CCS 109)	7,328 (1.9)	73.4 (± 13.9)	3,982 (54.3)

#### *Assessment of the Consistency and Timing of Data Capture*

Patient characteristics, specifically age and gender, are captured in the administrative database, a part of the KPNC-integrated Epic software system. These two variables were captured on all patients in this cohort.

The consistency of capture for other data elements was slightly more variable and required consideration of timing. Unlike patient characteristics, which are static, the timing of capture for vital signs and laboratory results, which can take place multiple times during an admission, is critical for risk adjustment of hospital outcome measures. Values captured at or near the time of first arrival for each hospital admission best reflect a patient's clinical status before treatment is initiated. For this analysis, it was important to identify first captured vital signs and laboratory results and to eliminate all other repeated values from each admission.

In the patient management system, a time stamp is captured for each patient when they first register for care in the ED. In addition, the patient management system captures a time stamp when a patient first enters an inpatient location other than the ED, such as the ICU, the operating room, inpatient floor, and transitional care units. We identified the time of arrival for each hospital admission and each entry location using these time stamps.

Next, we assessed the proportion of hospital admissions with each vital sign finding captured at 2 hours, 6 hours, 12 hours, and 24 hours from the time of arrival. This analysis was repeated for each of the 8 selected medical conditions. We also evaluated the proportion of hospital admissions with each laboratory test results at 2 hours, 6 hours, 12 hours, and 24 hours from the time of arrival at the hospital. This analysis was performed for all 8 medical conditions.

#### *Variation across Hospitals and Hospital Units*

Variation in the consistency and timing of data capture was evaluated across different first hospital entry locations and across different hospitals. These evaluations show how differences in patterns of clinical care among hospitals and departments impact the consistency and timing of capture for

laboratory test results and vital sign findings.

First, the proportion of hospital admissions beginning at each hospital entry location was calculated for the overall cohort and each condition specific cohort. Then we evaluated the proportion of hospital admissions with first captured laboratory test results and vital signs in each of the hospital entry locations in each of the 21 KPNC hospitals at 2 hours, 6 hours, 12 hours, and 24 hours.

#### *Assessment of Data Accuracy*

The distribution of values for individual data elements was assessed in the full cohort to determine the accuracy of the data. We examined the range of values between the 1st and 99th percentiles relative to normal values to ensure that they fell within physiologically plausible values. We also examined values at the extremes of the distribution ( $\leq 1$ st and  $\geq 99$ th percentiles) compared to normal ranges for each data element to determine if they were physiologically plausible or likely to be errors of measurement, documentation, or transfer of data. This analysis was performed for each of the 8 conditions, each hospital, and each hospital entry location to determine if errant values were more likely to occur in certain hospitals, certain hospital locations, or for certain conditions.

### **4.3 Results**

#### *4.3.1. Consistency and Timing of First-Captured Value*

Basic vital signs, including blood pressure, heart rate, respiratory rate, and temperature, were consistently captured in >90% of inpatient admission within 2 hours of first arrival ([Table 4.5](#)). For this reason, 2 hours was selected as the timeframe for obtaining first vital signs for risk adjustment. Patient weight was not captured in 90% of admissions until 24 hours after first arrival. Since weight is not expected to change dramatically during the first day of most admissions, this was deemed an acceptable window for capture. Oxygen saturation was also not captured in 90% of all admissions within 2 hours; however, it was captured within 2 hours in 90% of admissions for each of the 8 conditions ([Table 4.6](#)). Other extracted vital signs, related to aspects of respiratory support, did not meet the 90% threshold for capture within 2 hours of arrival for the overall cohort or for all eight medical conditions.

There was no significant variation in consistency or timing of capture of the basic vital signs across conditions ([Table 4.6](#)).

**Table 4.5: Proportion of Episodes with Captured Vital Signs at Various Time points**

Vital Sign Finding – Full Cohort	Total with Finding and Timestamp %	Within 2 Hours %	Within 6 Hours %	Within 12 Hours %
<b>Basic vital signs</b>				
Heart rate	99.7	96.8	99.4	99.6
Systolic blood pressure	99.7	96.7	99.3	99.6
Diastolic blood pressure	99.7	96.7	99.3	99.6
Respiratory rate	99.7	95.8	99.1	99.6
Temperature	99.7	93.7	98.5	99.5
Oxygen saturation	98.2	86.0	92.6	95.4
Weight	92.5	80.2	85.2	88.8
<b>Other vital signs</b>				
Pulse pressure	99.7	96.7	99.3	99.6

Vital Sign Finding – Full Cohort	Total with Finding and Timestamp %	Within 2 Hours %	Within 6 Hours %	Within 12 Hours %
Room air	96.3	65.8	74.7	81.7
Urine output	92.6	7.4	39.6	72.7
Height	80.2	47.2	58.1	68.3
Flow	71.7	21.8	44.8	58.0
Nasal cannula	59.8	18.9	36.0	48.4
Mask	30.5	0.9	12.6	21.4
FiO2	17.7	3.5	5.6	8.5
Nasal continuous positive airway pressure	6.3	1.3	2.0	3.0
High flow nasal cannula	5.3	2.1	2.9	3.4
Vent	3.6	0.7	1.2	1.9

**Table 4.6: Proportion of Admissions with at least 90% Captured Vital Signs at 2 Hours for the Common Medical Conditions**

	CHF	Pneum	AMI	Athero	Arrhyth	Sepsis	DM	Stroke
Heart rate	98.3	98.2	97.7	96.8	97.9	98.4	97.4	98.2
Diastolic BP	98.2	98.0	97.7	97.1	97.9	98.2	97.3	98.0
Systolic BP	98.2	98.0	97.7	97.1	97.9	98.2	97.3	98.0
Respiratory rate	97.6	97.3	96.5	96.1	97.1	97.4	96.5	97.1
Temp	94.0	95.8	92.9	93.9	93.2	96.0	94.7	92.5
Weight <sup>‡</sup>	95.4	92.4	96.3	95.5	94.5	93.2	94.0	92.9
Oxygen saturation	97.2	96.8	96.3	95.3	96.5	96.6	95.5	96.7

Laboratory results were captured in a slightly longer time frame after admission than basic vital signs in this cohort of hospital admissions. This is likely due to the fact that it often takes time for clinicians to evaluate which labs need to be ordered based on the patient’s symptoms as well as for laboratory tests to be collected and results to become available. Laboratory tests also take more time because they are often performed in a separate area of the hospital, whereas vital signs can be measured and resulted immediately. We found that only a limited set of laboratory test results were captured during the majority of hospital admissions.

The most frequently available test results were from the complete blood count (CBC) and basic chemistry panel ([Table 4.7](#)). These results only met the 90% threshold for capture at 24 hours after arrival at the hospital. The CORE team determined that 24 hours was a reasonable timeframe for obtaining first laboratory results for risk adjustment.

The proportion of admissions with laboratory test results from the CBC and chemistry panels captured at 24 hours varied only slightly by condition ([Table 4.8](#)).

<sup>‡</sup> Capture for weight is within the first 24 hours of admission because it is not likely to change substantially during that timeframe.

**Table 4.7: Proportion of Admissions with Laboratory Results at Various Time Points**

Lab Test Result – Full Cohort	Total with Result and Timestamp (%)	Within 2 Hours (%)	Within 6 Hours (%)	Within 12 Hours (%)	Within 24 Hours (%)
CSF culture	0.8	0.1	0.4	0.5	0.5
Blood culture	26.2	15.0	20.3	22.3	23.4
C. difficile	4.0	0.1	0.5	0.9	1.5
<b>Complete blood count (CBC)</b>					
Hemoglobin	92.7	61.2	72.7	77.3	90.6
Hematocrit	92.8	61.6	73.8	78.0	90.8
Platelets	92.0	61.1	72.4	76.5	89.8
WBC count	92.0	61.1	72.4	76.5	89.8
Lactate	26.0	16.0	21.8	24.0	24.7
Troponin 1	32.2	25.6	28.7	30.0	30.6
Troponin 2	29.9	21.4	24.4	27.0	28.0
<b>Basic chemistry panel</b>					
Potassium	71.3	49.3	57.2	60.2	69.4
Sodium	71.6	49.3	57.3	60.3	69.6
Chloride	71.1	49.3	56.1	59.4	69.0
Bicarbonate	71.2	49.2	56.8	59.8	69.2
Anion gap	71.1	48.8	55.9	59.3	69.0
BUN	71.0	48.9	56.0	59.3	69.0
Creatinine	75.2	50.8	58.7	62.2	72.8
Glucose	72.0	49.7	57.6	60.6	70.0
<b>Coagulation panel</b>					
INR	40.1	22.8	27.7	31.7	36.4
Prothrombin time (PT)	40.1	22.8	27.6	31.7	36.4
Partial thromboplastin time (PTT)	4.8	1.5	2.3	2.9	3.3
<b>Liver function tests</b>					
ALT	34.5	18.3	21.6	25.1	30.6
AST	35.9	20.2	24.1	27.4	32.2
Alkaline phosphatase	0.4	0.2	0.2	0.3	0.3
Total protein	1.1	0.0	0.1	0.2	0.5
Albumin	17.3	1.6	3.0	6.1	10.1
Bilirubin	31.4	18.0	21.1	23.9	28.1
<b>Arterial blood gas panel</b>					
Arterial pH	11.0	2.9	5.9	7.3	8.2
PaO2	11.6	3.1	6.4	7.9	8.9
PaCO2	11.2	3.0	6.0	7.5	8.4
FiO2	9.8	2.4	5.1	6.5	7.3



**Table 4.8: Proportion of Admissions with at least 90% Capture of Laboratory Result at 24 hours for the Common Medical Conditions**

	CHF	Pneum	AMI	Athero	Arrhyth	Sepsis	DM	Stroke
<b>Hemoglobin</b>	98.0	97.7	98.0	95.3	90.3	97.8	96.7	96.7
<b>Hematocrit</b>	98.0	97.7	98.1	95.3	90.4	97.9	97.0	96.8
<b>Platelets</b>	97.9	97.5	97.9	95.0	90.0	97.7	96.6	96.6
<b>WBC count</b>	97.9	97.5	97.9	95.0	90.0	97.7	96.6	96.6
<b>Potassium</b>	98.3	97.2	97.7	93.5	90.1	97.2	96.5	96.7
<b>Sodium</b>	98.3	97.3	97.7	93.7	90.1	97.2	96.6	96.7
<b>Chloride</b>	98.2	97.2	97.5	93.4	90.0	97.1	96.2	96.6
<b>Bicarbonate</b>	98.3	97.3	97.6	93.4	90.1	97.2	96.3	96.6
<b>Anion Gap</b>	98.2	97.1	97.5	93.4	89.9	97.1	96.2	96.3
<b>BUN</b>	98.2	97.1	97.5	93.6	90.0	97.1	96.2	96.6
<b>Creatinine</b>	98.3	97.3	97.6	93.9	90.2	97.3	96.5	96.7
<b>Glucose</b>	98.3	97.3	97.6	93.5	90.1	97.1	96.6	96.6

#### 4.3.2. Variation across Hospitals and Hospital Units

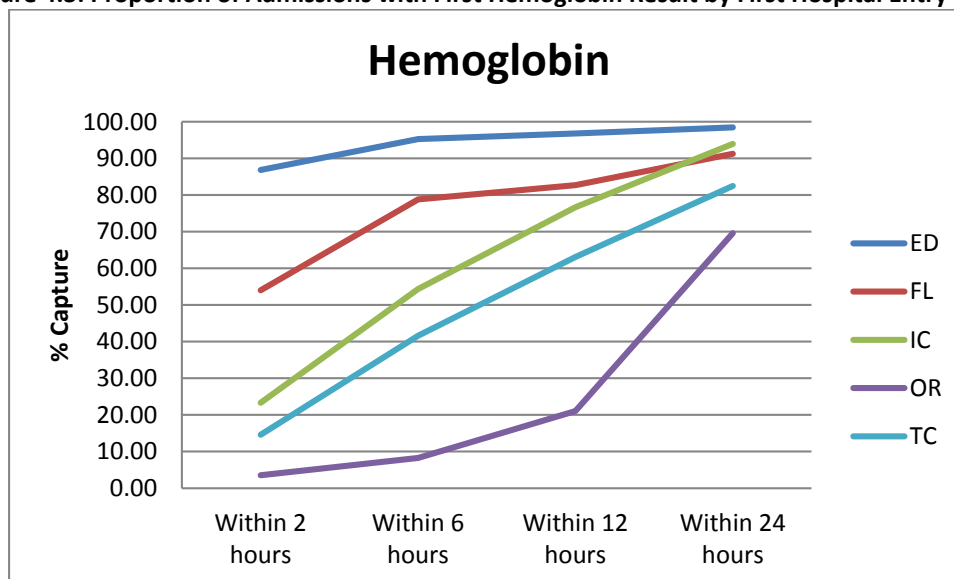
The total proportion of admissions with laboratory results captured at 24 hours is much higher in the conditions compared with the overall cohort due to the low capture of laboratory results among surgical patients during the first 24 hours of admission. To explore this in greater detail, we examined how the capture of laboratory results varies based on the hospital entry location.

