

MEDICARE PAYMENT ADVISORY COMMISSION

PUBLIC MEETING

Ronald Reagan Building  
International Trade Center  
Horizon Ballroom  
1300 13th Street, N.W.  
Washington, D.C.

Friday, October 29, 2004  
9:05 a.m.

COMMISSIONERS PRESENT:

GLENN M. HACKBARTH, Chair  
ROBERT D. REISCHAUER, Ph.D., Vice Chair  
JOHN M. BERTKO  
SHEILA P. BURKE  
FRANCIS J. CROSSON, M.D.  
AUTRY O.V. "PETE" DeBUSK  
NANCY-ANN DePARLE  
DAVID F. DURENBERGER  
ARNOLD MILSTEIN, M.D.  
RALPH W. MULLER  
ALAN R. NELSON, M.D.  
CAROL RAPHAEL  
WILLIAM J. SCANLON, Ph.D.  
DAVID A. SMITH  
RAY E. STOWERS, D.O.  
MARY K. WAKEFIELD, Ph.D.  
NICHOLAS J. WOLTER, M.D.

## **AGENDA**

### **Hospital pay for performance -- Karen Milgate, Sharon Cheng**

MS. MILGATE: Yes. In this session we're going to discuss whether it's feasible, given the status of quality measures for hospitals, for Medicare to base a small portion of hospital payment on quality. First we'll summarize briefly past Commission discussion on the topic and then through the body of our analysis on the quality measures for hospitals that are available.

Two years ago, in evaluating incentives to improve quality the Commission found that one of the most important incentives, payment to individual settings, was either neutral or negative towards quality. To address this problem, the Commission supported the concept of tying some portion of payment to quality and develop criteria for determining whether settings were ready for this type of initiative.

In March of 2004, the Commission determined that quality measurement for health plans and facilities and physicians who treat dialysis patients in fact was ready and recommended that Congress should establish a pay-for-performance program for those settings of care.

This is the first discussion to consider whether measures and measurement activities for hospitals meet that criteria. We are not suggesting the Commission identify a specific set of measures, but rather to determine whether a sufficient number and type of measures are appropriate for this use.

To assist MedPAC staff in this evaluation we convened a hospital measurement expert panel. The type of organizations that came to that panel were in your mailing. This presentation is based on staff research on measures as well as the discussion of the panel.

So again, the purpose of this analysis to try to answer the question of whether it's feasible to base a small portion of hospital payment on quality. There's really two broad considerations. First, the criteria that I've described and that Sharon will go in some detail next. And second, to think of that in the context of the cost of not moving forward balanced with the potential cost of moving forward with pay-for-performance. The cost of not moving forward can be measured in dollars as well as patient lives through complications, longer lengths of stay, readmissions and unnecessary pain and suffering for some Medicare beneficiaries.

This needs to be balanced with the cost of moving forward which primarily there are two that we've identified.

First, the administrative cost to hospitals and CMS to collect and analyze the data as well as to develop a ranking system, as well as the potential unintended consequences if in fact the criteria the Commission has laid out are not met in the hospital world.

MS. CHENG: After looking at a couple dozen private attempts to link payment and performance, MedPAC developed four criteria that help us to gauge whether or not a sector is ready to move to pay-for-performance. The first criterion is whether there is a set of well-accepted evidence-based measures. By this we mean, is there a set of measures that are familiar to providers that are going to be measured? Are they evidence based? Are processed measures based on clinically proven standards of practice? Are outcome measures based on an aspect of performance that has been linked to the outcome that we are measuring? And are they reliably measured?

The second criterion is whether there's a standardized mechanism for data collection. We look at this one to determine whether or not this measurement would pose an undue burden on either the providers or on CMS. We are also looking for standardized data collection to make sure that when we get this data we have something that we can compare from hospital to hospital.

Our third criterion is risk adjustment. If it's necessary, it should be adequate to maintain equity for providers and access for beneficiaries.

And our fourth criterion is whether or not providers can improve on these measures. This has a couple of aspects. First of all, are we getting a measurement that we can use on as many providers as possible? Are we measuring something that the providers believe is under their control? And is it an area that's been identified that needs improvement?

All of these criteria add up, hopefully, to whether or not a sector is ready for pay-for-performance and whether or not moving to pay-for-performance is going to lead to a substantial improvement for a substantial number of beneficiaries.

MS. MILGATE: To assess the measures that are available and in use for hospitals we divided them into four types of measures. Those would be process measures, outcomes measures, structural measures, and then patient experience of care as a separate measure. We're going to discuss each type in turn.

First, process measures are probably the most well used in the hospital sector, and they try to answer the question of whether patients in the hospital are receiving clinically appropriate care. That is, does the hospital have in place processes known, and are they used, that are known to produce better outcomes?

The strength of these measures, at least as discussed in our expert panel, was that at the same time it measures the quality of hospital care, it also helps identify what needs to be done to improve that care. In addition, because generally they are evidence based, they're well accepted by providers.

Examples of process measures include aspirin on arrival and discharge from a hospital for those with a heart attack, assessment for left ventricular function for patients with heart failure. For patients with pneumonia, whether they received an antibiotic within four hours of coming to the hospital. And for surgery patients, whether they received an antibiotic as a prophylactic one hour before surgery, and then was that antibiotic discontinued within 24 hours after surgery. Hopefully that gives you some sense of what the measures are like.

So who uses process measures and how are they used? As I said, they're widely used. The Joint Commission on Accreditation of Healthcare Organization uses them in their accreditation process where hospitals have to report on some process measures to be accredited. The Leapfrog Group uses some process measures in their public reporting and pay-for-performance initiatives. The National Quality Forum is not an organization that uses measures but is a consensus building organization and they have endorsed a fairly broad set of process measures in their consensus process.

