

# **Compilation of Small Projects to Improve MEPS Medical Provider Component (MPC) Quality**

Prepared for:

Agency for Healthcare Research and Quality  
Center for Financing, Access, and Cost Trends  
5600 Fishers Lane  
Rockville, Maryland 20857

Prepared by:

RTI International  
P.O. Box 12194  
Research Triangle Park, NC 27709

Disclaimer: The views expressed in this article are those of the authors, and no official endorsement by the U.S. Department of Health and Human Services or Agency for Healthcare Research and Quality.

---

**Compilation of Small Projects to Improve  
MEPS Medical Provider Component (MPC) Quality**

**Table of Contents**

Foreword ..... ii

Project #1: Thoughts on Not Fielding All SBD Nodes (August 2, 2012) ..... 1-1

Project #2: Special Project Task 1: A Preliminary Study on the Ability of Hospitals to  
Extract and Electronically Submit MPC Data Final Report (June 30, 2016)..... 2-1

Project #3: Unified Approach Summary (May 5, 2017) ..... 3-1

Project #4: SBD Special Project – Interviews on SBDs – Summary Report (June 23, 2017) ... 4-1

Project #5: MEPS MPC OP3 special project – review of SBD data mining – findings  
(July 20, 2017) ..... 5-1

---

## Foreword

This Working Paper is a compilation of the findings from five separate projects—conducted between 2012 and 2017—that examined approaches to improving the quality of MEPS Medical Provider Component (MPC) data. Each project was undertaken to identify and assess areas to improve data collection strategies and to ensure efficiencies. The primary focus was to address ways to quickly identify separately billing doctors (SBD) in the MPC. These efforts were intended to increase accuracy and reduce resources needed during the process of identifying eligible SBDs with a view to ensure quality and timeliness in the data.

---

This page intentionally left blank.

— **Project #1: Thoughts on Not Fielding All SBD Nodes** —

**Thoughts on Not Fielding All SBD Nodes**  
**MEPS MPC**  
**(August 2, 2012)**

Staff from AHRQ, RTI, and SSS have briefly discussed the idea of not fielding all separately billing doctor (SBD) nodes for the 2012 Medical Provider Component (MPC) in order to reduce the cost of the MPC data collection effort and minimize burden on potential SBD respondents. The following set of notes reflects some initial thoughts on this topic.

To begin, we first note that currently not all SBD nodes are fielded. Specifically, certain SBD nodes that have been collected among events in the office-based doctor (OBD) sample are not fielded again. Data from the OBD sample are essentially transcribed over to the SBD files. Also, if a firm refusal has occurred during OBD data collection, the case is not fielded in SBD data collection. However, if the OBD data collection result was unlocatable or a disavowal or if the overlap of SBD nodes and OBD data is ambiguous, the nodes are fielded during SBD data collection to clear up ambiguity.

For the purpose of this discussion, we can partition the SBD sample into four sets. Thoughts on how to deal with each set are provided below.

**1. SBD/OBD Overlap: Clean Match**

This set includes the following:

- PDDIRID of the node is an exact match with the OBD.
- OBD is a final refusal.

*Action:* We propose to declare the SBD a refusal. We will not contact the OBD.

- Date of SBD node is within +/- 1 day of the OBD event.
- PDDIRID of the node is an exact match with the OBD.
- OBD event has an S-type location of service (i.e., inpatient, outpatient, or ER).

*Action:* We propose to match these OBD/SBD pairs with minimal professional staff review. We will not contact the OBD.

- Date of SBD node is within +/- 1 day of the OBD event.
- PDDIRID of the node is an exact match with the OBD.
- OBD event is not an S-type location of service.

*Action:* If the hospital event is an outpatient event with simple radiology and pathology codes, we will review the specialty and role of the SBD. If the SBD is a general practice, primary care, pediatrician, internist, or any other general specialty or if the role is referred or copied, the node has a high likelihood of becoming a referred or copied doctor disavowal. We will not contact the SBD and we will treat it as a disavowal. Otherwise, we will re-contact the OBD.

## 2. SBD/OBD Overlap: Possible Match

This set includes the following:

- PDDIRID of the node is not an exact match with the OBD but name, address and other information suggest an OBD could be the same as an SBD node we seek.

**Action:** We will review these cases by hand and, if need be, re-contact the OBD for clarification.

- Date of SBD node is not within +/- 1 day of the OBD event.
- PDDIRID of the node is an exact match with the OBD.
- OBD event has S-type location of service for other dates.

**Action:** We propose to create series of comparisons to align specialty and role of provider recorded during hospital with the Current Procedural Terminology, 4th Edition (CPT4) codes (or description of services) recorded during OBD. The specifications would indicate whether the event would fill the node, cause a disavowal, or require the node to be fielded. For example, if the CPT4 was for a surgery and the node was associated with a surgeon, the OBD event could be used to fill the node without fielding.

- Date of SBD node is not within +/- 1 day of the OBD event.
- PDDIRID of the node is an exact match with the OBD.
- OBD event has no S-type location of service for other dates.

**Action:** Same as above but specifications may have differences between accepting the event as a disavowal.

## 3. SBD Does Not Overlap with OBD: Hospital Outpatient or ER Events

This set includes those SBD nodes that have not been collected in the OBD sample. This set includes the following cases:

- Hospital outpatient events with simple radiology and pathology codes.
- SBD specialty is general practice, primary care, pediatrician, internist, or any other general specialty or the SBD role is referred or copied.

**Action:** The node has a high likelihood of becoming a referred or copied doctor disavowal. We will not field these SBDs and treat them as a disavowal.

- All other SBD nodes associated with Hospital Output or ER Events

**Action:** Consider sampling this set using a simple “on-the-fly,” Bernoulli-type sampling approach with selection probabilities pre-assigned and based on role and specialty. For example:

If role is XXX and specialty is YYY, then pick a random number between 0 and 1. If the number is less than ZZZ percent then field the SBD. Otherwise, code the SBD as "Not In Sample." The ZZZ percentage associated with each XXX and YYY pair is something we would want to determine with AHRQ staff.

#### **4. SBD Does Not Overlap with OBD: Hospital Inpatient, Long-Term Care, or “Other” Events**

In general, we propose fielding all these SBD nodes due to their likely significant contribution to the hospital event. We might consider looking at historical data and identifying special cases where the SBD node appears to contribute very little to the Hospital event. To identify these cases, we propose looking at:

- Specialty and role of the SBD.
- CPT4 codes (or description of services) associated hospital event.

We could consider developing some rules such as:

- If role is XXX, specialty is YYY, and one or more CPT codes are CCC, then treat the SBD node as a disavowal.

Or

- If role is XXX, specialty is YYY, and one or more CPT codes are CCC, then subsample the SBD node at a rate of ZZZ percent, since the charge/payment contribution of the SBD is likely to be small.

We suspect, in most cases, that we would like to field the SBD node. For example; if the specialty was “surgeon” and the node was associated with a surgery, we would want to field these nodes.

**— Project #2: Special Project Task 1: A Preliminary Study on the Ability of Hospitals to Extract and Electronically Submit MPC Data Final Report —**

June 30, 2016

Medical Expenditure Panel Survey (MEPS)

Special Project Task 1:

A Preliminary Study on the Ability of Hospitals to Extract and Electronically Submit MPC Data

Final Report

Prepared by

Donald T. Mon, PhD

Robert Bailey

RTI International

3040 E. Cornwallis Road

Research Triangle Park, NC 27709



**CONTENTS**

<b><u>Section</u></b>	<b><u>Page</u></b>
Section 1 Background.....	2-5
1.1 Purpose of Special Project 1 .....	2-5
1.2 Goals and Objectives .....	2-6
1.3 What Problems Are We Trying to Solve? Current Issues in Collecting MPC Data.....	2-7
1.3.1 No Hospitals Are Currently Submitting Financial Data in Discrete, Computable Form .....	2-7
1.3.2 Accurately Capturing Separately Billing Doctors and Their Cost and Payment Information Is Challenging .....	2-8
Section 2 Methodology.....	2-9
2.1 Develop a Discussion Document .....	2-9
2.2 Select Hospitals.....	2-10
2.3 Identify Contacts and the Individuals to Invite to the Meeting .....	2-11
2.4 Set Up a 60- to 90-Minute Meeting to Discuss Whether MPC Data Extraction and Submission Are Possible .....	2-11
2.5 Conduct the Meeting.....	2-12
Section 3 Results.....	2-12
3.1 Findings from Hospital 1 .....	2-12
3.1.1 General Topics of Discussion .....	2-12
3.1.2 Patient Identification and Matching.....	2-13
3.1.3 Patient Demographics .....	2-14
3.1.4 Patient Encounters .....	2-14
3.1.5 Diagnosis Data .....	2-14
3.1.6 Procedure Data.....	2-14
3.1.7 Data on Doctors Who Performed a Service or Procedure on the Patient AND Who Are Not included in Service/Professional Fees.....	2-15

