Summary Report of the Technical Expert Panel Meeting, April 2013

(Deliverable #20)

Development, Maintenance and Support of Hospital Clinical Quality Measures for ARRA HITECH

Contract # HHSM-500-2008-00023I,
Task Order # HHSM-500-T0001

April 18, 2013

Prepared for:
Deborah Krauss
Division of Health Information Technology
Quality Measurement & Health Assessment Group
Center for Clinical Standards & Quality
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Mail Stop S3-02-01
Baltimore, MD 21244-1850

Submitted by:
Abt Associates Inc.
55 Wheeler Street
Cambridge, MA 02138

Summary Report, TEP #3

1. Executive Summary

On April 9 and 10, 2013, the Hospital eMeasures team convened its third Technical Expert Panel (TEP). As part of the Hospital eMeasures contract, Abt Associates and our partners have been commissioned to develop five new (*de novo*) clinical quality measures (CQMs) to report directly from hospital electronic health records (EHRs). The third TEP meeting focused on the clinical concepts which were approved for development in the upcoming contract year, all of which are outcomes measures. This included a discussion of the feasibility of development, including finding the necessary data elements for the measures in EHRs, as well as ability to risk-adjust the measures in an electronic measure format.

The TEP meeting included an overview of the five concepts proposed for *de novo* development, presented by Dr. Mark Metersky, co-Investigator on the project. The bulk of the TEP meeting was spent reviewing risk-adjustment methodology, feasibility concerns and progress on data acquisition; the overarching theme in these discussions was the innovation and challenges of developing outcomes measures using EHR data. Despite the complex processes involved, the TEP was strongly in favor of this approach to measure development and offered suggestions and support to this process.

This document summarizes the proceedings of the April 9-10, 2013 TEP meeting and provides an overview of next steps for the Hospital eMeasures project team.

2. Outcomes of TEP Meeting

Day One

The TEP began with an overview of CMS' strategy on measure development. Project Officer Deborah Krauss highlighted the National Quality Strategy aims of better health; better care and lower costs as well as the priority domains and the quality improvement objectives outlined for all measures being developed by CMS. Project Director Terry Moore then provided an overview of project accomplishments over the past project year, highlighting the work to date on: developing two *de novo* measures, retooling of existing measures for HITECH Meaningful Use Stage 2; and testing both of these measure sets in the field.

Dr. Metersky provided an overview of the five concepts proposed for *de novo* development (Appendix 1) and solicited comments from the TEP. Considerations raised are highlighted below:

- Many of the proposed measure specifications presented issues with operationalization: for example, how to acquire data on denominator exclusions for the ICU readmission measure where data such as planned readmissions are not reported reliably, or in a consistent manner. Other measures, such as the acute kidney injury measure, raised issues of being able to separate iatrogenic effects from conditions present at admission.

- It will be important to ensure that behind every proposed outcomes measure are concrete processes that we can target for improvement to lead, in turn, to improved outcomes. We must understand what accounts for process variation and adjust for these in risk-modeling. A clear understanding of the relationship between processes and outcomes will also facilitate development of clinical decision support (CDS) guidance that can be used for quality improvement.

- The team may want to narrow our measure development focus from global outcomes—for example on all failure to rescue episodes—to elements of an outcome that are feasibly and reliably reported in electronic data and for which there is evidence that improvement can occur. It was agreed that the Failure-to-Rescue measure should be on a longer development timeline, due to outstanding issues about defining the construct.

- Measure development should be aligned with other parallel initiatives to improve quality using EHR data, including the incorporation of CDS, to improve clinical interventions.

- We should look at claims data to elucidate documentation related to a particular diagnosis to help define data elements that we can use to build eCQMs (the feasibility of finding these elements in an EHR will be a second step needed here).

Hospital eMeasures team members Ryan Fair and Chengjian Che presented further background and future directions on feasibility testing, underscoring the multiple iterations of feasibility testing needed to obtain valid results and various ways in which feasibility is defined. For example, feasibility includes the extent to which data elements are captured in an EHR, as well as how well

data elements overlap - or are reported in the same way- across EHRs to ensure measure can be nationally implementable.

The issue was raised regarding whether measures should be developed based on what currently exists in EHRs, versus an ideal state in measure development—Should CMS "lead or follow" the measure developers? The TEP, including panel members from CMS and ONC, provided feedback in response to the need to overcome the absence of existing data elements in the EHR. The team should use the measure development process to present data elements to MAT developers, vendors and other government authorities that are priorities for EHR capture. A similar issue was raised in the context of risk adjustment covariates: The TEP wondered whether measure developers should work primarily with covariates found in an EHR, or base their covariate choice on the literature first, then determine data element feasibility. CMS and ONC emphasized the use of "core" risk adjustment variables, currently being developed, as well as what exists in the QDM. However, since electronic measure development is so new, we have to treat the operationalization of the measures as an ongoing process which necessitates the support of the government to encourage the evolution of new data fields in EHRs and in the MAT, and other measure development tools.

Day Two

The second day of the TEP focused on a discussion of risk adjustment, both why it is important for outcomes measures, and how the Hospital eMeasures team is proposing to undertake risk adjustment in the context of its current work. Specifically, Abt's Dr. Alison Sexton presented on hierarchical risk models, which would account for random variation in hospital-specific characteristics, but not quality of care. This methodology is considered best practice for risk modeling and is used by many NQF-endorsed risk-adjusted measures. Some highlights from this discussion are listed below.