CMS uses process measures for a variety of reasons. They use them and have used them for quite a long period of time in the QIO program. They actually developed some of the measures and use them for feedback to hospitals to improve care. They also use process measures in the Premier demo, which is a demonstration they're conducting to look at the possibility of doing pay-for-performance in hospitals in Medicare. They also use process measures in the new initiative where they tied reporting on some process measures to getting the full update to hospitals last year.

One initiative that they work with the private sector on is the next one listed on the slide and that's the Hospital Quality Alliance. A few years ago CMS, the American Hospital Association, the AAMC, the Federation of American Hospitals, JCAHO, AARP and AFL-CIO -- I believe there's actual organizations that were also involved in the beginning of this initiative -- developed and identified a set of 10 that hospitals could report on voluntarily. So that's another way that CMS uses them and that's a whole other initiative that is also going on at this time. This initiative has also identified another set of measures beyond the 10 initial that they also are going to ask whether hospitals would voluntarily report on sometime in the next six months to a year.

The most visible and widely used of any set of process measures are the 10 I spoke of. These are the 10 that the voluntarily reporting initiative had as their initial set and the 10 that Congress said they wanted to tie to the update last year. Hospitals, in that particular initiative, it's expected that almost all hospitals that were eligible to be a part of the program will report and have reported on that. There's around 3,800 hospitals that CMS expects to put their individual scores on the process measures up on their web site in November. So that's next month.

So just to summarize, as I said it's used for a wide variety of processes; accreditation, internal feedback, public reporting, as well as pay-for-performance.

So are there process measures that meet our criteria? Just to lay out the criteria here. Many are well accepted and evidence based, in particular the 10 I spoke of as well as the seven or so that the voluntary initiative intends to use in the next few months. In particular, our expert panel thought that the surgical infection prevention measures that were included in that were particularly promising. There are seven different surgeries that they cover and they thought that would be a good effort to really work on patient safety across the organization, so it wasn't so condition specific, which most process measures are very condition specific.

While a burden to collect, most hospitals are currently reporting on some for multiple purposes. They are reporting for the update purpose, for QIOs, as well as for the Joint Commission. Providers emphasize to us that if they were to be measured on process measures that it would be very important that all of those that ask for information from them define the measures the same. That they not be similar, they not be in the same area, but they be the same so that they're collecting it once for multiple purposes.

Risk adjustment on these measures is not necessary, so that's not an issue that we deal with here.

Can hospitals improve on these measures? Clearly, more improvement is possible on many of the measures. However, we did see that on the reporting on the initial hospitals in a voluntary initiative there were several that are at a fairly high level, which point out the need to continue to evolve to new measures as hospitals do improve to certain high levels across the country.

Most hospitals do see patients with one or more of these conditions. For example, if a hospital doesn't see heart patients, they may see pneumonia patients. It also might be useful though to look at some crosscutting measures such as surgical infection so that you aren't limiting their incentives program to a certain set of hospitals.

And there has been some discussion that a broader

set of measures might be necessary for small and rural hospitals, and if critical access hospitals were included in the program, to recognize some of their core functions such as stabilization and transfer.

MS. CHENG: The next type of measures that we're going to discuss are outcome measures. The panel that we spoke with agreed that outcomes are really the bottom line. Payers would to know how the care that they purchased affected their patients. Outcomes can capture clinical effectiveness and patient safety. However, as strong as a consensus was on the importance of outcomes, it was less strong for this type than for the other types on the subject of the readiness of the available measures.

Broadly, there are two types of outcomes, mortality and adverse events. An example of mortality might be the rate of mortality following coronary artery bypass graft procedure or other procedure-specific mortality rates. Perhaps the rate of mortality of patients hospitalized for pneumonia or other condition-specific mortality rates. These could be in-hospital, 30-day after admission, or a number of various windows.

A third example of an outcome measure would be the percent of patients who developed decubitus ulcers during their hospital stay. This is an adverse event and we measure it because we believe it's reflective of patient safety conditions at the hospital.

Some outcome measures are currently very widely used. AHRQ uses mortality in adverse events in their national report on the quality of health care. The NQF launched its efforts to develop a consensus on measuring hospital quality with a set that included some mortality measures and patient safety measures. JCAHO uses a measure of mortality in their core set and also in the information on hospitals that they make available to the public on their web site. The Premier pay-for-performance demo also includes mortality measures.

Some mortality rates are also reported widely publicly. For example, the hospital-specific CABG mortality rate on every hospital in California is currently available to patients in that state. And Health Grades, which is a public database of information hospitals includes mortality, and also on the very first page they have patient safety indicators that are available to patients on every hospital in that database.

The criterion that we mentioned earlier about risk adjustment is especially important for outcomes measures. The reason is because some types of patients are much more likely than others to experience mortality or adverse events. To maintain equity among the providers that we're measuring and access for risky beneficiaries, risk adjustment for outcomes should be sufficient to identify the

relative complexity or severity of the hospital's patients.

The adequacy of risk adjustment was an important criterion for the groups that considered whether or not to include mortality and patient safety in their measure sets. As your paper discussed in some detail, AHRQ, NQF, and CMS all considered the adequacy of risk adjustment before putting together their measure sets, and they included some mortality and some patient safety, but not all.

Risk adjustment can be achieved currently with the administrative data that we already have. Alternatively, the Commission might wish to consider a recommendation about adding some information to the claims or the administrative flow of data that would give us a somewhat better level of risk adjustment.

Information such as a secondary diagnosis on admission would allow patient safety measures to better distinguish between something that happened before a hospital stay and something that happened during a hospital stay.

As another alternative, a risk adjustment method similar to the private benchmarking organizations that do similar quality measures could be considered. However, this level of risk adjustment requires record abstraction or other intensive data collection efforts. The Commission would need to consider the trade-off between the burden and the improvement in risk adjustment before you would consider that to be a feasible measure for our measure set.

So are there outcomes measures that meet all four criteria? There are generally familiar evidence-based outcome measures. Depending on the risk adjustment they can pose very little burden. Some outcome measures can be measured on most hospitals. Hospitals do have room for improvement. And a sufficient level of risk adjustment remains a question for many measures.