3.1.8	Charge and Fees Data .....	2-15
3.1.9	Payment Data .....	2-16
3.1.10	Data Submission .....	2-16
3.2	Findings from Hospital 2 .....	2-16
3.2.1	General Topics of Discussion .....	2-16
3.2.2	Patient Identification and Matching.....	2-17
3.2.3	Patient Demographics .....	2-17
3.2.4	Patient Encounters .....	2-18
3.2.5	Diagnosis Data .....	2-18
3.2.6	Procedure Data.....	2-18
3.2.7	Data on Doctors Who Performed a Service or Procedure on the Patient AND Who Are Not included in Service/Professional Fees .....	2-18
3.2.8	Charge and Fees Data .....	2-18
3.2.9	Payment Data .....	2-19
3.2.10	Data Submission .....	2-19
3.3	Recommendations.....	2-19
Section 4 Conclusions.....		2-23

**LIST OF TABLES**

<b><u>Number</u></b>	<b><u>Page</u></b>
Table 1. Goals and Objectives of Special Project 1 .....	2-6
Table 2. Tasks Performed in Special Project Task 1 .....	2-9

**LIST OF FIGURES**

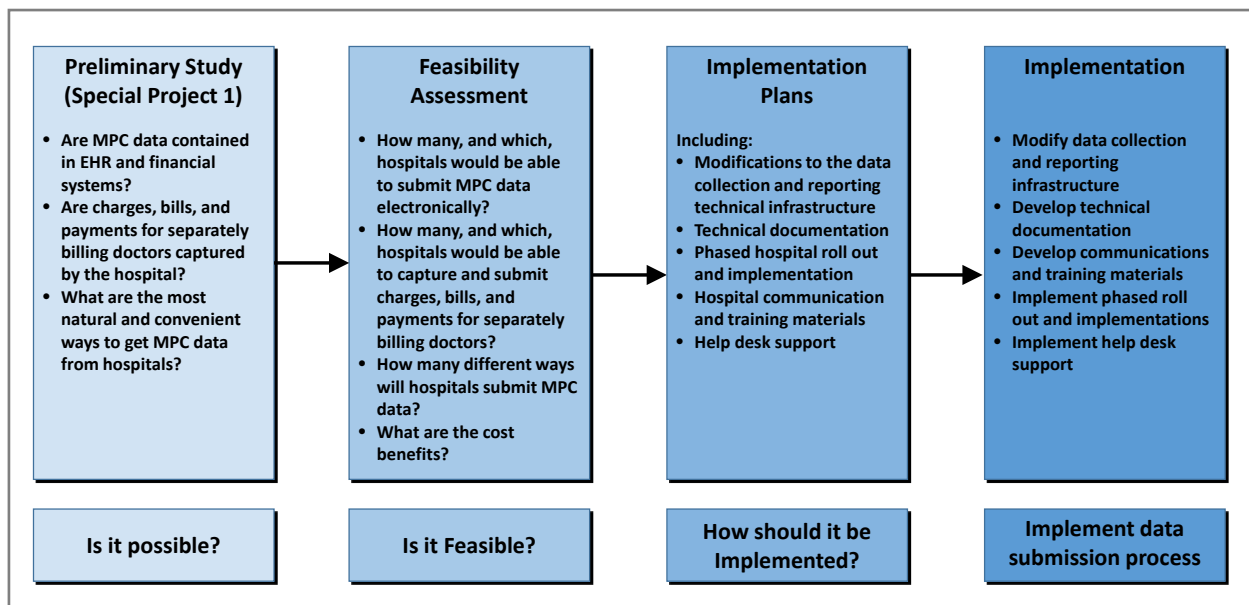
<b><u>Number</u></b>	<b><u>Page</u></b>
Figure 1-1. A Cautious, Incremental Approach to Implementing Large-scale Electronic Data Submission .....	2-5

## SECTION 1 Background

### 1.1 Purpose of Special Project 1

Currently, collecting Medical Expenditure Panel Survey (MEPS) Medical Provider Component (MPC) data through chart abstraction and follow-up conversations with hospitals and separately billing doctors (SBDs) consumes considerable time and effort. One way to reduce this time and effort may be through the electronic submission of data by the hospitals participating in the MPC. However, implementing a large-scale electronic data submission process requires careful planning and coordination. A cautious, incremental approach, such as that depicted in **Figure 1-1**, is recommended.

**Figure 1-1. A Cautious, Incremental Approach to Implementing Large-scale Electronic Data Submission**



Consistent with this measured approach, as authorized by the Agency for Healthcare Research and Quality (AHRQ), RTI International (RTI) conducted Special Project Task 1 as a preliminary study on the hospitals' ability to extract and submit MPC data electronically (see left-most box in **Figure 1-1**). The purpose of this preliminary study was to ascertain whether MPC data are contained in hospital electronic health record (EHR) and financial systems, and to identify the most convenient way for hospitals to extract and submit the data. Though there are no plans at this time to implement an electronic data submission process, nor any Feasibility Assessment or Implementation Plans (i.e., the middle two boxes in **Figure 1-1**), the information from this preliminary study is useful in a number of ways.

Ascertaining that MPC data are contained in hospital information systems as discrete data demonstrates that data extraction and submission is possible—at least for a portion of hospitals participating in the MPC. On the other hand, should it be found in this preliminary study that

MPC data in hospital information systems are largely text or difficult to extract, then the prospect of implementing an electronic data submission process no longer needs to be pursued because there would be no major advantage to it.

Obtaining early impressions on the challenges and issues hospitals may experience in extracting and submitting MPC data is equally important. Such information might provide insights into how the process can be made easier for the hospitals. However, if it appears that data extraction and submission is too arduous, then, as above, the notion of implementing the process can be dropped.

Finally, the findings of this preliminary study can inform the decision to conduct a more in-depth follow-up Feasibility Assessment (second box in **Figure 1-1**) that estimates the number of hospitals that may be able to submit MEPS MPC data electronically, targets specific hospitals to approach, and quantifies the projected impact of a data extraction and submission program (e.g., decreased hospital data collection burden, reduced cost, increased efficiency, increased data quality). In addition, the information from this preliminary study can inform the design of the data collection tools needed in the follow-up Feasibility Assessment, should it be pursued.

## 1.2 Goals and Objectives

Given the purpose of Special Project Task 1 (see Section 1.1), the following goals and objectives were specified (see **Table 1**).

**Table 1. Goals and Objectives of Special Project 1**

Goal	Objective
<p><b>Goal 1:</b> Confirm that MPC financial data are captured as discrete and computable data, and demographic data as separate text fields, in hospital EHR and financial systems.</p> <p><b>Rationale:</b> Data extraction and submission is achievable only if MPC data are contained in hospital electronic systems as discrete data. If confirmed, then implementing an electronic data submission program is a viable option that can be implemented at AHRQ’s discretion. If not confirmed, then the information of this preliminary study can be used to identify ways to make the current chart abstraction process easier (see Goal 3), or a decision can be made to stop pursuing electronic data submission.</p>	<p>Determine whether MPC charge, billing, and payment data are contained in hospital EHR and financial (e.g., billing, patient accounting, cost accounting) systems in discrete and computable form.</p> <p>Determine whether patient and physician demographic data are contained in hospital EHR and financial systems as separate text fields.</p>

Goal	Objective
<p><b>Goal 2:</b> Confirm that hospitals can identify separately billing doctors and provide their respective charge, billing, and payment data.</p> <p><b>Rationale:</b> Collecting SBD charge, billing, and payment data consumes considerable time and effort. Getting SBD charge, billing, and payment data from the hospital, or at least data on the doctors who performed a service or procedure on the patient and their billing services, will reduce time, effort, and cost for both the hospital and the MEPS project.</p>	<p>Determine whether data in hospital systems can identify:</p> <ul style="list-style-type: none"> <li>• The doctors who performed a service or procedure on the patient during a hospital-based event,</li> <li>• Which of the doctors’ fees are included in the Professional Fees of the hospital’s bill, and</li> <li>• The billing service each doctor uses.</li> </ul>
<p><b>Goal 3:</b> Identify ways in which MPC data collection can be made easier for the hospitals.</p> <p><b>Rationale:</b> Making data collection easier can reduce hospital data collection burden, increase data quality and efficiency, reduce cost, and potentially increase hospital participation in the MPC.</p>	<p>Document the most natural and convenient way(s) hospitals can extract and submit MPC data electronically.</p> <p>(From Goal 1) If MPC data are not contained in hospital EHR and financial systems in discrete or computable form, or if electronic data submission is not feasible, determine how the current chart abstraction process can be made easier for the hospital.</p>

### **1.3 What Problems Are We Trying to Solve? Current Issues in Collecting MPC Data**

To accomplish the above goals and objectives, it is important to address the major issues in the current data collection process. These issues are described below. The manner in which they were discussed with the hospitals participating in this project is contained in Section 2 – Methodology. The findings and recommendations from these discussions can be found in Section 3 – Results. Conclusions are discussed in Section 4.

#### **1.3.1 No Hospitals Are Currently Submitting Financial Data in Discrete, Computable Form**

To take full advantage of electronic data submission, MPC financial data must be in discrete, computable form, with demographic data contained in separate text fields. However, no hospitals are currently submitting MPC data in this manner. Hospitals that are submitting data in some electronic fashion (i.e., not paper) are sending portable document format (pdf) files on compact discs (CDs). Thus, though they are not sending paper charts, the scanned images on the CDs are simply electronic representations of their paper charts. As with paper records, the data from imaged records must still be abstracted and key-entered into the MEPS MPC IDCS system. Consequently, while imaged records might result in reduced burden to the hospital point of

contact and savings in other direct costs (e.g., copying, mailing)—the savings of which are minor compared to the total cost of the process—there are no real savings due to increased efficiency.