The majority of admissions in the dataset began in the ED ([Table 4.9](#)). However, a significant proportion began in the operating room, and this proportion varies further by condition. Laboratory results for admissions starting in the ED were more frequently captured within 24 hours as compared with admissions that began in the operating room. This was true for the full cohort and for each of the conditions. This is likely because it is common clinical practice to obtain laboratory test results for patients with planned surgical procedures several days prior to their surgery. This pattern is apparent when we examine the proportion of admissions with a captured value of a component of the CBC, such as hemoglobin, and a component of the chemistry panel, such as sodium, by hospital entry location ([Figure 4.3](#) and [Figure 4.4](#)).

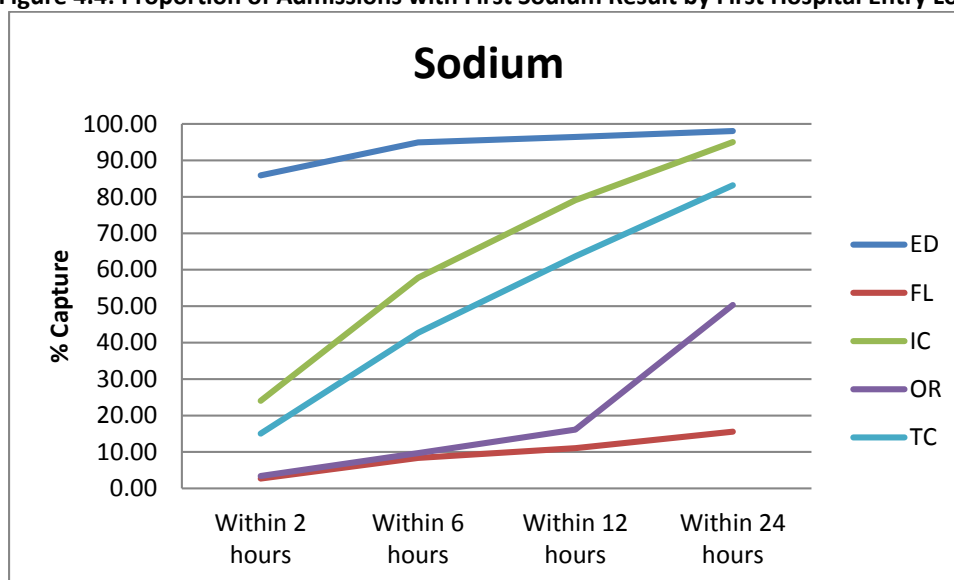
**Table 4.9: First Hospital Entry Locations by Condition Cohort**

Cohort	Emergency Department (%)	Inpatient Floor (%)	Intensive Care Unit (%)	Operating Room (%)	Step Down Unit (%)	Transitional Care Unit (%)
<b>All</b>	212,707 (55.7)	84,341 (22.1)	2,143 (0.6)	81,530 (21.3)	4 (0.0)	1,255 (0.3)
<b>Heart Failure</b>	9,521 (90.7)	426 (4.1)	87 (0.8)	397 (3.8)	0 (0.0)	63 (0.6)
<b>Pneumonia</b>	6,038 (90.4)	330 (4.9)	43 (0.6)	241 (3.6)	0 (0.0)	24 (0.4)
<b>Acute Myocardial Infarction</b>	5,471 (88.3)	160 (2.6)	108 (1.7)	426 (6.9)	0 (0.0)	28 (0.5)
<b>Coronary Atherosclerosis</b>	3,896 (58.8)	217 (3.3)	30 (0.5)	2,400 (36.2)	0 (0.0)	89 (1.3)
<b>Cardiac Dysrhythmias</b>	5,165 (90.4)	253 (3.9)	58 (0.9)	910 (14.2)	0 (0.0)	41 (0.6)
<b>Septicemia</b>	27,390 (90.9)	1,240 (4.1)	235 (0.8)	1,195 (4.0)	0 (0.0)	67 (0.2)
<b>Diabetes Mellitus with Complications</b>	4,952 (81.2)	472 (7.7)	52 (0.9)	609 (10.0)	0 (0.0)	13 (0.2)
<b>Acute Cerebrovascular Disease</b>	6,275 (85.6)	320 (4.4)	378 (5.2)	245 (3.3)	0 (0.0)	110 (1.5)

**Figure 4.3: Proportion of Admissions with First Hemoglobin Result by First Hospital Entry Location<sup>§</sup>**



**Figure 4.4: Proportion of Admissions with First Sodium Result by First Hospital Entry Location**



Similarly, an evaluation of the timing of capture across the 21 KPNC hospitals shows a consistent pattern in which most laboratory values are captured within the first 24 hours of an admission and relatively few are captured after that time window. However, some hospitals capture these laboratory results for a smaller proportion of their patients. For example, 4 of the 21 hospitals did not capture hemoglobin values on more than 10% of their admitted patients. Only 2 of the 21 hospitals captured sodium on more than 90% of all admitted patients. The remaining 19 hospitals had missing sodium values between 11% and 36% of all admissions. Our analyses did not reveal the cause of this finding. For example, we did not find that hospitals with fewer laboratory tests captured among admissions had a greater proportion of admissions directly through the operating

<sup>§</sup> ED: Emergency department, FL: Inpatient floor, IC: Intensive care unit, OR: Operating room, TC: Transfer of care unit

room. Further studies in a broader and more diverse population of hospitals will need to be done to better understand these apparent differences in patterns of care.

Basic vital signs, including blood pressure, heart rate, respiratory rate, and temperature, were captured in the first 2 hours after arrival at the hospital for more than 90% of admissions through the ED, operating room, ICU, inpatient floor, and transitional care unit. Basic vital signs were also captured in more than 94% of patients at each of the 21 hospitals within 2 hours of arrival at the hospital. Oxygen saturation was captured in the first 2 hours for more than 90% of admissions through the ED, OR and ICU.

#### 4.3.3. Data Accuracy

The ranges of vital sign and laboratory results values were plotted in the KPNC dataset to assess the accuracy of the data. Most values fell within a clinically reasonable range when taking into account that hospitalized patients will likely have vital signs and laboratory results that fall outside of the normal range for healthy individuals ([Table 4.10](#) and [Table 4.11](#)). There were a few aberrant values at the extremes of each distribution that appear to be errors in measurement or data entry. For example, the minimum value for temperature appears to have been recorded in Celsius rather than Fahrenheit. Similarly, the maximum value for weight was likely missing a decimal place.

We saw relatively little variation in the distribution of vital sign and laboratory result values across first hospital entry location. The values from the ED had a slightly larger range. This finding makes sense given that the population of patients entering the ED is heterogeneous and has conditions that vary broadly in severity. [Figure 4.5](#) shows the range of values by hospital entry location for hemoglobin as an example.

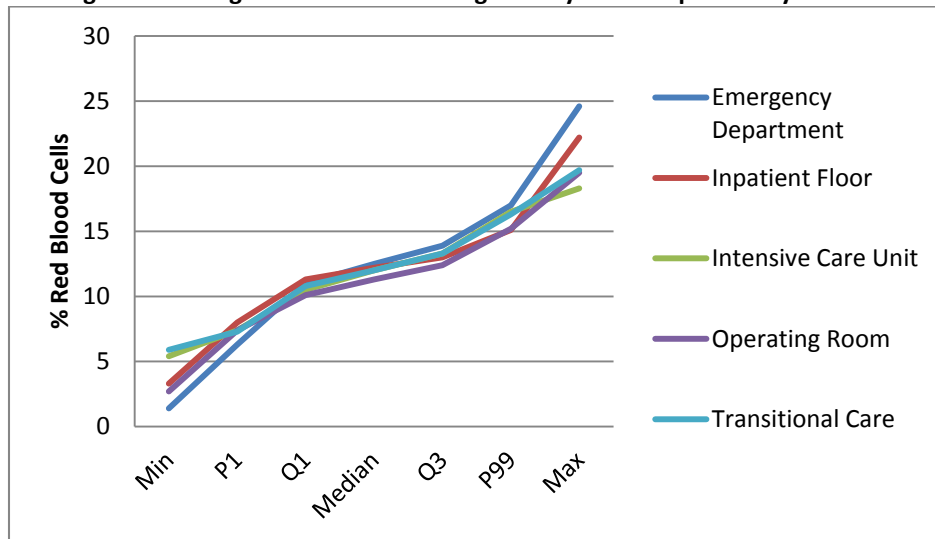
**Table 4.10: Range of Values for Vital Signs in Full Cohort**

Vital Sign – Full Cohort	Min	P1	Q1	Median	Q3	P99	Max	Normal Range
Heart Rate (bpm)	0	48	70	82	97	145	1125	60-100
Systolic Blood Pressure (mmHg)	0	84	118	131	148	205	313	90-140
Diastolic Blood Pressure (mmHg)	0	40	65	74	84	117	221	60-90
Respiratory Rate (breath per minute)	0	12	17	18	20	34	189	12-20
Temperature (°F)	32	96	97.7	98.2	98.6	102.7	127.4	96-100
Weight (pounds)**	0	92.3	145	172	205	337	19306	---
Oxygen saturation (%)	0	83	96	98	99	100	100	95-100

\*\* Weight in the KPNC system was collected and exported in ounces.

**Table 4.11: Range of Values for Laboratory Results in Full Cohort**

Lab Test Result – Full Cohort	Min	P1	Q1	Median	Q3	P99	Max	Normal Range
<b>Complete blood count (CBC)</b>								
Hemoglobin (g/dL)	1.4	6.8	10.9	12.2	13.4	16.7	24.6	12.1-17.2
Hematocrit (% red blood cells)	5.3	21	32.8	36.5	40	49.2	72.3	36.1-50.3
Platelets (count)	0	56	170	213	266	531	3737	---
WBC count (cells/mL)	0	2.8	7.4	9.7	12.6	26.8	494.2	4-10
<b>Basic chemistry panel</b>								
Potassium (mEq/L)	0.4	2.9	3.9	4.2	4.6	6.3	11.7	3.7-5.2
Sodium (mEq/L)	26	122	136	138	141	148	197	135-144
Chloride (mEq/L)	4	84	99	102	105	114	401	96-106
Bicarbonate (mmol/L)	1.5	13.7	24	27	29	38	5440	---
Anion gap (mEq/L)	0	2	7	9	12	23	289	---
BUN (mg/dL)	1	5	13	18	28	102	454	7-20
Creatinine (mg/dL)	0.05	0.44	0.71	0.9	1.23	8.3	58.27	0.8-1.5
Glucose (mg/dL)	11	62	97	116	148	485	1833	64-128

**Figure 4.5: Range of Values for Hemoglobin by First Hospital Entry Location**

#### 4.3.4. Confirmation of Feasibility Testing in Additional Hospitals

A hospital survey of data elements in the potentially feasible and relevant data categories was developed based on the expertise of the TEP, literature review, and extractable KPNC dataset. Given the diversity of electronic health systems available to hospitals and the ability to customize interfaces to align with provider preferences, it was important to ensure that the data elements identified through this process were feasible in multiple hospitals and geographic regions.