MS. MILGATE: Structural measures are measures that are used to ensure that the hospital is capable of delivering quality care. They often address systemwide problems rather than specific condition problems.

Examples of structural measures, there's really a wide variety as you can see from this slide. Accreditation was discussed in our hospital panel as a structural measure. Implementation of computerized provider order entry, also another example of putting a system in place to avoid medication errors.

If a hospital puts in place systems such as just having more sinks in the hospital that are available to the health care providers. that is one structural way that they could encourage handwashing. Or if a hospital puts in place a program to try to avoid similar medication abbreviations to alleviate some medication errors, that's also considered a structural measure. Those two are both a part of a

broader set that we'll talk about in a little bit more detail that the National Quality Forum developed their safe practices list.

Another example of a structural measure which was discussed in our hospital panel was volume as a structural measure. That would be, for example, the number of CABG procedures that are performed at the hospital. There's some literature that shows a relationship for some procedures between the volume of procedures and hospital quality.

So who uses structural measures and how do they use them? Probably the organization that supports structural measures the most is the Leapfrog Group, and they're also well known for this. When they began their program a few years ago, two of their first leaps, as they called them, were whether a hospital had implemented a computerized provider order entry. This was as a structural measure to look at whether a hospital was trying to reduce medication errors. The other was whether ICU units used intensivists. There had been some literature that showed that length of stay in ICUs was shortened when they used intensivists.

They also look at volume and have in the last year endorsed the use of the NQF-endorsed safe practices list, which includes the three that they had put in place at first and then 27 others.

The National Quality Forum, as I said, developed this consensus list and endorsed 30 safe practices. CMS uses structural measures primarily through its deemed status relationship with accreditation, but also has in place in their QIO program some safe practices measures that -- they overlap with their surgical infection measures, but I would think those would be considered structural as well. Then JCAHO, their accreditation product is a structural component itself, and within that there are quite a few different structural standards that they look at. In particular, the safe practices, they have their own questions about whether a hospital does safe practices or not.

So these have been used for pay-for-performance, in particular by Leapfrog, public reporting, and for certification processes.

So are there structural measures that meet our criteria? Some, such as the safe practices list and accreditation, are well accepted; others less so. There is a debate over whether implementation of the CPOE in the hospital is something that is perhaps too expensive for enough hospitals that it should not be used as a measure. There are discussions back and forth on that and I won't go into the detail of that now, but suffice to say there's some controversy about that.

In terms of intensivists, again the debate is one about whether there are enough intensivists available for

hospitals to actually use them all over the country. The proponents of using intensivists suggest that if in fact this were used as a quality measure the supply might increase. So the debate goes both ways.

The burden of collecting the data varies, but many hospitals are already doing it. For example, for accreditation many hospitals are also assessing some of the safe practices. The Leapfrog Group told us that they expect within this reporting cycle to have 1,200 hospitals who have filled out or are filling out the survey on whether they are using, and the degree to which they are using safe practices identified on that list.

Risk adjustment is not necessary and our panel was pretty much in agreement that improvement is possible for all hospitals on many of these measures. They were particularly supportive of measures that moved into the area of trying to improve safety by looking at what the practice should be to improve safety rather than counting the adverse events that were the result of unsafe practices.

MS. CHENG: The final measure type that we'll discuss this morning are patient experience measures. This type of indicator measures whether or not the goals of the patient were met during their hospital care. They may reflect whether or not the patient was truly at the center of care, did doctors and nurses and other professionals listen to the patient and try to understand what he or she was trying to achieve during their hospital care? Did the patient receive adequate information to be an active participant in his or her care while they were in the hospital?

These measures are cross-cutting in a couple of ways. First, they can apply to almost all types of patients. They're not restricted just to surgical patients or patients with a particular kind of condition. You can measure pretty much anybody who walks in the door of the hospital for care.

Second, they can transcend hospital care a little bit and break out of the silo by asking patients how well prepared did you feel for going back home or going to your next setting of care? So in that way they can be somewhat cross-cutting.

Some examples of patient experience measures are, how often did a doctor listen to you carefully? How often did nurses explain things in a way that you could understand? And did you get information about symptoms or health problems that you should look for after you were discharged from the hospital?

Many different hospitals use many different tools to measure some aspect of patient experience. CMS and AHRQ looked at all this measurement activity going on and realized that it would be good to develop a standardized set

so that hospitals would have comparable information so they could compare patient experience. They worked to develop HCAHPS, which would be a standardized tool that hospitals could use. They also looked at the tools that are already in use and they looked at the CAHPS survey that's used by health plans to measure patients' experience within health plans.

They used a broad stakeholder process to design a tool that would be relevant to as many information users as possible. They included hospitals, the American Medical Association, and AARP, among others, as they designed their tool. They streamlined the tool working with this group and with researchers, and they only retained items that passed tests for reliability and validity of measurement. Their inclusive approach to designing the HCAHPS tool will continue as they submit it to NQF also for their input on the tool. This tool has already been field-tested at over 100 hospitals in three states. It will go through additional dry runs in the field and is expected to be in final form for voluntary use by hospitals in the summer of 2005.

So are there patient experience measures that meet our criteria? Generally measuring patient criteria is well accepted. It may pose a small burden on hospital's depending on the tool. Risk adjustment for patient experience may not be necessary, but in fact with HCAHPS will be available. And improvement is possible on these aspects of patient experience.

MS. MILGATE: So we've covered a lot of ground here and in the background materials and we'd like to ask two things from you. First, we'd ask you to identify concepts or measures you think that would enhance the discussion that we may not have covered here, and to react to the analysis we've laid out. What we've found through interviews, evaluation of measures and their use, and the opinion of our expert panel is that the most promising type of measures for pay-for-performance are probably process measures. There's one set, the 10 we spoke of, which is already being reported to CMS on a widespread basis. And the same initiative that encouraged reporting on these specific 10 also intends to roll out a variety of others in the next six months to a year, including promising cross-cutting measures such as surgical infection.