To achieve these savings, a significant number of hospitals would need to extract and submit MPC financial data in discrete and computable form, and demographic data in separate text fields. It is unreasonable to expect that 100 percent of participating hospitals will be able to do so. However, even a 25 percent to 40 percent electronic submission participation rate would result in reduced data collection burden on the hospitals, increased data quality, and substantial savings in cost and efficiency. Confirming in this preliminary study that MPC data are available in these forms in hospital EHR and financial systems is the first step towards a successful electronic data submission solution.

### **1.3.2 Accurately Capturing Separately Billing Doctors and Their Cost and Payment Information Is Challenging**

The current process starts with collecting information about hospital events reported by patients in the Household Component. Once the event type (e.g., inpatient, outpatient, ER) and the dates of service for the event from hospital records have been captured (i.e., there is a correct match between an event reported by a patient and an event in the hospital's records), the names and other demographic information of SBDs are abstracted from the medical records.<sup>1</sup> In many instances (as in the case of a cardiac specialist), charges from these SBDs may account for a significant portion of total medical cost. Therefore, for analyses comparing hospital costs and professional fees, it is important to tease out the charges from SBDs. At the same time, charges from hospitalists, residents, interns, or other doctors-in-training, which are included in the hospital bill, must be excluded to eliminate duplicate charges.

This is the point at which the current data collection process often becomes problematic. Despite using precisely worded questions:

- All the SBDs associated with an event may not be captured, resulting in an incomplete record of total cost and payment.
- Doctors' charges already included in the hospital's bill may inadvertently be captured as charges from SBDs, resulting in an erroneous inflation of total cost and payment, as well as tremendous inefficiencies. If, after investing considerable time and effort, the unbeknownst erroneous cost and payment data are collected, additional time and effort are expended to correct the data when the error has been found.
- Doctors who are only copied on the record—that is, those who are not actively providing services and, therefore, do not have charges to report—may sometimes be captured as SBDs. As a result, significant time and effort is spent in two ways: (1) capturing charge and payment data from physicians who should have nothing to report, and (2) recording the disavowal.

---

<sup>1</sup> Separately Billing Doctors are defined as physicians who provided services during the event whose charges might not be included in the hospital bill, including attending physicians, surgeons, radiologists, pathologists, anesthesiologists, Emergency Department physicians, and consulting specialists.

- At times, SBDs, but not their billing service contacting information, may be captured, which makes contacting them to collect cost and payment data more difficult, time consuming, and costly than it otherwise needs to be.

Accurately associating SBDs with an event and determining which to pursue for additional cost and payment information is challenging. Confirming that most, if not all, of the information needed to clarify the relationships between hospitals and SBDs is contained in the hospitals' EHR and financial systems is key. This preliminary study attempts to confirm this case.

## SECTION 2 Methodology

RTI conducted the project according to the tasks summarized in **Table 2** and described more fully below.

**Table 2. Tasks Performed in Special Project Task 1**

Task	Description
Task 1	Develop a document to be used to facilitate a conversation with hospital staff on whether MPC data are contained in the hospital's EHR and financial systems, and the level of difficulty in extracting and electronically submitting the data
Task 2	Select hospital
Task 3	Identify contacts and individuals at the hospital to invite to the meeting
Task 4	Set up a 60- to 90-minute meeting to discuss whether MPC data extraction and submission is possible
Task 5	Conduct the meeting
	Provide the Final Report

### 2.1 Develop a Discussion Document

Dr. Donald Mon, Senior Director, Digital Health Policy and Standards, and Mr. Robert Bailey, Manager, Public Health Informatics, developed a conversation guide entitled *Medical Expenditure Panel Survey (MEPS) Special Project Task 1: Assessing the Possibility of Hospitals to Submit Data Electronically*. The contents of the document:

- Explained the MEPS MPC process.
- Briefly reviewed issues with the current data collection process.
- Conveyed the intent to:
  - Reduce data collection burden on the hospital.
  - Reduce the hospital staff's time and effort involved in the data collection process.
  - Make the process easier.



- Contained a set of discussion topics and questions designed to elicit responses from hospital staff regarding the goals and objectives of the project (see **Table 1**).

Kathryn Dowd, the MEPS MPC Project Director, provided management oversight during development of the document, reviewed and approved it, and presented it to the AHRQ Contracting Officer's Representative. The discussion document was reviewed and approved by AHRQ prior to the meetings with the hospitals selected for this preliminary study.

## **2.2 Select Hospitals**

Hospital selection was based on several factors listed below. It is important to note that a random sample of hospitals is not warranted for this preliminary study, nor any other part of the process described in **Figure 1.1**. Should AHRQ decide to move forward with electronic data submission, the goal would be to get as many hospitals as possible who are willing and have the capabilities to easily participate in the program, rather than a random sampling of them. The selection factors below: (1) may motivate hospitals to participate in a future electronic data submission program, or (2) demonstrate that hospitals have the structural attributes for successful data extraction and submission.

These same factors are important for this preliminary study. Using the factors below, if it can be demonstrated through discussion that it is possible for the small number of hospitals in this preliminary study to extract and submit MPC data, then it may be possible for a potentially larger number of hospitals fitting the same criteria to do so. The factors are as follows:

- **Hospitals that have historically submitted patient records in a number of consecutive years.** Rationale: From a programmatic standpoint, it would not be worth the time and effort for hospitals to set up, test, and implement MPC data extraction and submission if they are only submitting patient records sporadically over the years. The real gains are for those hospitals that submit multiple records each year over a number of years.
- **Hospitals with an IT, health information management (HIM), finance, or analytics staff that have the knowledge and expertise to set up, test, and implement MPC data extraction and submission.** Rationale: Even if the hospital has the requisite health IT infrastructure (see the bullet above), MPC data can only be extracted and submitted if the IT staff has the knowledge and skills to do so, or if the HIM, finance, or analytics staff can do so through their end-user computing or analytics capabilities.
- **Hospitals with a medium-to-high data collection burden—that is, those with a moderate (10 to 20) to a high number (50 or more) of patient-provider pairs and/or a moderate to a high number of contacts.** Rationale: A major goal is to reduce the hospital's data collection burden. Thus, it is important to identify those with medium-to-high burden and target them for a data extraction and submission program, should it be pursued.
- **Hospitals with whom the MEPS MPC has a good relationship.** Rationale: Given the time to conduct this preliminary study and the need to have a quality discussion, we chose hospitals with individuals we knew we could contact to set up the meeting in a

timely fashion and have a good discussion on the level of difficulty the hospital may have in extracting and submitting MPC data.

Based on the above factors, two hospitals were selected. The first—hereafter, Hospital 1—is a large (600+ beds), urban, academic medical center running Epic as its EHR system. Hospital 1 has an analytics group that is outside of, but works very closely with, the institution’s IT department. This analytics group uses Epic’s data warehouse, which is integrated with the EHR. Hospital 1 has submitted records in many consecutive years in the past. In 2014, Hospital 1 submitted records for 70 patient-provider pairs (high record volume), requiring follow-up contact with 3 hospital staff members in the HIM and finance departments in order to complete chart abstraction.

The second—hereafter, Hospital 2—is a medium sized (<400 beds) community hospital. Initially, it was thought that Hospital 2 was running an EHR system typical of medium-sized hospitals (e.g., Meditech). However, after successfully engaging with Hospital 2, we learned that, using the systems of the larger health system that acquired it, Hospital 2, too, was running Epic as its EHR system. Similarly, it has an analytics group that is outside of, but works very closely with, its IT department. This analytics group also uses Epic’s data warehouse. In 2014, Hospital 2 submitted records for 14 patient-provider pairs, requiring follow-up contact with 8 hospital staff members in the HIM and finance departments, as well as the Medical Staff Office, in order to complete chart abstraction. Even though the record volume is relatively low, this situation represents a fairly high burden for the hospital, since it required follow-up contacts with 8 staff members to complete just 14 records.

### **2.3 Identify Contacts and the Individuals to Invite to the Meeting**

By going through past contact records, project staff identified the hospital staff with whom they interacted during the abstraction of various records. The contact information for these individuals was passed on to Dr. Mon and Mr. Bailey securely over RTI’s Enterprise Secure Network. That information revealed that, as mentioned above, members from the HIM and finance departments were contacted for Hospital 1, while members from those two departments, plus the Medical Staff Office, were contacted for Hospital 2. In addition to these departments, staff persons from the hospital’s IT and analytics departments were identified using Dr. Mon’s network of IT professionals at those hospitals.

Dr. Mon spoke with the contacts from each department individually by phone. In the call, he thanked them for their past efforts, oriented them to the goals, objectives, and intent of this preliminary study, and requested a meeting with representatives from all departments in attendance. After hearing the goals and objectives of the meeting, the IT staff members and the analytics group at both hospitals agreed that the best entity to address our discussion topics and questions was the analytics group, rather than the IT department.