While the KPNC dataset is a rich source of information for feasibility testing, there was a concern that results may not be generalizable due to the fact that KPNC is an integrated health system on a capitated payment model and is advanced in its use of electronic health systems. In particular, one

caveat to using a KPNC dataset is that contiguous admissions to different hospitals within the system are treated as a single admission. Similarly, treatment in the ED is linked to the subsequent hospital admission. This continuity might not be true in other hospital systems. The fragmented documentation of emergency care and other inpatient care among other providers is an issue that was identified by the TEP as a potential barrier to identifying the start of care for hospital admissions and first captured laboratory tests and vital signs.

CORE partnered with Premier, Inc. to administer a survey to a handful of member hospitals using alternate EHR systems with specific questions about the feasibility of the data elements identified above. Among the four hospitals that responded to the survey, two were from the South-Atlantic region, one from the Mid-Atlantic region, and one from New England. Two of the four hospitals used a Cerner Millennium EHR, one used Meditech 6.1, and one used Allscripts. All were large hospitals. Two were hospital systems consisting of more than one hospital. Three of the hospitals had more than 500 beds and one had more than 300 beds.

All hospitals reported the ability to identify and extract all of the data elements demonstrated to be feasible and accurate in the KPNC dataset. In addition, most hospitals were already running code to extract these data from their EHR either to satisfy public reporting standards or to support internal quality improvement efforts. Two of the four could export the data in standard reports and two required some customization of reports to export some data elements.

All hospitals reported the ability to identify time stamps associated with first arrival in the ED or other inpatient locations for care although some had separate EHRs in the ED and inpatient settings that were not integrated. Only two of the four hospitals had EHRs that linked data captured in the ED with other data captured during the hospital admission. All hospitals reported capture of laboratory test results and vital sign findings as structured data elements in their EHR whether captured in the ED or in other inpatient locations. All laboratory test results and vital sign findings were reported to be associated with time and date stamps. A summary of the survey results are shown in [Table B.4](#), [Table B.5](#), [Table B.6](#), and [Table B.7](#) in [Appendix B2](#). This survey provides supporting evidence of the feasibility of the data elements presented in our report that can potentially be used as risk adjusters for mortality measures.

#### **4.4 Summary**

This section described a replicable process for identifying clinical data elements that are feasible and relevant to measures of mortality for adult hospitalized patients. This process involved comparing the list of categories of data elements identified as feasible and relevant by the TEP with a list of clinical variables that predict mortality for adult hospitalized patients identified through a systematic review of the literature. All of the data elements identified through this comparison were also mapped and extracted from the KPNC EHR system along with other tests commonly drawn in panels of tests included in the mortality indices. Next, a quantitative analysis of the KPNC dataset was performed to assess data format, capture, and accuracy.

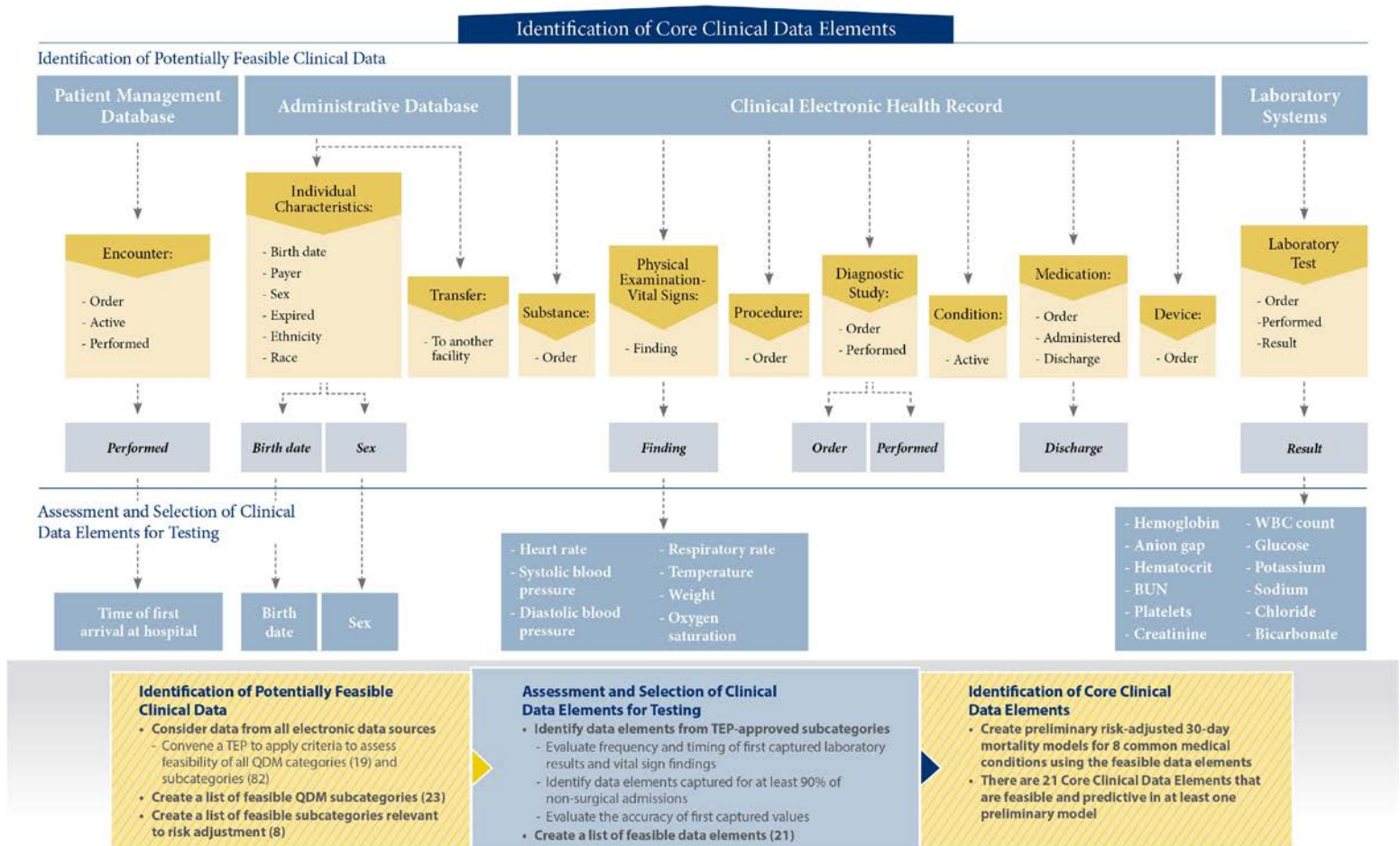
We also described an analysis of the additional data elements necessary to identify first captured data values for vital signs and laboratory tests relative to the time of arrival in the hospital. We demonstrated that timing, consistency, and accuracy of vital sign data do not vary by condition cohort, hospital, or hospital entry location. However, the capture of laboratory test results does vary among patients admitted for medical conditions in the cardiorespiratory, cardiovascular, medicine, and neurology specialty cohorts. We also found variation in frequency of laboratory tests across hospitals and hospital

entry locations. The latter was most likely due to infrequent capture among patients admitted following surgical procedures or directly from the operating room without treatment in the ED.

The goal of these analyses was to identify data elements that are captured during most hospital admissions close to the initiation of care. Ideally, data elements for risk adjustment should be available for nearly all admissions unless the absence of a result indicates a normal value. Since normal values may vary across different measures or circumstances, we chose to include only data elements that were captured in 90% of admissions within 2 hours for vital sign findings (with the exception of weight, which can be captured within the first 24 hours) and 24 hours for laboratory test results. Data elements that did not meet that threshold were not considered for testing in risk adjustment models.

These analyses indicate that 21 data elements currently meet our criteria for testing in risk adjustment models. For the full list of data elements, please refer to [Figure 4.6](#).

**Figure 4.6: Assessment and Selection of Clinical Data Elements for Testing**



## 5. PRELIMINARY TESTING OF RISK-ADJUSTED 30-DAY MORTALITY MODELS

### 5.1 Approach

This section describes testing of the data elements that we found to be consistently captured and extractable from electronic health systems in the previous sections. Specifically, we used logistic regression to determine whether these data elements are adequate predictors of 30-day mortality across a variety of medical conditions associated with hospitalization in adult patients. The statistical analyses presented in this section are not intended to be complete or comprehensive and do not represent fully developed mortality measures. For example, we did not examine the non-linear relationship between each element and mortality. The analyses are not intended to be sufficient for the development or modification of a hospital quality measure that could be used to compare hospitals' performance. Therefore, we do not present the calculation of risk-standardized mortality rates, hierarchical models, or sensitivity analyses that would be standard for measure development. The purpose here is to test the basic performance of the core clinical data elements for risk adjustment in preliminary statistical models predicting 30-day mortality.

The association between the data elements that were shown to be feasible in our previous analysis and 30-day mortality was examined using logistic regression in eight condition cohorts, as defined by principal discharge diagnoses ([Table 4.4](#)). Data elements that were significant predictors of 30-day mortality in any of the condition-specific models were included in the final set of clinical data elements recommended for use in quality measurement. Although we randomly excluded collinear variables from the models, all variables commonly drawn together in a single panel of tests were included in the final clinical dataset.

### 5.2 Methods

#### 5.2.1. Outcome

Mortality was selected as the most appropriate outcome to test the core clinical data elements' performance as risk-adjustment variables. It is one of the most commonly studied outcomes and has abundant literature on key clinical risk-adjusters in existing risk models. Additionally, stakeholders tend to have a strong preference for clinical data in risk adjustment of mortality prediction. Data elements that work well in mortality prediction models are likely to work well across other outcome measures. To align with the publicly reported mortality measures, a 30-day post-discharge timeframe for measurement of the mortality outcome was selected.

#### 5.2.2. Data Source and Cohort Derivation

The dataset used to derive the cohorts, risk-adjustment variables, and outcome was composed of data elements from administrative, patient management, clinical, and laboratory components of the KPNC electronic health systems. These data were extracted for all admissions of inpatient care that occurred at any of their 21 hospitals between January 1, 2010, and December 31, 2011. Refer to [Appendix B.2](#) for more information about how KPNC collected data in each system and mapped the data elements.

##### *Derivation of Eight Condition Cohorts*

The same eight cohorts that were identified in the [cohort derivation](#) section described in [Section 4.2](#)



were used to run the models. These cohorts are:

- Cardiorespiratory
  1. Congestive heart failure
  2. Pneumonia
- Cardiovascular
  3. Cardiac dysrhythmias
  4. Atherosclerosis and other heart disease
  5. Acute Myocardial Infarction (AMI)
- Medicine
  6. Septicemia
  7. Diabetes mellitus with complications
- Neurology
  8. Acute cerebrovascular disease

AHRQ CCS categories were used to select ICD-9 principal discharge diagnosis codes associated with each of these eight conditions. Three of these common condition cohorts correspond with measures that are currently publicly reported on Hospital Compare (AMI, heart failure, and pneumonia). However, the publicly reported measures use slightly different ICD-9 code groupings to define the cohorts than those defined using the AHRQ CCS categories.

### 5.2.3. Preliminary Model Development

#### *Logistic Regression Model*

For each model, we fit a generalized logistic regression model linking the outcome to the risk factors.<sup>1</sup> Let  $Y_{ij}$  denote the outcome (equal to 1 if patient dies within 30 days, zero otherwise) for the  $j$ th patient who presented with each of our eight selected conditions at the  $i$ th hospital;  $\mathbf{Z}_{ij}$  denotes a set of risk factors based on the clinical data. Let  $I$  denote the total number of hospitals and  $n_i$  the number of index admissions to hospital  $i$ . We assume the outcome is related linearly to the covariates via a known linked function,  $h$ , where

$$\text{LRM} \quad h(Y_{ij}) = \alpha + \theta \mathbf{Z}_{ij} \quad (1)$$

and  $\mathbf{Z}_{ij} = (Z_{1ij}, Z_{2ij}, \dots, Z_{pij})$  is a set of  $p$  patient-specific covariates. In our case,  $h$  = the logit link, which is the logistic regression model. We fit the logistic regression model described in Equation (1) using the logit link for the model development and model performance.

