MEDICARE PAYMENT ADVISORY COMMISSION

PUBLIC MEETING

The Horizon Ballroom
Ronald Reagan Building
International Trade Center
1300 Pennsylvania Avenue, N.W.
Washington, D.C.

Thursday, March 9, 2006
10:03 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.
NANCY-ANN DePARLE
JENNIE CHIN HANSEN
NANCY KANE, D.B.A.
ARNOLD MILSTEIN, M.D.
RALPH W. MULLER
ALAN R. NELSON, M.D.
WILLIAM J. SCANLON, Ph.D.
DAVID A. SMITH
RAY E. STOWERS, D.O.
NICHOLAS J. WOLTER, M.D.

AGENDA PAGE

P R O C E E D I N G S

MR. HACKBARTH: Welcome to our guests. We're beginning this morning with three sessions related to the broad topic of private plans serving Medicare beneficiaries, beginning with Part D plans, then moving onto Medicare Advantage and special needs plans.

Rachel, Niall, who's going to lead the way Niall?

* MR. BRENNAN: Thanks, Glenn. Good morning, everyone. We're here this morning to give you a work in progress report on analyses that are underway for a chapter about the prescription drug benefit in our June report to Congress. As you all know, the Part D program began on January 1 and the opening enrollment period continues through May 15. Last November we gave you an initial look at plans being offered under Part D, and now we're back, after having had a closer look at CMS data on plan benefit designs.

So this morning we'll try to give you a sense of what plans are offering and the benefit structure of those plans. We'll also give you some aggregate information about enrollment to date in Part D.

Let me summarize our findings first and then we'll

go through them in some more detail throughout the rest of the presentation. As we told you last December, despite initial fears that no more organizations would be willing to provide stand-alone drug coverage, there's been a significant amount of plan entry into Part D. Over 1,400 stand-alone PDPs are available across the 34 regions, with another 1,300 or so MA-PDs available in certain counties around the country.

Those counts of plans are a little different from what we told you about last November because they exclude plans that are set up for specific employers, plans in U.S. territories, and others such as special needs plans and cost HMOs. We excluded these groups because we wanted to give you a sense of the characteristics of plans that do not have any restrictions on eligibility or enrollment.

One key thing we found is that about 17 organizations account for the vast majority of stand-alone drug plans. In most cases, these organizations are offering the same two or three benefit designs across some or all regions of the country, and they typically use the same formulary.

The premiums of those plans differ by regions

though, typically on the order of $8 to $11 per month, and sometimes cost sharing requirements vary a little bit too. Another key thing we found is that most plans are not using Part D standard benefit design. They're using tiered cost sharing rather than straight coinsurance. This is probably because organizations believe that beneficiaries will want the predictability of fixed dollar copays rather than paying a straight 25 percent coinsurance. Also, many organizations have designed their benefits to avoid the standard benefit's $250 deductible, which again reflects the fact that Medicare beneficiaries generally prefer first-dollar coverage.

Finally, we found that prescriptions drug plans offered by MA organizations are more likely to offer enhanced benefits than stand-alone PDPs. They're also more likely to charge zero or a reduced premium for the drug portion of the benefit. This is because under the MMA, MA-PDs are allowed to apply a portion of the difference between payment rates and the plan's bid, called rebate dollars, towards lowering premiums or adding benefits.

I know you've all seen this chart before so we won't dwell on it for too long. We're showing it again though to remind you of some of the language we've used to

describe the standard benefit and how it's structured. This will help us think about how organizations can vary their benefit offerings yet keep the same actuarial value as the standard benefit. You know the standard benefit includes a $250 deductible, then a range of spending where the enrollee pays 25 percent coinsurance.

The point at which the 25 percent coinsurance ends and the coverage gap begins is called the initial coverage limit, which for 2006 is $2,250 in total drug spending. Then there's another range of catastrophic benefit coverage once an enrollee's out-of-pocket spending reaches $3,600 or total drug spending reaches $5,100.

For the purposes of talking about actuarial equivalents we're going to focus primarily on the lower part of this slide, the white area beneath the coverage gap. This is what many organizations are varying, yet keeping the actuarial value of their benefits the same as the standard benefit.

So what can an organization offer? Of course, they could offer the standard benefit, and some are. But a lot of beneficiaries are used to copays rather than coinsurance, and copays can be an effective way for a plan

to steer its members toward using preferred drugs. So one thing an organization could do is offer a plan that keeps the standard benefits deductible at $250 and its initial coverage limit, but swap tiered copays for the 25 percent coinsurance.

Another thing an organization could do is to charge no deductible but keep the same initial coverage limit and charge somewhat higher cost sharing. Similarly, a plan could have no deductible or a reduced deductible and keep cost sharing at about 25 percent of drug spending, but it would need to lower the initial coverage limit to have the same actuarial value.

All of these variations are called basic benefits. Once an organization has offered at least one basic plan in a region, it may also offer what's called an enhanced benefit. This means that the plan includes both basic and supplemental coverage. That supplemental coverage doesn't necessarily take the form of filling in the coverage gap. It could be anything that raises the actuarial value of the benefit beyond that for the standard benefit.

This next chart illustrates that a relatively small number of organizations account for the majority of

PDP offerings nationwide. The blue portion of the pie chart represents plans offered by the 10 organizations that have at least one plan in all 34 regions, what we call national plans. The yellow portion of the chart represents seven other organizations that offer 20 or more plans in many, but not all, regions; what we call near national plans. And the green portion of the chart represents all other plans.

I'd just note here that these percentages of PDPs are not weighted by enrollment, they're just simple shares of all 1429 PDPs on offer.

We also wanted to give you a sense of what types of plans these organizations are offering, so we've taken the previous pie chart and shown what types of plans are offered by the national, near national and other organizations. So you can see that the 62 percent of PDPs offered by organizations with a national presence is comprised of 34 percent that are actuarial equivalent to the standard benefit and 28 percent that are enhanced plans. So you can see here that no national plans chose to offer the defined standard benefit.

Similarly, for near national plans, you can see that their 27 percent share of the market is comprised of 6

percent standard benefit plans, 10 percent actuarial equivalent plans, and 11 percent enhanced plans.

Overall, 9 percent of plans are the standard defined benefit, 48 percent are actuarially equivalent to the standard, and 43 percent are enhanced.

Among all 1,429 PDPs about 57 percent are basic plans, whereby basic I mean either the standard benefit or plans with the same actuarial value but somewhat different benefit designs. The remainder are enhanced plans that include supplemental benefits. Again, these figures are not weighted by enrollment.

The bar chart of the right shows the distribution of premiums for basic plans in orange and for enhanced plans in gold. You can see that there are some enhanced plans with lower premiums than basic plans, and there's a fairly broad distribution of premiums for both. But you can also see that the mean and median monthly premiums for basic plans are about $10 to $12 lower than those for enhanced plans. Basic plan premiums are in the $30 to $35 range, whereas enhanced plan premiums are generally in the $40 to $50 range.

You may have heard recent reports by CMS that

average premiums for Part D are closer to $25 per month. The reason the premiums in this chart look higher are that they are not yet weighted by enrollment. We don't yet know how many beneficiaries have signed up for each plan. But if people are gravitating towards the lower premium plans then we'll see a lower weighted average premium than you would for the simple distribution we have here.

One interesting thing we found is that the median premiums for the standard benefit package are about $5 lower than for those benefits that are actuarially equivalent to the standard benefit. So among all of those orange bars on the chart, the ones on the left-hand side tend to be the standard benefit and the ones on the right-hand side tend to be actuarially equivalent. We're still looking into this but one explanation may be that actuarially equivalent plans tend to be structured with copays and they have higher premiums because beneficiaries are willing to pay more for the predictability of copays.

Now I'm going to turn it over to Rachel.

DR. SCHMIDT: This map gives you a sense of the geographic variation and average monthly premiums for basic Part D coverage around the 34 PDP regions. Regions that

have the highest average premiums for basic coverage, which are the ones in red, cost about $10 more per month than regions with the lowest premiums, which are the ones in yellow. This is somewhat less variation than we predicted last year where we were conducting a simulation of Part D premiums we developed using drug claims for privately-insured individuals who were also in Medicare. However, most of the specific regions with higher or lower premiums are the ones we expected and they correspond to where drug spending by Medicare beneficiaries tends to be higher or lower than average.

MR. HACKBARTH: Rachel, does this include the actuarially equivalent packages, or is this strictly the statutory?

DR. SCHMIDT: By basic benefits, I mean both the standard one and actuarially equivalent, but not enhanced. So this is excluding plans that have supplemental coverage.

MS. BURKE: But I'm interested in why you believe this pattern exists. You indicated that we had anticipated some of this, but I wondered about the sort of unique characteristics of those states that are on the high side. Is in statewide? Does it tend to be largely urban driven?

It's an interesting pattern, not entirely what I expected. They're not very urban areas as a general matter.

DR. SCHMIDT: I know John has an answered he's used before.

MR. BERTKO: Sheila, we had almost the same pattern show up on the FEHBP geographic stuff, which is the million or so federal retirees, and this was more or less an identical map to usage. It's probably weighted somewhat, like on the West Coast by the prevalence of managed care and the spillover of patterns. Then I think I personally referred to that middle section of the reds and blues as the fried food belt in terms of having a general higher prevalence of drug use. And it's not urban versus rural. It's regional as far as I can tell.

DR. SCHMIDT: Right. Again, it's consistent --

DR. REISCHAUER: In Louisiana they take a lot of prescription drugs.

MS. BURKE: That one didn't surprise me but so of the others did.

DR. REISCHAUER: It would be more interesting I think to look at what the numbers are just for the national plans because then you're washing out a whole lot of very --

and I see how close that is to what we predicted last year. What we predicted last year was much greater variation than has appeared here, but what you see going on here could have to do with the way they designed the particular plans that are not offered across the board everywhere.

DR. SCHMIDT: That's true, and there's kind of a hint of that sort of thing by looking at tables one and two in your mailing materials. You can see the range of premiums for the same benefit offered around -- and it does look fairly wide in some cases.

MS. BURKE: How does this track what we know about Medicare Advantage plans?

DR. SCHMIDT: Thank you for the segue because that leads to my next point. Some of the states in the West have average premiums for basic benefits that are lower than we predicted last year. One reason for this is probably that stand-alone drug plans have to compete with MA drug plans there, more extensively in that part of the country. And MA-PDs are able to buy down their Part D premiums with some of their rebate dollars, as Niall described earlier. So in parts of the country where MA penetration is higher, which includes the West largely, organizations offering stand-

alone drug plans probably felt a lot of competitive pressure to keep their premiums low.

Let me summarize some of the cost sharing requirements that we see among all of the stand-alone drug plans. Nearly 60 percent of all PDPs charge no deductible, and 91 percent of them use tiered cost sharing, typically with three or four tiers, rather than the standard benefit's 25 percent coinsurance. As Niall mentioned, organizations have probably done some market research and found that beneficiaries don't particularly deductibles and want the predictability of fixed dollar copays.

However, it's important to note that even when plans used tiered cost sharing they often use a combination of fixed dollar copays for the lower tiers, the ones that usually cover preferred generics and preferred brand name drugs, along with coinsurance on the higher tiers. Median copays for the lower tiers are what you might expect based on what you see among commercial plans, on the order of $5 for preferred generics, $23 to $29 for preferred brand name drugs, and something on the order of $50 to $55 for non-preferred brand prescriptions.

CMS allows plans to charge higher cost sharing for

specialty drugs and many plans are doing so. They are often using 25 percent to 31 percent coinsurance on a tier that covers biologicals or other higher cost specialty drugs.

Most of the PDPs offer mail-order pharmacy services, which is probably not surprising. As we told you last November, relatively few offer coverage in the coverage gap and such coverage is generally limited to generic drugs.

Now let me switch gears for a minute to talk about Part D's low income subsidy because it has important implications for both beneficiaries and plans. Under the low income subsidy, full duals are eligible for extra help that covers the entire premium for qualifying Part D plans as well as greatly reduced cost sharing and coverage in the coverage gap. Other beneficiaries with incomes of less than 150 percent of the federal poverty level who have limited assets may also qualify for the low income subsidy.

You can see the relevant income and asset levels for 2006 on this slide. Note that the asset test does not count a beneficiary's primary residence and vehicles, but it does count other assets such as the cash surrender value of life insurance policies.

Individuals with incomes of up to 135 percent of

the federal poverty level have their entire premium paid so long as it's with a qualifying Part D plan. Those individuals also only pay nominal copays and get coverage in the gap. People with limited assets and incomes between 135 percent and 150 percent of the federal poverty level get sliding scale premium assistance and reduced cost sharing. So Part D's low income subsidy is a very good deal for those individuals who qualify for it and enroll in a plan.

Beneficiaries who qualify for the low income subsidy may also be attractive enrollees to organizations who are offering Part D plans. One reason is that CMS is auto-enrolling these beneficiaries into qualifying plans, that is, randomly assigning them. They've already done this for full duals who officially lost their Medicaid drug coverage as of January 1, and may do so for other people who are enrolled in the low income subsidy this spring if those individuals have not yet picked a plan themselves.

As we've been seeing, there can be problems in transmitting data for auto enrollment. It's extremely important to transmit eligibility and enrollment information quickly and accurately to all the parties involved. And beneficiaries who are auto-enrolled are permitted to change

plans, so the timing of those changes and ensuring that CMS, beneficiaries, plans and pharmacies all know about those changes is very important.

Nevertheless, auto enrollment has some desirable features, particular for a population that can be hard to reach through a traditional enrollment process. Auto-enrollment is also desirable from a plan standpoint because it helps them avoid some marketing costs and helps assure them of a steady payment for premiums by Medicare. CMS also uses a special risk adjustment factor for plan payments on behalf of low income beneficiaries to provide more incentives to enroll them.

However, not every plan may qualify for auto-enrollees. In order to qualify, PDPs need to have premiums that are at or below certain regional threshold values that are calculated by CMS. Those threshold values are based on average bids from both PDPs and MA-PDs, but they are also at least as high as the lowest PDP premium in a region. In other words, the low income subsidy thresholds are designed to make sure that the eligible beneficiaries have access to at least one PDP.

In 2006, 29 percent of all PDPs qualified for

auto-enrollees, which is a total of 409 plans. There are at least six PDPs that qualify for auto-enrollees in each PDP region. Again, CMS assigns individuals randomly among those qualifying plans and enrollees are allowed to switch up to once a month if they prefer a different plan. If a beneficiary switches to another plan that does not qualify they must pay any difference in the premium between the plan they picked and the low income subsidy amount.

The PDPs that qualify for auto-enrollees are more likely to use Part D standard benefit design than the plans that didn't qualify. This might be cause for concern except, remember, that the low income subsidy covers much of the enrollee's cost sharing. So the really important question is whether the formularies of the plans that qualify for auto-enrollees are somehow different from the ones that did not qualify. We're taking a look at plan formularies in more detail and we'll be back to you in April with hopefully an answer to that question.

It's also important to note that since CMS will recalculate low income subsidy threshold amounts each year there could be likely turnover among plans that qualify for auto-enrollees. Using an auto-enrollment process again next

year means that some of the transition problems that we've seen this year could recur in future years.

Now let's move on to describe the Medicare Advantage prescription drug plans. This chart shows the distribution of drug plans offered by MA plans that are local HMOs, local PPOs, private fee-for-service plans and regional PPOs. Again, these are not weighted by enrollment. This shows a simple distribution of the plans.

As you can see a larger portion of the MA-PDs offer enhanced benefits than was the case for the stand-alone PDPs, 64 percent here versus 43 percent shown in the pie chart a few slides back. Also, a large proportion of MA-PDs charge no premium for the prescription drug portion of their benefit. Now to be fair, one should really look at the entire premium that a beneficiary would have to pay to join an MA plan. You can't just join the prescription drug part of an MA plan. Still it's clear that MA-PDs have used a portion of their rebate dollars to enhance their prescription drug plans or to buy down Part D premiums.

MA-PDs are also more likely than PDPs to charge no deductible in their benefit structure. About 80 percent of the MA-PDs are doing so versus 58 percent of PDPs. Like the

stand-alone plans, they tend to use tiered cost sharing instead of 25 percent coinsurance, and the MA-PDs are also using a combination of fixed dollar copays and coinsurance for higher tiers.

The levels of copays and coinsurance are comparable to those used by PDPs. MA-PDs are somewhat more likely to use a four-tiered cost sharing structure than PDPs. Again virtually all MA-PDs are offering mail-order pharmacy services. About 28 percent of them are coverage in the coverage gap, which is higher than what we observed among the stand-alone plans, 15 percent.

As Niall said, we do not yet have information about enrollment in individual Part D plans to learn about which types of plans are more attractive to beneficiaries. However, let me give you a sense of aggregate enrollment in Part D. CMS has released figures as of the middle of February, which are bit dated now, and at that time they said that about 25 out of 43 million beneficiaries have drug coverage either through Part D plans, through TriCare and FEHBP, or through former employers that are getting Part D retirees drug subsidy. CMS has talked about signing up 28 million to 30 million people in Part D's first year. That's

lower than the initial projections of enrollment by CBO and OACT but roughly in line with expectations that have been set by some investment research firms.

As you can see, there are groups of beneficiaries who don't necessarily need to make a choice about drug plans. For example, dual eligibles are auto-enrolled into plans, and individuals with retiree drug coverage can usually keep the same coverage they've had. However, there are a couple of groups of Medicare beneficiaries to which Part D plans need to actively market themselves: individuals who have no supplemental coverage to Medicare at all and those who have a Medigap plan since most Medigap policies don't cover prescription drugs. Those two groups made up about 30 percent of all the non-institutionalized Medicare beneficiaries in 2002.

One area where people have expressed concern is enrollment in Part D's low income subsidy. Enrolled in the low income subsidy is made up of two groups: beneficiaries who have Medicaid and Medicare coverage, so that includes the full duals, SLIMBs, QMBs, QI1s, and other individuals who have low income but no Medicaid. Enrollment in this latter group has been difficult. The Social Security

Administration announced in January it had determined that about 1.4 million such people qualified for the low income subsidy, after receiving about 4 million applications. Initial projections suggested that there could be as many as 3 million to 4 million enrollees in this category.

In conversations with beneficiary assistance groups we've learned that many of these individuals share some of the same characteristics as duals in terms of the socioeconomic problems that are associated with low income, and they can be hard to reach. Anecdotally, we've also heard that although the SSA has tried to streamline the process, documenting one's eligibility can be difficult. We've heard some speculation that, for example, some beneficiaries have a hard time finding their life insurance policies and figuring out the cash surrender value of those.

So there may be several reasons that enrollment or take-up of low income subsidy is lower than expected, whether it's simply a difficult-to-reach population, whether it's difficult to establish eligibility, or both.

We'll be back to you in April with more work on Part D. As Joan told you in January, she'll present findings from a nationwide survey we're sponsoring, as well

as focus groups and structured interviews we're conducting to find out how Medicare beneficiaries are gathering information about Part D.

As I alluded to earlier, we've also got an analysis of Part D formularies underway. We're taking a look at whether plans tend to use open versus closed formularies, the systems of therapeutic classes they're using, where particular categories of drugs are placed on cost sharing tiers, and to what extent plans are using management tools such as prior authorization and step therapy. We also hope to address the issue of whether plans that qualify for auto-enrollees have different formularies from ones that do not.

In addition, we will present any information we obtain about enrollment in specific plans as those data become available. That will help us understand whether beneficiaries or gravitating toward lower premium plans or plans that are offered by organizations that have broad name recognition or that sort of thing. Seeing plans that are the most popular for 2006 will help us to think about how the Part D program might look for next year when Part D subsidies begin to reflect bids that are weighted by plan

enrollment.

Now we're happy to take your questions.

MS. HANSEN: I wonder if in the next study that we have that we'll have a little bit more detail about the hard-to-reach, low income subsidy group in greater detail, whether more is forthcoming?

DR. SCHMIDT: In terms of our next steps you mean? There may be a bit of information perhaps associated with some of Joan's work in terms of some of the focus groups, we may be able to get a bit of information. We're discussing having some discussions with beneficiary advocates and that sort of thing. But I'm not sure that we'll address it as directly as you might like. We will be looking at the formularies, again, of the plans to take a look at that. But in terms of steps for trying to get at the hard-to-reach population, we will do our best but I'm not sure that it's maybe right as much on point as you might like.

MR. BRENNAN: But I do think that CMS is taking some steps on try to focus on alternative outreach strategies for that population.

MS. HANSEN: I realize it's not really our jurisdiction, per se, but I wonder if we could just have

that as, whatever the CMS side is doing, just to give the context to this enrollment Roman population. Thank you.

DR. SCANLON: Thank you very much. I found this incredibly helpful in terms of understanding where we are with Part D. I had spent time both last year and the beginning of this year working with a coalition of associations of people with chronic diseases about how they can help their members choose a drug plan. To be frank, I think what we came to was the idea that all you can do is slog through an incredible amount of information, try to be systematic about what you are seeing and make comparisons, but that is an incredibly laborious effort.

What emerged for me out of this was the idea that maybe there are some thing that can be done in terms of helping beneficiaries choose among plans. In particular the chart that you had which showed the distribution of premiums for basic and enhanced plans was kind of the trigger for this. I recognize this as a national chart and so anybody in a single region is facing a different distribution but that within a region there's still going to be a fairly wide distribution.

Maybe it's my economics background, my instinct

when I looked at a distribution of traditional plans would be, why would I look at the high-cost ones? This is a market. If they're all offering something comparable, why not get the best deal? But yet we see that there are enhanced plans out there. Just telling beneficiaries which are the enhanced plans so that they might actually consider, I'll look at some of the more expensive ones because maybe there's something more there that I'm going to get that is worth it for me. Now the problem is that enhanced plans can be enhanced in a wide range of ways and so it doesn't tell you a lot, but it tells you something.

The other thing is the issue of a basic benefit package versus the standard versus an actuarial equivalent. Just knowing that also helps you sort out when you're facing a choice of 40 plans.

So I think what we might want to consider is what we could encourage CMS to do in terms of providing additional types of information to beneficiaries than they have to date to help them narrow the choices to begin with and then do their shopping, because right now it's a very, very difficult task.

MR. HACKBARTH: I guess if I were looking at this,

if you had a graph like this for the plans in your market I'd say, one thing I want to do is look at those enhanced plans in the lower end of the distribution, there may be a particularly good value there. It's a starting point for your analysis.

DR. NELSON: A similar question. Are there differences in common between the actuarial equivalent plans and the standard plans? Among those actuarial equivalent plans are there certain characteristics that they all have in common?

DR. SCHMIDT: They tend to use tiered cost sharing. That's the primary characteristic they have. A large proportion of them have no deductible as well. So those seem to be the key features. Remember, the standard benefit has this $250 deductible, 25 percent coinsurance. Many of them use the same initial coverage limits, so they must be varying cost sharing by tiers in order to get at the same 25 percent equivalent.

DR. NELSON: I take it there aren't huge differences in the number of alternative drugs on the formulary and stuff like that?

DR. SCHMIDT: Again, our formulary work is

underway so that's another key dimension, you're right, that isn't really reflected in the data that you're seeing today. So have to ask you to told your question for next month.

MR. MULLER: Again I commend you for this very helpful work. If I can take you to slide 15. I have two questions. On slide 15, how many of those categories actually had to take a step to enroll as opposed to -- I understand the duals or auto-enrolled, but is that just the first and the third? Which of these 25 million actually had to take a specific action to enroll?

DR. SCHMIDT: Certainly the first are the ones that are for open enrollment. Some of those in MA-PDs also elected to go into MA plans. Some of that may have been MA enrollees in prior years and decided to stay with their plan. But yes, it's primary those first and third categories.

MR. MULLER: Do you have any sense then of the 18 percent still to go whether -- you probably would have the same take-up rate since some of these came more automatically. Is there some sense yet of how many of the 18 to go -- I'm taking the difference between 25 and 43.

DR. SCHMIDT: I hesitate to speculate. I know

there may be a last-minute rush to enroll depending on people's perceptions about whether the May 15 cutoff is going to stay around and their knowledge of the late enrollment penalty and that sort of thing. So I think it's difficult to speculate.

MR. MULLER: The second part is, remind me again in terms of the payments that CMS makes to the MA -- you pointed out that in the drug coverage they use some of the rebate to help write down some of the coverage. Remind me the payment that the MA plans get.

DR. SCHMIDT: You mean that enables them to buy down?

MR. MULLER: Yes.

DR. SCHMIDT: So now they're bidding, as of this year, on the package of A and B services, Medicare A and B services. There is a payment rate that's established in their operating area. They get to keep 75 -- not keep you, but use 75 percent of the difference between the plan's bid and that payment rate towards buying down Part D premiums, Part B premiums.

MR. MULLER: That was very clear the chapter. What's the payment rate though?

DR. SCHMIDT: What is the payment rate?

MR. MULLER: Where they're taking the 75 and 25 off?

MR. HACKBARTH: The benchmark.

DR. MILLER: Which we've discussed on -- right.

MS. DePARLE: This may be for Scott and it may be something just we don't know yet but is there any sense of how much the 25 percent amounts to in terms of money that's going back to the Medicare program from this payment calculation we've just been discussing?

DR. SCHMIDT: I don't think that we can say yet because we don't know enrollments in plans at this point.

DR. MILLER: The arrival of the enrollment data will allow us to do a lot of things to get a better sense of, as you looked at those premiums across the country and proportions of people in plans, but also to determine how much on the MA side we're spending relative to the benchmarks and exactly those kinds of things.

MS. DePARLE: When will we know what plans are going to do for next year? Is that a May kind of thing?