In terms of outcomes measures, both mortality and complications are widely used. However, the level of risk adjustment is at issue. Some have felt comfortable simply using claims for risk adjustment. Other initiatives require a deeper level of information. The Commission may wish to recommend additional data collection to improve risk adjustment methods. That said, two indicators are widely used that are outcomes and those are mortality for AMI and

mortality for CABG, and others are used for public reporting by one or more organizations as well.

In terms of structures, we find some disagreement on some of the measures, but also agreement on a few. For example, accreditation could be used as a good basic framework and as one measure for a set of measures used in pay-for-performance. Our panel felt that volume was something that would be useful information that should be included, for example, the number of cases in a measure, but not as a measure itself. The National Quality Forum safe practices were discussed as a good set of safe practices that could be measured. And while CPOE and use of intensivists may have positive benefits as measures, our analysis would find that they are less well accepted than some of the other structural ones.

In terms of patient experience, they appear to be very promising and our hospital panel felt they were a critical condition to the set of measures that would be used for hospitals, and that possibly they could be included when HCAHPS is final and in use by hospitals.

MR. MULLER: I know it's traditional to compliment you on the work, but this is exceptional. I think both this and the next topic on the agenda the staff really did extraordinary work. Mark and Glenn, this is just wonderful work.

I think we all have some comments where we might quibble a bit, but I think that the field has advanced so much in the last year or two. I think the fact that, as you've pointed out, a lot of the people who are working in this field have come together to try to get more standardized. I think even compared to our discussion two years ago when there were a lot of complaints about everybody's coming at in a different way, I think there's been exceptional progress made in terms of these initiatives at AHRQ and Leapfrog and all the people that you mentioned in your oral presentation. So I think for once we should say this is something that is moving quite well and aggressively, because oftentimes we say things aren't working as well as they should. So I think your chapter and your presentation lays that out.

Obviously, the more we standardize on this, the better we'll be able to get people to improve the performance, which is the ultimate goal here. So I think since the fact that enormous progress has been made we should note that and encourage all of the participants in this to keep trying to work in a common way so that in fact doctors, hospitals, payers, patients can all see what they're getting.

To paraphrase what our IT panel said yesterday, getting the tool out there, in some ways it's the start of the journey. And then obviously how we all in the field

respond to this is of critical importance. In the state that I'm in, Pennsylvania, this reporting has been going on for close to 15 years, largely on what you call outcome and safety measures. I think it's fair to say that the response to that public reporting has not been as dispositive of changes as one would perhaps like it to be. It's largely used there for evaluation. I think tying payment to make it is a critical step and in these recommendations we're moving more and more in that direction.

I know in the New York State CABG report over the same 15 years there have been reports that behaviors have changed in a very positive way. But I would say probably Pennsylvania has the most advanced system that I know about in the country. So just having the tool out there and having it public is not sufficient. Tying payment to it is of critical importance. I think moving to it in the kind of aggressive way that you're suggesting may not seem too aggressive, just 1 percent to 2 percent in the beginning, but I think moving in that direction is a very positive step.

I would also note that it's important to keep evaluating as we go, the responses as we start implementing. I think having both MedPAC and CMS and others get the learnings out there very quickly is of great importance, so that the best practices get both agreed upon as to what they are. Also the learnings as to what works in terms of making things better I think it's very important to get out there.

So in all I think this summarizes very well where the field has moved very aggressively, at least intellectually in the last few years. So I feel very good about the direction that we're going in.

DR. MILSTEIN: One's perspective on performance measurement and use of performance measurement for purposes of payment looks very different from the perspective of different stakeholders in the American health care system. It's not difficult for anybody to project, based on their place in the health care system as to how they feel about issues of how ambitious we should be about performance measurement and reporting now. I want to cite Ralph as an exemplar of providers that embrace performance measurement.

From a purchaser perspective, the world tends to look a little different and I'd just make a couple of comments along those lines. First, at the end of the day, once you immerse yourself in performance measurement in health care you realize you really have two broad choices. You can either measure a small number of narrow facets of care very perfectly and very cheaply, or you can measure care performance broadly, less perfectly and more expensively. Those are the two ends of the spectrum.

I think the staff suggestion about CMS requiring, in a judicious way, supplemental data on hospital bills in

order to support better performance measurement is absolutely essential to helping us resolve the dilemma I just cited. The quality work group of the National Committee on Vital and Health Statistics has actually done a very thoughtful piece recommending what is essentially the smallest increment to data that is currently submitted on hospital bills that if routinely submitted would allow the biggest increase in our ability to move forward more confidently on a broader set of more valid performance measures.

The structural measures discussion maybe would benefit from the additional following comment. If you look at structural measures, some structural measures are what I will call low bar structural measures. That is, these are things that you shouldn't be allowed to have your doors open if you're not doing, and that's what the JCAHO is focused on.

If one looks at the other end of structural measures, which I'll call more aspirational performance measures inspired by Quality Chasm visions as to what American health care should be, that's where you get into what the staff diplomatically described as areas of controversy and disagreement. I think the ICU staffing is a perfect example of that. In the NQF review of this they cited 12 published articles and the folks at Hopkins who have most recently published a review on this say that if we had the kind of staffing that are built into the NQF measures of ICU physician staffing, essentially something on the order of magnitude of 60,000 American hospital patients would not die every year. The majority of those would be Medicare beneficiaries. So it's controversial, but that doesn't mean we should shy away. It's not only length of stay and cost reduction issues but it saves a lot of lives.

Last, in terms of your question, what might we want to see on this list in the way of measures that we don't currently have, I would put near the top, measures of hospital longitudinal efficiency. Elliott Fisher and Jack Wennberg keep telling us that Medicare patients by and large tend to be well longitudinally to particular hospitals and their medical staffs. And those hospitals and medical staffs vary dramatically in the amount of Medicare benefits fuel burn associated with their managing a population of patients over time. So I would love to see a measure of longitudinal efficiency, which is one of the six IOM domains of quality, added to the list.