#### *Selection of Final Risk Adjustment Variables*

Data elements that were confirmed to be feasible in [Figure 4.6](#) were carried forward for model testing in each of the logistic regression models. The correlations among the data elements were examined and the linear relationship between these data elements and mortality was assumed. For data elements that were highly correlated, we only tested one in the model. For consistency across models, we made the determination to use creatinine over BUN, sodium over chloride, bicarbonate over anion gap or glucose, hematocrit over hemoglobin, and systolic blood pressure over diastolic blood pressure.

For each of the eight condition-specific models, data elements that were not shown to be predictive ( $p$  value greater than 0.05) were removed and the model was rerun until a final list of predictive



data elements remained. Data elements that were found to be predictive in any of the models were included in the final recommended list of data elements for risk adjustment. We also include laboratory tests and vital signs that were highly correlated and not included in risk models.

### 5.3 Results

The final logistic regression models for the eight condition cohorts performed very well, with C-statistics ranging from 0.748-0.843 and an adjusted R-square ranging from 0.1405-0.2826. The variables, odds ratios, and confidence intervals for the eight logistic regression models are shown in [Table 5.1](#). Odds ratios for continuous variables are shown for every one unit increase.

**Table 5.1: Odds Ratios and 95% Confidence Intervals for Data Elements Tested in Each Condition-Specific Risk-Adjustment Model**

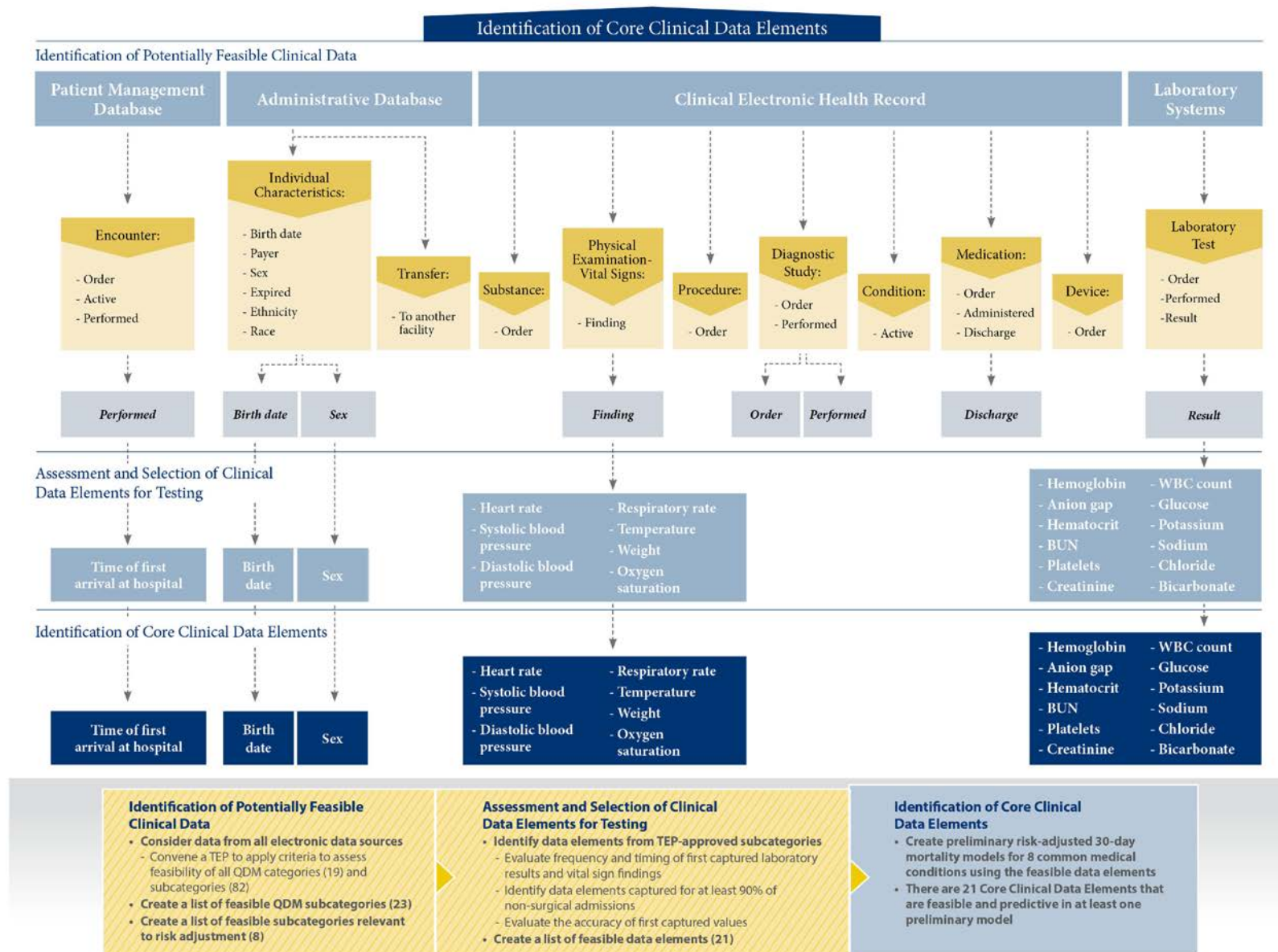
Name	CHF OR (95% CI)	Pneum OR (95% CI)	AMI OR (95% CI)	Athero OR (95% CI)	Arrhyth OR (95% CI)	Sepsis OR (95% CI)	DM OR (95% CI)	Stroke OR (95% CI)
<b>Patient Characteristics</b>								
Age	1.04 (1.04, 1.05)	1.05 (1.04, 1.06)	1.06 (1.05, 1.07)	1.07 (1.04, 1.09)	1.04 (1.02, 1.05)	1.03 (1.03, 1.04)	1.05 (1.04, 1.07)	1.04 (1.04, 1.05)
Gender (Male)	–	1.49 (1.25, 1.77)	–	–	–	1.12 (1.04, 1.20)	–	0.96 (0.95, 0.98)
<b>Vital Signs</b>								
Heart Rate	1.00 (1.00, 1.01)	1.01 (1.00, 1.01)	1.01 (1.01, 1.02)	–	1.01 (1.00, 1.01)	1.01 (1.01, 1.01)	–	1.01 (1.01, 1.02)
Systolic Blood Pressure	0.98 (0.98, 0.98)	0.99 (0.99, 0.99)	0.98 (0.98, 0.98)	–	0.98 (0.98, 0.99)	0.99 (0.98, 0.99)	0.99 (0.98, 0.99)	1.01 (1.00, 1.01)
Respiratory Rate	1.06 (1.05, 1.08)	1.06 (1.05, 1.08)	1.04 (1.02, 1.06)	1.12 (1.04, 1.20)	1.04 (1.00, 1.08)	1.06 (1.06, 1.07)	–	1.04 (1.02, 1.06)
Temperature	0.79 (0.73, 0.85)	0.78 (0.73, 0.84)	0.82 (0.73, 0.92)	–	0.81 (0.67, 0.98)	0.87 (0.85, 0.89)	0.74 (0.64, 0.84)	0.78 (0.72, 0.84)
Oxygen Saturation	0.98 (0.96, 0.99)	0.95 (0.93, 0.97)	0.96 (0.93, 0.98)	–	0.93 (0.89, 0.98)	0.96 (0.95, 0.96)	–	1.05 (1.03, 1.08)
Weight	–	0.99 (0.99, 1.00)	1.00 (0.99, 1.00)	–	0.99 (0.99, 1.00)	0.99 (0.99, 1.00)	0.99 (0.99, 1.00)	1.00 (1.00, 1.00)
<b>Laboratory Results</b>								
Bicarbonate	1.05 (1.03, 1.07)	1.03 (1.01, 1.05)	0.95 (0.92, 0.97)	–	–	0.97 (0.97, 0.98)	1.06 (1.03, 1.09)	–
Creatinine	1.20 (1.14, 1.26)	–	1.30 (1.22, 1.38)	1.60 (1.41, 1.82)	1.43 (1.28, 1.60)	1.09 (1.07, 1.12)	1.18 (1.10, 1.26)	–
Glucose	–	–	1.00 (1.00, 1.00)	–	–	–	–	1.00 (1.00, 1.00)
Hematocrit	0.99 (0.97, 1.00)	0.95 (0.93, 0.96)	–	–	0.97 (0.94, 0.99)	0.96 (0.96, 0.97)	–	0.98 (0.96, 0.99)
Platelet	1.00 (1.00, 1.00)	–	1.00 (1.00, 1.00)	–	–	1.00 (1.00, 1.00)	–	1.00 (1.00, 1.00)
Potassium	1.22 (1.10, 1.35)	1.34 (1.17, 1.53)	–	–	–	1.28 (1.22, 1.35)	–	–
WBC Count	–	1.02 (1.01, 1.04)	1.08 (1.05, 1.10)	1.15 (1.08, 1.23)	1.08 (1.04, 1.13)	1.04 (1.03, 1.04)	1.08 (1.05, 1.12)	1.13 (1.11, 1.15)
Sodium	0.96 (0.95, 0.97)	–	–	–	0.96 (0.93, 1.00)	1.01 (1.01, 1.02)	–	1.04 (1.02, 1.06)

## 5.4 Summary

All feasible data elements tested in the eight 30-day mortality models were found to be statistically significant risk adjusters in at least one model. Data elements that were found to be feasible but were not tested in the models due to collinearity with other data elements are assumed to perform similarly to those data elements that were tested. Additionally, the burden of retrieval for these additional data elements is very low because they are captured in test panels and should be easily extracted with the data elements that were tested in the models. For example, these analyses showed that systolic and diastolic blood pressure are consistently captured together, and diastolic blood pressure should be easily extracted with systolic blood pressure although it was not included in any of the models. Including these data elements will allow other measure developers the flexibility to make their own distinctions with regard to which collinear data elements to include in their models.

Therefore, the final list of clinical data elements recommended for risk adjustment of mortality measures includes all of the clinical data elements identified as feasible at the end of [Section 4](#). [Figure 5.1](#) contains the final list of core clinical data elements.

**Figure 5.1: Identification of Core Clinical Data Elements**

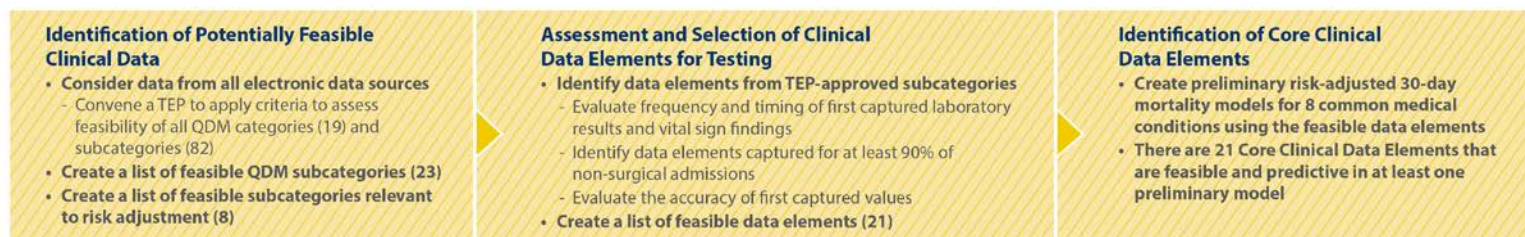


## 6. CORE CLINICAL DATA ELEMENTS AND POTENTIAL STRATEGIES FOR INTEGRATION INTO CMS PROGRAMS

### 6.1 Summary

The final core clinical data elements can be feasibly extracted from current health systems and can be used to risk adjust 30-day mortality models following admission for eight of the most common medical conditions causing hospitalization among adult patients. The derivation of these data elements described in this report is a structured and replicable process for establishing the feasibility and accuracy of EHR data elements as clinical practice changes and EHR systems evolve ([Figure 6.1](#)).