MR. BRENNAN: June.

DR. SCHMIDT: Plan bids are due in June, I think

June 5.

DR. REISCHAUER: I've got a couple of factual questions and then a more important consideration. If you're an MA-PD plan and you have enhanced benefits that you charge no premium for, which is the vast majority, do you also have to offer a standard benefit too?

MR. BRENNAN: Yes.

DR. REISCHAUER: So you can get the Cadillac free or the Chevrolet?

MR. BRENNAN: Every plan has to offer a standard benefit in order to be able to offer enhanced benefits but what you said is --

MR. HACKBARTH: It's an anomalous situation where you could either have a Cadillac for free or a Chevrolet for free which would you really like?

DR. SCANLON: But in terms of the accounting of the 75 percent that you have to offer back, you get a Chevrolet plus a dinner at a restaurant or something like that. They have an obligation to return money to you.

MR. HACKBARTH: So you have to use those dollars, if you're not using it for enhancing the drug benefit, you've got to use it for vision care or something else.

Good point.

DR. REISCHAUER: The extent to which these plan offerings have differed from the standard benefit design probably shouldn't be surprising because in two respects they are dimensions that try to appeal to healthy people, number one, where you're going to a reduced or zero deductible, what you're saying is folks who have very small expenditures, we're going to give you something. And the extent to which you go from the 25 percent coinsurance to tiered copayments of -- what you're saying is, people who use expensive drugs are going to pay more than otherwise would be the case. That these are two aspects where within that actuarial equivalence you're shifting the fraction of the benefit that goes to healthy people in the direction of people who are healthy as opposed to those who are sicker. I knew this was going to get your attention, John.

And so the question that I would like us to look at it is how good is the risk adjustment mechanism at offsetting what is otherwise an inherent bias towards attracting well people to these plans? I think we should devote some effort to that.

MR. BERTKO: Can I just respond to a part of Bob's

comment here? The $3,600 out-of-pocket max applies across the board to any plan. So in the case of the 25 percent part, people who are taking very expensive drugs will get to $3,600 rather quickly and then they go down to a 5 percent cost sharing. So they are protected, in the insurance catastrophic sense of the word, across all levels of plans.

DR. REISCHAUER: But the vast majority of people are below this level and what you're doing is redistributing the attractiveness among those people is all I'm saying.

MR. BERTKO: That's a true statement on the surface, but for the people who take the very expensive drugs those will be disproportionately into the catastrophic category.

MR. HACKBARTH: John.

MR. BERTKO: First of all, I wanted to congratulate Niall if this is your work on slide six. The graphics were amazing in terms of trying to put this in there.

MR. HACKBARTH: I think Joan has been passed in the animation derby. You have to pick your game up here to stay with them.

MR. BERTKO: Then I wanted to make at least one

observation. I think, first of all, you've done a good job by characterizing most of the plans as being national or near national plans. And then I'd say one more thing, from press reports, publicly available information on the PDPs only, there is in fact even more clustering. There's a report, I think it was in the L.A. Times, and if you added up the amount of membership in the top five plans -- those are big vendors national and near national -- you come up to perhaps about two-thirds, maybe even three-quarters of all of the category one and three enrollees that Ralph -- actually maybe one, two and three enrollees. Out of the 10.5 million people who are in stand-alone PDPs, not in MA, not other --

So people have voted with their pocketbooks and in fact it's actually maybe simpler to analyze then it seems with the bewildering number of 1,400. It's much more condensed than that.

DR. CROSSON: Just a thought about what we're going to be looking at in April around the formulary stuff, I had two questions. Is it going to be possible, or when would it be possible to correlate enrollment with the benefit design and formulary design, sort of as a triple

analysis? Is that going to be possible to say, we think formulary design is moving enrollment this way or to this degree, and the benefit design in terms of out-of-pocket costs is moving it? Or is that something that's just too complex?

DR. SCHMIDT: In terms of timing of data that we'll be able to obtain, I don't think it will be feasible for the June chapter. The open enrollment period ends May 15. We effectively have these chapters written in May so we're constrained in that manner.

In terms of the general issue of thinking about whether beneficiaries are looking at formularies versus benefit structure itself, I guess it depends on how enrollment works out and the degree to which relatively open formularies correlate differently in terms of cost sharing structure from others, and we don't know the answer to that yet.

DR. CROSSON: The question though, Rachel, is do you think we, say later in the year we would be able to have information like that to analyze?

DR. SCHMIDT: I certainly hope so. This is going to be a many year effort, we hope, in looking at patterns of

why beneficiaries are picking certain plans. We might get a little bit of information from Joan's focus groups, and I know David has been a part of those, to get a sense of whether they are thinking mostly about premiums and benefit structure versus formularies.

DR. CROSSON: The second part of the question is, somewhere in the chapter you talk briefly about restrictions on formulary design. CMS, I guess, has authority to approve or not approve formularies. Are we going to have some information about how that process is going, what criteria are being used, where the thought processes are?

DR. SCHMIDT: We hope to give you an overview in the chapter of what we are at least observing for this year. We've told you in the past about the USP approach, their therapeutic categories and the safe harbor provisions and the coverage of a couple drugs in each of the therapeutic classes. According to what CMS has put out, they don't necessarily follow the USP's therapeutic classes but they do tend to look at that, look to see whether USP categories are being used or not, and the number of drugs covered. We'll try to outline what we know about that process.

DR. MILLER: If I could just say something to

follow up on this. I don't want you to feel like you're hearing a reluctance to do this. The way I organize this problem in my mind is, one, we have to get the enrollment data, and it's happening in real-time. There been some complexities, and when we're going to get that and get it comprehensively is an issue.

Second, I think the notion of enrollment versus benefit design is probably within reach. Again, if we get the data, I'm not sure by June, but the notion of analyzing that, relatively more straightforward.

Then you get to formularies. The way I organize it in my mind, and you should object if this isn't right, the first thing we have to do is figure out, in a sense, a typology to capture what's happening before you can then correlate it with something. I think right there we're all a little nervous about what we're going to see and what we're -- not nervous. Just as analysts, we've never been in the middle of this before, so we're waiting to see that. Then we'll have to figure how do to even describe what's happening to relate it to something.

The other part of it is, depending on how dynamic this is I think that could be a little -- but we should be

able to capture something at a point in time and look at it, I would hope. So the reluctance or the hesitation you heard was no, we don't want to do it. It's, we don't know exactly what we're going to walk into on the formulary front.

DR. SCHMIDT: Yes, let me assure you, we're absolutely very interested in this stuff and we will be following it closely and keep coming back to you with more information as we get it.

MR. HACKBARTH: Thank you. Good job.

Next Scott is going to talk to us about Medicare Advantage plans.

* DR. HARRISON: Today we will quickly review the challenges the Medicare Advantage program has undergone for 2006 and show the resulting plan bidding and availability.

This was the first year plans bid to provide Medicare benefits. Their bids were compared with benchmarks and established by the MMA at the county rates that were previously used to pay plans. I'll go back over those details in just a minute.

New plan types were allowed this year. Regional plans were introduced. They are required to be PPOs and must serve entire regions built up from states. All other

plans are referred to as local plans. In return for the challenge of covering an entire region, the regional PPOs are allowed to have looser networks of providers than the local PPOs.

Another new type of plan is the special needs plan. They may restrict their enrollment to one of three types of beneficiaries: Medicare-Medicaid dual eligibles, beneficiaries living in institutions, and beneficiaries with certain chronic or disabling conditions.

A third big change is the introduction of the Medicare Part D benefit. Beginning in 2006, almost all MA plan sponsors have offer a plan that includes the Part D benefit or an equivalent or enhanced version, and they are paid for the Part D portion of the benefits by Medicare separately from their MA payments just as if they were providing a stand-alone plan.

The stand-alone PDPs represent a new form of competition for the MA plans, that have often provided drug benefits. The PDPs will offer a relatively affordable way for beneficiaries to remain in fee-for-service Medicare and still obtain prescription drug coverage. Rachel and Niall have just given you some idea of the drug offerings of the

MA plans and how they compare with the PDPs.

Bids for 2006 were submitted by plan sponsors last year. There were more than 2,000 bids submitted to provide Medicare coverage to beneficiaries in the plan service areas. Medicare non-drug payment to the plan is based on the plan's bid for the standard Medicare Part A and B benefits, or in other words, all Medicare benefits except for Part D. For this presentation the term did will mean the non-drug bid.

Payments to the plans are determined by the plan's bid and the payment area's benchmark. The benchmarks for 2006, as I said, were the 2005 rates updated by a national growth rate. The plan's bid is compared with a benchmark. Then for those plans that bid higher than the benchmark, the plan is paid the benchmark and the plan enrollees would have to make up the difference with a premium for the basic Medicare benefits.

If the bid is below the benchmark, the plan is paid its bid plus the 75 percent rebate. The plan must pass the rebate along to its members in the form of either reduced cost sharing, a reduction in premiums, or other supplemental benefits. Ninety-five percent of the plans

have bid below the benchmark and thus have rebates to distribute to their members.

Just for example, if a plan faced a benchmark of $1,000 per month and bid $900 per month it would receive its bid of $900 to provide the non-drug benefits plus $75 to rebate to its members in one of a few ways.

We've begun examining the 2006 bid data that has been provided by CMS. Unfortunately, we have not been able to obtain plan-level enrollment data so the analysis of the bids is unweighted. When we get enrollment data we will redo the analysis so that bids can be properly weighted.

For this analysis we divided plans into five groups: local HMOs, local PPOs, private fee-for-service plans, regional PPO plans and the special needs plans. We found that the bids tended to differ by plan type. Other than the special needs plans, the local HMOs were most able to bid below the benchmark and had the largest average rebates. Ninety-eight percent of local HMO bids came in below the benchmark, and when they did, the average rebate was about $80 per month.

Local PPOs were not as likely to beat the benchmark, and even when they were they received

substantially lower rebates than HMOs. Private fee-for-service plans were able to bid below the benchmark in most cases but their average rebates were about half that of HMOs. And regional PPOs had more trouble with the benchmarks, with only 69 percent of their bids being below the benchmarks.

Because the special needs plans target certain subsets of beneficiaries and are affected differently by the risk adjustment system they look different on these measures. Jennifer will discuss these plans in more detail in the next session.

I want to caution you, I've refined this chart from what you saw in your meeting materials so it's a little different, but mostly the same.

We examined the bids to see how the plans used their rebate funds. The bid data divided the rebates into five benefit groups. The plans could use their rebates to lower standard Medicare cost sharing, or to reduce the Part B premium or the Part D premium, or to enhance the drug benefit above the standard Part D benefit, or they could offer other supplemental benefits such as dental or vision coverage. We used the unweighted bids to see where plans

put their rebates. Again will redo the analysis once we get enrollment data.

Preliminarily, we found about two-thirds of the rebates would be used to lower cost sharing on Medicare non-drug benefits. The next largest use of rebates went to cover the supplemental services such as dental or vision services, but reducing the Part D premium and supplementing the Part D benefit also were used substantially to distribute the rebates. Rachel and Niall just showed you that it resulted in MA plans being able to offer lower premiums than the stand-alone PDPs.

Now let's look at how the bidding and rebates have translated into availability for Medicare beneficiaries.

2006 will be a record year for plan availability. Virtually all Medicare beneficiaries will have a Medicare Advantage plan available to them. Even though we saw that regional PPOs were not always able to bid below the benchmarks, they are the most widely available plan type, reaching 88 percent of the Medicare population.

We see here that many of the choices have zero premiums and provide enhanced benefits. Again terminology here, zero premium means no premium in addition to the

standard Part B premium. Zero premium MA plans are available to 84 percent of Medicare beneficiaries in 2006. The most widely available plan type is the zero premium HMO. Although premiums for the private fee-for-service plans and regional PPOs tend not to be as low as the premiums for the local HMOs, about one-third of beneficiaries do have access to zero premium private fee-for-service plans, and a similar share have access to zero premium regional plans.

Not all zero premium plans include Part D coverage but zero premium plans that provide drug coverage are also available. Either because plans are able to effectively manage benefits or because the benchmarks are high enough to support generous benefits, 67 percent of beneficiaries have access to zero premium plans that include Part D benefits with the most common plan type being the HMO, but zero premium private fee-for-service plans with Part D coverage are also available to 25 percent of Medicare beneficiaries. Even more generous, 25 percent of beneficiaries will have access to zero premium plans with Part D that offer some coverage in the coverage gap.

Now for April we will examine some other benefit characteristics. It is often difficult to categorize these

different benefits, but for right now I think we will look at out-of-pocket caps and copayments for a six-day hospital stay. If you have other benefits you'd like me to summarize please let me know.

But first I want to leave you with a final picture of how many MA choices beneficiaries now have. This chart shows the percentage of beneficiaries that have a different number of plan choices. For example, if you look at the bar above one to five plans, you will see that about 8 percent of beneficiaries have between one and five plan choices. We find that virtually all beneficiaries have a choice of two or more MA plans.

If we add the two bars on the left side of the graph we find that only about 10 percent of beneficiaries have five or fewer MA plan choices. If we look at the right side we see that 15 percent of beneficiaries have the opportunity or challenge to choose from over 31 plans. Beneficiaries in Broward County, Florida have the most choices available, 63 plans. Now bear in mind, these plan choices are in addition to the stand-alone PDP choice offerings discussed by Rachel and Niall.

Enrollment data will allow us to further examine

plan bid and availability. We hope to have that data shortly and we'll proceed to look at enrollment growth and look at Medicare payment costs relative to fee-for-service costs once we get that data.

Thank you.

MR. MULLER: This is a variant of the question I asked Niall and Rachel. With 95 percent of the plans bidding below the benchmark, in M+C we had an erosion of beneficiaries because -- they didn't call it benchmarks then but the payments were too low. What's our estimate of where the benchmarks are compared to the old M+C levels? I seem to remember we'd been estimating 7 percent to 10 percent but they may come in lower with these bids. What's our sense where the benchmarks are against the old comparable M+C number?

DR. HARRISON: We don't think the 7 percent number has changed much. However, the risk-adjusted portion, remember there was this budget neutrality policy that's been going on. That added 13 percent to the risk side scores for this year. We don't know what that would be in the future. That tends to vary this year, but for this year they added 13 percent.

DR. MILLER: But isn't the answer to the question, we'll be able to calculate that, where the bids are relative to the benchmark and where the payment rates are relative to fee-for-service once we get the enrollment?

DR. HARRISON: That's correct, we can give a summary figure to that.

DR. MILLER: So I think the way to come back to you is, we are headed toward being able to know that but it does require to know, across the country, how people are enrolled.

MS. DePARLE: When you think you're going to have enrollment data?

DR. MILLER: I really don't know. We're trying to work with the agency now to get it.

MR. HACKBARTH: Ralph's question was simply about the benchmarks, per se, and the starting point.

MR. MULLER: Obviously, with 95 percent bidding below, there's a signal there.

DR. REISCHAUER: I had basically the same question. I was wondering why, forgetting about enrollment, if you look at this chart that is preliminary percent of MA plans bid below, why for local plans you can't compare the

bids now to the AAPCC and see how many of them are below.

MS. BURKE: I had the same question.

DR. REISCHAUER: Does that depend on enrollment?

DR. HARRISON: Local plans don't only serve one county and they'll have enrollment for more than one county so the ratios may be slightly different. It gets a little tricky, but once we get enrollment we'll be able to do that.

MS. BURKE: I was going down the exact same road as Ralph. If we were looking at hospital margins that looked like this, I can imagine what our policy suggestions would be. If we're beginning to see a trend where they're all coming in below the benchmark or largely below the benchmark it does raise questions about the benchmark, I would assume.

DR. MILLER: But we've been pretty clear as a Commission what we think of the benchmarks. We've made that statement. Unless again I'm misunderstanding the question. We think that the benchmarks are set too high and that this is -- your point is this may be additional --

MS. BURKE: Exactly.

MR. MULLER: Your answer on the 13, is that seven plus 13 or 13?

DR. HARRISON: It could very well be seven plus 13 but the seven was a past enrollment-weighted number so that is hard to get at.

MR. BERTKO: Ralph, this 13 that Scott referenced is before coding intensity adjustments so it's less than that, and then it's phased in at a 75 percent level for 2006. So you can't just add them up. It's really 8 percent and then 75 percent of the 8 percent at the top end.

DR. REISCHAUER: Is there going to be a test on this at the end?

MR. HACKBARTH: Once we have the enrollment information we'll be able to do the next analysis which I think Mark started to talk about, which is what is the net effect on federal spending after you take into account that 25 percent of the difference between the bid and the benchmark goes back to the Treasury? So our old numbers of 107 percent, all that stuff is going to be outdated. There's a new dynamic and we need to have new metrics.

MR. MULLER: As pointed out, 75 percent is going into the rebates and all that so even if you start doing some simple math on this and even having the 13, you save 25 percent of the 15 or so, it's four or five -- still a

spending increase of 12 percent or whatever.

DR. CROSSON: I guess I just have to caution against simple numbers, as Ralph said, because in fact I don't think you can just say, the bids were this far below the benchmark, therefore that's an indication of what previous margins were, because there's a lot more that goes into that. For example, I think it's entirely likely that some plans right now are bidding below cost perhaps to get market share.

Secondly, I think some of this may be based on an intention to reduce payments to the provider side, perhaps even below some of the ideas that we have in the fee-for-service environment.

I think the analysis is complicated and I would probably caution against just jumping to a conclusion based on the bidding process in its first iteration.

MR. MULLER: But that may be, the bidding below cost may be the cost of an efficient plan, as we've use that term elsewhere.

MR. SMITH: Jay made the point I wanted to make. I think we need to be careful with 2006 behavior, both by enrollees and by plans. We're going to know a lot more in

May of 2007 than we're going to know in May of 2006. I don't know how many of the 2,700 options will still be around next year but it won't be 2,700. It will be less, precisely because -- partly because people are bidding for market share and partly because people didn't get enough market share even if they were bidding at a price they thought would work.

So I think, Ralph, I share your concern about the benchmark, and, Mark, you're absolutely right, we've been clear about our objections to the benchmark. It is bad no matter what behavior is. But behavior is not only influenced by the bad benchmarking but by marketing considerations as well.

DR. WOLTER: I don't know how possible this will be as we do our further analysis, but the penetration obviously of a rebate is very high, it looks like, across all types of plans. But I wonder if it will be possible to make some assessment eventually about whether or not there's a fair amount of inequity in terms of what benefit is available to beneficiaries in different parts of the country and that sort of thing, because the level of rebate will vary quite a bit depending on the benchmark, et cetera.

With the incredible diversity of what's happening here, obviously there's going to be quite a difference perhaps in what an enhanced benefits are available depending on what part of the country you live in. I think that would be worth tracking, if we could.

One other comment, I think -- I don't know when we would get to it -- the appropriate attention here is being paid to what's available to beneficiaries. I'm hearing a lot of concern in the provider community about not understanding these various plans. Some have to contract very specifically with the full network. Others do not. Many of the hospitals are not really sure what relationship they're in because they haven't necessarily been contacted by all of plans and yet for various reasons they will start seeing patients who are in plans.

A specific area I think where there's a lot of concern is critical access hospitals and how this is all going to play out in that world in terms of interim payment and those sorts of things. I don't even know if that's on our docket in this first wave of analysis but at some point it probably should be a little more --

MR. MULLER: If I can speech to that point. As we

know, a lot of states stepped in and I think it's up to 90 days now versus the first 30, to exactly that point, a lot of the beneficiaries who are coming are still under this broad state -- I don't know what the right term is -- waiver or whatever, transition. So you're basically saying, the state kind of gave a blanket on the stuff and you don't yet know what plans you're actually going to be contracting with and so forth. So I think a lot of this might become much more obvious once the 90-day transition period is over and you actually see who has communicated with whom on April 1.

DR. SCANLON: This in part relates to Jay's and Dave's point about being cautious, but at the same time I guess I would argue that there is a lot to learn from this experience, particularly is we take Nick's suggestion and go below the national level. In terms of our criticisms of the benchmark in the past, a lot of it's been focused on the floors. I think finding out what's happening in floor counties of different types, when you get enough of them together that some of the aberrations in terms of planned strategy and behavior average out. And it may help buttress some of the arguments we've made about the benchmarks in the past by doing the sub-national analysis, because I'm

assuming for like $80 average rebate for HMOs we've got a wide range that may exist. It could be very informative to know about that.

MR. HACKBARTH: One of the more striking numbers to me in Table 4 is the private fee-for-service column, where 93 percent of the bids were below the benchmark and an average rebate of $40. In general, and correct me if I'm wrong, John, we're talking about plans that by design, by definition are offering something very similar to fee-for-service Medicare, they don't have exclusive networks, limited networks, and they're typically, I would think, paying providers at or near the Medicare rate. A growing portion of the country is having the opportunity to join a private alternative to Medicare that's basically mimicking the Medicare program and getting expanded benefits for it. And they're doing that because they're going into areas where the Medicare Advantage rate is higher than the underlying fee-for-service costs; i.e. the floor counties.

So that particular column flashes to me, is this really what we want to encourage?

MR. BERTKO: May I respond to that just briefly? I agree with everything you said with the one add-on that

competing with fee-for-service Medicare in some ways is low hurdle. If there is any care coordination at all that can be put in, and we've got some comments on that, you then begin to have about that much money available to use. So just keep that in mind.

DR. REISCHAUER: Showing that I'm capable of being on both sides of the same issue within a short period of time let me just do some arithmetic that sort of supports Jay's point, which is the HMO column, the average rebate is $80 a month, and that's 75 percent presumably of the difference, per month. That's $1,280 a year would be the difference, which I don't know what the average Medicare beneficiary cost is but if it were $8,000 it's 15 percent, suggesting that the costs are about 15 percent below, which is then about where the Medicare fee-for-service average is. The suggestions that some of us might have laid out on the table that this is a whole lot different from what Medicare fee-for-service is might turn out not to be the case. That supports you.

MR. BERTKO: Just to respond to Bob again on this and to repeat my last comment, that in the HMO markets, which tend to have been squeezed earlier by the BBA, beating

Medicare fee-for-service, particularly in some of the high payment counties, is a low hurdle. There is a lot of use of inappropriate care. If you can have appropriate care, squeezing that 15 percent isn't that difficult.

DR. REISCHAUER: But the question is where should the benefit get from running a more efficient system, and shouldn't you encourage the private sector in a sense if it's doing that and shift the enrollment as a result? And then you can talk about ratcheting it down.

MR. BERTKO: Right, but in fact virtually all of the money goes to one of two places: to beneficiaries or back to Treasury.

MS. BURKE: Or it could stay with Treasury.

MR. BERTKO: But if it stayed with Treasury then you wouldn't have anybody enrolled in it because then you'd have a fee-for-service benefit, because the bid is on fee-for-service level benefits. So if 100 percent of the rebate was returned you'd have a fee-for-service benefit and nobody would enroll.

MS. BURKE: My point wasn't the rebate returning. My point was pricing it right the first time.

MR. BERTKO: I'm responding to that.

MR. HACKBARTH: John, you and I actually agree a lot on this and I specifically agree with your statement that at least in many areas of the country beating fee-for-service Medicare is a low hurdle because of the underlying utilization patterns. It's that very point that makes me question why we then have to have benchmarks that are higher than Medicare fee-for-service.

MR. BERTKO: I wasn't responding to that. But on the HMO side, they are pretty much level. There's no higher or very little higher that I know of.

MR. HACKBARTH: Any others on this? Nick?

DR. WOLTER: This is the discussion we had before. I guess my concern, again, for the rural low cost areas is that to the extent that the floors allow some reinvestment in other things, whether it's chronic disease management or other enhancements, that isn't possible if the benchmark is left at the low fee-for-service rate in those very low cost areas, whereas it's very possible when you can bid against the high fee-for-service rates in the high cost areas.

So the question is, do you try to bring some balance into this discussion and narrow the bell curve rather than stay at the county by county fee-for-service

level? That's been my concern. Then also this inequity issue in terms of what's available to beneficiaries in some parts of the country versus others I think is worthy of discussion.

MR. HACKBARTH: Those are important issues and you present them very well. If in fact though what we want to buy for Medicare beneficiaries in those rural areas is more of those good things, coordinated care and the like, I'm not sure why we shouldn't do that in fee-for-service Medicare as opposed to saying the only way you can get those things is through a private plan which we're supporting through floors. So I think there's an equity issue there in saying that to get this you have to go a certain route.

We've gone over this ground recently so we don't need to rehash it right now. I'm sure we will be back.

So thank you Scott. Well done.

We have one more presentation before lunch and that's on the special needs plans.

You can go ahead whenever you're ready, Jennifer.

* MS. PODULKA: I heard the magic words that I'm the last one before lunch so I will keep that in mind.

Today I'm here to provide you an update on our

examination of the special needs plans and provide some preliminary information from the first three of our four site visits. First I'd like to think Scott Harrison and Sarah Friedman as well as Jim Verdier and Melanie Au of Mathematica Policy Research for their help on this project.

As I told you back in January, SNPs are a new type of Medicare Advantage plan. They're targeted to beneficiaries who are either duly eligible for Medicare and Medicaid, residing in an institution, or chronically ill or disabled. SNPs offer the opportunity to improve the coordination of care for these special beneficiaries, and dual eligibles SNPs, in fact any SNP that covers Medicaid services offers the ability to improve the coordination of Medicare and Medicaid.

