Minutes
Initiation Work Group, HSCRC
Monday, August 28, 2006
9 am-10:30 am
Room 100, 4160 Patterson Avenue
Baltimore, MD 21215
As corrected: Oct. 2, 2006

IWG Members Present: Dr. Trudy Hall, Chair and HSCRC Commissioner; Dr. Charles Reuland, Johns Hopkins Medicine; Dr. Christian Jensen, and Ms. Deneen Richmond, Delmarva Foundation; Dr. Donald Steinwachs, Johns Hopkins Bloomberg School of Public Health; Dr. Vahé Kazandjian, Dr. Nikolas Matthes, Mr. Frank Pipesh, and Ms. Karol Wicker, Center for Performance Sciences; Dr. Grant Ritter, Brandeis University; Ms. Wendy Kronmiller and Ms Renee Webster, OHCQ; HSCRC: Mr. Robert Murray and Ms. Marva West Tan.
On conference call: Ms. Sylvia Daniels, University of Maryland Medical Center, Ms. Joanne Koterwas, St. Mary's Hospital; Mr. Gerald Macks, Medstar; Ms. Jan Bahner, Medstar. Ms. Susan Nieto, St. Agnes Hospital. (There may have been other persons on the audio conference. who did not identify themselves.) **Interested Parties Present**: Mr. Don Hillier, former Commission Chairman, Ms. Ing-Jye Cheng, MHA; Ms. Kristin Geissler, Mercy Medical Center; Ms. Charlotte Thompson, HSCRC; Mr. Craig Weller, Delmarva Foundation, Mr. Jim Miller, DHMH – OPF, Ms. Grace Zaczek, DHMH- OPOPS, Dr. Luis Mispireta, Union Memorial Hospital.

Welcome and Approval of Minutes- Dr. Hall welcomed the Work Group and attendees on the audio conference. She noted that there were some new representatives present from Delmarva Foundation, Dr. Christian Jensen, CEO and Ms. Deneen Richmond, Executive Director, DFDC
Senior Vice President, Interventions Team. Other Initiation Work Group members present and on the telephone introduced themselves. The minutes from the May 24, 2006 meeting were approved as distributed.

1. Status of Data Acquisition – Ms. Tan updated the Work Group on steps being taken to obtain the clinical measures data needed for the pilot. Piggy-backing data acquisition on the Maryland Health Care Commissions' data acquisition for the online Hospital Performance Guide proved to be not feasible. Ms. Tan said that HSCRC is now working with the Delmarva Foundation, in its role as the Quality Improvement Organization (QIO), to access the data. The QIO has access to quality measurement data already reported by the hospitals to the Center for Medicare and Medicaid Services (CMS) clinical data warehouse. Mr. Robert Murray, HSCRC Executive Director, recently sent a memo and data authorization form to all acute care hospital CEOs requesting consent to access each hospital's data. Ms. Tan is working with the QIO to develop the Data Use Authorization contract; and HSCRC with CPS's technical assistance is developing the very detailed data specification request required by CMS. Ms. Tan noted the CMS data access process may take some time. .

   In the meantime to keep the pilot moving along, Dr. Kazandjian has received consent from 18 hospitals to data obtained via the Quality Indicator Project (QIP), a data vendor for the majority of Maryland hospitals for required JCAHO and CMS reporting of Core Measures.. The data from the CMS warehouse will still be needed as it includes all Maryland hospitals but use of the QIP data permitted data analysis and statistical testing to begin. Dr. Kazandjian added that the 18 hospital sample was larger and richer than the five hospital sample originally envisioned for the Alpha pilot. (List of 18 hospitals attached.)

   Hospital Forum planning – Ms. Tan referred to the comments, attached to the agenda, received regarding the draft plan for a future Hospital Forum about the Quality Initiative. She did not review the comments in detail but pointed out that they emphasized that hospital

personnel will be very interested in learning that no extra data collection will be required. She also mentioned that Dr. Reuland identified several issues that will likely be raised about the methodology for rewards and incentives.