**Figure 6.1: Project Overview**



This process included both qualitative and quantitative assessments of clinical data. The qualitative assessment engaged experts to assess how clearly defined, consistently captured, and in what format specific categories and subcategories of data, as defined by the QDM, are available in current electronic health systems. We also asked these experts to assess the current ease of data extraction from electronic health system databases. This expert assessment provided us with a snapshot of current data feasibility.

This assessment of data feasibility was then validated through quantitative analyses of specific data elements in a large clinical dataset provided by KPNC. This assessment began with a systematic review of the published literature to identify specific data elements in validated indices that have been shown to predict mortality. Using the electronic data elements provided by the KPNC hospitals, we directly examined the consistency and timing of data capture, and accuracy of data elements from clinical data subcategories that were identified as feasible and that are commonly included in mortality prediction scales.

Data elements that were consistently captured for most adult hospitalized patients and found to be accurate were then tested in statistical models of 30-day mortality for eight common conditions associated with inpatient care. They were:

- Congestive Heart Failure
- Pneumonia
- AMI
- Coronary atherosclerosis and Other Heart Disease
- Cardiac dysrhythmia
- Septicemia
- Diabetes mellitus with complications
- Acute cerebrovascular disease

All feasible data elements tested in the logistic regression models were significant predictors of

mortality in at least one of the eight condition-specific models. Only age was a significant predictor in all eight models. These preliminary models, developed using the clinical data for risk adjustment, performed as well as models that use claims data alone for risk adjustment among conditions with NQF-endorsed mortality measures (pneumonia, heart failure, and AMI). [Table 6.1](#) lists the clinical data elements proposed for use in risk-adjusted mortality models.

**Table 6.1: Proposed Clinical Data Elements for Risk-Adjusted Mortality Measures**

Clinical Data Elements	Units of Measurement	Timing of First Capture
<b>Patient Characteristics</b>		
<b>Age</b>	Years	---
<b>Gender</b>	Male or female	---
<b>Vital Signs</b>		
<b>Heart Rate</b>	Beats per minute	0-2 hours
<b>Systolic Blood Pressure</b>	mmHg	0-2 hours
<b>Diastolic Blood Pressure</b>	mmHg	0-2 hours
<b>Respiratory Rate</b>	Breath per minute	0-2 hours
<b>Temperature</b>	Degrees Fahrenheit	0-2 hours
<b>Oxygen Saturation</b>	Percent	0-2 hours
<b>Weight</b>	Pounds	0-24 hours
<b>Laboratory Results</b>		
<b>Hemoglobin</b>	g/dL	0-24 hours
<b>Hematocrit</b>	% red blood cells	0-24 hours
<b>Platelet</b>	Count	0-24 hours
<b>WBC Count</b>	Cells/mL	0-24 hours
<b>Potassium</b>	mEq/L	0-24 hours
<b>Sodium</b>	mEq/L	0-24 hours
<b>Chloride</b>	mEq/L	0-24 hours
<b>Bicarbonate</b>	mmol/L	0-24 hours
<b>BUN</b>	mg/dL	0-24 hours
<b>Creatinine</b>	mg/dL	0-24 hours
<b>Glucose</b>	mg/dL	0-24 hours

## 6.2 Rationale for a Clinical Dataset

The electronic capture and storage of clinical data in EHRs and other integrated electronic databases presents a tremendous opportunity to advance quality measurement programs. Current outcome measurement relies heavily on administrative data. While these measures provide reliable information on hospital quality, the clinical community continues to express a preference for the use of clinical data to assess hospital performance.

The use of clinical data for risk adjustment also has broad stakeholder support in spite of the additional burden of retrieval compared to claims-based models. Similarly, we received overwhelming support for this work from nominees and members of the TEP representing hospital systems, EHR vendors, and specialty societies. Research from Kaiser Permanente, an early adopter of EHRs and eMeasures, shows that reporting EHR data is significantly less burdensome than comparable manual extraction once data are mapped and the extraction process is automated.<sup>2</sup>

CMS currently offers resources through the QDM and Measure Authoring Tool (MAT) that help measure developers identify value sets in EHRs that align with clinical concepts in their measures. However, until now, there have been no clear criteria or set of processes to establish which data are reliable and can be reported with minimal burden for use in quality measures.

These core clinical data elements, which were identified through the process outlined in this report, are a set of data elements that are captured for most adult inpatients and reflect their clinical status or severity of illness upon first arrival to the hospital, before the initiation of care. This dataset will provide measure developers with a standard set of reliable data elements that should be used as a starting place to build risk-adjustment models for a variety of outcome measures.

Furthermore, this work provides an additional tool for standardization that will improve the way data are captured, stored, and extracted. The benefits of standardizing common, discrete data elements extend beyond outcome measurement to other applications like EHR interoperability, real-time clinical decision support, research, and public health surveillance.

## 7. GLOSSARY OF TERMS

- *Administrative database:* 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, name, national and unique medical record identification numbers, dates of admission, dates of discharge, principal discharge diagnoses, and all hospital charges than might be included in a bill for care provided.
- *Clinical Classification Software (CCS) categories:* Groupings of related ICD-9 diagnosis and procedure codes in clinically relevant categories. These categories are defined by the Agency for Healthcare Research & Quality (AHRQ) and can be found at the [Healthcare Cost Utilization Project website](#).
- *Conditions:* Specific CCS categories related to clinical conditions.
- *Consistency of capture:* How often data element values are assessed and recorded for a cohort of hospital admissions.
- *Data accuracy:* The extent to which the extracted data element value is true. This includes fidelity between the value captured in the appropriate location in the EHR clinical interface (by the most accurate measurement method and the most appropriate healthcare professional) and electronic data values extracted from the EHR data warehouse. This also includes the absence of excessive errors in data values, for example, biologically implausible values or widely discrepant values.
- *Data availability:* The data are extractable as a structured value or in a format easily converted to numerical data (numerical, pseudo-numerical, and list) across individuals as well as EHR and hospital systems.
- *Data 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.
- *Data mapping:* Data mapping is the process by which two distinct data models are created and a link between these models is defined. It is most readily used in software engineering to describe the best way to access or represent some form of information. In this report the two data models are the EHR's clinical interface where clinical, laboratory, and other staff capture relevant data and the thousands of linked data tables that make up the EHR's permanent data warehouse where those data are transmitted and stored.
- [\*Electronic clinical quality measures \(eCQMs\)\*](#): Tools that help the Department of Health and Human Services (HHS) measure and track the quality of healthcare services provided by eligible professionals (EPs), eligible hospitals (EHs) and critical access hospitals (CAHs) within our healthcare system. These measures use a wide variety of data that are associated with a provider's ability to deliver high-quality care or relate to long-term goals for health care quality. CQMs measure many aspects of patient care including: health outcomes, clinical processes, patient safety, efficient use of healthcare resources, care coordination, patient engagements, population and public health, and clinical guidelines.
- *Electronic health records:* 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 healthcare settings.
- *Electronic health systems:* The complete set of digital environments that capture and store data in digital format generated from the provision of patient care within the health system. This includes patient management systems, administrative or billing systems, clinical laboratory systems, pharmacy systems, and clinical EHR systems.
- *eMeasure:* Measure that utilizes clinical data from electronic systems. This term often refers to measures that can be calculated entirely from EHR data using appropriate code that can be directly



applied to the EHR system.

- *Hybrid measure*: A quality measure that uses a more than one data source.
- *eSpecification*: The data elements, logic and definitions for a measure in an Health Level Seven (HL7) standard known as the Health Quality Measures Format (HQMF) which represents a clinical quality measure as an electronic Extensible Markup Language (XML) document that can be captured or stored in the EHR so that the data can be sent or shared electronically.
- *First captured values*: The data element values that are captured at or very near to the time of arrival at the hospital for care.
- *Hospital entry location*: The department in which a patient first enters the hospital to receive care, such as the ED, the operating room, or the inpatient floor.
- *Metadata*: The data providing information about one or more aspects of the data, such as the means of creation of the data, purpose of the data, time and date of creation, creator or author of the data, location on a computer network where the data were created, and standards used.
- *Meaningful Use*: A federal program that uses certified electronic health record (EHR) technology to: 1) improve quality, safety, efficiency, and reduce health disparities; 2) engage patients and family; 3) improve care coordination, and population and public health; and, 4) maintain privacy and security of patient health information. The program requires that healthcare providers comply with regulations in three stages over the next five years. These stages focus on data capture and sharing (Stage 1), advancing clinical processes (Stage 2), and improving outcomes (Stage 3). Compliance with these regulations is tied to reimbursement for healthcare services.
- *Natural language processing*: Computer programming that abstracts meaning from human speech or text. This can potentially be applied in medicine to extract data from clinical documentation that is recorded in free text rather than a structured format.
- *Patient management system*: Electronic system or software environment that manages certain administrative activities including allocating physicians, applying policies, and assigning beds. These systems also capture and store patient information, such as name, gender, date of birth, date of encounter visit, national ID or hospital identification number. These systems capture data about patient care workflow, including the registration of patient information, bed tracking, and discharge. The system might or might not be integrated with the clinical EHR.
- *Specialty cohorts*: A group of index admissions for patients with related condition or procedure categories that are likely to be cared for by specific teams of clinicians; there are five defined cohorts in this report (medicine, neurology, cardiorespiratory, cardiovascular, surgery/gynecology).
- *Structured data*: Data captured in a format that is numerical, such as integers or fractions; pseudo-numerical, such as dates; or list, such as “positive” or “negative.”
- *Subcategory*: This is a term that describes the second level in the hierarchy of logic used in the Quality Data Model to organize data elements into clinically coherent groupings. The QDM refers to this level as “data types,” defined as “the context expected for any given QDM element.”
- *Time of arrival*: The time stamp that is captured closest to the moment a patient first reaches the hospital for care.