The last comment is that I think performance measurement in health care is off by several orders of magnitude. If you think about the 10 measures that were agreed upon by that alliance -- I think it's called the alliance -- and you say, if you were to build denominator in the average American hospital of the number of important

processes that take place in that hospital that have to happen right if you're going to get a good outcome and say what percent of those important processes do those 10 measures represent, I would say it probably is less than one-hundredth of a percent.

For example, it's estimated for an average ICU patient each day something like 162 processes have to happen right. So 10 process measures is not even close to what we might need. Steve Jenks at CMS who does this research has basically said, if CMS continues -- CMS has used the QIOs to perfect these same 10 measures over the last several QIO scopes of work. He's basically said that if we keep working at and keep going are our current rate of performance improvement, by the year 2025 the Medicare program will have achieved near-perfect care for 10 process measures for three conditions. That is what we refer to as off by orders of magnitude.

So I guess I'll close by saying that people have said that one of the problems with the health care industry with respect to performance management has been what Don Berwick has called poverty of ambition. There's an equivalent danger on the buy side of poverty of ambition with respect to our purchasing and what we measure and what we reward. So I would hope that we would consider the broader end, the wider end of a measurement approach and not buy into what is in orders of magnitude accession to the difficulty of measurement. It's difficult but we have a lot of measures that are plenty good enough and it's a good time to move forward.

DR. NELSON: Terrific work. Great chapter. I really enjoyed it. I'd like to highlight a couple of the areas that I was struck by particularly. I think it's great to point out that there are some areas where we have improved. We are getting flagellated a lot and it's really nice to have some numbers that show improvement in areas where we've shined the light. I enjoyed that.

I like the emphasis on feasibility in terms of considering the burden and the cost of collection. One area that you might supplement that with would be to give some numbers on the current costs of record abstracts, maybe based on New York or Pennsylvania, just to get an idea of what it costs now. Perhaps you can do some extrapolation on what the additional cost might be for collecting data from chart abstracting.

You pointed out the need for coordination among the entities that are requiring data, to encourage standardization, and agree on a single set that can be reported to all of the various users. I'd really like to have that emphasized. There isn't any reason for successive hordes of folks coming into hospitals asking for the data to be arranged in a little different fashion for their purpose.

If they can all agree on what they want and how they'd like to have it delivered, it would be very helpful from the standpoint of cost and burden.

One area that you didn't mention with respect to data and it might be worth a sentence or two would be the importance of looking at the data with respect to racial and ethnic disparities. Data on race and ethnicity are being collected. The problem is that in quality measurements, performance measurement, oftentimes they aren't being looked at with consideration of whether we are making progress as a nation or individual facilities are making progress in reducing those disparities.

The final point has to do with the panel. It may be that you mentioned the composition of your expert panel before I came in. I didn't see it in writing. Unless they wish to remain anonymous, I think it would be helpful to have them identified because the validity of their comments depend on what they brought to the table.

DR. MILLER: I thought we named the groups that were represented in the mailing materials.

MS. MILGATE: Did you want to go through the groups? We didn't provide individual names, but I can you who the groups were.

DR. NELSON: I'm sorry, I must have missed it.

DR. MILLER: Generally our strategy in these things is we tell people when they come that we'll represent the views and not identify individual comments to individuals. So generally we put the organization, but we can tell you who was at the panel as well. We have no problem with that. We just want to attribute specific comments to specific people.

DR. NELSON: I understand that perfectly. I'd like to know who they are.

DR. MILLER: The list of the organizations that were represented were --

MS. THOMAS: It's at the top of page six.

MS. MILGATE: I'd be glad to go through it.

DR. NELSON: We can do it off-line. That's fine.

Thank you.

MS. RAPHAEL: There was one recommendation that your panel made that I thought was particularly important, and I don't know where it belongs in the way we're organizing process, structure, outcomes or patient experience. That was the hospitals capturing secondary diagnoses upon admission and also upon discharge. I thought that was something that really could have a lot of impact in terms of how care was delivered both within the hospital and after discharge. I think that is a problem, when someone comes in for one procedure and all you get is that one procedure, and they have hypertension and other things, cognitive impairment going on and you don't know it at all.

Both the people in the hospital don't necessarily know it and then you don't certainly don't know it upon discharge. I think that really creates a lot of gaps that contribute to unsafe practices.

Then my other thought is, as Ralph was saying, in New York State we have captured mortality data. Then the question becomes, how is it used? Because every year there's a flurry of activity when it comes out in the newspaper and then hospitals spends a lot of time on PR and how are they going to respond to this, either both to put out an ad saying, we are among the best or defend themselves if they're among the worst. And then lo and behold, it's over and nothing really happens after that.

So for me one of the questions is, let's say we get this right and we get the right order of magnitude. What then?

MR. HACKBARTH: Carol, to me, my response to that is that providing the information in and of itself is a good thing. But that's why we need to start moving towards paying for quality as well. We need to start acting on it. So the action can occur at many different levels but that, in a nutshell, is why I think it's important for Medicare and private payers to begin using the data and making a difference with it.

DR. MILLER: To go one step further, cautiously. I think we've pitched our whole approach to P-for-P and this is the next installment in that discussion that we've been having for a year now, as this is integrating these measures for the purposes of payment. We acknowledged at the outset of this discussion a year ago that other people were doing things like public reporting and CMS. I wouldn't say that we're excluding that from a possibility, but for our purposes and what we were headed towards ultimately recommending, I think we're talking about making it as part of the payment system. Is that fair?

DR. REISCHAUER: With respect to the patient experience measures, they make me very queasy when we get to pay-for-performance. I was wondering if in any of the literature they have examined whether there are systematic socioeconomic differences in responses? Because I have this feeling that the expectations of different groups are really quite different. If a doctor passes through the room of some group they're perfectly satisfied, and another group wants to intensively question the physician. That's their view of satisfactory interaction. That would be question number one.