2. Initial Statistical Analysis and Composite Score Construction- Dr. Kazandjian noted that the initial methodology recommended that the Alpha pilot involve a small number of hospitals but ultimately all hospitals involved in the Quality Indicator Project were invited to participate in the Alpha pilot. Dr. Kazandjian was pleased to announce that 18 hospitals did give consent for their data to be used in the Alpha pilot and this level of analysis is focused on what can be done with the data. He noted that the project is on time as noted in the timeline. He then introduced Dr. Grant Ritter, Brandeis University, to discuss his initial statistical analysis of data from the 18 hospitals. (Please refer to attachment for content of the presentation.) Dr. Ritter noted that his topic today is composite score construction and his challenge is to combine several patient-based quality indicators to create a single composite score to rank the performance at each hospital. Creating this composite score is quite challenging because it is intended to represent quality across all measurement categories not just for the individual disease. Dr. Ritter noted that it was very good to have some real data to analyze; that each indicator had significantly different numbers and that some statistical approaches discussed at the last IWG meeting seemed less applicable when real data were considered. Due to the great variation in numbers in each cell, and the very small numbers in some cells, Dr. Ritter does not recommend using the unweighted average, as demonstrated in this paper. He noted that the number of patients meeting AMI 7 (page 4 of handout) was only 32 out of 3280 patients (Although AMI 7 was identified as Adult Smoking Cessation Counseling during the meeting, this was a misstatement. AMI 7 in the handout represents JCAHO/CMS AMI 7a – Thrombolytic agent received within 30 minutes of hospital arrival. See attached key to numbers.) These data are existing data from 2005 and that over time, the numbers will likely increase with more data collection. Factor analysis will probably not be the appropriate way to weight these measures.

   Dr. Ritter discussed the CMS approach to weighting as outlined on page 10 and 11 of his handout. At this stage, Dr. Ritter suggested that the CMS method might be the approach for HSCRC to use also. The opportunity model fits well with the sample size available; 25-30 opportunities seem to be the minimum for use. The variance across the averages and the within hospital variance within this sample data set is great enough to conclude that the hospital ranking the highest is performing better. After combining within disease categories, Dr. Ritter had to combine scores across all categories. He looked at three diagnoses as there was not much data available for the Surgical Infection Prevention (SIP) measures. He looked at correlation across the three diagnoses. One would prefer to see correlation across the categories but this is not what Dr. Ritter found. There is correlation between Pneumonia (PN) and Heart Failure (HF). The Acute Myocardial Infarction (AMI) measure seems to be independent of both the HF and PN measures. How to take this into consideration is somewhat problematic and needs more analysis. Also as more data are available over time as hospitals gear up data collection, there may be more correlation. Dr. Ritter concluded with reiterating that having real data was very helpful.

   Dr. Kazandjian added that a traditional epidemiological method is to consider the ratio of observed to expected. Once we have that information, then the next step is how to build a profile and a single composite measure. He added that this step is feasible because the measures are evidence-based and the best practices are known. One area that will need further consideration is whether the expectation is that if a hospital does well in one area; it should do well in others. Experience tells us that hospitals are not equally good in all areas. So, many questions are still open. He concluded that the data from 18 hospitals is more robust than data from five hospitals would have been. The next step will be different proposals from CPS on

how a composite score could be constructed assuming that there is agreement on the correct approach to building correlations or whether some categories should be considered individually, and this is where weighted means would be used.

Dr. Hall then asked for questions and comments from the group. Dr. Steinwachs asked for an explanation of why the smoking cessation counseling number was so low. There was some discussion about definitions and exclusions as well as data collection issues. Dr. Steinwachs wondered if these issues meant that the measure should not be used in the composite. Dr. Kazandjian noted that these measures are continually validated by CMS and JCAHO. Dr. Hall noted that it important to continue to support smoking cessation counseling and the group should continue to look at these policy issues with measures with low numbers.

Dr. Reuland asked if all expected and observed measures across all three diagnostic categories were tallied in this presentation. Dr. Ritter answered that they were not in this example; the tallies were within a diagnostic category. He did look at the scoring in the manner suggested, and the AMI category was still independent. He also noted that measures with very few reported observations do not have much impact on the result.