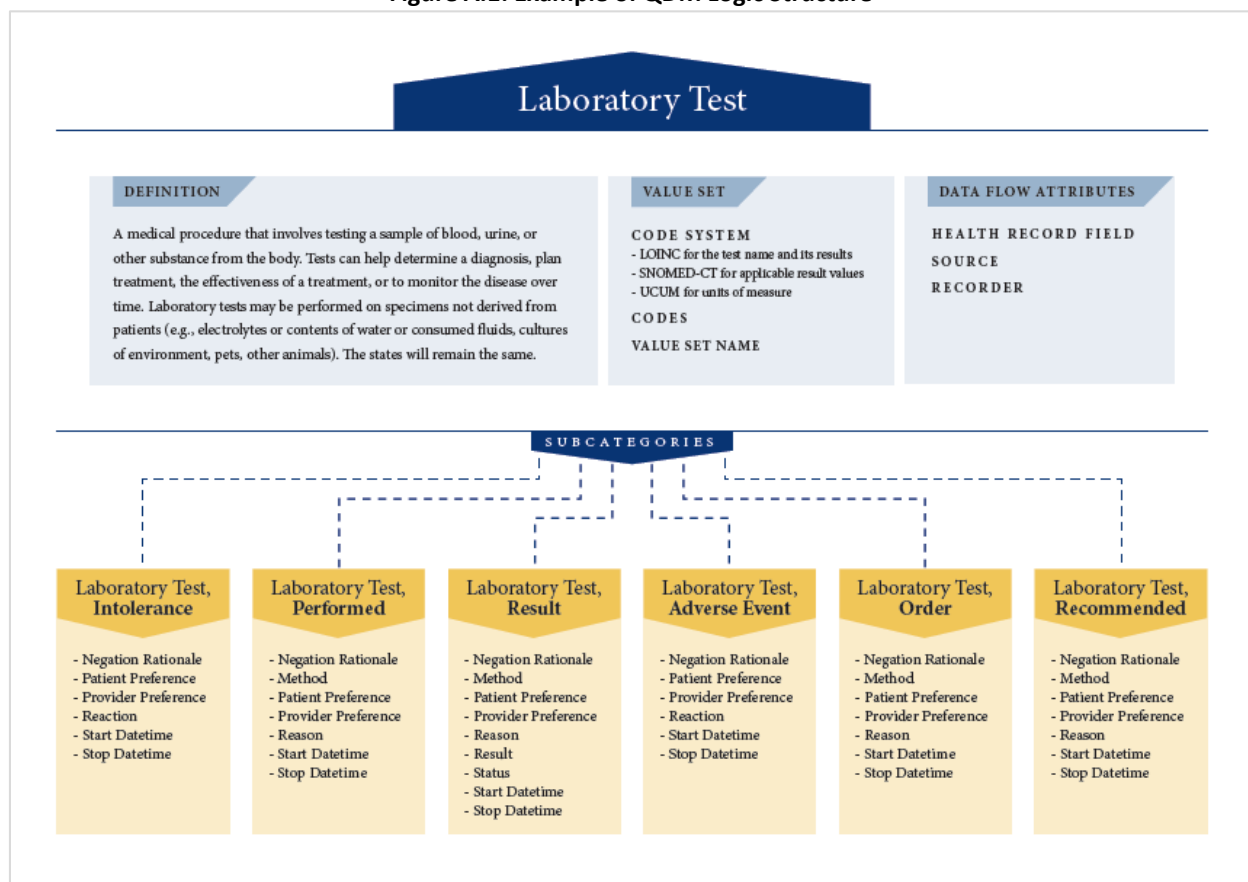
## REFERENCES

1. McCullagh P, N.J. Generalized Linear Models: Chapman and Hall; 1989.
2. Garrido T, Kumar S, Lekas J, et al. e-Measures: insight into the challenges and opportunities of automating publicly reported quality measures. *J Am Med Inform Assoc.* Aug 5 2013.

# Appendix A

## Appendix A1. QDM Categories and Subcategories in Feasibility Survey

Figure A.1: Example of QDM Logic Structure



### Full list of QDM Categories and Subcategories:

1. Encounter
  - 1.1. Encounter recommended
  - 1.2. Encounter order
  - 1.3. Encounter active
  - 1.4. Encounter performed
2. Individual Characteristics
  - 2.1 Patient characteristic, birth date
  - 2.2 Patient characteristic, expired
  - 2.3 Patient characteristic, clinical trial participant
  - 2.4 Patient characteristic, payer
  - 2.5 Patient characteristic, sex
  - 2.6 Patient characteristic, ethnicity
  - 2.7 Patient characteristic, race
  - 2.8 Patient characteristic
  - 2.9 Provider characteristic
3. Transfer of Care

- 3.1. Transfer to
- 3.2. Transfer from
- 4. Physical Examination: Vital Signs Only
  - 4.1. Vital sign recommended
  - 4.2. Vital sign order
  - 4.3. Vital sign finding
  - 4.4. Vital sign performed
- 5. Physical Examination: Neurological Assessment Only
  - 5.1. Neurological assessment recommended
  - 5.2. Neurological assessment order
  - 5.3. Neurological assessment finding
  - 5.4. Neurological assessment performed
- 6. Physical Examination: Other
  - 6.1. Physical examination recommended
  - 6.2. Physical examination order
  - 6.3. Physical examination finding
  - 6.4. Physical examination performed
- 7. Functional Status
  - 7.1. Functional status assessment recommended
  - 7.2. Functional status assessment order
  - 7.3. Functional status assessment result
- 8. Laboratory Test
  - 8.1. Test recommended
  - 8.2. Test order
  - 8.3. Test performed
  - 8.4. Test result
  - 8.5. Test intolerance
  - 8.6. Test adverse event
- 9. Diagnostic study
  - 9.1. Study recommended
  - 9.2. Study order
  - 9.3. Study performed
  - 9.4. Study result
  - 9.5. Study intolerance
  - 9.6. Study adverse event
- 10. Condition/Diagnosis/Problem
  - 10.1. Condition active
  - 10.2. Condition, family history
  - 10.3. Condition resolved
  - 10.4. Condition inactive
- 11. Procedure
  - 11.1. Procedure recommended
  - 11.2. Procedure order
  - 11.3. Procedure performed
  - 11.4. Procedure result
  - 11.5. Procedure intolerance
  - 11.6. Procedure adverse event
- 12. Device
  - 12.1. Device recommended
  - 12.2. Device order
  - 12.3. Device applied
  - 12.4. Device allergy

- 12.5. Device intolerance
- 12.6. Device adverse event
- 13. Medication
  - 13.1. Medication order
  - 13.2. Medication dispensed
  - 13.3. Medication administered
  - 13.4. Medication active
  - 13.5. Medication allergy
  - 13.6. Medication intolerance
  - 13.7. Medication adverse event
  - 13.8. Medication discharge
- 14. Substance
  - 14.1. Substance order
  - 14.2. Substance recommended
  - 14.3. Substance administered
  - 14.4. Substance allergy
  - 14.5. Substance intolerance
  - 14.6. Substance adverse event
- 15. System Characteristics
- 16. Symptom
  - 16.1. Symptom active
  - 16.2. Symptom resolved
  - 16.3. Symptom assessed
  - 16.4. Symptom inactive
- 17. Risk Category Assessment
- 18. Care Goal
- 19. Intervention
  - 19.1. Intervention recommended
  - 19.2. Intervention order
  - 19.3. Intervention performed
  - 19.4. Intervention result
  - 19.5. Intervention intolerance
  - 19.6. Intervention adverse event
- 20. Communication
  - 20.1. From provider to provider
  - 20.2. From provider to patient
  - 20.3. From patient to provider
- 21. Care Experience

## Appendix A2. TEP Member List

**Table A.1: Evaluating Electronic Health Record (EHR) Data Elements for use in Hospital Quality Measures:  
Technical Expert Panel Members**

Name	Organization (Title)	Location
<b>Howard Bregman, MD, MS</b>	Epic	Verona, WI
<b>Ralph Brindis, MD, MPH, MACC, FSCAI</b>	The American College of Cardiology National Cardiovascular Registry <i>Senior Medical Officer, External Affairs</i>	San Francisco, CA
<b>Zahid Butt, MD</b>	Medisolv, Inc. <i>CEO</i>	Columbia, MD
<b>Christopher Chute, MD, DrPH</b>	Mayo Clinic <i>Professor of Biomedical Informatics</i>	Rochester, MN
<b>Richard P. Dutton, MD, MBA</b>	Anesthesia Quality Institute <i>Executive Director</i>	Park Ridge, IL
<b>David Kaelber, MD, PhD, MPH, FAAP, FACP</b>	MetroHealth System <i>Chief Medical Informatics Officer</i>	Shaker Heights, OH
<b>Saul Kravitz, PhD</b>	MITRE <i>Principal Health IT Engineer</i>	McLean, VA
<b>Adam Landman, MD, MS, MIS, MHS</b>	Brigham and Women's Hospital <i>Chief Medical Information Officer for Health Information Innovation and Integration</i>	Boston, MA
<b>David Levine, MD</b>	University HealthSystem Consortium <i>Vice President of Informatics and Medical Director of Comparative Data and Informatics</i>	Chicago, IL
<b>Maggie Lohnes, RN</b>	McKesson Corporation <i>Clinical Quality Executive</i>	Fox Island, WA
<b>Rute Martins, MS</b>	The Joint Commission <i>Associate Project Director</i>	Oakbrook Terrace, IL
<b>Clement McDonald, MD</b>	Lister Hill National Center for Biomedical Communications <i>Director</i>	Bethesda, MD
<b>Meg McElroy, MBA, RHIA</b>	American Health Information Management Association (AHIMA) <i>System Program Manager</i>	Milwaukee, WI
<b>Mary Beth Mitchell, MSN, RN-BC, CPHIMS</b>	Texas Health Resources <i>Chief Nursing Informatics Officer</i>	Dallas, TX
<b>Karen Nielsen, MBA, MPA</b>	Siemens Medical Solutions USA, Inc. <i>R&amp;D, Analytics, and Business Intelligence</i>	Malvern, PA
<b>Kim Nolen, PharmD</b>	Pfizer, Inc. <i>Medical Outcomes Specialist</i>	Peachtree City, GA
<b>David Shahian, MD</b>	Massachusetts General Hospital Center for Quality and Safety <i>Vice President</i>	Boston, MA
<b>Christopher Snyder, DO</b>	Peninsula Regional Medical Center <i>Chief Medical Information Officer</i>	Ocean City, MD

### Appendix A3. Survey result tables – Infeasible subcategories

**Table A.2: Strong agreement on infeasibility (≥70%)**

Category	Subcategory	Met Data Capture Criteria (yes/no)
2. Individual Characteristics	2.3 Clinical Trial Participant	1/14
2. Individual Characteristics	2.9 Provider Characteristics	3/10
4. Physical Examination: Vital Signs	4.1 Recommended	2/14
6. Physical Examination: Other	6.1 Recommended	1/15
6. Physical Examination: Other	6.2 Order	4/11
7. Functional Status	7.1 Recommended	1/15
7. Functional Status	7.2 Order	4/12
8. Laboratory Test	8.1 Recommended	3/13
8. Laboratory Test	8.5 Intolerance	1/15
8. Laboratory Test	8.6 Adverse Event	2/14
9. Diagnostic Study	9.1 Recommended	2/14
9. Diagnostic Study	9.5 Intolerance	1/14
9. Diagnostic Study	9.6 Adverse Event	2/13
10. Condition/Diagnosis/Problem	10.2 Family History	2/14
10. Condition/Diagnosis/Problem	10.4 Inactive	4/12
11. Procedure	11.1 Recommended	2/14
11. Procedure	11.5 Intolerance	2/13
11. Procedure	11.6 Adverse Event	3/12
12. Device	12.1 Recommended	2/13
12. Device	12.4 Allergy	3/11
12. Device	12.5 Intolerance	2/13
12. Device	12.6 Adverse Event	3/12
13. Medication	13.7 Adverse Event	1/13
14. Substance	14.4 Administered	3/12
14. Substance	14.6 Intolerance	4/12
14. Substance	14.7 Adverse Event	3/13
19. Intervention	19.1 Recommended	2/8
19. Intervention	19.5 Intolerance	3/7
20. Communication	20.1 From Provider to Provider	2/8
20. Communication	20.3 From Patient to Provider	1/8
21. Care Experience	21.1 Care Experience	3/7

**Table A.3: Complete agreement on infeasibility (100%)**

Category	Subcategory	Met Data Capture Criteria (yes/no)
1. Encounter	1.1 Recommended	0/16
5. Physical Examination: Neurological Assessment	5.1 Recommended	0/16

#### Appendix A4. Summary of TEP discussion and comments

Category - Subcategory	Summary of TEP Scoring Discussion and Comments	Subcategory Should be Included in Feasibility Testing	
		Yes	No
Encounter – Performed	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.88</li> </ul> <p><b><u>Final Consensus</u></b></p> <ul style="list-style-type: none"> <li>Encounter Performed (documentation of a hospital admission) is best defined by the appearance of a bill in a hospital's administrative database with the accompanying length of stay. Information required for billing purposes is consistently captured and coded. However, exact start and stop times and dates for an admission lack a standard definition or method of capture. Such data would be structured as time and date stamps in patient registration or EHR systems which might or might not be integrated with billing or administrative databases.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>Timing of admission and discharge can be captured differently by patient management systems (time in a bed and out of bed).</li> <li>There are built-in incentives to capture these times with accuracy as payment is tied to length of stay, discharge timing, etc. This means that data are present. However, there is no standard definition for which data element (what pieces of metadata, what source of data) best captures the concepts of admission start and stop.</li> <li>Time and date stamps for discharge and admit orders are associated with coding systems. However, other data elements used to establish the start and stop of encounters might not be systematically coded (e.g., patient management data).</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>EHR or integrated patient management systems will always capture when a patient arrived in a bed and when they were discharged.</li> </ul>	v	
Transfer - To	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.86</li> </ul> <p><b><u>Final Consensus</u></b></p> <ul style="list-style-type: none"> <li>Transfer to meets all 3 data capture criteria although there is no coding structure representing this concept and important metadata (receiving facility) might be missing or not encoded.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>Transfer to is reliable and transfer from is not. Payment systems incentivize transfer to as capture of discharge.</li> <li>There is no standard code for a transfer.</li> <li>Transfer from data are only likely to be consistently captured if tied to a new incentive, such as quality reporting.</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>Target or receiving facility might not be coded (e.g., facility type).</li> <li>Receipt of a transfer (transfer from) is often only captured in notes as string or text data.</li> </ul>	v	