### **2.4 Set Up a 60- to 90-Minute Meeting to Discuss Whether MPC Data Extraction and Submission Are Possible**

A 90-minute in-person meeting was set up with each hospital during the week of June 20, 2016. Due to certain circumstances (e.g., vacation, retirement, scheduling difficulties, unable to leave

core responsibilities to attend the meeting), only representatives from the analytics group were able to attend the meeting.

Per Dr. Mon's and Mr. Bailey's request, the preapproved conversation guide, *Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC) Special Project Task 1: Assessing the Possibility of Hospitals to Submit Data Electronically*, was sent as an email attachment to the analytics staff members prior to the meeting. The advanced receipt of this conversation guide enabled them to identify the appropriate analytics staff members to attend the meeting and prepare for the discussion.

## 2.5 Conduct the Meeting

Following the conversation guide, Dr. Mon facilitated the discussion, while Mr. Bailey took notes on the points made by the analytics staff. At the beginning of, and at various points throughout the conversation, Dr. Mon and Mr. Bailey were careful to emphasize that:

- This was a preliminary study to determine whether MPC data extraction and submission were possible, focusing on ways that would be most convenient and natural to the analytics staff.
- No decisions about implementing a data extraction and submission program have been made.
- This study can help inform that decision.
- The ultimate objective is to collect MPC data with as little burden to the hospital as possible, and to make the process easier.

## SECTION 3 Results

This section contains the individual findings from Hospitals 1 and 2 (Sections 3.1 and 3.2, respectively), and a set of recommendations stemming from the findings (Section 3.3). Summary conclusions based on the findings and recommendations can be found in Section 4 – Conclusions.

The findings from the individual hospitals below follow the sequence of the discussion topics and questions in the *Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC) Special Project Task 1: Assessing the Possibility of Hospitals to Submit Data Electronically* conversation guide. Since conversation with each hospital was natural and free flowing, some questions were skipped so that the critical questions and topics could be discussed. Thus, for brevity, the major points provided by the analytics staff are presented below in summary and contextual form, rather than in a question/answer format.

### 3.1 Findings from Hospital 1

#### 3.1.1 General Topics of Discussion

The analytics staff at Hospital 1 confirmed that all of the transactional MPC data elements (i.e., patient and physician demographics, individual charges and payments) listed in the *Medical*

*Expenditure Panel Survey Medical Provider Component (MEPS MPC) Special Project Task 1: Assessing the Possibility of Hospitals to Submit Data Electronically* conversation guide (see **Table 1**) were captured in Epic’s patient demographic, revenue cycle, or physician directory modules. Some MPC data elements not used for patient care transactional purposes (e.g., Fully Established Charge, Charge Equivalents) were not in the hospital’s EHR. However, Hospital 1 has access to the Charge Master, which likely contains these data. Alternatively, these data elements may be found in the billing, patient accounting, or cost accounting systems operating within the finance department.

Many MPC data elements are already being extracted and reported for other purposes. Therefore, it is conceivable that Hospital 1 can leverage one of these extraction or reporting mechanisms to submit its MPC data. Of these, Hospital 1 identified the one mechanism it uses to submit similar data to a disease registry as the most natural and convenient mechanism. The analytics staff suggested, however, that, rather than reusing this mechanism and submitting MPC data directly to RTI, we request the registry to send Hospital 1’s data to us. This process, in fact, is already in use. For example, the registry has agreed to send Hospital 1’s data to the National Center for Health Statistics/Centers for Disease Control and Prevention (NCHS/CDC) to satisfy its Meaningful Use reporting requirements. Thus, by getting Hospital 1’s data through the registry, Hospital 1 does not have to expend any additional effort.

Recall, however, that the data received from the registry would only be a subset of MPC data—it would not contain both transactional and non-transactional payment data. A separate but perhaps less time-consuming process would need to be implemented to get the remaining data elements from Hospital 1. If it turned out that the time and effort in sending MPC data to the disease registry and submitting the remaining data to RTI was roughly equivalent to electronically submitting the entire MPC data set to RTI, Hospital 1 agreed that the latter was the better alternative and would be willing to build a data extraction and submission mechanism, pending business-owner approval and funding. The implications of receiving MPC data from the registry are discussed more fully in Section 3.3 – Summary of Findings and Recommendations.

Hospital 1 emphasized that, no matter which mechanism is used to submit MPC data to RTI, a very clear technical specification would be needed. The technical specification must contain definitions for each data element, as well as algorithms to compute certain data elements (e.g., total charges, total payments). In addition, Hospital 1 clearly stated that, while it had the expertise to do it, it would be easier not to have to do data transformations. However, Hospital 1 understood the need to transform the data and would be willing to provide documentation on its data elements so that project staff could accurately transform their data in back-end data processing.

Finally, Hospital 1 can easily use an Excel file, containing the list of patients whose records are requested, to begin the data extraction process.

### **3.1.2 Patient Identification and Matching**

The analytics staff at Hospital 1 stated that the patient name (last name, first name, middle initial), address (street address, city, state, zip code), race, gender, and date of birth fields, provided by project staff, should be sufficient to match patients in its data warehouse.

Not surprisingly, Hospital 1 noted that patient matching is the key to successful data extraction. To that end, the analytics staff stated that providing the data with unique identifiers, such as medical record numbers or social security numbers, would increase patient matching success rates. The implications of providing these identifiers are further discussed in Section 3.3 – Summary of Findings and Recommendations.

Wary of privacy violations, they emphasized that guidelines and validation procedures should be provided by RTI to prevent inadvertent submission of data on wrong patients due to imprecision in patient-matching. Currently, misspellings, synonyms (e.g., “Robert” versus “Bob” in the first name field), and differences in field values (e.g., date of birth) cause a rejection. Reviewing these rejections is a manual task. Having best practice information, such as the rate of incorrect spellings, would help build tolerance levels into the SQL-based patient-matching engine created for this initiative. The implications of providing these guidelines and validation procedures are further discussed in Section 3.3 – Summary of Findings and Recommendations.

In addition, the analytics staff inquired whether patients with HIV or behavioral health issues were automatically excluded from data submission. They agreed to submit such records when provided evidence of patient authorization and consent.

### **3.1.3 Patient Demographics**

The analytics staff at Hospital 1 confirmed that patient demographic data are stored in separate fields, consistent with Meaningful Use Stage 1 requirements. They also confirmed that fields, such as race and gender, are coded. They agreed to provide documentation on the codes they use in Epic so project staff can accurately transform the data on the back end.

### **3.1.4 Patient Encounters**

The analytics staff at Hospital 1 confirmed that they can track the Emergency Department as an admission source. They can separate the Emergency Department encounter from the inpatient encounter, but still link the two records to form one episode.

### **3.1.5 Diagnosis Data**

Hospital 1’s Epic EHR system is capable of storing hundreds of diagnosis codes, though the number of diagnosis codes it sends to the aforementioned disease registry is truncated at 99. Even with this truncation, Hospital 1 is able, and would prefer, to submit diagnosis codes in a one-to-many (patient to diagnosis codes) relational format, rather than a flat-file format. The implications of receiving diagnosis codes in a one-to-many file format are further discussed in Section 3.3 – Summary of Findings and Recommendations.

For 2015 inpatient records, diagnosis codes are captured in both ICD-9 and -10. This is consistent with the mandated October 15, 2015, implementation of ICD-10.

### **3.1.6 Procedure Data**

The analytics staff at Hospital 1 confirmed that they capture both ICD and CPT procedure codes, and can provide both the code set name and the code value in separate columns. As with

diagnosis codes, Hospital 1 is able, and would prefer, to submit procedure codes in a one-to-many (patient to diagnosis codes) relational format, rather than a flat-file format.

### **3.1.7 Data on Doctors Who Performed a Service or Procedure on the Patient AND Who Are Not Included in Service/Professional Fees**

Eighty to 90 percent of the doctors providing care to patients are hospital-employed physicians. Thus, since these physicians use Epic, and with the interoperability between Epic and hospitals' financial systems, they can: (1) identify each physician who performed a service or procedure on the patient, (2) capture each physician's professional fees, and (3) separate the professional fees from the hospital bill.

Of the 10 percent to 20 percent who are not hospital-employed physicians, Hospital 1 has relatively easy means of identifying SBDs. For example, by contractual agreement, physicians from a specific specialty practice have admitting and practice privileges. Since they are not hospital-employed physicians, every physician in this practice can be categorized as SBDs. From a systems perspective, tracing them as SBDs is fairly straightforward. When those doctors perform a service or procedure on the patient, they document their care using Epic. Since Epic's physician directory contains group practice and billing service demographic data, the service or procedure can be traced to the physician, who can be traced to the specialty practice and billing service.

A second scenario also demonstrates the relatively straightforward manner in which SBDs can be detected. All physicians providing care to their patients document their clinical treatment in Epic, but not all physicians use the hospital's billing system. Thus, in Epic, for those physicians who do not use the hospital's billing system, there would be a record for the physician in Epic's clinical care module, but not in its revenue cycle module. As the analytics staff stated, "It's clear cut. We either bill or we don't bill. There's no duplication." Consequently, a record in the clinical care module, but not one in the revenue cycle module, indicates an SBD. As above, the physician's practice can be traced using Epic's physician directory. The billing service for that physician can be traced either in Epic's physician directory if it is captured, or by contacting the now-known-physician's practice.