The second thing is an observation that comes from Arnie's comments. I take it at face value and look at where, notwithstanding the fact that we are going at warp speed compared to how we used to go, Arnie is saying we aren't going to get off the runway in 25 years. I'm

starting a different metaphor today. Everybody get their instructions on that? No farm stuff today.

[Laughter.]

DR. REISCHAUER: In a sense we might feel good and we might have some good PR, but we really aren't going to be changing the system unless there's a correlation between good behavior in the one-tenth of 1 percent we are measuring and everything else. I mean, do institutions which do the right thing on this little microcosm that we're measuring have different management styles, different operating styles that cause good behavior elsewhere? Somebody should be looking at that, the extent to which there are externalities and correlations here because maybe you don't have to measure everything and reward everything if you find that there's a very high correlation between some key things and almost everything else that goes on.

MR. MULLER: That in fact is the intent. Obviously, there are so many things that go on in these settings, you couldn't have 20,000 measures are hundreds of measures. So the question is, can you empirically show what the cross-cutting measures are? For example, one that has gotten more and more attention is infection control, for obvious reasons. If one is good at infection control that can therefore be correlated to a lot of other outcomes as well.

So I agree with Arnie. If you thought that you were just doing one-tenth of 1 percent, you're obviously missing the quality improvement opportunities. So the question is, are there cross-cutting measures that are correlated with good performance in general? I think, again, there's been a lot of progress made in understanding what those might be.

I would argue that you don't want to try to measure 200 things because that's exactly where confusion sets in. I do think you have to keep it to a simpler number. Whether that's 30 or 40 or 50, the way Leapfrog is moving, that's probably the right magnitude, even though as the analysis indicates quite well, here and there you will miss certain populations. Some things are just for kids, some are for adults, and so forth. But I think keeping those kind of cross-cutting measures is an empirical question and I think we should definitely look in that direction.

DR. REISCHAUER: You also have less hesitation putting a greater weight on the pay-for-performance if you're comfortable in that. You don't have to restrict it to the weight of the activities that you are measuring.

MS. CHENG: Just to quickly respond to your question on the patient experience. The folks that have put HCAHPS together also included in their research some of the effects of patient characteristics. So there is actually a

risk adjustment module that goes along with HCAHPS. It includes age, education, self-reported health status, whether English is spoken in the patient's home, what type of service area they received in the hospital, the interaction of age and service, and the mode of survey administration. So there is that module.

MR. HACKBARTH: Can I just leap in here and pursue this for a second? When Arnie was talking I found my heart beating faster, and I'm with you. I believe that we ought to be able to go faster.

But I want to really focus on this issue because in a sense what you're saying is a direct challenge, or arguably a direct challenge to one of our stated criteria about well accepted. It really raises the question of what is Medicare's role, what can Medicare do, being a public program run through a political process? We originally chose well accepted because we thought that that allowed us to build confidence and move forward in an orderly way with maximum political support and less friction that would slow down the process. It was sort of a step-by-step, cautious, very Washingtonian sort of process.

You're saying from your vantage point that that is mistaken. I think that is something that we need to hear from other commissioners about, because it is a challenge to one of our basic criteria.

DR. MILSTEIN: Maybe a 30-second response to say that the pivot is the question, is the term well accepted. Well accepted by whom? The narrow, inexpensive to measure process measures are very well accepted by providers. If you were to say to informed consumer leaders or purchasers, how do you feel about judging this segment of your supply-chain, if you're a purchaser -- to use the CMS 10 -- 10 measures of three conditions for everything you're paying for under Medicare? Any experience procurement person for a Fortune 500 company would say, you've got to be joking.

So well accepted has to be something that is arbitrated by not just the suppliers, the hospitals, but also the purchasers. I think the NQF, what's nice about the NQF is it's a place where multiple stakeholder views as to what needs to happen and what's reasonable to do gets arbitrated. In the NQF, for example, they said that there were 39 measures, not 10 measures, that would be a reasonable starting point, and also endorsed 30 safety practices. And I think they're about to endorse HCAHPS. So there we have a measure set of 70 that multiple stakeholders have come together, laid out on the table their own definition as to what's acceptable to them and come to some agreement. So I think that's an example of a richer set.

MR. HACKBARTH: Then maybe the gap is narrower then it sounded initially, because de facto that is what we're doing. We're looking to organizations like NQF to

validate, if you will, through their disciplined process that these are reasonable measures. We never explicitly said, we're just going to use the NQF, we're going to limit ourselves to that, but in fact that's what we've been doing.

MS. MILGATE: A couple of thoughts I'd add though is that the NQF hasn't done an analysis -- and it doesn't mean that these measures couldn't be used for this purpose. But they haven't done an analysis of whether they think these measures would be appropriate for pay-for-performance. So they said, we think these are good measures of hospital quality. Now that's a really gray area and who's really to decide? But that wasn't the purpose that they put them in the set.

Having said that, the 39 that they endorsed include the 10, include the seven that the voluntary initiative is going to go forward with, includes some of the others that CMS and JCAHO have also said are good measures. Just that they haven't evaluated them specifically for that purpose, so to say, let's just take theirs and those are fine is a little bit of a jump.

MR. HACKBARTH: But implicit is that there might be a higher standard yet for NQF to say that they're appropriate for pay-for-performance, if they say they're good valid measures but there's --

MS. CHENG: It could be that, yes. When you're putting money on the table -- and that's the other point I would add. I've heard Steve Jenks' comment and I think that's a really good comment that he makes. But he also was talking, I think, in the context of the QIO program where you didn't have public reporting, where you didn't have an actual dollar attached to the measure. Those were the only two other thoughts I'd add into the mix.

DR. MILLER: Could I just say one thing about this? This will be a little stylized and won't have the detail quite right, but what we did is when we pulled people together on the panel it represented these groups that we're talking about. It also had hospitals there in one form or the other. You obviously can't get everybody in that way. Again this is highly stylized.