Category - Subcategory	Summary of TEP Scoring Discussion and Comments	Subcategory Should be Included in Feasibility Testing	
		Yes	No
Physical Examination Vital Signs - Result	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.69</li> </ul> <p><b><u>Final Consensus</u></b></p> <ul style="list-style-type: none"> <li>Several instances of vital signs results are consistently captured at specific time points during the admission (triage in the ED, first on admission to the inpatient floor or ICU, etc.). The data are captured as numbers and usually in a structured field or convertible to structured data.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>The numerous variations in circumstances surrounding and methods of assessment of vital sign measurement make it difficult to conclude that there is a strict standard definition associated with any individual instance of a vital sign. For example, is a heart rate captured while sitting, standing, during an episode of pain or anxiety, by a machine, by clinical staff, in response to a patient's complaint about palpitations, or is the capture routine, etc.?</li> <li>Although there is a standard coding system for vital sign findings, syntax is not specified so that a code might correspond to the broad concept of a specific vital sign (heart rate) but will not specify how it is structured (e.g., as a 2- or 3-digit integer in units of beats per minute).</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>Machine-measured vital sign findings will not be fully integrated into many EHR systems although some instance of results from machine measurements are often entered manually.</li> </ul>	v	
Physical Examination Neurological Assessment - Result	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.60</li> </ul> <p><b><u>Final Consensus</u></b></p> <ul style="list-style-type: none"> <li>These data elements do not currently meet the data capture criteria.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>GCS might be captured routinely for ICU-admitted, trauma, or other patients with abnormal mental status.</li> <li>Clinicians' descriptions of neurological status are likely captured as text or string data requiring natural language processing to extract for analysis.</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>ICUs and trauma response teams might use standardized flow sheets to capture data which might result in structured GCS data.</li> <li>Although EHRs will have local coding systems for these data they might not yet be associated with standardized value sets, such as Logical Observation Identifiers Names and Codes (LOINC).</li> <li>There is no standard and routine neurological assessment currently captured across most or all inpatient settings or for most or all adult admitted patients.</li> </ul>		v



Category - Subcategory	Summary of TEP Scoring Discussion and Comments	Subcategory Should be Included in Feasibility Testing	
		Yes	No
Laboratory Test - Result	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>• 0.94</li> </ul> <p><b><u>Final Consensus</u></b></p> <ul style="list-style-type: none"> <li>• Data elements within this subcategory met all 3 data capture criteria.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>• Units of measurement associated with some test results data are not standardized.</li> <li>• Thresholds for and ranges of normal values are nearly always exported from clinical lab databases as text or string data. This text is sometimes bundled with the actual result and exported as one text field. Most common lab test units have universally standard ranges for normal values.</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>• Different panels of laboratory tests in different hospitals and clinical labs (e.g., chemistry panel) might have different individual components and be associated with different codes.</li> </ul>	√	
Laboratory Test - Order	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>• 0.81</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>• Results are more consistently encoded with standard value sets, such as LOINC compared with orders which might not yet be linked to standard value sets.</li> </ul>	√	
Laboratory Test - Performed	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>• 0.73</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>• Several TEP members expressed concern that this subcategory is not consistently captured or encoded across hospitals.</li> <li>• They also expressed the opinion that these data are not needed if the test result is what is desired (e.g., in risk adjustment).</li> </ul>		√
Diagnostic Study – Order	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>• 0.88</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>• Orders are not yet consistently linked with standard value sets, such as LOINC.</li> </ul>	√	
Diagnostic Study Result	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>• 0.47</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>• Results are not captured as structured data.</li> </ul>		√

Category - Subcategory	Summary of TEP Scoring Discussion and Comments	Subcategory Should be Included in Feasibility Testing	
		Yes	No
<b>Condition – Active</b> (included discussion of principal discharge diagnosis and secondary diagnoses for hospital admissions)	<p><u>Survey Consensus Score</u></p> <ul style="list-style-type: none"> <li>0.69</li> </ul> <p><u>Final Consensus</u></p> <ul style="list-style-type: none"> <li>Clinician documented data elements in this subcategory do not meet data capture criteria as they are not consistently captured and many instances of condition data elements, such as secondary diagnoses are not captured in structured fields (SNOMED or ICD-9 codes).</li> </ul> <p><u>TEP Discussion Points</u></p> <ul style="list-style-type: none"> <li>There are 2 problem lists. One is a clinical problem list assembled by the care team. The second is a billing or administrative problem list generated for payment via review of the medical record.</li> <li>The choice of conditions captured in administrative databases is strongly influenced by payment incentives.</li> <li>Clinicians have a poor understanding of the definition of a principal discharge diagnosis (what brought a patient in for care). Clinicians document their patients' most important problems, which could change over the course of their admission or from one clinician to the next.</li> <li>Clinical problem lists or secondary conditions are not standardized and tend to reflect each individual clinician's decisions about the problems that most affect patients in the moment. New regulatory standards might influence capture so that these lists become more standard and comprehensive. However, this will require significant change in clinician behavior/workflow.</li> <li>UB-04 claims forms cannot be filled out without diagnosis information so this is consistently available in administrative databases.</li> <li>For surgical coding, condition coding associated with specific procedures is likely to be more consistent and well defined</li> <li>The introduction of ICD-10 in 2014 and the Meaningful Use SNOMED mandate will likely make it more difficult to accurately and reliably pull this information in the coming years.</li> </ul> <p><u>Comments from Survey</u></p> <ul style="list-style-type: none"> <li>Unclear how emphasis on Meaningful Use will affect capture of clinical problem lists and admission diagnosis (which might be closest to primary discharge diagnosis)</li> </ul>		√
<b>Condition – Inactive</b>	<p><u>Survey Consensus Score</u></p> <ul style="list-style-type: none"> <li>0.31</li> </ul> <p><u>Comments from Survey</u></p> <ul style="list-style-type: none"> <li>Inactive conditions will not be consistently captured.</li> </ul>		√
<b>Condition - Resolved</b>	<p><u>Survey Consensus Score</u></p> <ul style="list-style-type: none"> <li>0.25</li> </ul> <p><u>Comments from Survey</u></p> <ul style="list-style-type: none"> <li>Resolved conditions will not be consistently captured.</li> </ul>		√

Category - Subcategory	Summary of TEP Scoring Discussion and Comments	Subcategory Should be Included in Feasibility Testing	
		Yes	No
Medication – Order	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.94</li> </ul> <p><b><u>TEP Consensus</u></b></p> <ul style="list-style-type: none"> <li>Data elements meet all 3 data capture criteria. Medication orders are not consistently encoded with a standard value set (RxNorm). However, the local codes assigned by the EHR software or code systems from pharmacy software vendors are easily mapped to RxNorm.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>Orders are more consistently captured and clearly defined compared with other medication subcategories such as administered.</li> <li>Orders are most likely to be associated with RxNorm codes compared with data generated from pharmacy (e.g., medication administered).</li> </ul>	√	
Medication - Administered	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.75</li> </ul> <p><b><u>TEP Consensus</u></b></p> <ul style="list-style-type: none"> <li>Data elements currently meet all 3 data capture criteria. However, they are difficult to interpret or apply in analyses as there is no consistent capture of the many nuances of medication administration (dose received, route, rate, starts and stops for drips, etc.).</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>Many pharmacies use proprietary coding that could be readily mapped to RxNorm but is not mappable at this time.</li> <li>Administration is problematic due to the complexity of drug dosing schedules, routes of delivery, dose timing delays/cancelations, etc. This a lot of data to wade through and define in a standardized way.</li> </ul> <p><b><u>Comments from Survey</u></b></p> <ul style="list-style-type: none"> <li>Increased capability to perform bar code scanning might help better integrate these data into EHRs.</li> </ul>		√
Medication - Discharge	<p><b><u>Survey Consensus Score</u></b></p> <ul style="list-style-type: none"> <li>0.73</li> </ul> <p><b><u>Final Consensus</u></b></p> <ul style="list-style-type: none"> <li>Discharge medications are not consistently captured. When captured, the data are often linked to local codes or pharmacy system proprietary coding structures. These can be readily mapped to RxNorm.</li> </ul> <p><b><u>TEP Discussion Points</u></b></p> <ul style="list-style-type: none"> <li>Meaningful Use will likely improve capture of complete lists of discharge medications although this aspect of transition of care is still evolving and falls short of being consistently performed.</li> <li>Currently these data are best extracted from records of prescriptions written in the EHR. These are not consistently encoded and must be mapped to RxNorm. In addition, what is prescribed might not represent all discharge medications.</li> </ul>		√

Category - Subcategory	Summary of TEP Scoring Discussion and Comments	Subcategory Should be Included in Feasibility Testing	
		Yes	No
<b>Laboratory Tests with Threshold Values</b>	<u><b>TEP Discussion Points</b></u> <ul style="list-style-type: none"> <li>Range limits and threshold values for laboratory tests are often text or string data and can only be used after some data parsing separate the numbers extracted in the text fields.</li> </ul>		√
<b>Discharge Disposition Data Elements</b>	<u><b>TEP Discussion Points</b></u> <ul style="list-style-type: none"> <li>Death in hospital is consistently captured with a standard definition in a structured field in most EHRs.</li> <li>There is routine capture of discharge disposition (including AMA and in-hospital death) although the location may vary in different EHR environments.</li> <li>These elements are not associated with specific codes.</li> </ul>	√	
<b>Care Directives and Palliation Data Elements</b>	<u><b>TEP Discussion Points</b></u> <ul style="list-style-type: none"> <li>Code status is captured in orders but is not associated with a standard code.</li> <li>Palliation is currently a poorly defined clinical concept.</li> <li>Other care directives (comfort measures only, plans underway to enter hospice) are not routinely captured in the EHR.</li> <li>SNOMED-CT codes for these concepts are now being integrated into EHR order sets and will likely be more available in near future.</li> </ul>		√

## **Appendix B**

### **Appendix B1. Literature review**

#### *Introduction*

A number of indices have been developed to predict mortality across a variety of medical and surgical conditions in adult hospitalized patients. These indices have been used for clinical research and to support clinical care in a variety of settings and patient populations, including in patients treated in EDs, admitted to ICUs, surgical inpatients, and inpatients cared for by general and specialty medical care teams. The clinical variables often included overlap between indices. Variables such as age, gender, common laboratory test results, vital signs measured, patients' comorbid conditions, procedures performed, and medications taken are all common to multiple mortality prediction indices. To ensure that these common variables were included in the feasibility analyses we conducted, we performed a systematic review of the literature describing mortality prediction indices that have been developed and validated.