### **3.1.8 Charge and Fees Data**

As previously mentioned, Hospital 1 has transaction-level charge, billing, and payment data within Epic's revenue cycle module, which is fed by the upstream financial systems. Global fees can be identified along with any associated adjustments (e.g., where an adjustment for a follow-up visit after surgery is shown as a deduction against the global fee). Fully established charges and charge equivalents were unfamiliar terms to the analytics staff. That may not be unusual, however, since these concepts are not typically included in the analytics they perform. These data elements are likely contained in their upstream charge master and financial systems. Since finance staff were unable to attend the meeting, there was no opportunity to confirm this point. However, the analytics staff stated they can confirm this point with the finance staff at a later date.

The predominant charge and payment model at Hospital 1 is fee for service. Its Epic EHR system can more easily identify charges and fees that are based on fee for service. It does not identify charges and fees based on capitation. However, its upstream financial systems may be

able to track charges and fees based on capitation. Since finance staff were not able to attend the meeting, there was no opportunity to confirm this point.

### **3.1.9 Payment Data**

As previously mentioned, Hospital 1 has payment transactions in its data warehouse. However, staff do not do much reporting on payment data. Thus, while they may have the data, they do not have existing data extraction or reporting mechanisms to reuse for MEPS MPC purposes. They would have to write a query to calculate total payment to supply that data element. Total payment is typically found in financial systems (as are total charges, total adjustments, etc.). Thus, the financial systems are perhaps better sources to retrieve such data. However, since finance staff were not able to attend the meeting, there was no opportunity to confirm this point.

Lump-sum payments can be identified for Medicare as the payer. However, the analytics staff did not believe that lump-sum payments can be identified from other payment sources.

The analytics staff acknowledged the importance of payment data analysis. They stated that while they do not do much reporting on payment data today (which is expected, given that they largely support clinical studies), this can change by the time an MPC data submission program is implemented, should an implementation program be pursued.

### **3.1.10 Data Submission**

Hospital 1 generally prefers to submit files via secure file transfer protocol (SFTP). In addition, it is able to submit multiple files in a relational schema, rather than as a large flat file. Hospital 1 staff have done so for other reporting agencies. Given the option of submitting data three times a year versus once for all sample increments, Hospital 1 opted for the latter.

## **3.2 Findings from Hospital 2**

### **3.2.1 General Topics of Discussion**

In general, the analytics staff person found the MPC data extraction and submission process as described in the *Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC) Special Project Task 1: Assessing the Possibility of Hospitals to Submit Data Electronically* conversation guide to be a rather typical request. Having reviewed the conversation guide prior to the meeting, the analyst provided his reaction in the following email:

Looks good, I don't see anything that should be a problem per the specifications you provided.

The only question I did have was around how you'd like to handle capturing account class changes (for example ED -> IP) as you mentioned. Would you need these 2 'encounters' to be merged as 1?

Other than that, I'm not seeing anything that would present an issue.

As with Hospital 1, all of the MPC data elements listed in the *Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC) Special Project Task 1: Assessing the Possibility of Hospitals to Submit Data Electronically* conversation guide (see **Table 1**) can be found in Hospital 2's Epic EHR and financial systems.

Similarly, the analyst acknowledged that Hospital 2 has existing data extractions and reporting mechanisms that contain most of the MPC data elements. However, Hospital 2 differs from Hospital 1 regarding whether and how to leverage these mechanisms. The analyst at Hospital 2 stated that, since existing data extractions and reporting mechanisms are used for different purposes, there are variations in data definitions and programming logic. The analyst stated that it would be best to develop a new data extraction and submission program based on specifications provided to him. He emphasized on numerous occasions that clear and concise technical specifications were essential to successful data extraction and submission. Since this is a typical request, he estimated that it would take approximately 2 weeks to a month to develop, test, and implement this computer program. He could only work on data extraction and submission with business-owner approval and funding.

As with Hospital 1, the analyst stated that it would be easier for data transformations to occur at the back-end MEPS MPC database. In addition, Hospital 2 can handle either an Excel or a comma separated values (CSV) file containing the patients for whom we are requesting records to begin data extraction.

### **3.2.2 Patient Identification and Matching**

As with Hospital 1, the analyst at Hospital 2 stated that the hardest part of the entire process will be patient matching. A patient-matching engine will need to be developed. However, he has built patient-matching engines before and believes building one for MPC data will not be difficult.

The most challenging part in developing a patient-matching engine will be in deciding how to address false positives or partial matches. The analyst stated that the MPC would need to establish the confidence threshold to prevent sending Protected Health Information (PHI) the hospital has not been authorized by the patient to share. By his definition, a confidence threshold is the level at which a match is deemed successful. For example, is matching on 50 percent of the patient-matching data elements (name, address, race, gender, and date of birth) a successful match? Or, should the threshold be set at 80 percent, 90 percent, or 95 percent? The implications of the MPC providing such thresholds are further discussed in Section 3.3 – Summary of Findings and Recommendations.

As with Hospital 1, the analyst at Hospital 2 stated that providing him with the patient's medical record number and/or social security number could increase patient-matching success rates. The implications of providing these identifiers are further discussed in Section 3.3 – Summary of Findings and Recommendations.

### **3.2.3 Patient Demographics**

As with Hospital 1, patient demographic data elements are collected as separate text fields. Race and gender are captured as coded fields. In addition, Hospital 2 can provide its technical documentation so that race, gender, and other typically coded fields can be accurately transformed by project staff at the back-end database.



### **3.2.4 Patient Encounters**

The analyst at Hospital 2 confirmed that staff can track the Emergency Department as an admission source. They can separate the Emergency Department encounter from the inpatient encounter, but still link the two records to form one episode. However, the linking can be represented in one of two ways. The patient can have one hospital account with two encounter IDs (one each for the emergency and inpatient encounters), or one hospital account and one encounter ID where the Emergency Department charges are reversed and added to the inpatient encounter. The implications of these linking options are further discussed in Section 3.3 – Summary of Findings and Recommendations.

Aside from the above linking logic, teasing out a patient’s events in Hospital 2’s medical records would not present much difficulty since Epic treats inpatient, outpatient, and Emergency Department events as different classes of data.

### **3.2.5 Diagnosis Data**

As with Hospital 1, Hospital 2’s Epic EHR system is capable of storing hundreds of diagnosis codes. Similarly, Hospital 2 is able, and would prefer, to submit diagnosis codes in a one-to-many (patient to diagnosis codes) relational format, rather than a flat-file format. The implications of receiving diagnosis codes in a one-to-many file format are further discussed in Section 3.3 – Summary of Findings and Recommendations.

For 2015 inpatient records, diagnosis codes are captured in both ICD-9 and -10. This is consistent with the mandated October 15, 2015, implementation of ICD-10.

### **3.2.6 Procedure Data**

The analyst at Hospital 2 confirmed that staff capture both ICD and CPT procedure codes, and can provide both the code set name and the code value in separate columns. As with diagnosis codes, Hospital 2 is able, and would prefer, to submit procedure codes in a one-to-many (patient to diagnosis codes) relational format, rather than a flat-file format.

Procedures can be tied to a global fee. Procedures tied to a global fee are displayed as an adjustment (deduction) against it. In addition, a code displaying “denied unbundled payment,” while not displaying a dollar amount, indicates that a global fee exists.

### **3.2.7 Data on Doctors Who Performed a Service or Procedure on the Patient AND Who Are Not Included in Service/Professional Fees**

Due to time constraints, this topic was not discussed in sequential order. However, in the course of the free-flowing discussion, the analyst at Hospital 2 confirmed that the Epic installation is able to provide physician demographic data through its physician directory and track which physicians performed a service or procedure on the patient in much the same way as Hospital 1.

### **3.2.8 Charge and Fees Data**

The analyst at Hospital 2 stated that charge and adjustment data are generated in Hospital 2’s upstream cost accounting system and tracked in Epic. As previously mentioned, he confirmed that they can capture global fees, as well as charges for individual procedures.

Fully established charges and charge equivalents were unfamiliar terms to the analyst. That may not be unusual, however, since these concepts are not typically included in the analytics his group performs. These data elements are likely contained in the upstream charge master and financial systems. Since finance staff were unable to attend the meeting, there was no opportunity to confirm this point.

As with Hospital 1, the predominant model at Hospital 2 is fee for service. Its Epic EHR system can more easily identify charges and fees based on fee for service. It does not well-identify charges and fees based on capitation. However, its upstream financial systems may be better able to track charges and fees based on capitation. Since finance staff were not able to attend the meeting, there was no opportunity to confirm this point.

### **3.2.9 Payment Data**

Due to time constraints, this topic area was not discussed at length. While discussing the background, the analyst described the way the staff handle global or bundled payments. Later encounters included in the global fee show an adjustment indicating a bundled payment write-off.

### **3.2.10 Data Submission**

As with Hospital 1, Hospital 2 generally prefers to submit files via SFTP. In addition, it is able to submit multiple files in a relational schema, rather than as a large flat file. Hospital 2 staff have done so for other reporting agencies.

Unlike Hospital 1, given the option of submitting data three times a year versus once all sample increments are available, Hospital 2 opted for the former. Submitting incremental data on a patient, such as payments that may have occurred after the previous data submission, can be easily done. The encounter ID would identify the patient's record as the same encounter over multiple data submissions, with the data in subsequent submissions indicating an incremental update to the record.