When the MMA created SNPs it established few additional requirements for them compared to regular MA plans. SNP must cover drugs plus additional services tailored to their population, and SNPs are allowed to limit their enrollment to their targeted population. SNPs are paid on the same basis as regular MA plans, including the same risk adjustment method to account for differences in expected beneficiary costs. In 2007 payments will be fully

risk adjusted using the CMS HCC model, and risk adjustment generally results in plans being paid more for special needs beneficiaries than for the general Medicare population.

The Commission in the past has expressed a desire to seek out opportunities for delivering high quality coordinated health care for dual eligible and other special needs Medicare beneficiaries. To describe how SNPs are taking advantage of this opportunity we chose to conduct site visits in four locations: Baltimore, Boston, Phoenix and Miami. As a whole, these areas show us SNPs in markets where there are many competing SNPs, there are existing special plans that converted into SNPs, Medicare managed-care enrollees were passively enrolled into dual eligible SNPs, organizations chose to offer multiple dual eligible plans, and there are all three types of SNPs, dual eligible, institutional and chronic care.

SNPs' goals and strategies for the future vary. Some SNPs plan to gain more experience before attempting to significantly increase their enrollment, also their benefit packages, or expand their service areas. Other SNPs are considering expanding their service areas, adding new plans, pursuing partnerships with states, and increasing their

marketing efforts.

SNP organizations can be characterized as falling into one of two groups. First, organizations that have experience providing services to special needs beneficiaries through a Medicare demonstration, Medicaid plan, or similar specialized plan. These organizations view SNPs as a natural extension of their mission.

Secondly are organizations that have experience operating MA plans and view SNPs as an opportunity to expand their selection of products for their members.

SNP relationships with states varied. Some have very close and long-standing established relationship with states while others have none at all. It is important to note that SNPs, even dual eligible SNPs, are not required to contract with states, and in fact CMS does not consider or track which ones do.

In our interviews we found that some dual eligible SNPs receive payment from states to include Medicaid benefits in their benefit package, but many do not. States may have little incentive to partner with SNPs, especially now that prescription drugs are covered under Part D, and about one-third of states have chosen to set their Medicaid

rates at or below 80 percent of the Medicare fee schedule to limit their cost sharing liability.

When we spoke with SNPs that do contract with Medicaid they noted many conflicts between the Medicare and Medicaid rules. They are eager for CMS and states to work to reduce these administrative barriers to better integration of the two programs. However, to date it appears that the bulk of any integration is occurring at the plan level. For example, several plans told us that they had to deal with separate Medicare and Medicaid officials at CMS and that rarely did they find that these two groups know what the other one was doing.

Specific to the coordination of separate Medicare and Medicaid funding streams, some dual eligible SNPs indicated that it was somewhat burdensome but, surprisingly to us, several SNPs told us that it was not a problem at all. SNPs all agree that the accounting requirements had no effect on their clinical care, coordination efforts, or on their relationships with providers.

CMS central office is primarily responsible for reviewing and approving MA plan applications, but because SNPs, especially dual eligible SNPs, are significantly

affected by state and local conditions it may be appropriate for regional offices to have a more active role in this process. SNPs generally said that CMS approved most applications with few changes. However, in contrast, SNPs expressed frustration over CMS's ongoing guidance for the program's rollout.

SNPs have mostly opted for targeted marketing with little emphasis so far on broader efforts. SNPs' approaches to outreach and enrollment differ significantly depending on their target populations and whether they receive passive enrollment.

Of course, individual SNP's marketing strategies varied but generally we heard that dual eligible SNPs had the broadest marketing strategies aimed at physicians, hospitals, community organizations and advocacy groups. Institutional SNPs market primarily to nursing facilities and families of residents. Chronic condition SNPs focused primarily on physicians, other chronic care providers and beneficiary advocacy groups.

SNPs with passive enrollment focus on retaining their current enrollees. You may recall that Medicaid managed care plans that converted into Medicare SNPs were

allowed to passively enroll their members. These beneficiaries then had to choose to either remain in the new dual eligible SNP, switch to another type of MA plan or return to fee-for-service.

Organizations that offer SNPs along with other MA products may be focusing on shifting members into the new product. And if they offer a commercial product line they may also focus on marketing to beneficiaries who are aging in and gaining eligibility for Medicare.

We've heard that the CMS web-based plan finder tool is difficult for the SNPs to take advantage of since their specialized focus and benefits do not fit well into the plan finder format. For example, SNPs who contracted with Medicaid to cover the plan premium, so that in effect beneficiaries were getting a zero premium plan and they weren't paying out-of-pocket, were still required to list the premium amount in the plan finder, so it was indistinguishable from other plans.

Congress must act to extend the SNP authorization beyond 2008. The MMA mandated that CMS report to Congress by 2007 on the impact of SNPs on the cost and quality of services provided to enrollees. However, there may be

limited data available upon which to evaluate SNPs. 2006 data may be muddied by startup issues, including incorrect enrollment data, and plans designed to improve care coordination and quality while reducing unnecessary costs may not exhibit measurable differences within just a year.

SNPs told us that they recognize the importance of quality monitoring to demonstrate that they add value, but several expressed concern that CMS's existing MA quality monitoring and reporting system is not as applicable to their special population. Some SNPs already have additional significant quality monitoring and reporting systems in place, either because they are Medicare demonstration plans in the past or because they have state Medicaid requirements. However, other SNPs do not appear to have any special quality efforts underway at this point.

Based on our very preliminary information we've focused our interest on a few key issues going forward. One is how many eligible beneficiaries will enroll in the SNPs? Will SNPs actually attract new beneficiaries or will they be shifting members from other plans and other product lines? In addition, will more SNPs establish relationships with states, and which Medicaid services will they cover in their

benefit package? Finally, how successful will SNPs be at streamlining conflicting Medicare and Medicaid processes?

As I mentioned we have one more site visit to conduct and we'll be coming back with more information in April, but I appreciate any questions and comments on the state of the work.

MS. HANSEN: Thank you and thank you also for the invitation to participate. I'm sorry that I couldn't attend.

Relative to the key issues, the relationships with states and realizing the complexities of the dual eligible, the two forms of both Medicare and Medicaid. One of the thoughts that I would suggest is, many of the PACE programs throughout the country have dealt with about 20 states already dealing with both the Medicare side and the Medicaid side. That may be just useful as a backdrop perhaps to talk to some of the national PACE association organization staff to learn a little bit about that whole format.

Going back to the other area though of enrollment, the passive enrollment of dual eligibles to some of the plans, that's a little bit different for a Medicare piece because usually on the Medicaid side there is enrollment

that's required by the state. But since there was passive enrollment in this case for the dual eligibles, as I understand, are the beneficiaries really informed about what that process is? I know they have the ability to opt out, but just to understand what this program is fully about.

MS. PODULKA: Unlike auto-enrollment in the prescription drug area, passive enrollment is a little different. These were beneficiaries who had actively opted to join a Medicaid managed care plan and when that exact plan converted and took advantage of the new Medicare SNP opportunity, rather than making those beneficiaries re-enroll in what to them is essentially the same product, it's just offered by a new government now, federal rather than state, they wanted to streamline the process.

DR. SCANLON: On that last point, I guess I had a different impression of passive enrollment. It was that you were in a Medicaid managed care plan, which may not have been a choice because in order to get your Medicaid benefit you had to be in managed care in some states. Then you were transferred --

MS. PODULKA: That's absolute correct. In some instances you may have been assigned at the state level, but

you had been in that plan. So you had at least a year's worth of experience in that setting.

MS. BURKE: But not necessarily getting your Medicare benefits.

MS. PODULKA: Correct.

DR. SCANLON: That's the issue.

Two different points. One actually goes back to Scott's presentation and the number and the table that shows that the special needs plans, 100 percent are below the benchmark, and $130 is the average. I guess what that raised for me was a question of, is there a problem with the risk adjusters here that we haven't fully recognized. That we know how risk adjusters are performing on average, but for the kinds of targeted populations that are being brought into special needs is the predictive value of the risk adjuster as good? I don't know if you've looked at that yet or if we could look at that at some point.

DR. MILLER: Yes, we can look at it and actually this thought has occurred to us in our own conversations, when you see both the growth in the plans, the number of special needs plans that are being offered, this question that we're asking ourselves -- we're not saying that all

plans are engaged in this but there does seem to be some differentiating among populations which then raises that question.

Then finally the benchmark point. All of this has come up in our conversation. In the past, we have looked -- this predates you -- we have looked at this issue of the risk adjuster and how well it does. We have not circled back to it in a year or more but it's certainly something that this is starting to raise the question on. So, yes, we can look at that.

DR. SCANLON: The second point was with respect to the relationship with the states. Since Dave Durenberger isn't here today I'll talk for a second about a conference that he ran three weeks ago in Minnesota about long-term care and the future of long-term care. At least in the upper Midwest there is interest in Medicaid managed long-term care, very strong in both Minnesota and Wisconsin, and the idea of integrating the two through special needs plans I think is something that is worth following for the future. It's not a dimension for the very short-term but as we move out with these plans.

DR. MILSTEIN: Not all providers are likely to be

equally skilled in managing special needs patients and one of the ways in which special needs plans might be able to provide better value and perform better would be by more aggressively narrowing their networks than regular Medicare Advantage plans. When CMS was reviewing applications from special needs plans, were those special needs plans held to the same standard of network-width as regular Medicare Advantage plans or were they given a little bit of leeway in terms of narrowing the network to focus on providers in their communities that were able to demonstrate superior capability or skill in managing special needs patients?

MS. PODULKA: It's my understanding that the special needs plans were still required to meet network adequacy requirements. From speaking to several of the plans, what they actually opted to do was take their existing network, if they had an existing MA plan, and augment their network with additional providers. But one point I'd like to get across about our work is that I'm coming to the conclusion that when you've seen one SNP, you've seen one SNP, and so I don't know how generalizable those findings are.

DR. MILSTEIN: My question was more about how CMS

administered the network access requirement and whether more leeway was given rather than the plan results.

MS. PODULKA: As I said, I believe they are still subject to the same network adequacy. Although they get to tailor services, they still have to fulfill all MA services. So therefore, they need a complete network. But I'll check more with CMS on this.

MR. MULLER: Just following on Bills's question on risk adjustment, how did the cognitive impaired institutional beneficiaries make -- did they make a choice or were they selected against? Do you know anything about that?

MS. PODULKA: It's not something that we've looked at specifically yet.

MR. MULLER: Because they tend to be the higher cost, institutional members, and if they -- I'm just going back to our specialty hospital stuff, if they're the ones that are selected against because you can't figure out how to move them over and you get the payment on the average.

MR. HACKBARTH: Okay, good job, Jennifer.

We'll have a brief public comment period before lunch. Please keep your comments brief.

* MS. WILBUR: I'm Valerie Wilbur. I'm the co-chair of the Special Needs Alliance. We represent about half of the special needs plans that have been approved to date including virtually all of the demonstrations, the Wisconsin partnership, the Minnesota senior health options, Evercare which used to be a demonstration, the social HMOs, the whole gamut. I wanted to just make a couple comments.

First, I wanted to compliment Jennifer on her presentation. I think she did a really nice job summarizing, and was very interested in a number of the comments that have been made around the table. I specifically wanted to address one that was raised about the integration of Medicare and Medicaid through the SNPs.

Our alliance happens to think that the SNP is a great vehicle for doing that. It's important to understand though that although about three-quarters of the SNPs are duals, very few of them are dually decapitated. So most of them just have Medicare capitation, they don't have a Medicaid capitation. Therefore they're only responsible for Medicare and acute care risk, not for long-term risk. So there's only a handful of the SNPs that are dually capitated like the demonstrations that have the ability to really

coordinate the whole package of service and be at risk for it.

So some of the members of our alliance, like the Wisconsin and Minnesota folks, actually have had somewhat different experience with respect to the accounting issues and the separation of the funding streams. It was mentioned that this didn't appear to be a problem. But for plans that historically have been able to take Medicare and Medicaid capitation, put it in one pool so to speak, and then allocate those resources based on individual patient needs that's a challenge that they're facing now. They're able to use their waiver authority to continue doing what they've been doing pretty much, although some things have changed, but they're real concerned about what happens in 2008 when they lose their demonstration authority. So it's really important to think about the dual issues in terms of the funding streams and the capitation.

What we have suggested to CMS who, by the way, has had a great open-door policy in working through some of these issues with us, is a couple of things. One, if you could have -- for the programs that are dually capitated, if you could have an integrated bidding process so that you

could take into account all of the different services that are being covered and the funding that comes from Medicaid as well as Medicare, that would be very helpful.

The other thing is, if you look at the accounting rules and the audit process in particular and take something like care management, where it's not always easy to figure out whether you put your dollars on the Medicare side or the Medicaid side of the ledger, and look at the plan's historical experience. So that if historically they've spent about 60 percent of their resources on Medicare services and maybe 40 percent on Medicaid services, go ahead and use that standard when you're doing the audit process instead of taking each particular care management item and trying to allocate it to one side or the other.

We're talking about some other things too. I'll move on quickly to the second issue I wanted to raise and that has to do with the performance evaluation, which you all know CMS has to report to Congress at the end of 2007 on the SNPs.

We're concerned about the requirement that the SNPs be evaluated for cost effectiveness and quality within this brief period of time because most of them didn't even

come online until January of this year. CMS acknowledges that we actually need to use a different set of performance measures because it's a different population, and if you want to distinguish whether these SNPs are doing a different or better job than regular MA plans, CMS acknowledges you need to use some different measures. But yet we don't have them and they don't have time to put them in place before the evaluation starts.

The other thing is, even if we had the evaluation measures they would need to collect most of the data before the end of the year in order to develop their report and vet it through CMS before it got to the Hill at the end of the year. So there isn't really an adequate time to measure performance and cost effectiveness in the data collection period, especially when they're all starting up and they don't even have all their clinical systems and data systems in place.

The third point about the evaluation is, we are concerned about looking at cost effectiveness in relation to the current bidding process for two reasons. A number of the demonstrations in particular that have the two funding streams have the advantage of having Medicare and Medicaid.

They're under demonstration authority. They can do things a little bit differently.

Also, they still get the frailty adjuster, they still have the full budget neutrality adjuster. So just because they came in under the bid this year doesn't mean that they're going to be able to continue to do those things once the budget neutrality goes away, and if they don't get the frailty adjuster, because that's still an open issue. So we think to try to do cost effectiveness evaluation in this year is premature.

We suggest the following. What if we go ahead and keep that report to Congress at the end of next year but do profiling, I think which is what MedPAC is going to do. There's a tremendous amount of information that could be gained in terms of looking at the plans, what incentivizes them to come into the market, what the benefit packages are, what the character the beneficiaries are. There's a whole series of information that would be very helpful to everyone in understanding this market. Then go ahead and get the new performance measures we need in place, spend this year working on that, collect data for a couple years and maybe have a cost effectiveness and quality report at the end of

2009 or 2010 when we've had time to do it right.

Thank you very much.

MR. HACKBARTH: Thank you. We will reconvene and 1:15.

[Whereupon, at 12:15 p.m., the meeting was recessed, to reconvene at 1:15 p.m., this same day.]

AFTERNOON SESSION [1:23 p.m.]

MR. HACKBARTH: Next up on our agenda is care coordination. Karen.

* MS. MILGATE: Good afternoon.

Improving value in the Medicare program requires a focus on care coordination settings and over time. Currently most efforts to improve quality and decrease costs are focused on individual providers. Yet efficiency at the provider level does not necessarily lead to efficiency at the program or the beneficiary level. If providers do not coordinate across settings or assist beneficiaries in managing their conditions between visits, overall cost of care may be unnecessarily high and the quality low.

While all beneficiaries benefit from efforts to coordinate care, the population most in need of these services is those with multiple chronic conditions.

In this session we present a draft chapter which pulls together all our discussions and research on how Medicare could support care coordination in the fee-for-service program. This chapter identifies the need for care coordination, key tools, and lays out two potential models. We do not anticipate recommendations in this chapter but

hope that the chapter will stimulate further discussion on the topic.

So why is care coordination needed? Because beneficiaries see multiple providers, the opportunity for poor coordination is great in the Medicare program. Also, because of improvements in diagnostic testing and treatment for those with chronic conditions, beneficiaries are living longer with those conditions and that means that the prevalence of those conditions are increasing in the Medicare program. And those with chronic conditions are a high proportion of Medicare expenditures, and we know that evidence continues to mount that many do not receive high quality care.

So why do these probes persist? One of the primary barriers is the payment system. The current fee-for-service payment design focuses on acute illness and injury, not care planning over time, is focused on providing payment directly to individual providers and not looking across patient settings.

It's also face-to-face reimbursement. It doesn't reimburse physicians or others for the care that they may deliver between visits such as education or patient self-

management training.

In addition, given the multitude of services complex patients need, it may not be possible for physicians to do all that they need to in the office visit as it's currently designed and some of the services that these beneficiaries need are not services that physicians have been trained to provide such as patient education.

And finally, clinical information systems so key for keeping track across settings are not widely used by the health system.

So our research in the last six months to a year has been to identify key care coordination tools and strategies that Medicare could use to support their use in the fee-for-service program. Our analysis has been based on interviews with those that have care coordination programs and others who have developed tools to coordinate care, as well as those that have measured the success of programs and those at CMS that are working with their various programs to coordinate care. We also have performed several claims analysis to look at patterns of care for those with chronic conditions and to look at the effectiveness of care coordination, looked at the published literature.

We found that there were two key tools for care coordination. The first was a person, often called a care manager and usually a nurse, who would monitor patient progress and educate the patient for self-management.

And then the second key tool was an information system, and there are a variety of uses for information systems. First, the programs would use the information systems to identify the most needy patients. And then the care manager would use the information system to track their progress and share information with physicians or other settings of care that may need it for clinical care.

We found that programs are more effective if the patient's primary physician is involved with the care management program. And we also found through our interviews that most programs are paid on a risk basis. That is not to say that they have any insurance risk, that is risk for the overall health services of the patient. But they do have risk for the cost of the interventions and usually they have to guarantee some level of savings in order to get paid. So because the programs need to show savings, the programs were very careful about who they actually target their services to. And so they often target

complex patients, often those with multiple chronic conditions.

Based on a literature review as well as our interviews, we found that cost savings were difficult to quantify but when cost savings were achieved they often varied, depending on the type of patient, the intervention used, and the time frame used for measurement. We did find, however, that in general both the literature review as well as interviews said that quality did improve on a variety of different process and outcomes measures due to the care coordination programs.

So what is Medicare currently doing to encourage care coordination? There actually are some efforts under way, actually they have been underway for a while, but then there are some new thoughts on how it might be supported in the fee-for-service program. First, the Medicare program has the Medicare Advantage program and there you have a capitated payment. Because the plans are at risk for all of the health services, there are incentives within that program for care coordination.

There's a new type of program, as you heard described this morning, the special needs plan, and one of

those can be targeted at those with chronic conditions. So this is another option in the Medicare program for care coordination.

On the right hand side of this table you see physician pay for performance. And while that's not yet in place officially as a program, many of the clinical measures that are contemplated for that program would improve care for those with chronic conditions. However, it's not really focused on the most complex patients or necessarily expected to improve care across settings.

The two in the middle, the Physician Group Practice demonstration and the Medicare Health Support pilot are other models that are being tested currently and I'm just going to briefly describe them because it's a nice basis to two potential models we're going to describe in a moment.

The Physician Group Practice demonstration is a demonstration where CMS contracts with a group of providers and it could be a group practice or it could also include a larger system which might have a hospital in it. That group of providers takes responsibly for coordinating the care of their patients. If, as a result of their care coordination

activities they achieving savings for the program, the group of providers can share in those savings.

The Medicare Health Support pilot was mandated in the MMA, and in that model CMS contracts with organizations whose sole focus is care management. They don't necessarily have any formal affiliations with providers. In that model, CMS identifies beneficiaries with certain chronic conditions for whom the program will be responsible. They are paid an up front per member/per month fee, but if they don't achieve savings for the population they have to pay some or all of that fee back to the program.

For the rest of the presentation we will describe two potential models that draw from those two in between models, and then we seek your input on the design features we describe in the two potential models.

MS. BOCCUTI: So for these potential future models, the first of these we'll call the provider-based organization model. In this model, providers are really large enough to be able to maintain their own care coordination programs. Specifically, group practices and integrated health systems have the infrastructure needed to employ the nurse case managers and other staff and purchase

information technologies.

These main components of care coordination, therefore, would be housed within the provider organization.

Payments for the care coordination program could be at risk or in the shared savings model to the provider organization. But smaller fees for the physician activities related to care coordination could be paid to the group or the health system.

In the second model, we examine ways for care management organizations to work collaboratively with smaller physician offices. In this model, the same kinds of risk-based payments would be paid to the external care management organization but physicians and nurse practitioners could also receive monthly fees for their interactions with that care management organization. These interactions could include regular communications, referrals and forwarding test results, for example.

So in both these models we also need to discuss ways that patients could designate a personal medical home. This designation would imply an agreement between the patient and the physician that the physician's office would serve as the patient's central source of medical care and

case management.

MS. MILGATE: So here we want to talk just a little more detail about the financial incentives of the two models that Cristina just laid out. Both models assume the care management program, whether it's a group of providers or a stand-alone program, would be paid on an at-risk program. There's really two reasons for this and it primarily comes out of our interviews.

The interviewees said that it was important for the care management programs to have "skin in the game" to ensure cost effective interventions and that they thought that that gave them also the flexibility to design interventions and change interventions as they saw they needed to be more effective and to also perhaps change who they targeted the interventions to.

So we saw, through the pilots, two potential ways of having at-risk care management performed. The first was shared savings, and that was the example that we gave that the Physician Group Practice model is using. Here again there's no up front fee to the organization but they can share in any savings that they generate for the Medicare program or at least they're eligible for those savings.

In the second, you could pay an at-risk care management fee. And again this is based on the Medicare Health Support pilot. And there is a fee that's paid up front to the care management program, but if they don't meet their savings targets they would have to return some or all of those.

Lastly, as Cristina mentioned, to provide incentives for physician involvement and to pay for their time involved with interacting with the program there could also be a fee paid by CMS to physicians for such things as their referrals, entering information into the information system, as well as returning phone calls to the care management program.

We envision that in model one that fee would go to the group of providers and the group would determine how to distribute it further to the individual physicians within that group. And in the second model that the payment would go directly to individual physicians. The physicians would have to have contacts with an organization to provide the services and it would be limited to patients that were eligible for these types of services.

Another question is how eligibility and enrollment

would be determined in both models. Currently, the programs we found that CMS is contracting with both rely on CMS, first of all, identifying beneficiaries that are eligible for the program. However, it's done in a couple of different ways which could apply here as well.

In the Physician Group Practice demonstration basically what CMS does is identify which beneficiaries use that group of providers as their primary home essentially for care. And then the program is really responsible for that overall population. However, underneath that the organization can choose to target their efforts in a much more targeted way. But in the end the savings calculations are done on the whole population.

In the Medicare Health Support pilot, as well as another demonstration, the high cost demonstration, CMS's efforts to identify beneficiaries are more focused on certain complexity level of patients. But even underneath that identification of a population, again the organization can further target their efforts if they so choose.

In addition in the program we see that physicians could identify and refer additional eligible patients in either model, that would be in either the group of providers

or if they were working with an external care management organization. As Cristina said, it would also be important for beneficiaries to designate the physician office in both models as their medical home.

Accountability would be important in both programs. For the care management program accountability for savings is built directly into the risk-based payment mechanisms. It doesn't seem there would need to be any separate mechanism for accountability on the cost factors. We would also expect though that the organizations would report information on quality measures, both process and outcomes, to CMS. There are several different patient experience of care surveys being developed or have been developed for these types of programs, and those could also be used.

In model two we think it would also be useful for physician offices to report on additional clinical quality measures that would be associated with care for these beneficiaries.

MS. BOCCUTI: And then to step away a little bit from those models on this last slide, we've brought up some issues related just to the fee schedule. When we're

thinking about what we were just discussing before, we were looking a little bit more at the non-face-to-face kind of coverage and care coordination activities. But if we look also for chronic care management, we want to think about also valuing the face-to-face time that the patient and the physician are having.

So for care associated with face-to-face visits, current E&M codes technically do cover the care coordination activities but may not adequately account for the needs of the complex patients. That concern is really compounded for practices with high shares of complex patients so that would occur repeatedly.

Two mechanisms that we can discuss within the fee schedule that could address these issues are to first, increase E&M payments for selected codes say for high-level codes or for codes associated with prolonged face-to-face visits.

A second mechanism could be to establish new fee-for-service billing codes for face-to-face time spent specifically with complex patients.

That concludes what we have here today. We can answer questions certainly on this and you may want to

discuss other issues.

DR. NELSON: I think this is really good work and I appreciate where you're going with it.

But I want to urge us to think more broadly about care coordination and go beyond just conceptualizing it as reminding a diabetic patient to measure their blood sugar or a patient with congestive heart failure to weigh themselves every day with the accompanying education that goes with that, and acknowledge that a lot of care coordination that Nick does and that I did involves advising patients on when they should get an imaging study done or when they should see a surgeon, under what circumstances, and matching them up with a surgeon that's best suited to their personality and so forth.

So care coordination is the kind of thing that happens in the diad between the patient and the doctor in the offices every day. We don't want to you lose that. We certainly want to make it better than it is now, but we don't want to lose it.

It seems to me that the two essential features of a care coordination effective program, and not only the information technology that you mentioned, but think more

broadly than just a care manager and think in terms of a care team. It seems to me then that since most care in this country is still conducted in practices of five or less that we need to conceptualize a model that utilizes that and think in terms of one model being a virtual group, which the IOM is talking about, where small practices all decide to get together and invest together in information technology and hiring the monitoring and education capability that is currently being conducted by disease management firms, perhaps having disease management firms contract for that or perhaps hiring the group itself, building that capability.