The take on it was, if you're thinking of pay-for-performance, there's probably a lot of process stuff that's ready to go, and lots of already agreement on sets of 10, 13, 39, depending on what level of outcomes. Probably a couple of them or some of them ready to go, but risk adjustment remains an issue. The structural stuff I'm a little less clear on. Then finally, the patient experience, everybody's doing it but not a lot of gelling across the industry and maybe not so much. But it was specifically to bring those different points of view together.

If you think I'm off-base here you need to say. I think in your walk-through and trying to explain each of the

pieces and where they thought the places were ready, in your summary, that's in a sense what you were trying to represent across those groups.

MS. MILGATE: Yes, I would say we wouldn't have suggested from what we saw that those 10 are the only ones. Maybe that's what you're trying to get it. There certainly are some others, and I guess that's what we were hoping you would help with. This is the direction we saw in our research on the measures, the use of the measures and the hospital panel. But we didn't mean to suggest that only those 10, for example, would be the only ones. If the Commission feels like there's some other areas it's important to push in, there do appear to be measures that are used for public reporting, for example.

MS. DePARLE: Just on the narrow point of the National Quality Forum. I sit on the board of that and others here have been involved in it. It may be true that the development of those hospital criteria did not explicitly talk about pay-for-performance. But I just want to emphasize something that they've said, which is that was a very lengthy process with a lot of stakeholders, and it was a difficult process. I don't think anyone who was involved in that was unmindful that eventually that's where this was going.

That's why, frankly, everyone was at the table, duking it out, making sure that we could move in one direction. There was a lot of concern about having multiple different groups coming up with all these requirements. But I don't think anyone would say, oh, I'm shocked that someone might think these could be used for that.

MR. HACKBARTH: Thanks. That's helpful.

MS. BURKE: Following in Bob's analogy of the day, the airport analogy, if I think of the hospital as the airplane, I'd like to talk for just a minute about the pilot.

In the course of this discussion, which was spectacular and I understand that the purpose here was to inform us about what was occurring specifically with respect to hospitals and measures, I wondered at what point we ought to also again opine on the importance of the relationship between the physician and the hospital in terms of the achievement of these activities. There is a reference in the text, in some cases the hospitals are concerned that they have little control in some respects because of the role of the physician.

I was questioning Nancy-Ann because I had this in the back of my mind and couldn't recall specifically. There was the activity in New Jersey, which has subsequently been halted as a result of lawsuits, relating to the hospital's capacity to share in the benefits as a result, with the physicians. I wondered if in the course of this there isn't

reason to talk about, more directly, the things that would have to occur, or whether it makes sense to begin to look at to what extent physicians play a role in any of these outcome measures, whether they are the process measures or whatever they happen to be, and how one might begin to think about the relationship and how we would provide the opportunity for that to occur.

I don't mean to get into the middle of a lawsuit in this case where the demonstration has been halted, but I think we ought to, in the course of talking about this, continue to reference the importance of that linkage and how over time it ought to be a system, just as we talked about yesterday, the need to particularly tie the A and B side so in fact the rewards are linked, and that there is a benefit to both hospital, but as well there is opportunity to influence the physician or engage them in these outcomes. In the course of this it seems to me, talking about what we know or the importance of that linkage more directly, may make a lot of sense.

MS. MILGATE: I am anticipating in the discussion we're supposed to have next month on physician pay-for-performance that we'll have a discussion of that, but it certainly could be included in this one as while.

MS. BURKE: I think we need to create the linkage. They are, at the moment, distinct in a sense, but in fact we do reference it briefly in the text about the concern that without that linkage hospitals are somewhat impeded in terms of what they can achieve. Similarly, physicians' success will be tied in part to the hospital also putting in place the systems that allow the physicians to succeed. So they have to be supportive systems. And the measures are clearly linked. But I think we ought to look at it in both cases so they don't continue to appear to be distinct activities because they're really not.

MR. HACKBARTH: I have three more on the list and then we're going to have to move on.

DR. WOLTER: Just a few thoughts. Karen, your verbal summary of the presentation I thought was quite nice and I assume we'll see some version of that in writing when the final chapter comes out. But in your comments on the process measures I'm glad to see how our thinking about that is evolving, because those are, right now at least, a very important way to tackle quality if they are evidence based. They have one other virtue, there a little bit less dependent on volume than some other things. So I think even organizations with lower volumes who are working on those process measures contribute to quality. Also in and of themselves, to measure them you have to put system approaches to care in place. So in a way it incents the behaviors and the changes organizationally that need to happen.

I'm going to pick up on a couple other things, just say them a different way. One of the things I believe we need to do is be somewhat bold in our tone about this issue, and to push pretty hard, that this work needs to move very quickly. I think we need to be very specific that we will be moving very quickly beyond 10 measures. I think also if we link this, at least conceptually, to the episode profiling that we discussed yesterday, if those two initiatives were pushed, at least in parallel, aggressively I think it would create huge beneficial changes in how health care is delivered in this country, and perhaps push changes in how we are organized and how we deliver health care. So I see a linkage there that we might want to explore. So some sense of urgency -- and I know we talked about that last year as a commission -- on this particular topic.

The issue of trying to create coalescence of the different groups creating these measures is important, not only so that it's easier to be doing one set of things and not different sets for different groups. But there's another issue. If these measures will continue to evolve, which I believe they will, we need a place where experts can help that evolution and make decisions fairly quickly. When is ACE inhibitor no longer a measure? When is ventilator bundle measurement a new measure? If those things are very fragmented it's going to be hard for this whole effort to move along as efficiently as it might.

So I don't know what that means, but should we as a commission be looking at recommendations about the process of oversight that ultimately should come into place so that we can really push these initiatives very hard, but also have an organized way to get those changes adjudicated as we move forward, and as Arnie said, to make sure that the right stakeholders are a part of that conversation. I think that could be an important contribution as well.