## **3.3 Summary of Findings and Recommendations**

This section integrates the findings from the two hospitals participating in this preliminary study and provides some detailed level recommendations based on those findings. (Higher-level recommendations are discussed in Section 4 – Conclusions.)

**Summary Finding #1: Hospitals have the ability to capture MPC data. (Goal 1 Met.)** Both hospitals have the ability to capture MPC financial data as discrete and computable data, and condition procedure/treatment and demographic data as separate text fields. Thus, Goal 1 is confirmed. Since MPC data can be captured in the EHR and financial systems of these two hospitals, it can be assumed that other hospitals running Epic or other leading EHR systems will have the ability to do so as well. Given the current level of EHR adoption, it is likely that a significant portion of hospitals participating in the MPC program will have an EHR system capable of collecting MPC data. Though it is unreasonable to expect 100 percent of hospitals participating in the MPC program will have the ability to do so, even 25 percent to 40 percent of the hospitals may result in substantial benefits, addressing the issues discussed in Section 1.3.1.

**Summary Finding #2: Though there are existing data extraction and reporting mechanisms, those mechanisms did not appear to result in their easy reuse for MEPS MPC purposes.** Both hospitals had existing data extraction and reporting mechanisms that appeared to extract and submit a high proportion of MPC data—only payment data was in question. However, neither of the two hospitals suggested that they would use those mechanisms to directly submit MPC data.

Hospital 1 suggested that the project gain access to its data from the disease registry in which it has submitted all but MPC payment data. This approach has some interesting implications for MPC data collection. First, while it may reduce hospital burden to extract and submit data for all but payment MPC data, a supplemental process would have to be implemented in which Hospital 1 would extract and submit payment and non-transactional data (i.e., fully established charges, equivalent charges), and subsequently be merged on the back end with the data submitted by the hospital registry. Should this Preliminary Study proceed to the Feasibility stage (**Figure 1-1**), determining whether this approach is more feasible than requesting the entire set of MPC data from Hospital 1 would have to be assessed.

This assessment may be worthwhile because there will likely be other hospitals participating in this registry for which their data can be accessed as well. Thus, by going to the registry, it may be possible to obtain MPC data for multiple hospitals from a single source, offloading the data collection burden from all those hospitals, but albeit on to the registry.

- **Recommendation #1: Identify single sources where they exist and, should Feasibility Assessment proceed (see Figure 1.1), determine the feasibility of using these single sources.**

Both hospitals agreed to develop data extraction and submission computer programs to support the MPC initiative. Hospital 1 agreed to do so if it turned out that the time and effort in sending MPC data to the registry and submitting the remaining data to RTI was roughly equivalent to electronically submitting the entire MPC data set. Hospital 2, however, immediately took the position that developing data extraction and submission computer programs to support the MPC initiative was a more natural and convenient approach for them. In both instances, clear and concise technical specifications provided by the project was essential, and business-owner approval and funding was needed.

- **Recommendation #2: Should the data extraction and submission program proceed, clear and concise technical specifications should be developed in the Implementation Plans stage (see Figure 1.1).** The technical specification must at least contain unambiguous definitions of the data, algorithms to compute appropriate data elements (e.g., total payments), and guidance and validation procedures for patient matching. Since the hospitals preferred that data transformations be handled at the back-end database, the technical specification would not include data cleansing/transformation rules aside from the validation procedures needed for patient matching.
- **Recommendation #3: Should the data extraction and submission program proceed, the project should consider developed of materials in the Implementation Plans stage (see Figure 1.1) to help the IT department or analytics group secure business owner approval and funding.**

**Summary Finding #3: Patient matching is the most challenging aspect of the data extraction and submission process.** Both hospitals stated that patient matching will present the biggest challenge in the entire process. Both stated that providing them with unique identifiers, such as medical record numbers or social security numbers, would result in a higher match rate.

Obtaining medical record numbers from the patients during Household Component data collection may be challenging, but possible.

- **Recommendation #4: Should Feasibility Assessment proceed (see Figure 1.1), consider obtaining medical record numbers from patients as a feasibility factor.**

Providing the hospitals with the patient's social security number is a tenuous alternative at best. Patients may be reluctant to provide their social security numbers, and the resulting unease may reduce Household Component data-collection completion rate. In addition, doing so would require AHRQ to obtain approval from the Office of Management and Budget, which may result in additional oversight for minimal gains as described in the previous sentence.

**Summary Finding #4: Different approaches in linking events may inadvertently impact charge, billing, and payment analyses.** Hospital 2 provided two approaches to linking the emergency encounter with an inpatient encounter, when the patient was admitted through the Emergency Department. In one of those approaches, one hospital account and one encounter ID are created, wherein the Emergency Department charges are reversed and added to the inpatient encounter. In this approach, the charges, bills, and payments for the emergency event may be artificially low, while that for the inpatient event may be artificially high. The technical specification given to the hospitals must provide clear guidance on the allocation of charges, billings, and payments so that such impacts are avoided.

- **Recommendation #5: Should the data extraction and submission program proceed, clear guidance on the allocation of charges, billings, and payments materials should be developed during in the Implementation Plans stage (see Figure 1.1) and included in the technical specification provided to hospitals.**

**Summary Finding #5: It is easiest to submit diagnosis and procedure codes in a one-to-many relational schema.** Both hospitals stated that they preferred to submit diagnosis and procedure codes in a one-to-many relational schema, rather than in a flat file. These codes are stored in a relational schema in their EHR systems, and thus, is a natural and convenient way for them to extract and submit the data. The MEPS MPC database currently stores diagnosis and procedure codes using a name-value pair technique. However, the routines to load data from the relational schema submitted by the hospitals to the name-value pairs in the MEPS MPC database is trivial. However, some hospitals or registries may be more familiar with flat-file data submission. Thus, the technical specification provided to the hospital must contain clear instructions on the layouts of each file in the relational schema (see Recommendation #2).

**Summary Finding #6: Terms such as “fully established charge” and “full equivalent fee” may be unfamiliar to the analytics staff tasked with the MPC data extraction and submission process.** The analytics staff at both hospitals were unfamiliar with the terms “fully established charges” and “full charge equivalents.” These and other unfamiliar terms will need to

be clearly defined in the technical specification (see Recommendation #2) to avoid confusion and potential inaccurate submission of such data. This is also a clear indication that the combined efforts of the finance staff, who may better understand these terms, and the IT or analytics staff is essential.

**Summary Finding #7: Separately billing doctors can be identified through EHR and financial systems. (Goal 2 Met.)** Through their EHR and financial systems, both hospitals were able to: (1) identify each physician who performed a service or procedure on the patient, (2) capture each physician's professional fees, and (3) separate the professional fees from the hospital bill. It is highly likely that other hospitals using Epic or other leading EHR systems will be able to do so, as well. Thus, one of the major issues in the current MPC data collection process (see Section 1.3.2) can be addressed through an electronic data submission initiative.

**Summary Finding #8: Payment data is the one segment of the MPC data set in which there is less familiarity among analytics staff in this preliminary study, which may cause difficulties in performing data extraction and submission.** The analytics staff at both hospitals were less familiar with payment data since they have not needed to use such data for clinical research, analyses, or reporting. With small sample size in this preliminary study, however, it is difficult to extrapolate this finding to population of hospitals participating in the MPC program. In addition, since finance staff were not able to attend the meeting at either hospital, it could not be confirmed that the finance staff had sufficient experience with payment data that meets the MPC program's needs.

- **Recommendation #6: Should Feasibility Assessment proceed (see Figure 1.1), additional investigation on the hospital's ability to extract and submit payment data according to the MPC program's needs should be conducted.**

**Summary Finding #9: Each option to submit MPC data by time period was selected, which may require the MEPS program to offer both options.** When presented with the option to submit MPC data three times a year versus once all sample increments were available, Hospital 2 opted for the former, while Hospital 1 opted for the latter. It is highly likely that this pattern will continue when the options are offered to a wider set of hospitals participating in the MPC program.

- **Recommendation #7: Unless there is an analytical reason why one option or the other should be prohibited, both options should be offered to hospitals and the MEPS MPC staff should be prepared to support these options.**

**Summary Finding #10: A data extraction and submission program appears to be a viable solution in making MPC data collection easier. (Goal 3 Met.)** Given the feedback from the analytical staff, it appears that a data extraction and submission program can be a viable solution in making MPC data collection easier. Neither hospital identified insurmountable obstacles in the program. Both hospitals stated there would be some additional, but not difficult, effort to develop the data extraction and submission computer programs and process to support their business owners. Conceptually, the benefits from electronic data submission (e.g., reduced time and effort, increased data quality) should therefore accrue. However, such benefits should be quantified at some future time.

- **Recommendation #8: The anticipated benefits of data extraction and submission should be quantified using the suggested metrics in the Feasibility Assessment stage (see Figure 1.1), should it proceed.**

## **SECTION 4**

### **Conclusions**

This preliminary study was to assess whether data extraction and submission represented a viable solution for making MPC data collection easier on participating hospitals. Given the results of this study, it has been ascertained that: (1) MPC data can be collected in EHR and financial systems, (2) separately billing doctors can be identified and there are concrete ways to discern their professional fees from the hospital bill, and (3) the data extraction and submission process potentially would not be difficult to implement for those hospitals operating EHR and financial systems similar to the two hospitals in this study. All three goals of this preliminary study were met.