So let's think beyond just disease management and the way that's conducted now and think about new organizational models.

The third piece of which, besides information technology and a team approach, would be some sort of certification or credentialing -- let's see, that's too strong -- some means of determining that the physicians in their practices have the capability and have established the team capability to carry this out. And perhaps of reward through pay for performance then would go to the physicians, as long as they created that other capability.

DR. CROSSON: I would just like to congratulate both of you, too, for continuing to advance this ball. Every iteration of this is more thoughtful and helpful. We're still working in this netherworld between Medicare fee-for-service and then prepaid Medicare, particularly Medicare prepayment where there's a delivery system organized in the way that you have described it. And obviously among your two models I have an inherent bias towards model one.

Two of the things that made prepayment to delivery systems in the past difficult have been number one, the actual ability to bear risk and manage risk because of not having capital reserves or sophistication or the like. And then another one is just in terms of modeling it is the fact that in fee-for-service you don't really have enrolled patients. You don't have members, as we would say.

It sounds to me like you're getting close to the second one with what you're calling medical homes, so I'd be some more interested in to what extent is that agreement that you talked about between the patient and the medical group or integrated system like a lock-in or not? Because that ties back into the risk piece. Obviously if you're not

in charge of everything you can't really be at risk for it.

To just go back to the first point, have you begun to think at all or model about the amount of financial risk inherent here? And is that likely to be within the ability of target delivery systems to manage?

MS. BOCCUTI: We'll start with the medical home issue that you brought up. I don't know that it would just be limited to model two. Model one, with a group or a system, if there's going to be some sort of designation going on potentially with the beneficiary doing the designating, so you can think of it has a responsibility that the provider has as well as the beneficiary to be seeking that organization first say, or to discuss care management.

Not the lock-in, whether soft or hard, I think we need to discuss that. We haven't really brought that up for the Commission and I don't think we're going to make that decision but perhaps you want to comment on that. I think we can see the pros and cons that Medicare has been dealing with that on other issues, on how hard the lock-in needs to be. It's easier for planning purposes, but it may not be exactly what beneficiaries want.

We should also mention that the ACP, the American College of Physicians, has been working on the advanced medical home model. They have issues that -- and these are related similarly to that. They are now working out all the details too, but they are discussing the issue that that relates to, too.

MS. MILGATE: I just wanted to comment on his second question.

MR. HACKBARTH: If I could just chime in on the lock-in issue. I was making a note so I missed the first part of what you said, Jay, and stop me if I'm wandering off into the wilderness.

As I think about this general area, I've been anxious that we explore non-lock-in models that are in keeping with the basic format and guiding principles of traditional fee-for-service. We have Medicare Advantage for beneficiaries who are willing to choose a more restrictive system and they've got a wide array of options to choose there now, at least in many markets. That's not perfect, but we've got some action over there on the lock-in side.

I think the void right now is traditional fee-for-service Medicare, which one of its guiding principles is no

lock-in. And so that's why I've been thinking we ought to be focusing here principally on no lock-in models.

Now the middle ground is a voluntary designation by the beneficiary of a medical home, some people call it sort of a soft lock-in, where they retain their freedom of choice. But they've made a voluntary decision that I'm going to go to this physician or this organization as my counsel on where to go for my medical care.

MS. MILGATE: Could I add to that? One of the ways that came up wasn't so much as to make it less risky for the organization managing it. It actually came up in a sense of making sure the beneficiary was really committed to the process, because there was a lot of discussion on how important beneficiary commitment to the program was for effectiveness of the program.

So I don't think that even in that context they saw it as a lock-in at all, that they could only go to that physician for care related to that condition or for anything else.

But the concept was really that it was very important for the beneficiary to actually be committed to the program and committed to that physician for it to work

effectively.

DR. CROSSON: So I think I can understand the concept of a moral commitment on the part of the patient. But what I still can't figure out, and maybe I'm just missing it here, is if you're going to say okay, we'd like you to manage these 100 diabetic patients for a year and we're going to pay you in the end or reward you or whatever, based upon the total cost of care for those 100 diabetics. But the 100 diabetics can go anywhere they want for care services, including other physicians, hospitals or entities that aren't part of this operation.

How do you then managed that risk?

MR. HACKBARTH: It depends in part of the nature of the risk. Nick, help me out here but as I understand the Group Practice demo, there's an opportunity to share in savings but I don't recall there being a penalty if the costs are higher than expected. So it's an asymmetrical risk. It's an opportunity that's being taken, as opposed to the sort of insurance risk that Kaiser Permanente bears.

DR. CROSSON: They're not really risks.

MS. MILGATE: You're at risk for the dollars you put in as investment to manage the beneficiaries. That's

what you're at risk for.

DR. REISCHAUER: You're at risk of winning the lottery.

MR. HACKBARTH: Did I describe that correctly, Nick?

DR. WOLTER: Yes, I think though that I would say, in terms of the Physician Group Practice demo, I think they're some real flaws in the financial design which would take a little longer to explain than we have here. But in terms of the question you're asking, we're assigned a panel of patients. Many of them get a good part of their care a couple of hundred miles away. They might have hospitalizations that are outside of our organization.

And so the issue there is what Jay is raising, which is that whatever protocols or evidence-based medicine standards we put in place to reduce variation, don't manage to penetrate the whole population that we're then responsible for. So that's an issue that I think in future designs needs to be addressed.

MR. HACKBARTH: But it is -- you have an opportunity to gain, as opposed to an exposure to loss if projected expenditures are higher -- if actual expenditures

are higher than projected; is that right?

DR. WOLTER: That's theoretically the case.

MR. HACKBARTH: Do you want to elaborate?

DR. WOLTER: There's issues with how it was designed because you have to, first of all, save 2 percent compared to the increase in cost to a comparator population, but you don't share in that 2 percent. You only share in savings beyond the 2 percent. This is over three years.

So if you net good performance against the cost of the interventions and then compare that to what you would have made had you just gone on in standard fee-for-service with regular numbers of admissions, the guarantee in my view -- and we've modeled this -- is we probably will see less reimbursement than if we had not participated in the program.

That's because we have a hospital in our system. I think if you don't have a hospital in your clinic and you reduce admissions, then it's a different -- so that's why I say it's complex.

MR. HACKBARTH: Do you want to address Jay's other question?

MS. MILGATE: I think it's related, I think, to

this discussion. Because you asked about risk and the ability to bear. Just to say again, the risk is generally for the cost of the interventions, not that that's a small risk but it's smaller than costs for all of the health services like insurance risk.

And I think that that's one of the interesting questions as to how many patients need to be eligible for the program for it to be reasonable for either a group of physicians or maybe Alan's virtual group to actually take on the risk of the care management functions themselves. And that, to me, is sort of another analysis of how many patients need to have to feel like it's worth it for you to take on even the risk of the care management costs?

The other thing I'd like to say though, in terms of groups of providers taking this function on, I think there's also kind of a fine line between investment that would be useful for good clinical care anyway and then the extra investment we may be talking about here. In some of the PGP demos, for example, they told us that -- these were practices that were fairly evolved -- that they were planning on doing some of these things anyway. But this just really gave an extra incentive to go beyond where they

were planning on going with their care management functions that they were thinking about expanding into.

I don't know if that gives you enough information. We don't really know what's the right size but I think that's an area for further discussion.

DR. MILLER: Isn't there also one other mechanism, just by way of risk, just to complete the picture that Jay is asking about is in this situation where you have the care management organization, at least the way it's working in the disease management model, those organizations have some risk for their administrative fee. So it's not an insurance benefit risk. All your issue still attach, now do I manage it.

But they have been assigned populations, or at least given populations that they can go after. Then they have some kind of targeting even within those populations that they may use to sort through who is going to go into the program. And what's at risk for them is the administrative fee that they're getting, not the insurance benefit. Is that all correct; guys?

MS. MILGATE: Yes, definitely. And you can see also, in that model, they even have less control over where

beneficiaries go in terms of providers. So they have a very strong incentive to try to work with physicians to the extent possible that's even, I would think, harder for them than it would be for your organization for example.

DR. CROSSON: But the risk to the administrative fee, which is why those organizations are in it in the first place, isn't that in the end a function of how much the care costs for the patients that they're supposed to be managing?

MS. MILGATE: That's right. They are at risk. Basically there's a 5 percent target. Let's just throw out the Medicare Health Support pilot as the example here. CMS and the programs negotiated. I don't know how much power was on the either side. But there was a target that was set at 5 percent of savings.

So CMS hands the organization what they have designated in the particular region as an eligible group of patients and then the organization can target below that. But they are measured on cost savings different of actual costs versus expected costs for that population of 5 percent.

I'm sorry, I probably didn't explain that as clearly as I needed to.

MR. HACKBARTH: Just a clarification on that. The target against which the organization is measured is 5 percent of the total eligible population assigned by CMS?

MS. MILGATE: Right.

MR. HACKBARTH: If they choose to go to a smaller subset, then they've got to proportionately save more.

MS. MILGATE: Exactly but it's still calculated on the whole population. And it's compared to a control population in that region. So they're hoping to actually have some really robust findings from that design.

DR. CROSSON: One last point and then I'll stop. So I can understand, in that model, how the care management organization, the disease management company, which presumably has the ability to influence the care of that patient no matter where the patient is taken care of, I can understand how that might make sense.

It's somewhat diluted if, in fact, you have a model where the care management organization works cooperatively with some physicians but not others. So I don't think it's a perfect play.

But I still have problems understanding how that applies to the first model, where you have a designated

physician group but that physician group or integrated delivery system doesn't actually have the ability to influence all of the care, even though their risk is limited to the care management fee and not insurance risk.

MS. BURKE: Unless I'm missing something, why wouldn't they have the same authority or relationship any other physician group would have in terms of where they admit patients, what they order?

MR. HACKBARTH: The way I would envision this is this is a voluntary program, and so there may be many physician organizations or integrated delivery systems that say looking at how our practice works, how this community's referral patterns work, this is not a business that we want to get into.

But Wennberg and company have found that when you look at Medicare claims data that there are, at least in some places, de facto delivery systems where even though there's no lock-in the referral patterns in the community are such that they are pretty tight in terms of where people get their care. If you go to so-and-so as a primary care physician, the probability that you're going to use this group of specialists and this hospital is pretty high.

DR. CROSSON: It's a long drive from Duluth anywhere else.

MR. HACKBARTH: Right. So in that circumstance, and I don't know how common that circumstance is around the country, it may feel like hey, this is a risk that we can reasonably take on, especially if our risk is properly constrained. We're not talking about being on the hook for the total overage of Medicare expenditures but a much smaller piece.

And all of these variables, whether it's 5 percent or exactly what the characteristics of the Group Practice demo, they're all subject to negotiation and adjustment. There they are continuous variables. There's nothing set in stone about any of them.

MS. BURKE: Can I ask a question so that I understand why there would be a question about this? As I understand the patients that they envision participating these are fairly complex chronically ill patients; correct? Who arguably have a series of comorbidities that are managed.

And the savings arguably comes from either the avoidance of institutional services, the better coordination

of testing and of treatment that would largely be done by either the coordination among a group of specialists who are managing the patient, or through an internist or a primary care physician who essentially is sort of moving this person around.

So to a certain extent, a physician in a physician group has, as we have at least traditionally believed, an enormous amount of control over a number of those decisions. Now they don't have control over the costs in the hospital, per se, that is what the hospital's base is. But they do in terms of the admissions, they do in terms of managing the patient and keeping them out of the hospital, they do in terms of the follow-up care, they do in terms of the testing or the referral patterns that they have with their colleagues.

So I guess, Jay, what I'm trying to understand is why you don't imagine that, even in that not locked in non-traditional HMO system, there isn't some authority? Because Wennberg tells us, in fact, that there is.

DR. CROSSON: And I would agree with you but that's a qualitative argument you're saying. Isn't it likely that those physicians or that group will have a

significant amount of ability to influence where those patients go for their care? and I would agree with you.

And if you pick the places where for geographic reasons or traditional reasons a particular group or integrated practice does, in fact, hold onto all those patients it holds true. But it doesn't necessarily hold true.

MS. BURKE: Sure it does.

DR. CROSSON: No, it doesn't.

MS. BURKE: More than it does that you don't lock a person into an HMO for life. They can walk.

DR. CROSSON: I'm sorry if we're getting off here. But if you're talking about relatively small improvements in costs it doesn't take very many of your 100 diabetic patients to decide to go off to the next state to get some services, which are fully paid by Medicare, which then change the numbers for that group significantly. Five patients, three patients, two patients out of 100 with renal transplants or -- that's a bad example, but other high-cost examples, then change the dynamics.

So the at-risk administrative fee on any given year disappears. But completely out of the control -- and

I'm using a small example but a larger example would hold true, too.

So I'm doing sort of a purist quantitative approach. I don't disagree with you more qualitative approach.

DR. SCANLON: It seems to me that between these two models that the independent care management company is dealing 100 percent with strangers. They've got no control over these people and they're at risk for something for it.

The group practice starts off dealing at least with some people that they know. You may lose some. But when they start off, they're at an advantage relative to the independent care management organization.

The real issue comes down to what Glenn was saying is what's the formula for this risk in terms of how much you're at risk and what you have to accomplish in order to keep the fee? Because the group practice, you've got the advantage of knowing and influencing the care of some of those patients directly. Whereas the care management organization as to work through moral assuasion with every one of the physicians that these people deal with.

MR. HACKBARTH: And I think that the beneficiary

designation may also play a role in here. We don't know to what extent but it seems logical that if the beneficiary is involved in designating this group practice as my medical home that may also alter their tendency to go outside even without a lock-in.

MS. BOCCUTI: Can I say one short small thing? I think also that the second part which is the fee to the physician group, the second fee that's not really at risk I think is related a little bit more to the medical home designation. It's that that physician is getting the payment. So the relationship between the beneficiary and that physician, if they're going to get the monthly payments, rather than the external care organization that may not know the patient as you're saying.

So I see the connection as relating to the second payment stream rather than the first.

DR. KANE: One of the questions I had was what's the time frame? Because to me if it's a one year settling of accountability for the risk of -- your savings, you may actually cost more because you're doing the right thing.

So I was getting confused as to what risks are we talking about? And it's talked about in the form of savings

but it could actually be that you're doing the right thing and it costs more.

So the whole thing of tying it to savings seemed kind of confusing to me, unless you're just going to limit yourself to COPD at the last stages of life before you hit the ER.

MR. HACKBARTH: I think that's an important point and again it's a dimension of this that needs to be worked out. We're talking a very high conceptual level, whether the right duration is a year or longer, I don't know the answer to that.

DR. MILLER: I think that was one of the points that Nick didn't go into in his comments, is that when you look at the effects over the course of this, whether it's one year or three years or five years that you expect to look it. I know in other conversations that's one of the points that Nick has made and he just didn't want to detail it here. But certainly the time frame is one of the issues.

DR. KANE: I think that means maybe we shouldn't just call it cost savings but maybe there should be other parameters by which people get rewarded around hitting protocols or compliance rather than costs.

DR. REISCHAUER: Isn't there a quality threshold you have to meet? I mean, in Nick's thing, and it's three minimum, I think.

DR. KANE: And if you hit it you get paid, whether the costs went up or down.

MR. HACKBARTH: And there are a variety of different ways that you might factor in the quality. But I think it's very important to have that as part of the evaluation process. In the Group Practice demo, not that it's a perfect model but it's one that's developing. As I recall Nick, you have to hit certain quality targets to be eligible for a financial bonus? Is that the way it works?

DR. WOLTER: Actually it's the opposite. You have to hit the financial targets. Once you hit those then a portion of the savings is given to you based on the savings and a portion is given to you based on the quality measures. And the percentage that's based on quality increases over the three years.

MR. HACKBARTH: Okay.

MS. MILGATE: One more comment on the time factor. There clearly is evidence that for some chronic conditions savings can be achieved much more quickly than others. And

so that is, in some ways, the key factor, I think, for CHF. Programs told us over and over again they felt like they could show savings within a year. For diabetes more like five years.

So the two programs we talked to most about are looking at over a three year period but then measuring it each year at least to give some benchmarks to the programs.

MS. HANSEN: First of all, I appreciate the education in the whole aspect of care management in the fee-for-service world because this is something that is relatively new to me. But a couple of things that were raised earlier, one is what's the right number in which to get some impact? And again I just offer the ROI to some of the PACE programs that have clinically complex people with a profile of say eight comorbidities and about eight medications, polypharmacy along with cognitive.

It took like 150 people to really make it work to include kind of the whole case management model. But there are financials that you could get from the National PACE Association that would offer that.

I have three areas. One is quickly the whole aspect of the E&M aspect of coding and the incentives. I

appreciated the chart that you had about the non-face-to-face reimbursement. And now that we're moving more toward the knowledge assessment of complexity and how to judge that rather than strictly the face-to-face, because I'm thinking about how we use that and whether or not there is some coding that goes to that weightedness of complexity. So when people have those many comorbidities, is there either an embellished code or a new code issue we talked about. I don't know quite how to figure that out and other people can do that, but it just seemed to merit that kind of added weight.

The other one is looking at the -- so it has to do with electronic.

The other thing I wanted to just talk about when Alan spoke about let's look at care coordination more broadly, and I would like to kind of stretch beyond the physician, as well, and look at what happens to the beneficiary with the issues that require that.

There are three things that I think about. One are the things that become care transition points. That's when things go wrong and how things happen. So how to build that into a process and then eventually maybe an outcome

measure.

The other one is medication management, which does fall squarely on the physician. But there are patients, I think I may have mentioned before, that get 20 to 25 medications which clearly any commonsense way of looking at it is an issue for quality and potential poor management.

The third one is when people start developing symptoms, you mentioned congestive heart failure is a very easy one to oftentimes catch. But there are other symptoms that we can catch early. And how does care management in any model, whether it is one or two, capture this so that people get the treatment they need quickly to basically avoid that?

So those are just some of the textural issues of look at it, whether it's one or two. It seems like the relationship with the physicians would be stronger just because of the degree of influence. But I know that the disease management companies have been quite effective even with "strangers."

MR. BERTKO: I'm going to go back to what Jenny and Alan said about looking at this more broadly. I was originally thinking of it in terms of the timing of data

availability but I would also say that it could be looked at as an issue of is this really a kind of disease management registry model? Or could it be a broader model, including management of acute conditions such as discharge planning and prevention of readmissions and reductions of ER use.

You either may want to say something about it or say no, this is just a chronic care management one.

I think, and maybe Jay and Nick would agree, that in integrated delivery systems, whether they're groups of providers or whether a plan or group practices you may have some opportunity for this. In a totally fee-for-service one where there is no personal contact, as we heard earlier today, the timing of the data coming in on these people is so late -- days, months, even years -- that you might not be able to do that if they weren't already in a chronic circumstance.

And yet this is fairly important in managing care and keeping people basically out of the hospital to save money.

MS. MILGATE: On your first point, the various programs we talked to some of them, those that particular were centered or at least had in the mix a hospital actually

identified people that they could really do a lot of good for within the hospital and then would actually integrate the care management right into the discharge planning and then on beyond.

So I don't know that that was available to every single person that was discharged but they certainly did target some people like that as well as then we talked to models that were similar to your other examples. So I wouldn't want to limit it to one or the other really. So just make that clear.

MR. BERTKO: I agree, and flesh that out, make it clear. It's almost like model one and model 1A. Or model one applies to two streams or two kinds of patients -- I'll call them acute or acute episodes within chronic -- and then pure chronic.

MS. MILGATE: So just be a little clear about that?

MR. BERTKO: Yes.

MS. MILGATE: In terms of timing of data, I would actually ask Nick this. It's my understanding that there's a fairly quick turnaround on administrative data that's given to these programs so that they can kind of update how

to target their interventions and their progress. As that true?

DR. WOLTER: I think they're trying to be quicker. We were a half-year go live into the program before we saw the base year data, for example.

Just quickly, I kind of agree with the distinction John and Alan were drawing. There's a coordination care that around patients that maybe aren't as deeply complicated, whether that's preventive care or acute care. And that may be something we'd look at slightly differently in terms of payment mechanisms. Maybe it's through the E&M codes. Maybe it's a way to address some of the primary care manpower issues that have come up here. But that is maybe a slightly different bucket.

My comments are more addressed to the more complex chronically ill patients. As many times as we've seen the concentration of Medicare payments into a small group of patients, it still struck me in your paper -- which I agree was very excellent by the way -- that 61 percent of inpatient payments were for three diagnoses or some combination of the three. Which just strikes me as a huge opportunity.

And that's why I feel that, for example, in pay for performance we have a great opportunity to focus our efforts in these early years rather than go the broadly applicable to every physician or every diagnosis. If we're serious about tackling where the high concentration of chronically ill patients and high costs are. And I think that would be a very helpful message to CMS and others, that we should create some focused efforts in pay for performance around these kinds of patients because I think there will be a lot of early successes as opposed to 300 measures that cover plastic surgery and allergy and everything under the sun. That's just a bias that I have.

Also, I think this is an area where there's a tremendous as opportunity, as Alan said, to create incentives that create new organizational approaches to health care delivery, whether that be virtual physician groups or physician groups that now don't work with hospitals, perhaps the eligibility for per member/per month payment for chronic disease management in these conditions is made available only to physicians in hospitals that come together and agree that they're going to tackle these issues.

And maybe that's part of looking down the road 10 or 20 years from now when these kinds of incentives do create models of health care delivery that are more synergistic.

The IT issue is a big one. You emphasize that very nicely in your paper. I think though that the state of the art in IT is very immature. Even for group practices who have their own patients assigned to them in these projects, creating disease registries that get all of your diabetics enrolled is somewhat of a heroic effort. It's amazing what's not currently the state of the art in IT, even when you've made the commitment to put those systems in place.

And then how do you sort out which of your 1,800 CHF patients would be good to enroll in the program? Even when they're your own patients you have to enroll them in a way, you have to seek them out, identify them, get them to participate. That's not a simple task, even in the case of a group practice.

The IT issue I think, in terms of looking at creating linkages between physicians and hospitals, we have the sort of countercurrent things going on right now in

health care where hospitals are prohibited, for the most part, in terms of placement of IT in physician offices because some of the Stark kickback and those kinds of issues.

I know there's a conversation going on about trying to relax those things. But maybe there's a way to create dollar caps or transparency that would allow those kinds of things to happen so the IT portion of this could move more quickly.

So those are my thoughts.

MR. HACKBARTH: I'll go back to Nick's first point and the one made by Alan at the outset. There is, in fact, as we speak, a lot of care coordination that goes on largely uncompensated. And one idea is well, let's develop new codes that identify that and reward it and maybe we'll get more of it.

But looking at this from a budgetary perspective, that means paying for stuff that we now get for free as opposed to other models where if you put the administrative fee at risk you're only paying out the new dollars if, in fact, you get offsetting program savings.

So there's a different -- as we look through these

options we need to be sensitive to the fact that there are very different sorts of budget implications among them.

DR. SCANLON: This relates somewhat to what you were just saying. In looking at the payment mechanisms that you described, the goal of involving physicians is obviously key. Though I guess I'm somewhat in the context of what Glenn just said, I worry about creating kind of an identifiable payment for this involvement.

In the first model, in some ways, I think the bigger question is the risk issue that we talked about. What's going to be the trade-off in terms of how much is being paid versus the risk that the organization is going to take? Because if I pay a group practice to involve their physicians the money gets lost. There are already financial flows within that group. And where these dollars impact is not totally clear to me as the Medicare program as the payer.

So the bigger issue is the money that goes to the group and what the group has been asked to do for it and how accountable it's going to be for that and then whether they can accomplish that.

In the second model, yes, there is no relationship

between the care management organization and the physicians and there potentially needs to be one in terms of the physicians being responsive. But there's a question of whether that's best engendered by Medicare making a payment with a set of requirements for the physician to respond to or it's better to think about changing again the risk/reward relationship with the care management organization and allowing them to make payments to the physician so that they have a direct relationship and they have better control and there's more accountability for their cooperation in terms of the care management that the organization is trying to accomplish.

I think it would be good if we, in some respects, talk about these options here as a range of things, that we don't know which one might be most effective, and that there are, in some ways, pros and cons to different choices within this.

I'm particularly interested in knowing what we will learn from the Medicare Health Support as well as the Group Practice demo in terms of answering some of these questions because it's not obvious how, particularly the model where the care organization -- I've kind of always

thought, physicians, the last thing in the world they're going to want to hear is this care organization is on the phone saying what are you doing for these people that I've enrolled?

So how we're going to be effective in that is something that is very challenging, given that the care coordination, if we could get make it work, would be very beneficial.

MS. HANSEN: Bill, if I could just build on that, before I left in San Francisco that's one of the things we did as the care management provider. We paid, we had a small demo with private physicians in the community. And we paid for their time. They would actually participate on a case-by-case basis for that.

What was more difficult though in this care management, and I don't know how to solve this, is when the care management system or your geriatrician specialty people who perhaps know that a practice should be done differently, there's a best practice in medication treatment but the local physician may not be up on the latest, is how to influence that level of practice because of the sensitivities involved from physician to physician to bring

it to another level of quality. That we never solved easily.

DR. MILSTEIN: Some of these comments at this point build on prior comments. Maybe they can be thought of as a reinforcement.

First, while it's clear that any form of care coordination that doesn't provide for physician input is doomed, that said if one of our collateral goals in making any program change is not to stimulate innovation in health care, particularly in the methods of health care delivery, I think we're in trouble. So with that in mind, I just wanted to really endorse this idea of widening eligibility for what might constitute the primary medical home or the care management organization beyond the range of organizations we've cited so far.