DR. WAKEFIELD: Just a request. You give it a nod in a few places, and to just ask you as you keep doing your work that you try to call out, when it makes sense, some special attention to small rural facilities because of some of the unique circumstances that they face. Most of the folks I speak with have no interest in standing outside of the work that's underway in quality. That is, critical access hospitals, small nursing homes, et cetera. As a matter of fact I think a lot of them would feel that that would do them ultimately a disservice. That is if they're not part of this, and reporting and providing information to the public that sends very much the wrong message.

So how we reconcile that when you're talking about payment and given the way some of our payment structures currently exist is a challenge. It's also a challenge because of low volume. So where there are measures that are

tied to volume, if we get a lot of empty cells on the reporting, that's problematic too. If there are a lot of asterisks there, there's plenty of concern in the field about the message that inadvertently sends. If you can't report, then what's going on in this facility? So those are some of the challenges.

I guess some of us who work in this field are really looking for breakthroughs, hoping that it comes from the person sitting next to us because we don't have the immediate answers. But if there are ways of reporting, rolling up data, aggregating information that would allow more data to be put into those cells that otherwise would remain blank -- there's got to be some additional thinking and hopefully some breakthrough that occurs on that front. So that's one issue, engaging everyone in this, trying to find ways to engage everyone and making sure that we're collecting data on areas that make sense for those small facilities.

Some of us have been working on patient safety issues and patient safety practices in rural hospitals now, I've been part of an initiative for over a year, and clearly there's a lot of good overlap on areas of focus between what's coming out of -- associated with urban hospitals and what seems to be quite relevant for rural hospitals. But there is also some variation around the edges in what I think are fairly important ways. You mentioned them; i.e., issues of transfers and referrals and patient stabilization. So to really try to track on those areas that might make the most sense, especially to the smallest facility, we need to be every bit as concerned about quality there as we are, obviously, in those large facilities.

The last point I want to make, just to stay with Bob's, and hopefully I'll be the last one that makes this comment, but to stay with this aviation analogy and extend Sheila's remarks. You can fly someplace in a 757. You can fly someplace in a Supercub. Some of us prefer to be in Supercubs over being herded into 757s. But the point of it is that the structures and the processes are a little bit the same. Both of those planes are trying to accomplish the very same outcome, but the way they're configured is a little bit different, how you move the controls is a little bit different, who's flying them and on and on and on.

So the point of it is to say, even if you're applying something like an intensivist standard to intensive care units -- UNC is doing some really good work right now and they'll be able to report it pretty quickly about what intensive care units look like in rural hospitals. And I can tell you, being an old intensive care unit nurse, when I worked urban hospitals there wasn't a huge difference between one critical care unit to another; same equipment, et cetera. You look at intensive care units as defined in

rural hospitals, it can be anything from one monitored bed at the end of a med-surg wing to a free-standing, patients on ventilators, full bore wraparound sets of services.

So we've got to be thinking about what we're applying those measures to. Standardization is absolutely critical, but there's going to be a little variation on the theme that we'll want to be sensitive to as well. That's going to be hard but hopefully we can pay a little bit of attention to that.

MR. DURENBERGER: Mr. Chairman, I'm back with your comments about your heart rate. I think it's probably an appropriate analogy, but I'm going to start testing my pulse from now on when I read this material, because I agree with everybody else that just in the two years plus that I've been on MedPAC, the strides that we've been making are really tremendous.

One comment about a reference that was made here, just by way of an observation from our part of the world, and then I have a question.

The reference is made at the end of the material to Health Partners and their decision to deny patient for hospital care resulting in serious avoidable events. That came about quality only because Mary Brainerd, the CEO, was also chair of several patient safety commissions and committees and things like that. And in some setting she made a comment about why should we be paying for seriously avoidable events. All of a sudden that got into the newspaper. Of course, all hell broke loose because of all of her colleagues in the hospital business, who work quietly behind-the-scenes reporting all of their serious events for the first time in history, got very upset with her.

So that's sort of like a comment to clarify, this was not a really deliberate strategy on the part of some payer to take us to the next level. It was just the logic - I remember my wife and I sitting at 5:30 in the morning, reading our St. Paul paper, and we said gee whiz, why do we pay for errors?

And then the next question is why do bill for errors? Which is a question I asked of the head of the Minnesota Hospital Association, how many of your hospitals are actually billing for errors, how many doctors bill, et cetera.

That's our little background from Lake Wobegon.

My question is sort of like Bill's on something I think Sheila and Nick were talking about. And that is developing measures of labor productivity. I know it isn't quite right on point of the outcomes approach and things like that. I guess the hospital people here can speak better to this than I. But changes in clinical care processes are so critical to achieving the quality goals and the performance goals but they are also a great value to the

organization in enhancing the productivity, the efficiency and so forth of the process of delivering care.

It would just strike me that it would be worthwhile, as we develop our work on this, and I'll leave it to others to comment, to encourage labor productivity and to recognize the various ways in which people are taking on the connection between quality, clinical care, operations, satisfaction within the organization, as well as satisfaction from those who are the beneficiaries of the organization.

DR. MILSTEIN: I wanted to follow up on the question that Alan asked. Alan said has anyone priced out what supplying the information from the medical record would cost if we wanted a better set of performance measures.

If you look at the recommendation from the Quality Work Group of the National Committee on Vital Health Statistics and let's say order of magnitude, what are they suggesting CMS and other payers require as a condition of payment going forward? You sort of say about how much would it cost a hospital if it got into production mode to routinely collect those data elements and report them on a hospital bill?

That's not been priced out specifically, but a very, very similar market basket of about the same magnitude and estimated workload has been up and collected in Pennsylvania routinely as a condition of payment for about the last five to 10 years. And that has been estimated at about \$18 per hospital discharge for a very substantial improvement in our ability to measure not just processes but risk-adjusted outcomes.

DR. NELSON: And that's down from \$34. They got that started in late '80s or something like that. And the first couple of years it was \$34 an abstract. So it's come down substantially then, if you take into account inflation.

MR. HACKBARTH: Okay, thank you. Well done.