The topic of how the current manual chart abstraction process could be made easier (see the second parts of Goals 1 and 3, **Table 1**) was not discussed for two reasons. The audience with whom we discussed data extraction and submission consisted of the analytics staff at both hospitals. Thus, these individuals were more focused on the technical aspects of data extraction and submission, rather than the current manual chart abstraction process. In addition, since HIM and finance staff were not able to attend the meeting, there was no opportunity to compare whether the tenets of a data extraction and submission approach represented a solution that could be truly easier than the current chart abstraction process.

Based on the encouraging information gleaned from this preliminary study, the next stage—Feasibility Assessment—can be strongly considered.

**— Project #3: Unified Approach Summary —**

Date: May 5, 2017  
 To: Karen Davis  
 From: Katy Dowd, Erica Saleska, and Rob McCracken  
**Subject: Unified Approach Summary**

For the 2015 cycle, we reviewed 13 Large Provider Hospital bands as potential candidates for obtaining complete Separately Billing Doctor (SBD) information from Hospital Patient Account (PA) records. For each of these 13 bands, initial discussions with our Points of Contact (POCs) included review of what type of SBD information was included in the PA records.

Unfortunately, we were able to collect complete SBD information for only 1 of these 13 bands. For HD246, we collected complete SBD data for all 18 pairs completed for this band. For another band, HD502, we collected complete SBD information for 10 of 22 pairs (46 percent) using PA alone.

For the remaining bands, after review of PA and/or discussion with the POC, we requested Medical Records (MR) as complete SBD information was not included in the PA records. Of the 442 completed pairs represented across these 13 bands, 93.7 percent required MR (n=414) and 6.3 percent included complete SBD information in the PA records (n=28).

Overall, we needed to obtain MR for the majority of the 442 pairs included in these 13 bands. In the 2016 cycle, we are continuing to evaluate these and other Large Provider bands for PA only candidacy.

The attached file includes a band by band detailed summary of these findings.

# Bands	# Yes – PA Included Complete SBD Data	%	# No – PA Did Not Include Complete SBD Data	%	# Mixed (Some with SBD, Some Without)	%
13	1	7.69%	11	84.62%	1	7.69%

# Completed Pairs (13 Bands)	# Yes – PA Included Complete SBD Data	%	# No – PA Did Not Include Complete SBD Data	%
442	28	6.33%	414	93.67%

## SUMMARY OF FINDINGS

Generally, these 13 bands were identified as possible Unified Approach candidates because we had received both facility and professional billing during the 2014 cycle, and/or had reason to believe we could obtain both during the 2015 cycle. For many of these bands, we had to determine during the 2015 cycle whether those facility and professional fees provided by the POCs accounted for all charges a patient could incur from a Hospital event.

The main reason we found that the Unified Approach could not be used with a particular band was that the billing records provided by the POC could not account for all possible charges associated with Hospital events. While a POC or POCs for a band may have been able to provide records for all facility and professional billing they handled, there were SBD charges from outside their health system that they did not handle. The only way to identify those SBDs was through review of the MR. For a few of the 13 bands, initial conversation with a POC led us to believe that the POC could provide all of the necessary information, only to find through review of records received or from later conversations with the same POCs that some SBDs from outside the health system were used.

We also found with one band, that billing within the band varied by facility. While some facilities used no outside SBDs (and a Unified Approach could be used for pairs from those facilities), other facilities did use outside SBDs. For events at facilities using outside SBDs, MR was required in order to identify when those SBDs were used. Note that when a patient is in multiple pairs within the band, this approach works only if the billing allows identification of the facility where each event occurred.

For the 2016 cycle, we've found six bands that are possible candidates for the Unified Approach. Two of the bands have been identified early in the 2016 cycle as possible candidates. The other four are university providers that, based on online information, indicate they could provide complete billing information with no outside SBDs and no MR information necessary. These six bands account for 41 hospital pairs in the 2016 Wave 1 sample. The next step is to review Wave 1 records to identify whether these bands appear to be legitimate candidates for the Unified Approach. If they appear to be, a recommendation will be made to AHRQ to proceed with the Unified Approach (on a band-by-band basis). If approved, we can start the Unified Approach with Wave 3 of the 2016 cycle. There are more university groups that can be reviewed; however, there are no additional category 1 or 2 providers at this time that we think are candidates for the Unified Approach.

We will continue to attempt to find efficiencies during data collection. For example, when a POC handles records for both a MEPS MPC Hospital and SBD group, we will request all facility and professional billing from the POC during Hospital data collection, even though we may still need to collect MR to identify SBDs for whom the POC does not handle records. While this is not the Unified Approach, it does consolidate requests from that particular POC to one portion of the year.

We will continue to add to our data collection toolbox. Part of this is recognizing the situation where a Unified Approach is the right way to move forward. We will also continue to create and adapt other approaches to work SBD data collection more efficiently.



**— Project #4: SBD Special Project – Interviews on SBDs – Summary Report —**

Date: June 23, 2017  
To: Karen Davis, AHRQ  
From: Diana Greene and Christine Carr  
**Subject: SBD Special Project – Interviews on SBDs - Summary Report**

## BACKGROUND

The objective for this special project was to have in-depth discussions with executives at national professional medical organizations, or practicing doctors associated with these organizations, to better understand the process of Separately Billing Doctor (SBD) billing. We completed in-person, phone, and email discussions with medical executives and professionals to better understand SBD billing as it applies to the MEPS MPC. Each discussion guide was tailored to apply directly to the type of specialty of the professional organization/practicing doctor.

We attempted contact with 21 organizations and/or individuals. Seven interviews were completed (2 in-person, 4 phone, 1 email). The interviewees ranged in expertise specialty from pathology, radiology, surgery, diagnostic services, to general medical. In-person and phone meetings were between 30—60 minutes, addressing topics associated with SBDs. No information about individual MEPS respondents or organizations was shared with the interviewees.

## OVERALL THEMES

### Data from Payors

Multiple interviewees suggested that we contact payors. This was discussed amongst the RTI MEPS team. Previous internal discussions had suggested that insurance companies are not likely to provide payment data because they work out special deals with various healthcare systems. Insurance companies do not want the minimum/maximum payments known as that is confidential business information. To the argument that the payor data are anonymized, the large insurance companies know each other so well that they could name their competitors with the slightest characteristic type data. A small number of existing MPC Hospital bands have refused as they say their contracts with payors forbid it for these reasons. There is potential in terms of Medicare and Medicaid payors, given the availability of Medicare and Medicaid claims data. Though payors may not agree to provide data on charges and payments, they may be willing to provide the names of providers that billed them.

### Motivating Hospitals

Another common theme was that hospitals need to be motivated to share all possible information. Interviewees agreed that it is not in the hospitals' core objectives or business model to provide these data. For SBDs, they need to see a direct benefit from the research before agreeing to participate. Members generally participate in endeavors they believe will assist them in managing and operating their practices optimally. If possible, we should communicate how their participation would result in a positive experience and outcome for them. Perhaps we could share MEPS Statistical Briefs with providers to show recent data on healthcare that might be of interest and use to them.

## Professional Fee and Technical Components Identification

Professional fees versus technical components was a topic of significant discussion. For there to be a separate bill, there must be a professional component (PC). The physician fee schedule would note if there is a professional component to be billed. If a doctor provides services in a hospital there must be an indication of this somewhere in the medical records (a report, an order, etc). Someone must bill the PC; otherwise it looks like a kickback. There is a separate fee schedule for labs; this would also indicate if there is a separate bill.

Additional comments on diagnosis, procedure and billing codes:

- For CPT-4s, the Modifier 26 is used when only the professional component is being billed. Modifier TC is used when only the technical component is being billed.
- Use revenue codes to make assumptions on professional components.
- If surgery is done in an inpatient facility, DRGs are used. There are probably 10 or 15 DRGs that cover most orthopedic services.
- If outpatient or ambulatory surgical center, they have APCs (Ambulatory Payment Classifications) set up for different outpatient and ambulatory care services to bill for their facility.
- Example of DRG versus APC: A total knee replacement in an inpatient setting – DRG would be used. If done in outpatient setting, APC would be used.
- Many CPT-4 and ICD-10 codes are associated with orthopedic services. CPT-4 codes in the twenty thousand series is the Musculoskeletal System section. Most are orthopedic. There are a couple of other sections of CPT-4s that have surgical service – the sixty thousand series includes the neuro related services. The radiology section is used a lot by orthopedic surgeons because they are all there at the same time for certain procedures. When looking at ICD-10, a pathologist estimated that there are approximately 18,000 possible orthopedic codes.
- The universe is much smaller for DRGs and APCs than for ICD-10 codes. We could track when hospital bills a DRG; there are relatively few of those. ICD-10 can be tracked to types of surgeries, but there are many of those.

## CONSOLIDATION

All interviewees had comments on consolidation of medical practices/groups.