Just to give an example, community pharmacists. We have some very nice examples in the private sector now, the Asheville experiment being one, in which other categories of health care personnel have been shown to be very successful in being the lead primary manager -- primary care coordinator.

I'm not sure we need to limit the list to them.

I'm thinking medical social workers probably. When you think about care coordination, it's something quite above and beyond anything I was ever taught in medical school.

The second point is a reinforcement about this new code for face-to-face coordination. I think again, if our interest is in stimulating innovation, I think acknowledging Glenn's point about paying for things that we're not currently paying for, that said I think by not paying for these things we're not getting enough of this stuff.

And so I personally would be supportive of considering the expansion of the basis of this new code to include non-face-to-face care, whether it's via e-mail or telephone or whatever, because the longitudinal management of patients has got to enable the physician or care coordinator to expand beyond the 0.01 percent of a patient's waking time that's face-to-face with the physician. And these other modes are already proving successful and in some cases really a documentation of reduction in total PM/PM spending associated with some of them.

Last but not least, to weigh in on this last question about how do we deal with this level of risk or lock-in. Could we consider, in the next draft, the pros and

cons of a multilevel patient designation in which patients would have the ability to designate various degrees of delegation to their care coordination manager, including selecting the specialist and hospitals who they might see but not limiting to them if they prefer more freedom than that.

DR. MILLER: I know we're out of time so I'm going to say this really fast. Remember on the fee thing and getting the physician involved, you do have the ability, and even under the demonstration now, the disease management organizations do have the flexibility to do something with the physician. And so that is certainly one mechanism.

The second thought is this new fee and the inherent risk of paying for what we're already getting and the budget implications and that. Remember, assuming an adequate payment -- and I recognize there's an SGR issue out there -- but we make recommendations across the board. One could talk about within the fee schedule and moving money around within the fee schedule.

Last thing on the face-to-face point, another way to think about that issue is if you make this payment as in okay, here is something tied to the patients that you are

managing, it doesn't have to be reimbursing for every e-mail as much as it's sort of saying here is a fee that covers that stuff. That way you're not at as much financial risk.

MR. HACKBARTH: Okay, much more on this later. Good job.

Next is physician resource use.

* MR. BRENNAN: Today we are presenting the latest in a series of presentations on our work in the area of physician resource use and our use of two commercially available episode groupers which group claims data into clinically distinct episodes of care on a set of Medicare claims.

The two groupers we're using are Episode Treat Groups, created by Symmetry Data Systems and the Medstat Episode Grouper created by Medstat.

In addition to the resource use component of the analysis, we're also calculating a set of claims-based quality indicators for the same population on the same set of claims.

At the March and April meetings we'll be presenting the results of our analysis using a 5 percent sample of Medicare claims and once this report cycle

concludes we'll begin analysis of 100 percent of claims in selected geographic areas, permitting us to build on the lessons we've learned from the 5 percent analysis and begin to constructive physician level case loads, resource use scores, and quality scores.

I just want to quickly go over some of the technical results from the analysis. We ended up grouping approximately 204 million claims from calendar years 2001, 2002 and 2003. This is a 5 percent sample. The ETG group assigned approximately 90 percent of these claims to episodes while the MEG grouper assigned approximately 80 percent of these claims to episodes.

While this represents a not insignificant proportion of claims, upon further examination we found that the claims that could not be grouped were ancillary services such as tests and they did not represent a large proportion of dollars. With The ETG grouper the group's claims represented 94 percent of all dollars and with the MEG grouper the group claims represented 96 percent of all dollars.

In addition, both groupers had some trouble grouping home health records, although again they represent

a small share of both overall claims and dollars.

Once the episodes were created, we subsequently deleted any episodes that did not have a clean start or a clean finish. The clean period concept essentially means that a certain period of time, for example 60 days, needs to have elapsed before an episode can be considered to be closed. It's important to only have clean complete episodes in your analysis because you don't want to bias the results of your analysis by including potentially low resource use non-complete episodes.

Finally, we deleted any episode that had resource in the top or bottom 1 percent or had total payments that were $30 or less.

As we outlined to you in November, we're focusing our analysis on a subset of episodes that are particularly relevant to the Medicare population and we're also standardizing payments in order to facilitate comparison across geographic areas.

For the purposes of this presentation we don't intend doing an exhaustive comparison of the two groupers. Instead, we'll present some high level comparisons in the next few sides, but for simplicity we'll focus on the

results from the MEG grouper for the remainder of the presentation. We have, however, generated the same analysis using both groupers and where appropriate we'll note any differences or similarities between the two.

This table presents a comparison of the ETG and MEG groupers for some of our selected conditions. Going from left to right the first two columns indicate the number of episodes created by the MEG and the ETG groupers respectively, while the second two columns indicate the average number of dollars associated with each episode for the two groupers.

As you can see, for certain episode such as hypertension -- which we've abbreviated to HBP in the table -- and breast cancer there's broad agreement between the two groupers both in terms of the number of episodes created and the average resource use in those episodes. However, for other conditions some anomalies exist.

For example, congestive heart failure, while the number of episodes created is broadly similar between the two groupers, the average resource use for CHF episodes created by the ETG grouper is more than twice that of the MEG grouper.

Now obviously these differences in results between the two groupers for these conditions are enough to warrant further examination and we're looking into these differences. We've also spoken with the people at both Symmetry and Medstat to find out a little more about the clinical underpinning of their two groups and under what circumstances claims could group differently and lead to result such as the one you've just seen.

However, it's also important to note that the ETG and MEG groupers do differ in some very fundamental ways which may make explicit comparisons between the two groupers difficult. Perhaps the biggest difference between the two groupers is in how they create episodes. The MEG grouper relies solely on ICD-9 codes to create episodes, whereas the ETG grouper relies on both ICD-9 codes and procedure codes to create episodes.

To go back to the congestive heart failure example again, and on the last side there was a difference on average costs, I can say you that CHF episodes created by the ETG grouper have a much higher proportion of costs attributable to inpatient hospital stays than the MEG grouper. Additionally, in the MEG grouper CHF is found in

over 40 episode groups in addition to the stand-alone CHF episode group, reflecting the fact that it's not a disease but a condition that can be the result of many diseases.

In other research, Medstat has found that among all patients with CHF about 20 percent were found in severity stages of other episodes and it's possible that these are more likely to be related to inpatient stays which could account for the cost discrepancies between ETGs and MEGs, although as I said we're continuing to check into this.

Ultimately the test will be less about absolute differences between the two groupers and more about whether or not the groupers rank physicians differently.

This table illustrates in some more detail some episodes created by the MEG grouper that present each episode by disease stage. Disease staging is a concept used by the MEG grouper that assigns different stages to episodes depending on the overall severity of the episode.

Stage zero or one represents the lowest severity stage and stage three represents the highest. It's an important concept because ideally you don't want to compare physicians who predominantly treat patients with stage one

of a particular episode with physicians who predominantly treat patients with stage three of a particular episode.

Going from left to right the first column represents the selected episode, the second the episode stage, the third the percentage of episodes that fall into that stage, the fourth the percentage of payments that fall into that stage, and the fifth is the coefficient of variation associated with each stage.

DR. NELSON: I have a point of clarification. The staging is according to the temporal stage? That is, along a time sequence? Or is it on a severity stage?

MR. BRENNAN: It's severity, so based on specific ICD-9 codes and subcodes and the like.

As you can see, stage three episodes tend to account for a disproportionate amount of resource use relative to their size. For example, stage three coronary artery disease, or CAD, accounts for only 19 percent of CAD episodes but 53 percent of total payments associated with CAD.

Similarly, stage three colon cancer episodes account for 41 percent of payments but only 16 percent of colon cancer episodes.

You can also see that the coefficient of variation also tends to decline with the progression in severity of an episode. For example, the coefficient of variation for stage one CAD is 262 compared to 109 for stage three CAD. We think that this may be because there's more discretion in treatment options during less severe stages of an episode, although we'll be discussing this with our expert panel in the near future in order to solicit their opinions.

We also examined episodes according to the types of services that accounted for all of the resource use within an episode. This table presents for selected episodes the percentage of resource use that was associated with hospital inpatient care, evaluation and management care, post-acute care, procedures, imaging, tests or other not classified. Again, the results are quite interesting and again they confirm that the groupers do appear to be grouping claims appropriately. I say groupers plural because we have created a similar table using the ETG grouper and the results are broadly consistent with the exception of those CHF episodes that I mentioned earlier.

The table highlights particular areas of interest for some episodes. As you can see, CAD and pneumonia

episodes feature high levels of inpatient use. In contrast, more than 50 percent of resource use for hypertension and sinusitis is associated with evaluation and management care.

With that I'll turn it over to Karen, who will walk you through some of the results associated with attribution to physicians, both in terms of resource use and quality.

MS. MILGATE: So now we're going to switch gears a little. Niall has just described what the groupers can tell us about physician resource use, but in order to reach the goal of differentiating among physicians based on resource use we first have to be able to identify the physician most responsible for that use. And in that analysis, we'll be attributing episodes to individual physicians.

And further, because the ultimate goal is to also tie quality indicators to the analysis, we'll also be looking at how we would attribute performance on quality indicators to physicians.

This is critical when we turn to our 100 percent analysis later this year where we'll need to actually choose an attribution method, so we used our 5 percent sample to explore the various methods.

To do so we talked with our expert panel who have run these groupers, as well as the panel has two clinical experts, to ask them about attribution methods for both resource use and quality. And then we also looked at the variety of different programs in the private sector that have created there own attribution methods to get advice from them and identified the following issues.

First, it's important to decide if you want to use dollars versus contacts with physicians as the unit of analysis. The advantages of dollars is you can get a sense of the intensity of the visits that the beneficiary had with the physician. The advantage of contacts is you can really look at the physician who saw the patient the most and maybe more likely the one to have actually managed that patient's care.

After you decide whether you want to use dollars or contacts, there's also questions of whether you want to look at all dollars, which could include hospital dollars, procedures, tests, labs, et cetera, or if you should just limit it to evaluation and management codes given that they might be more likely to identify the physician who actually had some responsibility for managing the patient's care.

And further, particularly for episodes where a lot of the care is delivered in the ambulatory setting, you might actually want to limit your attribution rules to evaluation and management that occurs outside of a hospital setting.

One of the key questions is what is the appropriate threshold? And by that we mean what percentage of visits or dollars are you talking about is enough to attribute the actual episode to any single physician? And there the range we looked at were anywhere from 30 percent to 50 percent, which seemed to cover most of what various programs do.

In addition, beneficiaries often see more than one physician in an episode of care, so should the episode be allowed to go to more than one physician? And we looked at that, as well.

And finally, are the methods the same for resource use and quality? And you'll see, as we talked this through, we found the answer is no, they are slightly different.

So these are our findings on resource use. What you see here is a table that looks at the percent of episodes that are attributed to an individual physician.

And this is using the MEG grouper, as Niall said we were going to talk about those results.

Down the left-hand side you have the various attribution methods we looked at. Just one note, we first of all, took off looking at all types of dollars or contacts. The expert panel said that in most cases most programs really look at E&M dollars or contacts, so we took that out from the beginning.

And then across the columns, the column we're going to focus in on here the most is the all column where we basically have the percentage of episodes across all the selected episodes that we chose that could be attributed to a single physician. The other episode types there are really to give you a sense of the variation but we're not going to go through those in any great detail.

On the first row you see the evaluation and management visits where we set a threshold of 30 percent. In that case we found that 90 percent of all selected episodes could be attributed to a single physician. So that means we found that 90 percent of all episodes you could identify one physician that was involved in that episode for 30 percent of the visits in the episode.

I'll keep going through this and I'll say it more clearly as we go along.

When we increased the threshold to 50 percent, that is that for an episode to be attributed one physician had to be responsible for 50 percent or more of the E&M visits, that number went down -- which you would expect, that's a more conservative test -- to 75 percent.

When we looked at dollars to see if you used dollars if it would change the percentage of episodes that could be attributed we found that, in fact, it didn't do much to change the percentage that could be attributed, particularly at the 30 percent threshold. Still we found that 90 percent of episodes could be attributed to an individual physician who had 30 percent or more of those dollars.

When we move that threshold up to 50 percent the number again went down, but it did not go down as much as it did when we looked at visits. So that went down to 82 percent of all episodes.

We then looked at if we used evaluation and management visits or dollars outside of the hospital setting only, if that would change our attribution results. We

found again a fairly high number, at a threshold of 30 percent, of the episodes could be attributed to a single physician. The 86 percent there is a little lower than the 90 but it didn't go down that much.

We also, although it's not on the chart, did look at multiple attribution and again found that it didn't matter that much, that there were some episodes that could be attributed to more than one physician but still it was a very high percentage were attributed to a single physician.

So we found really across the board that we could attribute a high percentage of all selected episodes to an individual physician.

So we found that the episodes could be attributed to physicians, but did they get attributed to the right physicians was the next question we tried to get some information on. here we looked at the percent of episodes that are attributed to a physician by specialty and in general found that the type of specialty to whom the episodes were attributed seemed to make clinical sense. Again we wanted to go back to our expert panel to what they thought. But for example, we find that 38 percent of coronary artery disease episodes were attributed to a

cardiologist. And if you look at prostate cancer, going on down the side there, that 64 percent of those episodes were attributed to a urologist.

I want to make one note about a category here because it shows up in a lot of our rows, even in the broader charts that were attached to your mailing materials, and that's the outpatient specialist. That refers to any care that was delivered by a physician in the outpatient setting. There's not a designation on the claim that tells us what type of physician, so it just becomes an outpatient specialist.

Now we turn to our quality analysis and our ability to identify individual physicians who were responsible for the quality of care. So remember again that the ultimate goal is to put together both measures of resource use and quality in the end of the analysis. For this we're using a set of claim-based quality indicators that has been developed and revised over the years for MedPAC. And we've talked about this set in previous meetings but it's the Medicare Ambulatory Care Indicator Set for the Elderly.

Some examples, just to get you oriented to the

type of indicators we're talking about, is for example for diabetes the percent of eligible beneficiaries that received eye exams or A1c tests within a certain time frame. And for CHF whether eligible beneficiaries receive appropriate lab tests or get timely follow up after a hospitalization for CHF.

So here the goal is to identify the physician that is most able to affect the beneficiary quality for specified indicators. Remember that the indicators are all associated with a specific condition so here what we do is physicians are assigned based on the level of involvement with the beneficiary for that specific condition.

So for example, for the beneficiaries that are eligible for A1c tests for diabetics, then we look at all of their evaluation and management visits and contacts that were associated with their care for diabetes. And then we test our various attribution methods on those dollars and contacts.

So for example, we found that for 91 percent of diabetics needing an A1c test, we could identify a single physician who delivered 35 percent or more of that beneficiary's evaluation and management for that condition.

So we only looked at the care that was related to the condition that made them eligible for the indicator, not all of their care.

The overall results here were that when we used a method of using 35 percent as our threshold, and we came upon that by looking at what we found in the resource use the analysis frankly, of E&M visits that 78 percent of all the indicators could be attributed to an individual physician. When we increased that threshold to 50 percent, that number went down to 63 percent.

Again, we wanted to look at whether this method of attribution was assigning the beneficiaries' care to appropriate type of physician. Again, we found similar results. It seemed to make clinical sense. And further, we found that both the resource use analysis attribution method and the quality rules assigned beneficiaries to the same types of physician. For these particular conditions the top four that you see there were exactly the same in both the resource use analysis as well as the quality analysis.

MR. BRENNAN: Over the next few weeks, we'll be going over some of these results in conjunction with an expert panel, as we've mentioned, and at the April we'll

present you with further results from the 5 percent analysis including variation in resource use by MSA and variation by specialty.

We'll also have incorporated additional analyses that will permit us to risk adjust episodes and we will examine specific procedures that appear to drive resource use. Once the 5 percent analysis is completed, we will move on to the 100 percent analysis where, as I noted at the beginning of the presentation, we'll be able to build physician-level case loads and deal with technical issues like what is the appropriate number of cases a physician needs to have in order to qualify to be counted.

We'd be happy to answer any questions on these matters.

MR. BERTKO: A couple of questions or comments here. I think, Niall, the first one goes to just what you talked, the number of episodes being real important here. You had the 5 percent sample rather than 100 percent sample, so it means for a given physician in a given state or GPCI, wherever you did this, you don't really know this.

I guess what I would comment on here with the threshold is a 50 percent threshold is better to use if you

have enough episodes. Maybe that number is 100, which would be really good. But it's better to get enough episodes. So if you're not getting 100, then you settle as a good enough 30 percent threshold for that. At least that's my bias on this.

MS. MILGATE: 100 per physician?

MR. BERTKO: That's what I've heard from some other stuff. I don't know, Arnie, if you'd agree with that or not.

Then the second comment is -- well, first of all, let me say I'm extremely appreciative of the amount of work you've done to get to this particular set. It's enormous amounts of data being spun through.

And then, unfortunately, I'm going to suggest more work.

If I've understood your efficiency one on page nine, your expert panel said to only evaluate docs, group them in episodes, using E&M visits or E&M dollars. And my bias would've been to use all dollars because on behalf of -- I'll call it MedPAC -- it's important for us to look at all dollars.

Now the problem with this may be that in infers

delivery systems and there might be objections to that. Maybe that's what the panel was objecting to.

MR. BRENNAN: I think part of it may also be a technical limitation in that when we use E&M visits or dollars they are, in almost all cases, explicitly linked with a physician UPIN, whereas if you start to pull dollars from other settings, particularly the inpatient setting, there's not necessarily now a UPIN associated with that. Now you could attribute an episodes to a hospital, I guess, or to the physician who sees the patient in the hospital. But the actual inpatient care is not necessarily explicitly linked to a UPIN.

MR. BERTKO: No, but it could be explicitly linked through the grouper. I don't know the Medstat one, but the ETG grouper will, in fact, drag in inpatient care. And I guess that's what I'm advocating for is to look at inpatient, outpatient and physician costs as you evaluate how somebody does this.

DR. MILLER: Can I clarify something? The episode drags everything in.

MR. BRENNAN: Right.

DR. MILLER: This question of E&M visit versus

dollar is within that episode you have physicians, hospitals and everything. It's how to say which physician was most --

MR. BERTKO: So maybe that's my confusion. I should have asked better. If you're strictly using the visits and the dollars to create the attribution but not the efficiency measurement, then I'm --

MR. BRENNAN: Absolutely. It's just to assign it to a responsible provider. We will count all the dollars.

MR. BERTKO: Okay, then I'm fine on that one.

DR. MILLER: I just wanted to avoid the additional work, which is why I was listening very careful to the question.

DR. WOLTER: I find this very interesting also and it got me thinking about where we are headed with it and I suppose we'll talk about that in the future in terms of what would we do with this? Would it be linked to some physician incentives at some point in time? Would it be just an information reported back in hopes that that might help change how some practice standards are set?

And of course, my bias is that if we could use it to incentive the coordination of care, since if an episode is defined by being 35 percent of E&M visits, by definition

two-thirds of the care is somewhere else. And it may well be that if we're thinking about virtual networks and creating incentives around those to form, maybe there's something in that as well, in addition just to staying with the focus on an individual physician.

So whether that makes any sense to where this is headed I don't know today, but it got me thinking as you presented the information.

MR. HACKBARTH: Just as a reminder, to this point what we've said it is the purpose for developing this tool is confidential feedback to physicians, although we have opened the door to the possibility that based on development and exploration that we ultimately may recommend that it somehow be incorporated in the payment, including pay for performance. But the first step is feedback.

DR. MILSTEIN: This report is so good I have no comments on the core of it. My comments are really ancillary and really relate to maybe a parking lot issue as we begin to confront related issues on different topics.

That is at the end of the day, as per Nick's comments, we begin to transform this into any kind of policy decisions that are going to make this count in the market,

make this matter. The issue of accuracy of attribution is getting to get to be more and more important.

We have an opportunity at this point to begin to set in motion collateral changes that would enable much less cloudiness regarding attribution. And what I have in mind specifically is -- and I defer to you, Glenn, as to what we hook this to -- is to begin to move forward with I'll call it the transparency discipline, making sure that at some point in the future when we're paying for a lab test or we're paying for medication we know the provider ID number of the physician ordering such. That's the pathway to having much more confidence in the future about our attribution algorithms.

MS. BURKE: I think this really is kind of touching on where Nick started to go. And that is if we look at this chapter and then reflect back on the conversation we just had in terms of coordination of care it seems to me again that our ability to begin to track an individual physician's impact on a particular patient, not only for purposes of feedback in terms of their resource utilization for purposes of comparisons with their colleagues for best practice purposes, it also -- unless I'm

sort of not fully appreciating what we're hearing -- I think it underscores once again the impact and the role that an individual physician has in terms of that particular patient and how they're managed.

Because again what you see here is in a number of these cases we're picking patients who are chronically ill, who are managed in complex situations over a very long period of time.

And so I think it underscores for me again that as we begin to develop more of this understanding, although I think again to the point made, the more we can understand about attribution and the further that goes beyond the sort of direct impact but to the hospitalization and whatever it happens to be, that we have of growing set of tools available to us. Again, the accuracy, as Arnie has suggested, is going to be critical to us.

But it seems to underscore once again that there is a way to do this through individual physicians, which has been the stumbling block for a very long time, is that once you get outside of a group setting how do you begin to incentivize individual physicians who are largely entrepreneurial in nature? And I think again this

underscores, as I understand it as we gather this information, again our capacity growing to be able to do that.

And so again, unless I misunderstand it, I think we are now putting together the tools that will allow us to incentivize in different kinds of ways. So the accuracy, the broader we can get this in terms of capturing attribution, I think will be very important to us.

MR. HACKBARTH: How do we evaluate when an attribution rule is good enough and compare them? The approach you used here I found interesting and somewhat comforting that if you vary the threshold you weren't getting dramatically different results when you moved from 30 to 50 percent. That was reassuring to me.

Do you have other thoughts about how you determine when you're good enough?

MS. MILGATE: This is a little off topic on that but what's been interesting to me, in thinking through what's the right threshold, actually is central, is John's point. There's kind of a balance between having it high enough so you really feel comfortable that you've got the right physician who is responsible and having enough sample

size.

So if we can get a high enough sample size I think John is right, that we would want to ultimately set that as high as we can.

And remember those episodes, like the 10 that are not assigned in the first method, and the 25 percent that are not assigned in the 50 percent, would not be assigned. So if you didn't have some sense that you could clearly assign it, then those would just be thrown out so you wouldn't have a concern about those you still had left, I guess is all I'm saying.

DR. MILLER: Plus, even within your analysis, I think we also looked at after you made the attribution you looked at the type of physician it ended up getting attributed to for the given condition. And at least there was some face validity and not to push the question off from us.

I think part of this is going to be us sitting around looking at it and saying this feels about right.

DR. REISCHAUER: My comment was going to be on this point, and I guess I'm much more comfortable with lower numbers than John and Arnie happen to be. For me the

question is not just with the threshold is but what's the next largest person you could attribute it to? And judging from the 30 percent threshold versus the 50 and the fact that the episodes go from 90 percent only down to 75, you realize that in a sense the next person is pretty darn small on average. And my guess is it's very small.

And so you'll have somebody that you attribute 30 percent or more to one person and the next one is 6 percent or below. And so, in a relative sense, you can be pretty comfortable about this.

MR. BRENNAN: That's right. Just to add one more data point for you all, when we did test multiple attribution using a 30 percent threshold. We went from --

DR. REISCHAUER: In theory, you could have three people with 30 percent. But in fact, you don't. You have one with 47 and another with six.

MR. BRENNAN: So of the 90 percent, 78 percent were still only attributed to one doctor and 12 percent were attributed to two or more. So it seems to be -- most care seems to be fairly concentrated around one doctor.

MR. HACKBARTH: What you say, Bob, makes sense but in trying to look at this from the perspective of a

physician, okay 30 percent of the E&M visits, and now we're saying you're responsible for all this specialty stuff and all this imaging that the specialist may have ordered and the inpatient stay, for a lot of people, for a lot of individual practicing physicians, that may feel like a stretch.

DR. REISCHAUER: But remember when were setting the threshold at 30 percent, still 75 percent are above 50 percent. So there will be aggrieved parties but there will be quite few of them.

DR. SCANLON: Also, you have that patient at 30 but you've got a lot of others at 60 and 70. We're talking about attributing the patient, an individual patient, to a physician. The physician is going to have a distribution. Part of the key here is going to be that there are enough patients that a physician has that there are some risks spread.

MR. BERTKO: I just want to add that Mark made a comment, face validity. And I would add transparency to that. So as long as it's not a black box, we've been working on us with the Milwaukee Business Coalition. And with our big brother, 25 or so employers, behind us, the

docs have said okay, now that we understand it, it might be good enough. I think I'm putting it in the right paraphrase.

MR. HACKBARTH: So in that conversation you do different iterations and let them see how different rules might affect attribution?

MR. BERTKO: No. We do it a way that we have found that our Ph.D.'s have said sounds pretty good, and then we explain it to them and go to the medical societies and such. This whole discussion of what these guys did would be, believe me, well beyond what an average county medical society would want to hear.

MR. HACKBARTH: Yes, although this one variable, the attribution rule, seems -- these are bright people. That's pretty basic.

MR. BERTKO: They may have gotten an A in calculus but not in stat.

DR. MILSTEIN: My earlier point about -- and again, as the consequences that pivot on this begin to get more significant, the increasing importance of confidence in the attribution. Earlier I mentioned that the way to remediate this is to make sure the drugs and lab tests, and

for that matter imaging tests, are coded as the to the ordering physician. The same certainly applies to physician specialist care that originates in a referral from another physician.