#### *Methods: Environmental scan*

To identify potential mortality prediction indices, we consulted with a number of experts and identified 13 common mortality prediction indices that use clinical data:

1. Original National Surgical Quality Improvement Program
2. Revised National Surgical Quality Improvement Program
3. Simplified Acute Physiology Score (SAPS) III
4. Mortality Probability Model (MPMo) III
5. The mortality prediction model for emergency medical admissions developed by Goodacre, et al.
6. Sequential Organ Failure Assessment (SOFA) score
7. The mortality prediction model developed by Tabak, et al.
8. Logistic Organ Dysfunction System (LODS) score
9. Laboratory Acute Physiology Score (LAPS)
10. Acute Physiology and Chronic Health Evaluation (APACHE) IV score
11. Physiologic and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM)
12. Elixhauser Comorbidity Index
13. Charlson Comorbidity Index<sup>5-17</sup>

We used Medical Subject Heading (MeSH) and key terms from source documents describing the derivation and validation of these scales to develop our search strategy for the systematic literature review.

We ranked the MeSH terms and keywords by how often they appeared as search terms in these documents. We chose the ten most common terms, omitting those related to patient characteristics (e.g., humans, middle aged, female, male) and to care setting, to build our search strategy.

#### *Methods: Literature Search*

We performed a search using MEDLINE using the following search strategy. We performed each of the following MeSH term and keyword searches:

- \*Hospital mortality OR \*Outcome assessment (health care)
- Severity of illness
- \*Risk adjustment OR prognosis
- Logistic models OR Predictive value of tests OR Reproducibility of results
- Prospective studies OR Cohort studies

Search terms marked with an asterisk (\*) indicate MeSH terms. Search terms that do not have an asterisk indicate keywords. We combined each of these searches using the Boolean operator “AND” and limited final publications to English publications involving adult, human patients, and those published in the last five years (2008-2013).

### *Study Selection*

We identified inclusion and exclusion criteria *a priori* and then applied them to the identified publications. We wanted to include only indices that were developed for hospitalized, adult patients, not those with chronic or acute illnesses treated in outpatient settings. We were also only interested in scales predicting mortality across a variety of conditions and adult inpatient populations.

For the purposes of tabulation, exclusions were applied sequentially. We excluded publications:

- That did not have an adult population.
- That did not use a hospitalized population of patients.
- In which there were no identifiable predictors, scales, or indices for mortality tested or used.
- That used a duplicate scale (one we had already identified in our environmental scan).
- That used a predictor or group of predictor variables, but not a scale or index.
- That used a condition-specific prediction scale (e.g., a scale to predict mortality among adults hospitalized for pneumonia).

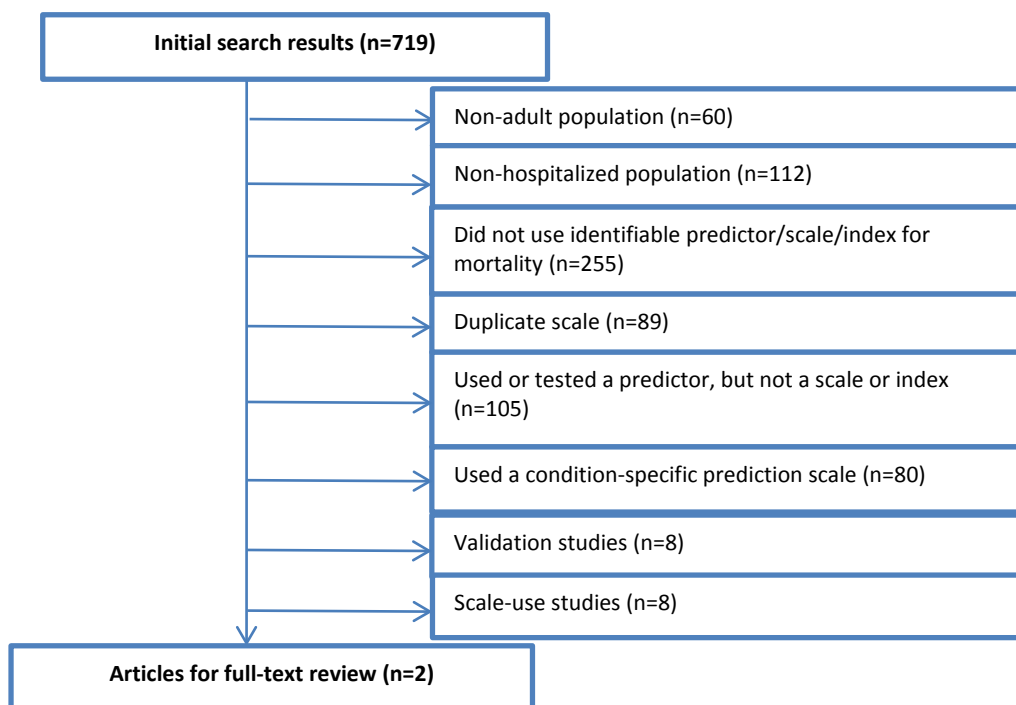
Additionally, validation studies and scale-use studies were identified during the initial review to track the validation of the identified scales, or for reference in case a validation study needed to be retrieved. These publications were not full-text reviewed.

We expected our search to produce articles related to the development and validation of well-known, condition-nonspecific mortality risk measures, such as those 12 identified above.

### *Results*

The electronic literature search identified 719 unique publications. Two reviewers independently examined the title and abstract of each of the publications and rejected 709 publications for meeting at least 1 of the 6 exclusions listed above. This initial abstract review left 10 articles for full-text review. Upon detailed review of these publications, 8 additional articles were excluded. The most common reason publications were excluded was that they “did not use an identifiable predictor/scale/index for mortality” ([Figure B.1](#)).

**Figure B.1: Flow diagram of literature search initial results to articles that were full-text reviewed**



**Table B.1: Data Elements for Patient Characteristics in Mortality Indices**

	NSQIP original	NSQIP revised	SAP S III	MPM O III	ED Mort	SOF A	Tabak	LOD S	LAP S	APACH E IV	POSSU M	Elixhauser	Charlson
Age	√	√	√	√	√	-	√	-	-	√	√	-	-
Gender	-	-	-	√	-	-	-	-	-	-	-	-	-
Full code	-	-	-	√	-	-	-	-	-	-	-	-	-
American Society of Anesthesiologists class	√	-	-	-	-	-	-	-	-	-	-	-	-

**Table B.2: Data Elements for Physical Examination in Mortality Indices**

	NSQIP original	NSQIP revised	SAPS III	MPMO III	ED Mort	SOFA	Tabak	LODS	LAPS	APACHE IV	POSSUM	Elixhauser	Charlson
GCS	-	√	√	√	√	√	√	√	-	√	√	-	-
Blood pressure *mean/ **systolic	-	√*	√*	√**	√*	-	√	√**	-	√*	√	-	-
Heart rate	-	√	√	√	-	-	-	√	-	√	√	-	-
Temperature	-	√	-	√	-	-	-	-	-	√	-	-	-
Respiratory rate	-	-	-	√	-	-	-	-	-	√	-	-	-
Oxygen saturation	-	-	-	√	-	-	-	-	-	-	-	-	-
Urine output	-	-	-	-	√	-	-	√	-	√	-	-	-

**Table B.3: Data Elements for Laboratory Results in Mortality Indices**

	NSQIP original	NSQIP revised	SAPS III	MPMO III	ED Mort	SOFA	Tabak	LODS	LAPS	APACHE IV	POSSUM	Elixhauser	Charlson
<b>Creatinine</b>	√	√	-	√	√	√	√	√	√	√	-	-	-
<b>WBC</b>	√	√	-	√	-	√	√	√	√	√	√	-	-
<b>BUN</b>	√	-	-	√	-	√	√	√	√	√	√	-	-
<b>Bilirubin</b>	√	√	-	-	√	√	√	-	√	√	-	-	-
<b>Sodium</b>	√	-	-	√	-	√	-	-	√	√	√	-	-
<b>Arterial pH</b>	√	√	-	-	-	-	√	-	√	√	-	-	-
<b>Albumin</b>	√	-	-	-	-	√	√	-	√	√	-	-	-
<b>Platelet</b>	-	√	-	√	√	√	-	√	-	-	-	-	-
<b>Glucose</b>	√	-	-	√	-	-	-	-	√	√	-	-	-
<b>Hematocrit</b>	√	-	-	-	-	-	-	-	√	√	√	-	-
<b>Potassium</b>	-	-	-	√	-	-	-	-	-	-	√	-	-
<b>PaCO<sub>2</sub></b>	√	-	-	-	-	-	-	-	√	-	-	-	-
<b>FiO<sub>2</sub> or PaO<sub>2</sub></b>	√	√	-	-	√	-	-	√	√	√	-	-	-
<b>Bicarbonate</b>	-	-	-	-	-	-	-	-	√	√	-	-	-
<b>Anion Gap</b>	-	-	-	-	-	-	-	-	√	-	-	-	-
<b>PTT</b>	-	-	-	-	-	-	√	-	-	-	-	-	-
<b>Hemoglobin</b>	-	-	-	√	-	-	-	-	-	-	-	-	-
<b>Troponin-1</b>	-	-	-	-	-	-	-	-	√	-	-	-	-



## Appendix B2: Results of the Premier member hospital surveys

**Table B.4: Hospital System Characteristics**

	Hospital A	Hospital B	Hospital C	Hospital D
Census Region	South Atlantic	Mid-Atlantic	South Atlantic	New England
Hospitals in System	3	2	1	1
Teaching Hospital	No	Yes	No	Yes
Licensed Beds	>500 beds	>500 beds	300-399 beds	>500 beds
EHR System	Cerner Millenium	Cerner Millenium	Meditech	Allscripts

**Table B.5: Identification of Hospital Entry Time**

	Hospital A	Hospital B	Hospital C	Hospital D
Identify transfer from another facility	Yes	Yes	Yes	No
Distinguish between inpatient admission and ED admission from the transferring hospital	Yes	No	Yes	n/a
Time of first arrival in ED	Yes (whether admitted or not)	Yes (whether admitted or not)	Yes	Yes (whether admitted or not)
Time of first arrival on inpatient unit	Yes	Yes	Yes	Yes
Time of admission	Yes	Yes	Yes	Yes
Different depending upon initial arrival location	Yes	Yes	Yes	No
Link ED and admission data	Yes	No	Yes	No

**Table B.6: Capture and Extraction Capabilities of Core Clinical Data Elements**

Data Element	Captured In Inpatient EHR	Captured in ED EHR	Structured Format	Extracted for Reporting	Extracted for Other Purposes	Time and Date Stamps Captured
<b>Vital Signs</b>						
Blood Pressure	4/4	4/4	4/4	4/4	4/4	4/4
Heart Rate	4/4	4/4	4/4	4/4	4/4	4/4
Respiratory Rate	4/4	4/4	4/4	4/4	4/4	4/4
Temperature	4/4	4/4	4/4	4/4	4/4	4/4
Oxygen Saturation (SaO2)	4/4	4/4	4/4	4/4	4/4	4/4
Weight	4/4	4/4	4/4	4/4	4/4	4/4
<b>Laboratory Tests</b>						
Hemoglobin	4/4	4/4	4/4	4/4	4/4	4/4
Hematocrit	4/4	4/4	4/4	4/4	4/4	4/4
White Blood Cells	4/4	4/4	4/4	4/4	4/4	4/4
Platelets	4/4	4/4	4/4	4/4	4/4	4/4
Sodium	4/4	4/4	4/4	4/4	4/4	4/4
Potassium	4/4	4/4	4/4	4/4	4/4	4/4
Chloride	4/4	4/4	4/4	4/4	4/4	4/4
Bicarbonate	4/4	4/4	4/4	4/4	4/4	4/4
BUN	4/4	4/4	4/4	4/4	4/4	4/4
Creatinine	4/4	4/4	4/4	4/4	4/4	4/4
Glucose	4/4	4/4	4/4	4/4	4/4	4/4

**Table B.7: Additional Considerations for Laboratory Test Results**

	Hospital A	Hospital B	Hospital C	Hospital D
<b>Laboratory data directly entered into EHR</b>	Yes	Yes	No	No
<b>Laboratory data imported into EHR</b>	No (Lab integrated into the EHR)	Yes	Yes	Yes
<b>Store specific upper reference limits</b>	No	Yes (HL7)	Yes	No