- There are multiple challenges inherent in measuring outcomes using EHR data in way that allows fair comparisons among providers. Specific methodological questions include:
    - Will missing observations be handled differently in risk modeling from how they will be addressed in measure specification? Current criteria require positive proof that the condition is met.
    - How to address correlation of high outcomes to one another? There are certain structural factors that hospitals cannot alter.
    - Should measures be stratified? This will be informed in part by the proposed propensity scoring analysis.
    - Socio-economic factors cannot and should not be used to adjust observed outcomes, but they will be correlated with outcomes and are not within a hospital's control.
    - How much data/episodes of care are necessary to develop and test the risk models? How many hospitals should be included in the test bed?
    - How do we define "good", or "better" performance of the risk model?

  - Applying hierarchical risk models will ultimately require data from a wider array of hospitals than those that will be used to develop the model, due to the need to develop hospital-specific intercepts.

- There is currently a lack of functionality and precedent for incorporating risk adjustment into an eMeasure: the present system/environment does adequately support risk adjustment for eCQMs. However changes to HQMF and the Measure Authoring Tool (MAT) to address this issue have been proposed and are expected to be implemented in this calendar year.

- Team members emphasized that because the Hospital eMeasures team will be undertaking one of the first efforts to develop and operationalize risk adjustment in an EHR environment, this will be a proof-of-concept that will require multiple iterations. In addition, models will require ongoing refinement as national data become available. It was recommended that the team collaborate with and learn from other measure initiatives reporting risk-adjusted outcomes, although some EHR-specific issues are unique.

Dr. Metersky then led a discussion of the potential covariates for each of the measure constructs. TEP members provided comments and questions that will be addressed in the individual clinical work groups. One major issue for consideration was the available data sources for defining comorbidities, and the use of problem lists versus diagnosis codes.

The meeting ended with an overview of data needs for developing and testing risk adjustment models, potential data sources and the status of data acquisition efforts. To develop and test valid risk adjustment models for the proposed measure constructs, the team requires access to a large volume of de-identified ER and inpatient data that includes all the potential data elements for measure and model specifications. Options include administrative/claims data, registry data and EHR data; advantages and disadvantages of each were reviewed.

The following issues were raised in discussion:

- EHR data are the most appropriate for developing and testing risk models for eMeasures, because ultimately these measures will be calculated using EHR data elements.

- The lack of an available national dataset for research and testing is problematic. While efforts are underway to create such a dataset, the timeline does not align with Meaningful Use Stage 3. It may be appropriate to start with a smaller, non-representative dataset, for example, from a set of hospitals in one health system, with the understanding that multiple iterations will be required to refine the final model for any measure used to compare facilities.

- The TEP endorsed using the existing eICU research dataset, which includes data from more than 400 ICUs, to develop and test the models for the ICU readmission measure.

- The team should leverage ongoing conversations between ONC and private sector efforts to create EHR databases for secondary use purposes in their search for potential data sources.

3. Next Steps

Based on TEP feedback, the Hospital eMeasures team will continue to pursue data acquisition and development of an analytic plan for risk modeling. In addition, we will convene clinical work groups of relevant experts who can provide guidance around the specific questions and concerns raised about specifying the five proposed *de novo* measure constructs. The first of these work group meetings, for trauma mortality, was scheduled for April 11, 2013. The team will keep the TEP updated on all of these activities, and may access the expertise of individual measures as needed.

4. Appendices

4.1 Appendix 1: Five proposed *de novo* measures

- *Rate of respiratory failure developing 48 hours or more after hospital admission*: This measure will look at non-surgical patients who require ventilation beginning greater than 48 hours after admission to the hospital. This common, costly, and high morbidity/mortality event is frequently an end result of lapses in care (e.g., opiate overdose, excessive volume repletion, aspiration). This will be a risk-adjusted outcomes measure.
- *Rate of hospitalized patients without acute renal failure at the time of admission who meet established criteria for acute kidney injury 48 hours or longer after admission*: This is also a common, costly, and high morbidity/mortality event. This measure will be restricted to patients who develop acute kidney injury, most likely defined by changes in serum creatinine and diminished urine output 48 hours or more after hospital admission and will also be specified as a risk-adjusted, outcomes measure.
- *Failure to rescue mortality in-hospital*: This measure will look at mortality resulting from 'failure to rescue' in hospital patients- specifically General Surgery patients, Orthopedic and Vascular patients with complications, plus patients who died in the hospital without complications. This measure will be a risk-adjusted outcomes measure, similar to the concept of endorsed paper-based measure NQF 0352. Further, the measure may be further narrowed in scope to include a smaller subset of patients, depending on the availability of EHR data and clinical work group guidance.
- *Rate of readmission to ICU within 48 hours of discharge from the same ICU*: Readmission can be easily measured from EHRs, is a common quality metric and is associated with mortality and morbidity. This measure would require risk adjustment.
- *In-hospital mortality rate for selected trauma ED admissions*: This measure will look at all deaths among adolescent and adult trauma patients. This outcomes measure requires a robust risk-adjustment model.