If those visits were coded with the identifier of the referring physician, it would also take us a big step towards confidence in attribution algorithms.

MR. BRENNAN: The claims data does have information on the referring physician UPIN but it's not very highly regarded at the moment, which ties directly to your point that we should improve coding of it.

DR. MILSTEIN: It would get better if there were consequences associated with the attribution.

MR. HACKBARTH: Others? Jay, last comment.

DR. CROSSON: I know we've decided we're talking about producing this information for comparative, education, and all the rest of that. But we've also had a lot of discussions about impacting the costs. So if you sort of think about the point you brought up which is what about the percentage here which would be a percentage of impact on the attribution which would be viewed let's say by the primary care physician as unfair. So what dynamic would that, in

fact, likely create in the fee-for-service community?

If you know you're a primary care physician and over time your income is going to be in some way related to what you do but to some portion of those downstream referral-based costs and that you have the ability to direct that and you have some more judicious use of resources available to you than others, you might in fact think that the dynamic they could be created by that same unease that you described might not be a bad one.

MR. HACKBARTH: That's where you'd want to go. I guess my point is simply it's a long way from where we are.

[Laughter].

MR. HACKBARTH: Thank you. Good work and we look forward to the next installment.

Next is quality measurement for hospital care.

* MS. CHENG: This is actually the second presentation in a series that we started at the end of last year and I was up here with Jack Ashby and Anne Mutti to talk to you about measuring inpatient hospital quality and resource use.

What we're after is really trying to make a tool for you. What we'd would like to do is to try to get our

hands a little bit closer, if not around, the notion of an efficient provider. We've said now for a while efficiency is a combination of two things, neither of which are particularly easy to measure, one of which is resource use and the other is quality.

You've heard a little bit about our work to develop resource use measures. We just talked about physician. That's running in parallel with hospital resource use measurement. And what I'm going to talk with you about this afternoon then is the quality part of the inpatient hospital resource use measure.

Obviously, the first step of getting toward a quality measure is choosing the indicators that are going to be a part of this. Because this is a tool that we want to be able to use here on the staff, our first criteria was it had to be run off of data that we could either collect and manipulate ourselves or that we could collect from another source.

We set a couple of other criteria to make sure that we had a robust set of indicators. The first one was we wanted to measure an indicator that we had a sufficient sample size for at most hospitals. We're contemplating

being able to describe 4000-plus hospitals. So we set an initial threshold. We'd like it to be something that about 3000 or more hospitals would have a sufficient sample size so we could get a score for them.

Our second idea was we wanted to measure things that occurred with some frequency. This is not a very scientific threshold obviously, but our notion was for a couple of the measures that we could measure with sufficient sample size at more than 3000 hospitals. They were such rare events that they were only occurring at 1000 hospitals or less. So most of the hospitals in our sample had zeros. We thought maybe that wasn't where we wanted to look first.

So at least for the time being we've set about three indicators aside because they occurred very, very infrequently.

And then our final idea was we wanted to get indicators that had some evidence of variation. To the extent to which we think that quality varies from hospital group to hospital group we wanted to have indicators that gave us some evidence of variation. If all of the scores were really tightly grouped, if there are quality differences they would be harder to see. So we wanted to

see some variation.

So applying those ideas we looked out across quality measurement for hospitals and, broadly speaking, there are three sets. We have two that were developed by AHRQ and these have been used very widely in the research literature. Those are the mortality indicators and the safety indicators.

The third large set we had are the process measures. Now these are the result of the work that CMS has done. They feed directly into the Hospital Compare, and these are the measures that were linked to the voluntary participation. Hospitals voluntarily submitted their scores for these processes and CMS collects it and develops these scores.

So when we applied our criteria, we could measure these at a lot of hospitals, they happen with some frequency, and they had some variation behind them. We came up with a set of 37 quality indicators. And I've put them up in this matrix here because I think we can be really satisfied with a set that has a fair bit of breadth and a fair bit of depth. What I've done is I've described them in terms of the mortality, safety and process kinds of quality

that we can measure with this set. And then down the other side the kinds of patients that we're going to be able to include.

So we've got some focus, an ability to focus on surgical patients. We also have sets that let us look at three different condition-grouped patients types. And we have four adverse events that we can measure on all patients in the hospital.

So when you take this set of 37 and you look at them together, I think you've got a pretty nice picture of what's going on in the hospital to the extent that we can measure it today.

The good news is we've got a lot of indicators. The bad news is we've got a lot of indicators. So what I'm going to do now is walk you through some of the work that we have done to start to contemplate how we can make some sense. When we start looking at hospital groups and we want to be able to make some comparisons about their comparative quality, I could come back to you with 37 different comparisons of three or four or five hospital groups and I think we'd have a bit of a hard time making sense out of it.

So what I'd like to suggest is that we should

think about ways to group these indicators together to make some patterns out of them and then, to the extent that we can, maybe start to summarize the scores. So that rather than having 37 comparisons we can break that number down a little bit. We may be driving toward a single measure. We may be driving toward a small set, three or five measures.

Let's see what the data looks like and I wanted to get you feedback on where we're going with this and grouping and summarizing the data that we can collect.

The next part of my presentation then are going to be some initial work that we've done on not so much testing as illustrating several methods of grouping the indicators together and summarizing the results. One way we could do it would be the weight each indicator in the set the same, just sort of take them as they come.

Alternatives to equal weights would be assigning some kind of weight to the scores that would make some contribute more than others to the final score of a hospital. We could do that by some estimate of their relative importance. We could do it perhaps by some estimate of a number of patients or what have you. But there are some ways that we can manipulate the data other

than equal weighting.

But as I progress through here I'm going to start with more or less taking each one equally.

So here are two groups that I thought and that our group thought sort of came out to us as natural ways to group this information. The first one is by the type of indicator. So put all the safety ones together, we've got about a dozen of those, put the mortality together, put the process together. So we've been discussing it by indicator type.

The other way that came to us was to look at patient play. So surgical, the three diagnoses and then the all patient.

I don't think anybody can read this but that's okay, just hang with me for two seconds. This is just an example. This is less than half of the data that we're going to be able to bring to you.

So what I'm going to do is I'm going to grab his first row that we have and I'm going to blow that up. So we're looking at the first row of that illegible table so thanks for hanging with me.

What we've got here then are five of the 12 safety

measures if we were to group the data in that fashion. Let me help you read this real quick. In the first row, the first column, we took two regions of hospitals. I pulled them out of the hat. I just wanted to show you some data.

So we picked hospitals in the South Central region and we're comparing them to hospitals in New England region. So that first number, 28, is the rate of accidental puncture per 10,000 discharges aggregated across hospitals in the South Central region. All of the numbers that we are looking at here are failure rates. I had to do a little bit of a transformation there because of our concepts. We've got safety as a concept, mortality and process. You do want processes to happen. You don't want mortality to happen, you don't want adverse events to happen. But I've translated them all so they're all failures. These are all rate at which things you don't want to have happen happen per 10,000 discharges.

The next step then was to compare our two groups. So I've calculated a ratio, and that's just the ratio of the score for the South Central group to the score for the New England group. So on that first indicator your ratio is 0.77. That's lower than one. Low is good because these are

things you don't want to have happen. So the way you would read that is South Central on that indicator is better that the group to which we are comparing it, hospitals in New England.

You could calculate that ratio for each one of the ones on the screen. Just in the ones on the screen you find that for the hospital groups that we've got, in some cases South Central is better. In some, it's worse. We could take this one more step and we could say that the two groups are the same if we couldn't find a statistically significant difference between the two.

That's the kind of thing then. We've grouped the data so we've got the concepts the same. We've put them into a group that maybe is going to give us a little bit of information. So our next step is going to be can we make a summary score rather than give you a big column of 37 of these?

So I'm going to move to the next slide. I'm seeing heads nodding.

This would be one way to summarize the 37 indicators after we group them. The first number on this slide, 1.33, is the average of the ratios for these two

groups in the safety group. I've taken 12 of the indicators and I've made a single ratio. Remember behind that ratio then South Central does better on some, worse on others. When you take the average then you can say on the whole, for safety, South Central's performance is worse. Do the same thing with mortality and do the same thing with process.

If we stopped right here then we could bring you three answers to which group is better.

You could take it one more step if you wanted a single score and you could say well all right, what happens if you take safety, mortality and process, given equal weight to each type of indicator, what would you come up with as a summary answer? And here what you see is that you would still come to the conclusion that South Central hospitals' performance was worse on the whole for our quality set. So that's that 1.61.

For illustration, if you took each of the 37 indicators, you didn't group them by type, you just took each one of the 37, took the average ratio, you get 1.71. This is the punch line that's going to be on the next slide, too. So you come to the same conclusion whether you'd group the data or not in this case.

Now if you had different relationships between the two groups of hospitals, that might not be the answer you'd come to but in this case you get the same answer.

So what I'm going to show you is one more way to take the same data, regroup it, reweight it, and then see if you come up with the same conclusion.

What I've done here is calculate, I've taken those 37 indicators, I've put them into groups by patient type now. So within each one of these groups, there's a variety of indicator types. For example, the heart failure group has two mortality measures and then process measures in it. And then I've compare the hospital groups again.

I told you the punch line already. You come up with the same answer then when you ask about the comparison between these two groups. But the reason that we're working through this data is that I'm trying to get some input and some feedback about how you feel about some of the ways that we've tried to test, grouping them together, about the summary scores that we're coming up, and give you a sense of where we're going to go from here to bring you this tool and develop it further.

So we've got some work to do. One of the

questions that I think grouping this data together brings up right away is if we can think of another way to group it would what we get different results when we're comparing hospital groups? And what would that mean if we got different results?

How is our answer being driven by the kinds of indicators that we have available? If some of the measures dropped out or some other measures came in, would we be getting different results? A very basic question is grouping and summarizing this kind of data, different indicators, different things going on, is that the right way to go? And then once we start getting toward a final iteration of this tool, what does it tell us about other factors that might relate to hospital quality?

Just as we were working through these two regions of hospitals, a lot of questions came right to the fore. Are we really comparing two regions or are we comparing hospital characteristics that are not necessarily randomly sorted into those two regions? So we would start looking at questions like that once we had a little bit of an idea of what this tool is going to look like.

That's a lot of data. And we're early in the

process. So to the extent that I can, I'll answer questions about what we've done. And please give me some ideas about how you'd like to go forward.

MS. BURKE: This is terrific work. I want to make sure I understand where weighting occurs and where it doesn't occur and just ask a question.

For example, if you were to go to your blowup slide of the first row, although this, I must say, underscores what I was always told which is if you get sick you go to the airport if you live in Washington, and you fly north. So that's reassuring. Or take a train.

If, for example, I were to look at within the safety category among the five areas that you have listed, are there weightings that occur within them? For example, is there a determination made in the calculations that a collapse of your lung is a more serious issue then a decubitus ulcer? Do you weight within the weightings within a category? That's one question.

The other question is there is an acknowledgment here, for example, the infection due to care. That's a presumption that that occurs in the context of your current treatment.

There's the issue, for example, with a decubitus ulcer which is a presentation question. Is one admitted or does this occur in the course of your treatment?

Some of these are things that occur there. Respiratory failure, arguably there. Infection due to the care, presumption it occurs there. The decubitus ulcer could be a presenting issue or it could be one that occurs at the time.

But there are clearly variables within these groupings as their relative importance.

I mean, if I had to choose, I'd rather get punctured than my lung collapse, depending on whose needle punctured me. It depends on whose needle it is and where the puncture is. And I'd put mortality right at the bottom of my list, but maybe somebody else would vote differently.

[Laughter.]

MS. BURKE: But I just wondered, as you build them within where the values arise, I can understand the weighting, and I think you're approaching it exactly the right way.

There are groupings and there are weightings within but how does that weighting structure work?

MS. CHENG: Right now within the groups that I've discussed each measure is weighted equally. So we haven't tried to make what I would suggest is a somewhat more qualitative judgment: punctures really bad, respiratory failures not as bad. So we haven't done that.

We could certainly contemplate, if that's the direction we're going, sitting down and putting some qualitative weights. To make that work mathematically, you can't just suggest that infection due to care is really bad. You'd have to be willing to say it's 4.5 times worse than -- and so you can see where that impulse would come from but the math could get a little hairy.

MS. BURKE: The math could be complicated but I think for credibility purposes, going forward, the more refined this is -- I mean, I think anyone looking at it would suggest that dead was worse than an ulcer. And all things being equal, you really ought not equate one with the other. There ought to be some variation.

But you're right, the complication will be how much worse is being dead? Well, it's probably substantially worse. But how you vary within those areas, I think, is a complicated one.

It depends on your religion, that's true, whether you're going on to a better world.

[Laughter.]

MS. BURKE: But I think that that will be a question that -- yes, there are Medicare savings. It's like subsidizing cigarettes in nursing homes. It was always a good idea.

[Laughter.]

MS. BURKE: But I think that will be something that we'd want to think about, is within those measures.

MS. CHENG: We tried looking across other people that have tried scorecards and there's a lot of scorecards that weight them equally because acknowledging that you would like to give them relative weights is difficult. So some scorecard went that direction. Other scorecards have gone the direction of trying to give them relative importance. And one of the first places they go would be to something that you could quantify. So maybe the rate at which these occur would be one way you could -- or the number of patients that are in the denominator or something like that. If there would be a way to assign them a weight like that, that would be another thing we could think about.

DR. REISCHAUER: There's an assumption that if you weight everything equally you really aren't making a judgment when, in fact, you are making a judgment and you're making one that you know is wrong.

DR. MILLER: Let me give you a different way to think about this problem, because if you were to get into the business of let's decide what the weights -- I mean, we could put together experts, ask people to do stuff . We can do that.

But another thing is, and I think this is part of what Sharon is trying to illustrate. She organized it once by condition. And within the condition there were measures that were safety. But she also organized it once by the type of measure. And then it might be that you would conclude that -- since I think death is, Sheila thinks it's the most important, you might want to present the information by these categories either to weight them or even not, just to say I think it's important that this be held out separately because this may be intellectually, without assigning a weight, important to know how they performed on this relative to other things. And there was

some slide that did that.

DR. REISCHAUER: But the consistency of the outcomes here, no matter how we weight it one way the other, is really a function of the fact that on almost all of these individual measures the South Central is worse than New England. So the more difficult thing would be if we were doing New England versus Pacific Coast or wherever there are equivalently good hospitals. And then each one of these would have come out with a different one maybe, a different worse/better.

MS. BURKE: I presume. The presumption is this ultimately is not South Central versus New England. It's hospital A versus hospital B. So it's going to get up close and personal real quick. So it isn't going to be California versus the world. It's going to be the MGH versus the Brigham.

MS. CHENG: We are trying to crawl before we walk here.

MS. BURKE: I understand.

MS. CHENG: And you're absolutely right, that's the direction we're going. But boy, I'd like to try to get something that we could use on groups of hospitals and not

necessarily hold it to the standard of could we use this A versus B. If we can get it to work on the group level then we'll have something that we can use at these to compare groups of hospitals.

MS. DePARLE: I'm trying to think about -- this may not be a fair question given what you just said, but what came to my mind when I was looking at this were the issues surrounding the hospital mortality data that Glenn worked on when he was at HCFA. You sort of raised this, Sheila, when you said do person present with the beginnings of a decubitus ulcer or does it develop at the hospital? So to what extent are these measures or indicators risk adjusted for demographics and for the presentation of the patients?

MS. CHENG: One of the cuts that we used, all of the indicators that we put into this set, are ones that at least most of the Commission had a chance to at least think about a couple of years ago when we applied the Commission's criteria for good measures.

MR. HACKBARTH: The safety measures that AHRQ developed ones?

MS. CHENG: Right. So to the extent that we had

to chance to look at them, these are the ones that we said they have risk adjustment behind them. They have evidence that suggests that they're reliable and valid. The safety ones are ones that make people a little less comfortable. They have a lot of exclusions that go in front of them.

So just off the top of my head for decubitus ulcer, because we don't know exactly what people present with, a large group of diagnoses that are likely to have come to the hospital, whether we know whether they did or not, with an ulcer are excluded. If you came from a nursing home, if that was a source of admission, you're excluded. So we don't know whether that patient did or not, but we don't even put them in the denominator.

So the comparison that I've shown you and the comparisons we'll be able to make with this set are risk adjusted and have exclusions that at least give us some comfort that we're making valid comparisons.

MR. MULLER: I commend you for this. I think trying to have this comparative information available is of major gain and import. Obviously, as the comments and questions from Nancy-Ann and Sheila have already said, this gets very juicy when you start getting down to lower levels

of comparison such as on the hospital basis. And then, of course, all the usual caveats about risk adjustment become so important. We started that dialogue. For example, in my hospital, one of the other hospitals, all of the deaths occur in our hospital because they transfer right at the time that they're ready to go.

MS. DePARLE: Yeah, yeah, yeah.

MR. MULLER: That's what they all say, right?

[Laughter.]

MR. MULLER: But I think the risk adjustment is therefore of critical importance.

But I think when you think about -- I've said this in different settings. When you think about it, there's 50 years of financial information that are available on hospitals but the measurement of quality in hospitals is still relatively new in the last five or 10 years. It's so much easier for almost all of us to talk about the finances of a hospital than to kind of say here's the quality of care. It's an issue I deal with with my board. You always want to say you have these great doctors in cancer, heart, et cetera, and so forth. But to have these kind of qualitative measurements that can really stand out there in

public and be available I think is something that we all have to keep moving towards.

And obviously one of the challenges has been, in the last several years there have been so many measures put out there, it's kind of hard to figure out -- you know, with the 57 measures, the 84 measures and so forth, that we've described even in our own work over the course of the last year or two, it's sometimes hard to figure out how to put them all together.

So using some of the categories that you've used here to try to group them, I think is a major advantage. So I commend you for that.

I think continuing to think in those directions, as to how to group them -- I mean, I could start giving you comments on some of them right now but I'm not going to get into that because some of them are such smaller weight. Not just joking about the mortality one versus smoking cessation, but you want some -- either you need some agreement on weighting, which I think is very difficult to secure, or you need to have some that are close enough that weighting them equally is not as big a random event.

So you don't want things that have such major

consequence as pneumonia in a hospital and so forth, and smoking cessation, which by and large very few do as well as one should.

So I think continuing to go in this direction is a good way to go and I do think keeping it perhaps, obviously from these big regions you really want to start getting it down to county and metropolitan levels. Inexorably you're going to get down to the hospital level because that's where people want to really -- that's where the levers for most places of improvement can in fact be pushed. So I think it's good to go in that way.

I think just having the ratios is intuitively ingenious because I think it allows people to come to quicker judgments. One of the real difficulties with the various dashboards and scorecards that are out there right now is most people who aren't in the field have a hard time knowing what's the right number.

So therefore, having have this kind of comparator around one I think is a very ingenious and clever way of getting that kind of comparison quickly into that without forcing people to know exactly what the rate might be. So in that sense it could serve considerable public purpose

because people do understand ratios of more or less than one. So I commend it.

I'm sure you'll get a lot of comments, not just from myself but others, as to which ones should go into it. But I think, especially in the first two categories, those are all, I think, by and large pretty consequential ones that will withstand further scrutiny.

MS. CHENG: Just real quickly to one of Ralph's points, because Jack and Anne and the whole team that's been working on this, we've been asking ourselves. The reason we have two mortality measures for most of these events is because we measure it once in the hospital and then once 30-day. We've been wondering if that's double counting or if, as you suggested, there really are differences in hospitals' decisions to retain a patient or to send them home or to another setting that it's fair to use both of those.

MR. MULLER: You really need the 30-day. Obviously we've dealt with this in the past in other settings. You start having border issues about what information gets reported by what states. For example, being in a state that's right on a border, many of our patients are from New Jersey and they don't report the 30-

day numbers to Pennsylvania. So you have 40 percent of your patient base taken out of the denominator but they're still in the numerator.

Those kind of things always make life a little bit more complicated when you start getting --

DR. REISCHAUER: The Soprano effect?

MR. MULLER: Yes. Roseanne Rosannadanna.

[Laughter.]

MR. SMITH: Following up a little bit on what Ralph said, I think you're right about the ratios, Ralph. On the other hand it's interesting. If Sheila's right and mortality is 4.5 times worse than any of the safety or process measures, you look at this chart and it really doesn't make much difference whether you stay home or go north. The ratio changes dramatically if you exclude the process and safety measures. So as Ralph said, I think figuring out the weights is terribly important.

And if some of these are appropriately weighted significantly higher than the others, then the 1.6/1.7 disappears in a flash.

You ended, Sharon, by saying what factors. It strikes me we know some of them. Staffing matters, volume

matters, hospital type matters. And region may matter. Maybe the region is not a good grouper but it's a good factor and we ought to think about maybe substituting where are you located rather than where are you located as a grouper, but where are you located as a factor along with staffing volume and hospital type as a way to see if we can't come up with something that discriminates more finely.

MR. MULLER: You have the Dartmouth 306 groupers and one could go in that direction and some of the RAND people have done work off that, too. So I think once you have the database you can start figuring out -- you can look at states, you can look ta counties, you can look at the 306 hospital regions and with computer time you can start seeing which ones make more sense by displaying it that way.

DR. MILSTEIN: A couple of comments. Since we're building something now for the future one of the things we may want to think about is integrating into our scorecard that measures flow that we can now count on based on what the Deficit Reduction Act has required. For example, the Deficit Reduction Act requires, I don't know whether it's 2008 or 2009, but for hospitals to report their status on three relatively highly important safety, of the NQF safe

practices. Are they present or not?

So if we're building a scorecard now for something for implementation in the future, we could begin now holding space available for those measures that we know are going to flow based on the Deficit Reduction Act.

Let me make another point and then finish with a question. The issue of weighting in the relative disutility or whatever you want to call it, different bad outcomes, there has been a fair amount of work in that area and there is research at Wharton that's already been published on weighting of relative types of complications, including how much you weight death versus a non-serious versus serious complication that's already been published. Robert Kaplan at UCLA, who has been one of the leading thinkers in this so-called Quality Adjusted Life Year, has also done research in the acute area.

Let me close with a question and that is if I remember the AHRQ specifications, especially with respect to the safety measures, because those are based on complications and whether or not hospitals are -- hospitals have been shown to vary quite a bit on their inclination to code complications. I remember when AHRQ came out with that

list of so-called patient safety indicators. They came with a warning label that basically said in order to be able to use this you have to make sure that a given state's level of discipline and monitoring and management of hospital discharge data reporting is up to a certain level.

And so maybe I'm just asking you to elaborate on your earlier comment that these so-called safety measures, which at the end of the day are complication measures, are clean and reliable and we don't have to worry. None of the measures have to be all that good but the basis on which you feel they are good enough. Maybe you could elaborate.

MS. CHENG: You're absolutely right and I think that would be something that -- we picked two regions because it seemed like a way to group hospitals that we could sort of get our heads around. And I think that to take AHRQ's caveat, you wouldn't want to have tried to do this with Louisiana versus Massachusetts because you know that there's going to probably be consistent differences in coding between those two states.

I blew it up to a region, there may very well be regional differences as well but I was hoping at least by putting several states together you might want to look

behind this and see the comparative rate of coded complications versus the known health of the population or something like that to get a feel for whether you're measuring coding differences here or complication differences.

DR. MILLER: To get more comfortable with it, as you churn through looking at these things if you're finding that specific measures or collections of measures, the safety category, just seems to move all over the place each time you move to a different level of aggregation. That might tell you maybe that one's not a good one to work with or to put very much weight on it or something like that. I think some of that can fall out from the data analysis. You put the categories together, you can look at how it runs across the data. If you're getting very different result it may tell you that.

DR. SCANLON: This, in some ways, is reminiscent of the development of DRGs. When we were developing DRGs the goal was to explain something, the costs per admission. And statistics were applied but they were done in a constraint way. It had to be done with an outcome that was going to be understandable to clinicians and that there was

kind of agreement that this made sense. You had to be able to sell it.

This is similar in that statistics might help you. It's handicapped because there is no dependent variable. You don't have costs. Well, we want to measure quality. There is no single quality metric. We're actually looking for something that's akin to that.

But we may think about going through the same process, which would be to think about statistical methods that might help us summarize the variation that we see in these 34 variables. Or let's say that we get better measures and we have 64 variables.

There's a question of how many of those variables are redundant? When you're talking about differences in hospitals that certain things just move together and that you really only need to focus on a core set of five or 10. That will give you sort of a much more manageable problem because it won't solve the weighting problem for you because the weighting problem which involves values is something that you're going to have to confront. But it's a whole lot easier to think about that if you're dealing with this relatively small set than if you are with the 34 you've got

today or the much bigger number you're going to have tomorrow.

And that's going to, I think, help increase your confidence about applying this at the individual hospital level when you can talk about -- you can understand the relationships among these things and you can see that things move together and that there's no need to measure all of them, you only need to measure a certain number of them.

MR. HACKBARTH: Has there been any research on the correlation among the quality measures? I think I heard something about that.

DR. MILSTEIN: There is research that's relevant but it suggests that intrahospital correlation of quality measures is not a good validator because the overwhelming evidence is that hospital performance varies substantially by service line. can be some service lines in a hospital that are excellent and others that are quite subpar.

DR. SCANLON: But it may be that you can get to a more parsimonious set than the 64 or whatever we're going to ultimately end up with.