Ms. Ing-Jye Cheng noted that it was good to have a dynamic data-driven process, but it also means that the weights will change year by year as the data changes or has the potential to change. She raised some concerns for ongoing consideration: 1.) How often would reweighting occur? 2.) How do we know what the impact is of improved reporting versus improved quality? 3.) What happens when the measures change? Would there be a reweighting process and how would that be made transparent? 4.) Should some measures be weighted more or less because of State or other strategy issues or because the hospitals are still implementing processes to report the measures? 5.) Is this an initial process for the hospitals and the IWG to have input on what the weights should be? 6.) How do you set thresholds for measures when you will be combining measures and combining thresholds? 7.) How does HSCRC stay on top of this fluid process and keep the hospitals informed when the results will be impacting reimbursement? As a final technical question, she wondered if there was an opportunity to count one patient multiple times across measures and across diagnostic categories and how this would impact the composite score. Dr. Ritter responded that a patient could be counted twice, for example if patient came in with an AMI and developed pneumonia. Dr. Matthes said that CPS had to look into this issue for sampling for the Quality Indicator Project, and the good news is that it is a very uncommon finding to have patients who fall into more than one category. Regarding how often the measures would be reweighted, Dr. Ritter noted that each opportunity has a weight of one (1), so as the number of opportunities increased, the weights would increase. But the weight would stay one, so the change would be automatic.

Dr. Reuland wondered why the numbers varied so much for each measure within a diagnostic category and what the hospital drivers for data collection and reporting were. Dr. Ritter said this was not known and he added there was a potential for gaming. Dr. Kazandjian added that incentives can change reporting. Depending on how the initiative is structured, there may be different results with voluntary reporting versus reporting with consequences. There is a policy component to this process which differs from the statistical component. The policy issue is should weights or prioritization be given to certain measures as a strategy? He added that there is a sequence to development of the methodology. First, he suggested that the group focus on the quantitative before going to the policy issues. We are describing the data, the gaps, and then the importance of certain measures to the State. He noted that Ing-Jye Cheng has raised important questions for ongoing consideration.

Dr. Matthes added that the data may change over time. For example, in the Premier/CMS demo, he heard a recent presentation about year two data which indicated that there is not much statistical difference between the top two deciles. This was not the case in the first year, but the change is already appearing in the second year for the hospitals who self-selected to participate in the demo. (Premier is basing graduated rewards on the top two deciles.) Right now there is good differentiation in the Maryland data that was analyzed, but that may change over time and the intent is that it will change over time. He also spoke briefly about the fact that patients may not be represented well in all measures for a category and noted that the good news is that a recent Harvard study shows good correlation between AMI process measures and AMI mortality.

Mr. Murray asked if different options for a composite measure would be presented at the next meeting and Dr. Ritter answered positively.

Dr. Ritter responded to Ms. Cheng's question about thresholds. He noted that if you combined categories that did not have good correlation, it pushed the score for hospitals closer together, which may not be a desired result. In that example, a hospital that ranked at the top would not be statistically better than other hospitals.

Dr, Kazandjian added that in a project such as this one, with very narrow margins in quantitative measurement across hospitals, the goal is to bring everyone up to the same level. Because we are working with evidence-based measures, we know what is expected, and we want the observed to be the same as the expected. If we are successful, there will be a time when the present cohort of measures will be useless because there will be no variation. So, we need to plan from day one for the second cohort of measures. This initiative will have policy and research issues on an ongoing basis; otherwise the initiative will become obsolete after awhile.

Ms. Cheng noted that she is concerned about volatility in the combining process. We should consider if stratification into peer groups would help this issue.

Dr. Mispirata stated that he did not believe that the differences between the observed numbers for the measures were due to gaming. He felt that the exclusions for individual measures were a factor. He also noted that the lack of correlation between disease categories could be due to hospital administrative goals and the different clinical teams given responsibility for meeting the diagnosis-related measures. He gave an illustration from the hospital where he practices. He wondered if it is going to be possible to construct an overall composite measure. (See also attached copy of email received after the meeting with comments from Ms. Geissler who was a meeting attendee.)

Dr. Kazandjian concluded that we are not into interpretation yet. CPS will continue with the descriptive analysis and we will proceed to consider the construction of composite measures. HSCRC staff will provide the econometrics perspective.

3. Adjournment- Dr. Hall thanked the whole group for their participation and effort. The next meeting date will be Friday, September 29 at 9 am to 10:30 am at HSCRC. Dr. Hall adjourned the meeting.