- Hospital ownership of practices or direct employment of physicians has stalled. The percentage of physicians in those arrangements has remained the same, 32.8 percent, since 2014. This figure is probably lower for radiologists.  
[<https://www.healthexec.com/topics/leadership/physician-ownership-practices-falls-below-50>]
- “2016 marked the first year in which less than half of practicing physicians owned their own practice-47.1 percent.” [<https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/public/health-policy/PRP-2016-physician-benchmark-survey.pdf>]
- Consolidation has been a huge medical trend but it’s not happening as frequently in pathology. Pathologists are not being employed by hospitals as hospitals do not see pathology as an income source. Hospitals are only looking at increasing their income base.

- Increasingly, orthopedic surgeons are becoming hospital employees. This has grown in the last 10 years and is expected to continue. Right now, most orthopedic surgeons practice in private practice – either with other orthopedic surgeons or in a multi-practice (multiple specialties). Estimated 15–20 percent are directly employed by the hospital; the remaining would be private. Would estimate that 10 years ago it was probably only 6–7 percent that were hospital employees.
- The use of billing services is increasing, probably due to ICD-10 implementation and other government regulations that have increased the complexity of billing.
- Consolidation, especially in billing, is an important trend. It started by moving to local billing groups, then it went regional, next it went to pathology only billing, and now it’s moved to large billing groups.

#### POTENTIAL NEXT STEPS

1. Ideas on how to find or obtain list of doctors that bill separately
  - Directly ask large hospitals/systems for lists of the contracted SBD groups. The Medical Staffing department should provide it. A possible reason to refuse sharing the list is due to business intelligence/development. They’re in competition with other local hospitals.
  - For specific SBDs, ask “Are your pathologists independent contractors or are you billing for them?” “Do you provide and bill for the technical component?”
  - If the hospital won’t provide a list, check whether the hospital releases the clinical performance measures on physicians to the public. This a good way for us to gain a list of doctors providing care at the hospital. Expect some refusals if clinical performance measures are used in hospital contracting.
2. Ask the payors for information. For insurance company payors limit the request to the names (and possibly contact information) of doctors/practices/health systems that billed for services. If billing (charges/payment) data is the issue, perhaps they would be agreeable to a request where we are not asking for billing specifics.
3. Investigate the possibility of obtaining Research Identifiable Files (RIF files) of Medicare and Medicaid claims. To obtain these data we would need to formally request RIFs. The cost of obtaining these data is dependent on the type and amount of data needed. It is unclear whether pulling data for specific individuals is permissible, and if so, how that would affect costs.
4. EOB statements could be a good source of data on who provided services. Anything that was attempted to be billed would be included but they do not always include CPT-4 codes.
5. Ask our large provider POCs which labs they contract with, and whether the patient would ever receive a bill from the lab. That may help eliminate labs as SBDs for those Hospital cases if all billing is included in facility charges. Often labs contract with hospitals, but hospitals handle the billing.
6. “Medical records” is too general a term. We should put more emphasis on specifically what we want, such as pathology/lab reports. The medical records department should have easy access to these pieces of information, and should be able to export it. We currently have these items listed in the MR AF packet, but we need to emphasize that we want all

records, including reports and orders, associated with this patient. Perhaps the AF packet could be redesigned to draw attention to reports and orders. For example, we could add checkboxes for each patient on the confidential patient checklist. Also, we could add language to the contact guides, specifically asking for these specific reports.

For example, laboratory/radiology professional services should generate reports. Reports should list 2 doctors' names: a laboratory director (person with CLIA license) & a pathologist. These names are not in the progress report/discharge summary. Explicitly ask for all items, in this case laboratory/pathology/radiology reports.

7. Ask for EMR/EHR export. The typical EMR includes multiple tabs such as: labs, notes, ECG, Pathology, etc. We may be getting only the “NOTES” tab from the patient chart/EMR, which wouldn't include pathology, radiology, and other reports. Other common tabs to ask for: ECG tab (cardiology), Labs, Micro, Pathology, Imaging, Procedures, Referrals.
8. Gaining cooperation at a higher level, someone who has the authority to export full patient information/charts to us via electronic portal or fax/mail.
9. Propose a small study to an insurer, or a small group of insurers. Use a small subsample of the MEPS MPC respondents to compare the already collected MPC data to the information obtained from the insurer. This could guide us to better data collection techniques in terms of identifying SBDs.
10. Although there has been a recent review of the ICD-10 query used to identify missing SBDs, the CPT-4 query has not been reviewed recently. We recommend reviewing and updating the CPT-4 query. We could also add a DRG query to flag events that are associated with certain specialties. In addition, develop and fine-tune protocol for re-reviewing records and/or retrieving additional data from hospitals.
11. Maine and Wisconsin both have all payor claims databases; we could investigate the availability of these databases.
12. Training Point: The term “on staff” does not mean the doctor is employed by the hospital, or that the hospital bills for the doctor's services. It means the doctor has privileges at the hospital – s/he is allowed to provide medical services/is credentialed in the facility. Hospitals would not have a list of doctors by date but they must have a list of doctors credentialed in their facilities.

— **Project #5: MEPS MPC OP3 special project – review of SBD data mining – findings** —

Date: July 20, 2017

To: Karen Davis

From: Thom Miano, Erica Saleska, and Katy Dowd

**Subject: MEPS MPC OP3 special project – review of SBD data mining – findings**

This memo includes findings for the MEPS MPC OP3 special project related to the investigation of SBD data mining. The goal of this special project was to begin work that will help us assess our ability to quickly identify SBDs from patient records. The hypothesis is that these methods may lead to higher accuracy and reduced resources needed during the process of identifying eligible SBDs.

#### OVERALL RESULTS

The results of the work through June 30, 2017, have not yet indicated useful or characteristic patterns in the SBD data that can be used for *reducing burden in SBD identification and data collection* beyond the current existing methods; however, additional investigative analyses could be worthwhile. We have included recommended paths for future work on this topic in the “Suggestions for Next Steps” section listed at the bottom of this memo.

#### METHODS

The work completed through June 30, 2017 was conducted as detailed below.

1. RTI identified features and a plan for cleaning (e.g., approaches for handling categorical values).
2. We focused on a subset of the 2013 SBD dataset to begin development of code, constraining features to specialty, role, and contact priority status. We then wrote preliminary models using logistic regression and random forests.
3. RTI one-hot encoded (or created dummy variables) for “Specialty”, which has high cardinality, and used this for logistic regression.
4. The implementation of random forest used was unable to handle the high cardinality. In R, the H2O library was suggested as an alternative that had tree and forest algorithms that could handle high-cardinality categorical data. H2O is a high-performance machine learning computing library in R. When doing machine learning, one can get different results based on the implementation of various algorithms that they are using. Different libraries can implement algorithms differently, so the library one uses can have an impact on the final results. Quality controlling by evaluating the results of different libraries can be a useful step when appropriate.
5. We wrote code to produce scatterplots of dimensionally-reduced data using the T-distributed stochastic neighbor embedding (t-SNE) algorithm. t-SNE is a dimensionality reduction algorithm that attempts to reduce the feature space of a dataset while maintaining local and global distances between points in the full feature space (i.e., a dataset with all the

variables). As compared to other dimensionality reduction algorithms (e.g., principal component analysis), t-SNE can preserve the natural structure of data in lower dimensions, which is useful for qualitative observation of structure in data and verification of quantitative observation of structure in data. This analysis was done to evaluate support for possible structures, clusters, and persistence of structure over time.

We attempted to use two implementations of t-SNE. One was too slow and timed out before results could be completed (several attempts were made). The other could not be installed due to an inability to install the necessary software dependency tree (OpenSSL not available on node).

The following variables and feature transformation techniques were used during the analysis:

- Physician name, specialty, role, hospital name (where event occurred), practice name, billing service name, procedure or BETOS codes.
- For physician name, one could look at previous years to see what SBD categorization that physician received (if any).
- Specialty, hospital name, practice name, billing service name, and procedure and BETOS codes are all high-cardinality categorical values, so an appropriate strategy needs to be developed to determine how to handle them.
  - Some implementations of tree-based models (e.g., H2O) are supposed to be able to handle high-cardinality categorical variables.
  - One-hot-encoding (i.e., creating “dummy variables”) is an option, but doesn’t work well with tree-based models.
  - There are a variety of techniques for transforming categorical values to continuous, which are commonly used in applications like customer acquisition, response modeling, default prediction, and fraud detection. Example techniques:
    - *Terms:*
      - P = Positive instance; N = Negative instance
      - TP = Total positives; TN = Total negatives
    - *Ratio* of observations associated with an individual level
      - $P/N+P$
    - *Weight of evidence*
      - $\ln ( (P / TP) / (N / TN) )$
    - *Perlich ratio*
      - Appropriate when working with data that incorporates many-to-many relationships.

## SUGGESTIONS FOR NEXT STEPS

1. RTI would suggest revisiting any non-response modeling already completed in prior cycles and assess the use of the feature transformation techniques noted above. Measurement of

the impact of these feature transformation techniques could be insightful. In conducting this assessment, it would be useful to pair a data scientist with a MEPS-MPC statistician who is well-acquainted with the data and who can revisit any earlier non-response modeling.

2. It could be worthwhile to test some multilevel models (e.g., Bayesian Hierarchical Modeling), given that many of the variables of interests have high cardinality. This could be implemented in addition to or as an alternative to feature transformation of categorical variables to continue.