DR. WOLTER: As I looked at this there's very few institutions in the country that would have 10,000

discharges and if we're looking at failures per 10,000 discharges, what really struck me about this would be the importance of linking the process measures that we require for hospitals with this kind of information because to me the idea would be if more hospitals get to 100 percent in terms of implementing evidence-based protocols that deliver those process measures over time we ought to see in these rolled up measures here improvements on a regional basis.

And so I think there could be important linkages of this to the individual institutional process measure reporting. But it's going to be hard to take this down to the individual institutional level. If you only have 300 discharges for bypass surgery, one or two cases just changes everything.

So I think it's going to be important to think about the linkages between what we require at the individual institutional level and data like this.

And then I always keep wondering when we have these reports to look at where are we headed in terms of the overall coordination of what is decided should be looked at for hospitals or physicians or whatever? And what should be requested of them? IOM, I know, Alan, you and Bob are on

committees looking at reporting. Were looking at report. And I haven't yet got in my mind the picture of how this is unfolding so that at some point in time we have some sense of who's going to coordinate this and make it a little clearer to providers how the decisions will be made and how the adjustments will be made as the evidence changes, et cetera.

I don't known if that's an issue for us or not.

MR. HACKBARTH: It is an issue that we have flagged at various times in our chapters on pay for performance. We've said that we think that the process needs to be, among other things, streamlined and providers need to get some consistency in the measures used so that they are not overly burdened, and so that we have the maximum impact on their behavior. If their efforts are being diffused in all different directions we're less likely to get the sort of progress that we would like.

The IOM, in the first of their reports, has made some pretty specific proposals about how we might achieve some of those goals. When we come back to P4P in the future, I would be open to looking at those recommendations and seeing if we want to explicitly add our support to that

approach. I think it's a critical issue for the future of pay for performance.

MR. MULLER: I echo Bill's comments about these measures really have to be salient in terms of really describing the differences in these places. A lot of them, as you say, are correlated so they really don't make much difference. That's point one.

Secondly, there's obviously a great temptation to use measures that one can get off the claims database because we have it. Some of the ones that Arnie mentioned are not available in the claims database, so they're very hard to get it. They have a lot of persuasive power but there's no way of getting them into your analysis in any kind of comprehensive way.

So I think one of the things that we may want to speak to and that we've spoken to earlier in the day is to what extent are we willing to go with measures that aren't as good as possible because you can get them off of the claims database. And this argument about administrative claims database has been going on for 30 or 40 years, so it's not so they we're talking about just today. But there's a reason why people keep going back to those

databases, because you can get them.

At the same time, if you start making profound judgments about the quality of care in a hospital and a doctor's office based on them, understanding they may only get you 60 percent of the way there, then people are going to have real problems with it. So I think making judgments, you can get good descriptive data on physician practices and hospitals, maybe not at the division of a small grouping by the claims database. You can get even better information from medical records and other kinds of case descriptive information but you can't update it and get it in any kind of consistent way.

So I think speaking to that over the course of our work would be -- is claims database good enough in terms of what we're trying to get it? And also the question of saliency, I think, is a very important one as well.

And also Nick, I would say in terms of the volume I hope we don't get immediately into kind of measuring this year by year in a kind of a payment system and so forth, because obviously things don't change in most of these settings that quickly. So one can aggregate over a longer period of time and obviously have bigger samples by looking

at this two or three years at a time. Very few things change that quickly. So one can get a three-year rolling average of some of this data and that perhaps get your sample size up a little bit more.

But those are the kind of things I think we can keep going, but I think basically trying to put this into categories that allow this discussion to go forth in a way that we can get more agreement, because I do agree that there's just too many measures out there, it's too confusing, and we are not advancing as quickly as I thought we would in terms of agreeing on what are the measures that really make a difference in terms of the quality of outcome.

DR. KANE: I was just going to mention that financial data is produced like this all of the time and we might learn a few things from it. One thing that helps when you're looking at financial data, financial ratios, et cetera. One is to say how many hospitals are contributing to that measure. For instance, you say South Central, but if there's only five hospitals that are producing that measure it would be helpful to know that.

Also it would be helpful to the distribution overall of the ratios across the region. So you're showing

the average but are they skewed? How skewed is the distribution of values?

I guess the other thing in financial measures that's really useful for a benchmark is to show not necessarily one region next to the other, but to maybe take the best quartile and show everybody next to the best quartile for a benchmark. So that ratio isn't Massachusetts to Louisiana or New England to South Central, but it is everybody else to the best quartile. Or if you eventually are going to get down to the hospital level, that hospital up against the best quartile.

And then finally, for how to weight these, I agree with Arnie that I think there are ways that you can weight these that relate somewhat to the amount of damage they actually do either in QALYs or death being obviously the worst.

But another one is the likelihood that it will occur. Because some of these have a 0.00001 percent likelihood of occurring, and others have much higher likelihood of occurring. And maybe the ones that are more likely to occur are the ones we'd rather -- once you get past death, which is obviously not a great one to have occur

-- would be the ones you'd want to focus on.

I would agree that we should definitely look into weighting them in some way other than just smoking cessation is the same thing as a puncture of your lung.

MS. CHENG: We certainly can try that. Right off the top of our heads we didn't go to incidents. We can only see the full-size numbers on safety. But keep in mind the ones that are going to happen the most often are the process measures. And so you'd be giving a great deal of weight to aspirin and very little weight to mortality.

DR. KANE: I'm inclined to keep those three categories overall separate anyway, and then go into the frequency with which they occur because I think you really are mixing apples and oranges. That's like trying to mix a profitability index with a solvency -- I think you just get garbage.

Even Bill Cleverly stopped trying to do that after a while.

MS. HANSEN: Actually, just to think about adding -- it's not probably appropriate but I'll bring it up because Dave, I'll tag onto a comment you made about staffing, whether there's a structural measure here about

what the stabbing is like, what its retention is, what its perhaps potential use of temporary types of staff. That component.

I know in some states the whole ratio of staffing has become a factor of safety.

And then going back to safety, and it may be that these are some of the areas that Ralph had said are tough to measure because we don't have them. The frequency issued that, Nancy, you just brought up about medication errors that lead to untoward events like death or falls that lead to hip fractures or death or failure to rescue. These are some areas that seem to be pretty significant in terms of the safety aspects of hospitals right now.

So again, it doesn't help because I'm offering some other ways to think about it, but eventually how to ferret down to the most salient elements, including the structural components of the hospital.

MS. CHENG: Just by the way, failure to rescue is actually in the set. And we tried hip fracture, and we can put that back in, but that was one of the measures we could get a sufficient sample size. But the good news is it happens so infrequently that it's even a smaller rate per

10,000 discharges. If that's an important one, we can put it back in.

MS. HANSEN: No, I would value your analysis on it if it is that infrequent.

MR. HACKBARTH: Okay, when you're finished writing, Sharon, you can put on your home health hat and proceed with home health measures.

* MS. CHENG: This afternoon Sarah Friedman and I are going to launch a new topic for you, and these are process measures for care delivered in home health.

We don't anticipate doing a presentation in front of you again on this topic before we write this chapter for the June report. So that's a little bit of a heads up. We would like to get your comments on this material now so that we can incorporate that in the draft of the material that's going into the June report.

In this presentation, what we're going to do together here is discuss the need to evolve the quality measure set for home health. We're going to talk to you about the first step that we've taken on that path, which was our work to gather best practices in two areas of home health.

And then finally we're going to talk about the next step in moving from best practice as a concept to process measurement.

The focused our work on two areas, fall prevention practices and wound care practices. We did that for about three reasons. We looked at the expert consensus in the literature on home health and there was a lot of a sense behind the importance of both of these practices in home health among the experts.

We also found that there was a consensus that this is a pretty important part of the home health as a benefit, keeping patients safety at home is really central to the mission and what we're trying to achieve with home health.

And finally, they have the advantage of being applicable to pretty much every patient that's being cared for in the home health sending. That gives them an advantage over the measures that we have now. I'm going to go this in a second, but in five seconds, the measures we have now only apply to patients who have a potential for functional improvement. And that's a subset, and it leaves out a chronic care population, people who are trying to be maintained at home so they can avoid a nursing home or

another care institution. And we had a sense early on in this process, we wanted to get some measures that reflected the quality of care for those folks, and they were by and large being left out by the outcome measure sets that we had.

So why add process measures? We've said a couple of times, in fact we just hit on this a moment ago, quality measurement should not be a static thing. It should evolve as more measures become available, as we can reach for different concepts. So this has been a goal that we've had as we started off measuring quality in home health.

The Commission established an agenda to do so in 2003 and we were starting to contemplate pay for performance across the Medicare program. And then in 2005 we said specifically that home health was a setting where we thought we were ready to start thinking about implementing pay for performance.

We made that decision based on our assessment of the starter set that was available in 2005. It's still available now. And those were largely outcome measures. We said at the time that we had a goal to evolve the set, had a good place to start, but let's see what else we can reach

for. And process measures were one of the areas that we identified now a year or so ago as someplace that we wanted to go with this setting.

So process measures would allow us to hit a couple of these goals. First off, it would allow us to broaden the quality of measurement. We'd be able to add patients that are getting care that is not likely to lead to their functional improvement but could reflect the quality of the home health efforts to keep them safely at home and to care for wounds that they might have.

We and the NQF, at looking at the outcome measurement sets, said that we would like to evolve the set to measures that applied to more patients. We also wanted to be able to move from the concept of clinical effectiveness, which is really where our outcomes were geared toward, and see if we couldn't reach into safety, which is another important kind of quality that's been identified by the IOM as one of the six types of quality. So we wanted to see if we could broaden into another type of quality in our next generation.

There's another sort of intuitive appeal. Process measures are a very distinct, very practical tool that says

to get a better outcome for the patients that you care about, that you are caring for, here's a really great evidence-based thing to do. So the outcome measures have the benefit of setting a goal. We would like to see more people have less limitations due to shortness of breath and a process measure can say to all the providers in the Medicare program here's a tool, here's a clinical practice that we think will get you closer to that goal. So they have an intuitive appeal as well.

By developing process measures, you're encouraging the diffusion of evidence-based practice. One of the issues that we've talked about a number of times for home health is that there's a wide variation in the practice of home health. One of the members of our panel said quite pointedly she's in a national Organization and she routinely sees the same kind of patient in different parts of the country getting very different care. So to the extent that you could develop process measures and you could say this is a good clinical practice, what it might have the benefit of doing is pulling together some of that variation that is the result of deviation from evidence-based practice. I think that would be a good step forward in home health.

And finally, this is a broader goal that we have for quality measurement, but it is a tool to encourage better information systems. The Commission has identified quality reporting and attaching it to a pay for performance system as a way to incentive the development of better information tools. Collecting, managing and reporting on the content of home health visits, which might be one of the steps towards getting process measures, would require a higher level of information, technology information system, than most agencies are currently operating.

So by contemplating process measures in a pay for performance system what you're talking about is putting an incentive behind information innovations like putting point of care computers in the hands of nurses. You're contemplating having a system that takes evidence-based pathways and embedding it right in there in the assessment activity that the nurse is doing for each one of these patients. You're talking about putting an incentive behind maybe using telemonitoring to automate the regular collection of vital signs.

So all of those together I think make an argument for why we think it would be a good idea to reach for

process measures for home health as a setting.

So to start us on that path what we've done is convene a home health best practices panel. It had two parts. One was a group of people that met here in D.C. around a table and gave us their input. We also had a review group of a similar set of experts that helped us get more out of the results of our panel discussion group.

Both of these groups included nurses, academics and many people that have a long experience with home health. We had representatives of both for-profit and not-for-profit home health providers. And we had voices of both large and small agencies to give us input about these practices. The quality measurement experts included a member of a national measurement group, a representative of CMS, and a representative of a national quality accreditation organization.

So we got a good group of people together and we put several questions in front of them. On the screen is a list of the questions that they answered for us. First off we wanted to know what's the evidence behind the best practices that we were asking them to describe? We were after two things here. One, very technically, what is the

evidence base that links doing this practice to a measurable improvement in outcomes? And the second one was we wanted to at least hear some stories about a successful implementation of this practice in the field, that agencies that were maybe a little ahead of the curve have tried this and it's worked for them in their agencies and they've seen measurable improvements in their outcomes.

Second, we asked them about the impact of the diffusion. He wanted to make sure we were focused on high impact practices.

And the last one was to get a sense from them about what kind of data would be needed. We weren't at the data point yet but we wanted to see if we could sort of steer them towards things that could be collected without an undue data burden.

So the panelists in the first area of focus shared with us these fall prevention practices. One that they were very passionate about was the use of a standard multifactor fall risk assessment tool. It should include things like the patient's detailed fall history, which is a very good predictor of their ability to remain safely at home without a fall and a medication inventory because they could look at

medications that are known to increase the likelihood of a patient falling at home.

The second one that they discussed was the use of validated techniques to measure that patient's fall risk. Panelists felt that some of the assessments that were going on today had kind of devoted into a bit of a check box exercise and they felt that there were validated strong evidence-based tools that you could use so that you got a really good sense of this particular patient's risk of fall.

One that was measuring postural hypotension. The idea here is that you measure somebody's blood pressure and you see if it changes significantly when they're standing. It's a very good predictor of their risk of falling.

Another one was to ask patients who are able to stand on one foot for 10 seconds. This did two things. One, it allowed the nurse in the home to directly observe the patient's balance or any balance deficits. And the panelists also told us that this had a very good halo effect that it allowed the patient to really understand what their balance deficit was and to understand the importance of trying to work to alleviate that to the extent possible. So it got some good patient engagement and buy-in in the

process.

The final one was to link the assessment tool to appropriate follow-up, to put that appropriate follow-up right in with the same activity when the nurse is assessing that fall risk. That practice was based on a national study group convened gold star agencies and was talking to them. What makes your outcomes consistently beat the national average? And they found consistently that those gold star agencies were embedding the practices right in their assessments.

As we alluded to earlier, we asked them to look at fall prevention and they conceded that this is a very important area for improvement. This is a place where we could do a lot better for many patients at many agencies. Falls are a common cause of re-hospitalization among home health beneficiaries and some research suggests that falls not only can lead directly to an injury that's caused by the patient falling but they can also be the trigger for a really detrimental cascade when the patient decides that they're going to self-limit activity and then they might be exacerbating underlying chronic conditions.

And so it's a place where we felt that we needed

to look and the panel felt that way as well.

The panel's practices that they brought to us we found to be consistent with a comprehensive meta-analysis that included 62 different clinical trials, randomized controlled trials, of these practices and involved over 21,000 elderly adults. Many panelists also reported that they had had success with implementing these practices in the field. And some of our panelists felt that a prevention process measure would be better than an outcome measure.

The reasons that they gave to us on the panel was that measuring falls as an outcome relies a great deal on patient self-report which might not be reliably calibrated. And also, it was difficult for some agencies to really concede that a fall that happened on a Wednesday when their nurse hadn't been there all week could really be directly attributed to the quality of their fall prevention practices. So they wanted to reach to these fall prevention practices maybe as a better way to get at this concept.

With that, that's on our fall prevention. I'm now going to switch to Sarah Friedman and she's going to share with you our panel's results on wound care.

MS. FRIEDMAN: Here the panel has identified

several areas of practices where wound care can be improved. First, panel participants stressed the need for standardized wound assessment. This includes a regular head to toe assessment where nurses identify, count and state venous and pressure wounds. Surgical wounds should also be monitored regularly.

Panelists also recommended keeping updated images of wounds in the patient's record to supplement the medical record. The agencies represented on the panel use a variety of assessment tools and agree that a comprehensive tool should investigate the location, size, drainage, and margin of the wound, as well as inspect for signs of wound infection.

The next practices are ones that are triggered by the presence of wounds identified in the assessment. If a patient has a pressure wound, the following steps may be appropriate. Offload the wounded to relieve pressure from the wound area, turn the patient and instruct the regular caregivers to turn the patient on a turning schedule.

If a patient has any kind of wound, the nurse should implement an appropriate infection control strategy as well as educate the regular caregivers about infection

control. If a patient's wound requires treating the wound bed, the nurse should use a standardized wound bed preparation technique.

Finally, the panelists discussed the need for protocols on physician communication. If the wound does not respond within two weeks or shows signs of infection, the home health nurse should contact the patient's physician.

The next slide will explain how process measures based on these practices achieve the goals for process measures presented earlier in the presentation.

Measuring these processes would broaden the scope of current quality measure sets. Because the wounds compromise the safety of all home health patients, the panelists believe that regular head to toe assessments would benefit all patients, regardless of their diagnosis or potential for functional improvement. This is the rationale given by the panelists for giving all home health patients an initial wound assessment to be followed by the appropriate link to care.

As discussed earlier, another goal of process measures is to measure an action over which the provider has direct control. These actions should be specific tools that

home health agencies can use to improve outcomes. The fact that all of the practices discussed above are currently in use at home health agencies represented on the panel suggests that these are such tools.

Panelists indicated that measuring use of standardized protocols for wound care treatment would reduce the variation in care provided by home health agencies. Analysts believe that one reason for current high variation is that some doctors routinely prescribed outdated wound care treatment rather than treatment based on current evidence. One example familiar to panelists is preparation of the wound bed. The use of a wet-to-dry wound dressing technique is frequently prescribed by doctors even though evidence and nurses' experience suggests that in some cases it is preferable to keep the wound bed moist.

Measuring home health agencies use of evidence-based treatment protocols should reduce the variation that this causes as well as encourage the agencies to engage physicians in a broader examination of best practices.

Sharon will now continue the discussion of next steps for turning these practices into process measures.

MS. CHENG: There's yet one more step that you

have to take in getting from what is a best practice to what can be applied as a process measure. You have to define specifically the patient population to which this practice applies. You have to describe very precisely what time and how often it should occur, a very specific definition of the practice itself and if there are any exemptions for patients who should not receive this care.

The process measure, as you put it together then, could be tested against the Commission's criteria for good measures. Is it reliably specified? Is it a valid indicator of good practice? And would it require unduly burdensome data collection?

CMS right now is in the midst of a contract to do some similar work and they are developing condition-specific process measures. They have a contract with the University of Colorado and what they're doing is looking at best practices and process measures for practices that are related to such things as a care for a diabetic patient or the care for a patient with chronic obstructive pulmonary disease.

The next steps that we see would be to look at the feasibility of taking some of the best practices that our

panelists identified and translating them into process measures.

With that, we'd like your input on the process that we are under and this as a potential chapter for the June report.

DR. NELSON: I really don't want to sound like a wet blanket, and I'm aware of the fact that we asked for this, but in my five-and-a-half year experience with MedPAC, we haven't gotten into the business of developing process measures or practice guidelines. And it's the kind of thing that I'd like to see AHRQ do, or CMS do, or the home health community themselves do it.

It seems to me that what we would want to focus on is the application of performance measures when they were developed by somebody else with respect to how are the data collected? What is the burden in terms of documenting all of this? And what is the linkage with payment policy?

So I think what you've done is great. I personally think they're great measures. But I'm timid about adding, about publishing something that lays an additional documentation burden on the folks that are already coping with OASIS when I'm not so sure that it's

been validated or pilot tested or going through the other kinds of things that the National Quality Forum and others do before they put their stamp of approval on it.

DR. SCANLON: I could respond to Alan by saying that you can take it in the context that the chapter, as written, identifies the need for going beyond the measures that represent improvement, recovery, et cetera, because there are other types of home health patients. It was an outside process. Maybe it's a recommendation in the text, and not a bold-faced one, is that some other group continue to look at process measures because they recognize that there is variation in terms of home health patients.

I'm disappointed, though, from a different context which is the perspective of our discussions when we talk about the adequacy of payment and we look at the distribution of margins and we see they go from zero to 40 or 50 percent, I say we don't understand the home health benefit. And this panel, though that wasn't their charge, they could have come back with information in terms of a richness of process measures, a richness of quality measures that would have told us a lot about the home health benefit, but they didn't.

I come away from this still concerned that it's about the recovery, the rehabilitation patient. It's about skilled services. The role of the aide is still a complete back black box and certainly part of the margin variation is the fact that the aides have really diminished in terms of the frequency of their services. And probably there's great variation across agencies in terms of their services.

So I'm concerned that while we talk about pay for performance at the margin, we've got this fundamental problem about the base payment that we don't understand what we're buying. And when there's a 40 percent margin I don't think of it as all efficiency. I think that we have to ask ourselves what did we get? And should we be doing a whole lot better in terms of describing what we want and then being able to measure whether we got it?

MS. CHENG: I think this is a critical part for the tone of this work. We didn't ask the panelists to tell us all of the practices that were going on in home health and we specifically asked them about falls and about wounds. So this is not the universe of good practice and of benefit that the home health service can deliver.

This was done against two backdrops. We do have

outcomes that are looking at functional improvement, which is also an important part of the benefit. We were aware of work that CMS was doing about the care of diabetics, about the care of patients with COPD, and my sense was the best way to use our resources was to not duplicate that work.

So this is really important for tone and I really don't want to put these out suggesting that this is what we think is the breadth of the scope of what the home health benefit is about. It's just two places that are important that we found best practices in.

And so I'll make sure that's a critical part of the chapter.

DR. MILLER: I just want to reinforce that because if you're upset about that -- not upset, wrong word -- we have to take responsibility for that because we directed and tried to focus the group.

But that's not to dispute your point. Your point is well taken. There is still a broad misunderstanding of what's going on in the product. So I don't want to miss the point for the tone issue.

I also want to say, on Alan's point, just so everybody tracks on how we could proceed here as a

commission, one way to look it what Sharon has done so far is to say look, we've pulled these experts together, we've looked at these two areas, they agreed they were important, they helped us talk about how to specify. And we've actually taken this a certain distance and could say this is important for someone else to go and now make the linkage between practices to process and then, depending on that outcome, we could come back in behind it and say yes, now make this part of the pay for performance or not if it doesn't happen.

I think the question on the table, and you've been very clear on your opinion of it, is -- because I don't think Sharon is saying she would put these measures, recommend using these as they stand. These have to go to a process like an NQF-type of process. It's whether someone else does it or would we, as a commission, contract and take it through that process. I think that was the question on the table.

And I think I hear your answer to that. That's not our business. We should kick that to someone else.

MR. HACKBARTH: Others?

Good job, Sarah, Sharon.

Last for today, certainly not least, is clinical lab services.

* MS. KELLEY: In previous sessions Ariel and I have discussed concerns that Medicare is not paying accurately for clinical lab services. When we talk about inaccuracy here, we mean with regard to relative prices. Medicare's payment rates were initially set separately in each of 56 carrier markets based on what local labs charged in 1983.

It was thought at the time that charges were substantially higher than costs, so the fee schedule rate for each carrier was set at 62 percent of prevailing charges for hospital-based labs and 60 percent of charges for independent and physician office labs. 20 years later it would be surprising if relative payments were accurate.

This is especially true given that the method for determining payments for new services is likely to generate inaccurate rates.

Improving Medicare's payment methodology is important because uses of services is growing. The clinical lab benefit has grown an average of 9 percent per year since 1999, reaching almost $6 billion in 2004. This is despite the fact that payments have been updated only once during

that time period.

We can expect the rise in volume and complexity to continue in future years as the range of lab tests expands and as innovations and equipment and techniques make some testing more efficient and automated. The growing prevalence of clinical practice guidelines and advances in medical knowledge will also boost the use of screening and monitoring tests, as will the implementation of pay for performance programs.

The challenge for Medicare will be to improve its payments without cost data. CMS has had some success in overcoming that obstacle in developing payment systems for other providers, namely the RBRVS for physicians and competitive bidding for durable medical equipment suppliers. Both methods could be considered for lab services.

But encouraging efficient use of tests through payment mechanisms will be more difficult. Many lab tests that are important for preventing and treating disease are underused. But at the same type there's evidence that greater use of tests does not lead to improved outcomes at the population level. This lack of relationship raises questions about whether every lab test is of value to

Medicare beneficiaries and to the program. Ariel will talk more about that in a moment. But one thing to consider as we think about improving the payment system for lab services is that tests are ordered by physicians. So labs themselves can do little to control volume.

Bundling certain physician or hospital outpatient services with associated lab tests could help control the volume of some tests but this approach may not be broadly applicable. And at any rate, limiting growth across the board in the use of lab services would not be desirable, given the fact, as I said, that experts believe many screening and monitoring tests are underused.

A final issue is the fact that no coinsurance is required for lab services.

So Ariel and I are going to briefly review information about labs and the services they provided and then we'll discuss some options for improving the accuracy of Medicare's payments.

You've seen part of this slide before. Lab services are furnished, as you know, by labs in hospitals and physician offices, as well as independent labs. And then there are institutions such as nursing facilities,

dialysis facilities, that also have labs and they're included in that other category.

Frequently those services are covered under other Medicare benefits. As of August 2005, there were more than 192,000 labs in the U.S. and that number has grown on average about 2 percent per year over the last decade.

As you can see here, physician office labs account for about half of all labs but they furnish a much smaller proportion of ambulatory lab services paid under the lab fee schedule, about 17 percent. By comparison, hospital-based labs furnish about half of all ambulatory lab tests and independent labs furnish about 31 percent.

It's important to note that relatively few labs, even in hospitals, perform all types of tests. Most labs conduct some subset of test in-house and send out other tests to labs called reference labs, which provide a broader range of tests.

Although there are over 1000 items on Medicare's lab fee schedule, the volume of tests is fairly concentrated with the top 10 tests accounting for 38 percent of total volume and 45 percent of total payments. Venipuncture accounts for an additional 18 percent of volume under the

fee schedule and an additional 6 percent of payments. So 11 tests and services account for about half of all payments.

The tests on this slide with asterisks grew more than 10 percent between 2001 and 2003, with complete blood count growing at a rate of about 25 percent per year.

MR. WINTER: As we noted in December, many of the rapidly growing lab tests are recommended by clinical guidelines. For example, complete blood count tests and potassium tests are recommended at certain intervals for patients with congestive heart failure.

A study by researchers at RAND which was published in the New England Journal of Medicine a few years ago found that many recommended tests are underused. For example, only 34 percent of patients newly diagnosed with heart failure received a CBC test within the recommended time frame and only 24 percent of diabetics received hemoglobin tests as recommended.

On the other hand, there's evidence which we'll discuss on the next slide that regions that provide more lab tests do not have better outcomes. There's a tension between these two findings that may merit further exploration.

Researchers at Dartmouth Medical School co-authored a study three years ago that found that geographic regions that provide more health services overall do not provide better quality care or have better outcomes. At our request, two of the researchers, Eliot Fisher and Daniel Gottlieb, modified their analysis to look at whether areas providing more lab tests in general have improved outcomes. They used the same data and methodology as the original study.

First they ranked over 300 hospital referral regions by their intensity of outpatient lab testing, and intensity is based on per capita Medicare spending on outpatient tests standardized for geographic differences in payment rates. Their models adjusted for differences in demographic characteristics, patient comorbidities and other factors.

First, they looked at whether areas that provide more tests have greater long-term survival and fewer hospital readmissions for three cohorts of beneficiaries: patients with heart attacks, colon cancer and hip fracture. They included all lab tests, not just those that are used for these conditions. They found that areas with more tests

per capita do not have higher survival rates for beneficiaries with one of these conditions.

In addition, greater use of tests was not associated with lower rates of readmission after 90 days. In fact, high use areas had higher readmission rates.

Second, they examined whether high use areas had lower rates of hospital admissions for a representative sample of beneficiaries from the Medicare Current Beneficiary Survey. They found that patients in those regions were actually more likely to have at least one hospital admission in a one or two-year period.

What could explain the findings that regions using many tests had more hospital stays is that both lab tests and hospital admissions could be proxies for underlying practice patterns. Regions with more intensive practice patterns are likely to have both more tests and more hospital stays. Collectively, these findings raise questions about the marginal value of additional lab tests.

If many tests are underused, why don't areas that provide a lot of tests achieve better outcomes? One possible explanation is that regions providing more tests in general may not necessarily provide more clinically

recommended tests. The original Dartmouth study published three years ago suggests that regions that deliver more health services overall have a mixed record when it comes to providing clinically recommended tests. Patients in the high health care spending areas were less likely to receive Pap smear tests than patients in lower spending areas. On the other hand, higher spending regions provided more PSA tests, which are used to screen for prostate cancer, and more lipid panel tests for diabetics.

Another possible explanation is that in high use regions tests are being done more frequently than recommended by clinical guidelines. Finally, the frequency of testing is not the only thing that determines outcomes of care. How physicians interpret clinical information and manage their patients may be more important factors.

As part of our work on physician resource use, which Niall and Karen discussed earlier, we'd like to examine variations in the use of lab tests by physicians for similar episodes. This may help shed more light on the relationship between the use of tests and quality of care.

We could also look at whether physicians who order more tests use fewer of other services. And we might

examine whether an episode-based payment might create incentives for more efficient use of services.

A final thought here is that a pay for performance system could reward the greater use of clinically recommended tests.

Now we'll go back to Dana.

MS. KELLEY: Turning now to how to improve Medicare's payment system.

As I mentioned before, a stumbling block to setting accurate payments is the absence of provider cost data and that Medicare has had some success in overcoming that in some other areas, namely physician payment and durable medical equipment. Using technical expertise from the private sector, Medicare has established resource-based relative values for physician services and relying on supplier bids to approximate costs Medicare used competitive pricing to set payments for durable medical equipment.

A regulatory approach to laboratory services would involve the development of a new fee schedule based on recent data on the resources needed to furnish lab tests. This approach could improve Medicare's payment system by better aligning payments with the resources required to

furnish tests. A method for establishing relative values for existing and new tests and for reviewing the relative values over time would need to be developed. Establishing a resource-based payment system for lab services would, in some ways, be far simpler than developing the physician fee schedule was. Most importantly, there are about one-sixth as many codes for lab services as for physician services.

But as you know, developing and maintaining a system such as that is time consuming and costly. Indeed, to keep the RBRVS up to date, CMS has had to rely heavily on the American Medical Association and physician specialty societies and it's not known whether the various clinical lab associations would be able to undertake such a role.

Competitive bidding may be a more viable option. This market approach is based on the theory that competition among labs will result in a price for tests that more closely reflects their costs than other pricing methods. To implement such a program policymakers must design market and bidding incentives to achieve a balance among Medicare goals of access, quality, choice, equity and efficiency. A bidding process that focuses solely on price, for example, might compromise access and quality.

The MMA mandated that CMS conduct a competitive bidding demonstration for lab services and CMS is now in the design phase of the process.

The lab industry has been opposed to competitive bidding. Industry organizations argue that clinical lab services are complex medical services requiring significant training, expertise and supervision, as compared to health care equipment and supplies which are usually standard and interchangeable.

Both the American Clinical Laboratory Association and the College of American Pathologists maintain that competitive bidding will compromise the quality of lab services.

The competitive bidding design currently under consideration would include all tests and services paid under the clinical lab fee schedule with the exception of Pap smears and colorectal cancer screening tests, which Congress specifically excluded. All labs with $100,000 or more in annual Medicare lab payments in the demonstration area would be required to bid. This amount would be calculated for the lab company, including all affiliates. This requirement would exclude most physician-owned labs

from mandatory participation. The demonstration will run for three years in two sites, which have not yet been selected.

The bidding process under consideration is very similar to the one used for the DME demonstration. One important difference is that bidders would be required to bid for all services with the exception of Pap smears and colorectal screening tests. By comparison, the DME demo allowed bidders to bid only on selected categories of products, such as enteral nutrition or urological supplies.

CMS considered a design that would have included only the top 100 tests in the demo, but the industry was opposed to that plan. Many labs feared that larger reference labs would be able to underbid smaller labs by offering high volume tests at cut rates, subsidizing any losses with relatively high payments for more rare tests.

Requiring all bidders to bid on all services, however, may not eliminate this potential problem. Many labs will have to bid on a substantial number of services that they do not provide in-house. Some labs fear that reference labs still may undermine smaller labs by setting high prices for smaller labs that contract with them for

services.

For the purposes of bidding, small labs could ban together to create bidding consortiums, subject to review by the FTC, which could allow labs to extend their test menus, capacity and geographic coverage.

There's a lot of information on this slide. I'll just hit a few highlights.

After suppliers submit bids, the draft design plan would use a multi-step process to select the winners in each area. CMS would calculate a single composite bid for each lab, which would be a weighted average of a lab's prices for all tests using weights based on each tests share of totally expected demonstration volume. This would have the effect of weighting a composite bid more favorably if the bidder lowered prices for items that Medicare purchases frequently rather than discounting more rare tests.

CMS would use the composite bids to rank each lab from highest to lowest and then identify a cut-off composite bid price which must be lower than the composite bid that would result from current fee schedule prices. The cut-off price would be determined using criteria such as capacity, geographic coverage, quality, the number of winners and the

distribution of composite bids.

Bidders with composite bids less than or equal to the cut-off price would be winners. Losers would receive no reimbursement for Medicare tests under the demonstration. Required labs that chose not to bid would also be unable to receive Medicare reimbursement for lab tests.

CMS would then calculate a payment rate for each test. Winning labs would be paid the same price for each test, regardless of what they bid. Medicare's prices would be set to provide winners with total revenues for all labs services that were the same or greater than the revenues implied by their composite bid, assuming the lab furnishes the typical mix of lab services.

The demonstration would include structures and processes to monitor quality and access. Winning labs would be required to report data on six different measures of turnaround time and would also be monitored on the results of proficiency testing, survey inspections, log-in error rates and physician satisfaction surveys.

CMS would also monitor five different rates of lab tests per beneficiary, including monitoring specific lab tests to ensure that diabetics and other patients were

receiving tests as recommended by clinical guidelines.

The last issue we want to talk about today is that of beneficiary coinsurance. As you know, there is no coinsurance for lab services.

The Congress has, at times, considered applying a 20 percent coinsurance which would equalize cost-sharing between clinical lab and most other Part B services. In its June 2002 report, MedPAC estimated that such a change would reduce Medicare spending by $1.5 billion in 2002. At that time, the Commission concluded that, because beneficiaries do not initiate their use of lab services, adding coinsurance would probably not encourage more efficient use of care and might pose a financial barrier to low income beneficiaries who lack supplemental coverage.

In addition, the cost of billing and collecting coinsurance might exceed the coinsurance amount for low payment tests.

So to summarize, we're concerned that we're not paying accurately for lab services, especially at a time when use of tests has been growing and is likely to continue to do so. The absence of cost data poses a pricing problem for Medicare and Congress has asked CMS to explore whether

competitive bidding would help solve this problem.

I've taken you through some of the highlights of the proposed design for the competitive bidding demonstration and more detailed information is in your written materials.

One thing to think about is whether MedPAC wants to comment on any aspects of the proposed design and whether there are other payment methods that should be examined. Another issue you may want to explore is how to encourage more efficient use of lab tests.

And finally, you may wish to consider whether, for the sake of equity, coinsurance should be required for lab services, as it is for most other Part B services.

We look forward to your comments.

DR. SCANLON: We seem to go quickly to the competitive bidding option and I'd like to express some caution, because the one thing that was remarkable about the DME competitive bidding demos was, in some respects, the tender loving care that CMS gave to those demos. And care to the extent that they couldn't replicate it on a national basis. They invested a lot to make it work.

And now we're faced with under the MMA there's the

provision that in metropolitan areas we're going to have competitive bidding for DME. And my sense is they're going to have to do it differently, and there's the potential that we get somewhat different results. And so that's something we don't have the experience with yet to know whether that's going to happen.

The second thing about competitive bidding for DME was that they selected the items, too, that they were willing to put out for bid. So it wasn't a complete DME schedule that resulted. It was savings on particular items.

If we're thinking about the way this lab demo seems to be set up, you bid on all of the services. But I wonder what that does in terms of the competition among labs. How many labs are going to be disqualified because they're not going to be in a position to bid on everything that's required?

One of the things about competitive bidding that one has to worry about is maintaining your bidders over time. Because yes, you can maybe get savings in the first round but if you don't have a healthy market where there are people that are going to come in and challenge the former winners, and that you can have former winners become losers

in future years, that starts to deteriorate.

So those are all things I think that we need to be concerned about. I'd be more cautious about waiting for the experience of the demo is an important aspect of this. But then also the experience of the demo, it shouldn't be assumed that it can be duplicated nationwide. One needs to think about how it has to be adapted in order to do it nationwide.

The other fundamental challenge, and I don't have an answer to this at all, is the issue of volume growth and how does one address that? I'm not sure that the coinsurance is necessarily going to be the effective way to do it. It's a troubling aspect, but I don't have a policy answer for you.

MR. HACKBARTH: Bill, go back to the first issue for a second, how we price accurately. You expressed reservations about competitive bidding. How do you size up the alternatives?

DR. SCANLON: Well, I think that the alternative, in terms of trying to do something that's similar to the fee schedule, the one disadvantage that we pointed out is that it's somewhat expensive.

MR. HACKBARTH: Talk about tender loving care.

DR. SCANLON: But you do it once and you potentially can cover the universe. You do it once for -- we probably have more stability in this area than we do in physician services in terms of the kinds of things that we talk about the RUC dealing with. I think we would have a lesser challenge over time in terms of trying to maintain this schedule than we do with physician services. And that would be my not well informed judgment.

But it seems that that was our principal objection to that, was that it was going to be costly to do this.

I guess I'm saying I don't want us to underestimate the cost of competitive bidding, that we'd need to look at the cost of both very carefully first before we say that we want to make a choice based on the cost of implementation.

DR. KANE: I'm just reporting a little bit about what goes on in our marketplace in Boston, but we've got two or three integrated delivery systems, one of which is this Partners Health Care System that's kind of made all the hospitals and doctors send their lab tests into the mother hospital as part of their integrated delivery system. And

then you can get access to those lab tests wherever you are out there in the doctor's office as a part of the system.

What would happen if Partners lost the bid? I'm just trying to understand what that would do to the integrated delivery systems and electronic records where they have the lab tied in to the system, if you had places lose bids?

I just think it's kind of contrary to the notion that we're trying to foster systems of care.

DR. WOLTER: That was one of my concerns, as well, as I was kind of thinking about the clinical implications of this. First of all, from the provider standpoint, most hospitals are going to have to provide lab services of some kind for intensive care and emergency room and stats. Even in clinics, our oncologists wouldn't stand for not having same-day lab results prior to infusion therapy, for example.

So the implications of trying to maintain a smaller base of lab services, if you lost a contract, would be significant for the clinical delivery of care.

And that, as far as labs go, they are a little different I think that a DME commodity in that they are an integral part of the overall clinical care of the patient.

And so if you looked at the more advanced electronic medical records that are coming along, lab is a key part of that. It's right there. The physicians can look at that, then they can jump to imaging reports, they can jump to transcription, et cetera. And if your labs are off in somebody else's system, integrating those results with the other clinical care items in the patient's history is going to be much more difficult and I think much less effective.

Even patient access to their own labs is something that is now starting to happen. If they have access to their record, including their laboratory, and it's there in an integrated way I think that has value.

And then some of the new decision-support tools that are being embedded in the new electronic medical record allow you to look at lab result trends compared to when medications were started and stopped and some things like that that will be lost, I think, if we fragment where that care is provided.

So I guess I worry about all of those things and is competitive bidding he better approach, as opposed to some administrative pricing approach if we think we have an issue?

The last point, which I have made before, in the hospital world I think there or five or six service lines where theirs is profit: lab, imaging, certain surgical procedures. Much of the rest of it is neutral or negative. And so if we can't price right across the whole range of services and we just pick on ones where there may be more profitability, we're going to have some problems eventually.

And so that's why I think over time, not just where we think we have to price down but in outpatient surgery and some of those other areas where we're seeing negative margins, I think we have to address it as a package or we're going to run into trouble eventually.

MS. KELLEY: Can I address those comments? This would only be for lab services that are paid under the fee schedule. So hospital inpatient labs would be completely separate and continue to be paid under the hospital PPS. CMS has also been concerned about this. Congress specifically excluded from the demonstration labs that are the result of a face-to-face -- labs that are connected to a face-to-face encounter, which CMS has taken to be physician office labs. So if you see your doctor and then walk down the hall to his or her lab and have the blood drawn there

that's considered a face-to-face encounter and that would be excluded from the demonstration.

But CMS has also expanded that now to include hospital outpatient labs, as well. So hospital outpatient labs would continue to be paid under the current fee schedule.

What would not be paid under the current fee schedule are what they're calling hospital non-patients. So if you go to the hospital to get your blood drawn because the physician sent you there, then that would be paid under the demonstration.

DR. WOLTER: That would mitigate some of my concerns. However, if you have a patient that you're checking some lab on at two or three month intervals, not necessarily seeing them that day, or if you're an independent physician office and happen to use the hospital for those lab services, there could be instances where that volume change really is quite significant.

I think the integration into the electronic medical record, of course, is an issue that really isn't addressed very well by all of this. And I think there's a significant issue.

MR. HACKBARTH: Dana, before we leave that, do you rough numbers on what percentage of lab services would be subject to the demo versus the exclusions? The exclusions sound very, very large.

MR. WINTER: Hospital outpatient labs account for about half of volume and just under half of spending as a share of all lab services paid under the lab fee schedule, if you go back to that slide. So if you assume that all outpatient labs are excluded, then it's about half that --

MR. HACKBARTH: What about the exception for the face-to-face?

MS. DePARLE: What's in? What's left?

MR. HACKBARTH: So 34 percent. These don't exactly match up with the --

MS. KELLEY: We have not been able to sort out the number in the hospital-based row. Hospital-based labs account for half of all services paid under the clinical fee schedule. But some large proportion of that number, of the 49 percent, is going to be outpatient. What we don't know, what we haven't been able sort out from the claims is what proportion of that it is.

MR. HACKBARTH: She said if you go to the doctor's

office and the doctor says you need to go to the hospital and get your blood drawn and these tests done, that's included under the demo.

MS. KELLEY: That would be included.

MR. HACKBARTH: If it's a hospital outpatient department, it's not.

MS. KELLEY: Right. So when you go to the doctor and the doctor says you need these test drawn, and here's a list of labs you could go to, Quest, LabCorp, the hospital down the street, then that would be something covered under the demonstration. But if you go as an outpatient to the hospital to receive some medical care and get labs associated with that care face-to-face in the hospital, then that would continue to be covered under the fee schedule and not in the demonstration.

MS. BURKE: And the labs that are drawn but sent out in a docs' office?

MS. KELLEY: If it's drawn in the doctor's office that's considered part of the face-to-face encounter, even though it's sent to the hospital or Quest or wherever.

The independent line, everything provided by independent labs is included unless it's sent to them by a

hospital or a physician office lab.

But keep in mind that labs have direct billing. So for the most part these are services provided, the independent category are services that are from when you show up at a lab facility that just draws your blood because labs have direct billing. So a physician office, even if they send it out to an independent lab, the physician office generally -- not always but generally bills for it.

DR. REISCHAUER: Can I ask how does Aetna and BlueCross BlueShield determine how much they pay for these tests?

MR. WINTER: We have not looked at this ourselves but the Institute of Medicine did in its report from 2000. They hired a consultant to look at and compare Medicare rates to private payer rates and look at how private payers set their rates for lab tests. The private plans generally, by and large, base their fee schedules on the Medicare fee schedule although they adjusted it.

DR. REISCHAUER: So we should then do ours on theirs.

[Laughter].

MR. WINTER: There were some wrinkles so in some

cases they would pay a capitated rate to a lab to do all the tests for their enrollees. In some cases they would pay a hospital more than a physician-base lab, in some cases they would pay a physician lab more than an independent lab, so there are definitely variations.

For the most part, the private plans paid more than the Medicare rate, and this again is six years ago. We haven't updated this. The one exception where private HMOs were paid a little bit less, 2 percent less for 22 tests that they looked at, and Medicaid HMOs were even lower than that relative to Medicare.

DR. CROSSON: Getting back to the concerns that Nancy and Nick raised, I don't quite get the value in this model of forcing out the so-called loser bidders. What value does that create, other than terrifying the people who are making the bids? Because if you just pay them the demo rate you end up with the same outlay; right?

DR. REISCHAUER: Bid high? What's the incentive?

DR. CROSSON: All right, because there's no way to direct the business.

DR. REISCHAUER: You could pay them 2 percent less than the demo rate or something like that, but let them stay

in, sort of a penalty.

MS. BURKE: I'd like to get back to a point that Bill raised earlier on and get away from the payment piece of this, and that is to the more fundamental question of the volume issue. Essentially, what the chapter focuses on largely, and what we've talked about today, is how one sets up a method of payment for a fraction of the total test. One wonders at some point the value of doing this, if essentially you're excluding 70 percent of the tests from the demonstration, which at least in terms of the payments is what it looks like.

But setting that aside for a moment, I think there is also a critical question which is how are we going to essentially begin to influence people's use of tests in an appropriate way? I was troubled that we have this strange extreme where there is overuse but apparently underuse of what ought to appropriately be done and overuse of things that ought not. So it's not like they're just trying to get more because they're not getting more where they could in fact get more appropriately. They're getting more and not doing what they ought to do where they ought to do it.

So one of the things I'd like us to begin to think

about and figure out how we want to opine on is the question of how we also begin to encourage and influence behavior that is appropriate use of lab tests that are linked to quality and to outcomes. And I think whether we want to deal with that separately, whether we want to reference that point, but I think this has to be linked as well to incentivizing good behavior which is how all of our --

DR. REISCHAUER: But that might have nothing to do with what we pay labs.

MS. BURKE: I understand and appreciate that.

MR. HACKBARTH: In fact, one place that it might lead back to is physician resource measurement, and this is one of a number of --

MS. BURKE: That's exactly right. But I would hesitate, and Bob's exactly right. How we pay a lab, particularly since you're excluding all these other labs, won't impact what people order. That's an independent decision. But I would hate to have us talk about and get involved in the payment without mentioning that we are equally as concerned about the appropriateness of the test and the volume and that that's something we need to focus on as well, even though this won't influence it directly

because the labs are doing with the labs get, but they're not ordering it.

MR. HACKBARTH: Others?

DR. REISCHAUER: Should we have any kind of discussion on of coinsurance and how much appetite there is for coinsurance?

MR. HACKBARTH: Sure. Go ahead.

DR. REISCHAUER: Not much.

[Laughter.]

MR. HACKBARTH: For the same reasons as were covered in 2002.

Just for my education, can I just ask Ralph and Jay and Nick a question? You all have labs onsite and you do some of the work onsite. Do you also take some of it and ship it off to a reference lab?

MR. MULLER: We don't because we're so big but part of -- no. And obviously the hospital-based stuff, the volume is so high that really the marginal cost is incredibly low. But it's all the new tests that are being developed right now in and the next years that are going to get very expensive. The very specific, these genetic tests and all of the new drugs that are coming down that are much

more individually specified and so forth are much more complicated tests. They're not likely to be that Quests and LabCorps are not likely to get in that business because it's not the kind of volume stuff. They do very well on the high volume stuff, where again the marginal cost is basically zero when you look at it.

So I think, in terms of the discussion we're having, it would be interesting -- I would like to see this broken out a little bit more. While these volumes one here on the complete blood counts and stuff like that, you get up to 44 percent here. But I would hypothesize that a lot of these tests that are pretty small in number are going to get fairly big in terms of payment because these tests can be $1,000 tests in true costs to run, as opposed to fees and so forth.

So a place like us, the big places tend to have their own labs. But doctor's offices, by and large, send them out to the Quests and LabCorps and so forth. Smaller hospitals send them out to the Quests. That's the business that the Quests and LabCorps are in.

MR. HACKBARTH: I'm trying to get a sense of the cost curve here. When you say that you are large enough to

do it in-house and your marginal costs are low, they're competitive with Quest or one of the big reference labs?

MR. MULLER: Yes.

MR. HACKBARTH: So the level of scale required is not -- you're a big institution.

MR. MULLER: But then, we're one of the 10 or 20 biggest places in the country in terms of that kind of scale. But the community hospitals in the area, and others can speak to this, they tend to send a lot of theirs out, on the more routine stuff to the national competitors and on the stuff that's less used to places like us or Hopkins or Duke, et cetera. We are reference labs for regional places.

MR. HACKBARTH: So the community hospitals, I guess they're choosing between Quest and the other one or two competitors based on price?

MR. MULLER: They'll generally go to the Quests and so forth of the world because they're just more price competitive and they want that kind of business.

MR. HACKBARTH: So they're able to parcel out their business and say part of it we need to have integrated into our practice but another piece we can send outside for a good price? And so why can't Medicare do that?

MR. MULLER: As Nick said, some stuff you need right there in your emergency room, your ICUs and so forth. So you have some minimal capacity. But by and large, they're not going to make these -- these lab systems, I mean they're $10 million, $20 million investments. So the really big places can do that. The 50-bed hospital is not going to be able to afford that and so forth. So they keep a pretty much more modest thing.

So the Quests and LabCorps really, whatever the 5000 hospitals in the U.S., an awful lot of them send things out but I would think top 500 or 600, in terms of scale keep most of it in. They may still send something out at the margin, but the really big places hardly send anything out.

DR. CROSSON: It's essentially the same answer. We have three large reference labs, two very large ones in California and one in Oregon. And we don't send anything out. We do send tests from some of our other areas to those labs, but we essentially have internalized everything.

DR. WOLTER: We partner with the Mayo Clinic to provide reference lab services in our region and then we ourselves provide lab services, for example, to a number of the critical access hospitals that we support.

DR. MILLER: Just to say a few things about these comments and where to go, as we looked at this we're pretty convinced that the current pricing structure and how we've arrived at all of this and how we're maintaining it currently doesn't make a lot of sense. And you have this inherent tension of lots of volume but not so clear that it's the right volume. So that's what brought us to this topic.

Then leaning towards the competitive bidding I would say it just a little bit different. The Congress has kind of pushed people down this road, which is why we paid some attention to it. And one way to think about what's happened here is some of these questions about why is the demo being designed that way. Or how does the demo address this issue? These are things we could ask to be addressed if we think that there are anomalies or problems with the demo.

I'm not trying to push you into the demo. But by law they're going to do it. So we might want to say if you're going to do it this would be a better or worse way to design it.

So one way to think of all these comments is maybe

we should think through this and make some recommendations or some suggestions on how to run the demo to deal with these kinds of issues, although I think I heard pretty clearly not a lot of enthusiasm for competitive bidding, at least as it stands.

My last point was going to be a point that was just made, so I hate to be redundant. The volume issue we are at least so far thinking of that in the context of the physician episodes and looking at it through the measurement of physician resource. And that's sort of the way we were figuring we would chase that. And we'll be sure that these two things refer to each other.

MR. MULLER: I would also urge us, just like 10 or 12 years ago one could have anticipated that the devices would explode in the sense of use and forth. These new tests are going to do the same thing in the next five to 10 years. So the big money is not going to be in blood counts. It's going to be in these very specified, highly specific tests that give you a lot of advance in terms of therapy and treatment. They're going to be very highly valued by the patient and the doctors because they can tell you how to proceed with therapy.

And one could arguably then make cost arguments that by targeting therapy much more precisely they will save a lot of money for the system. That by and large most things we add to the system head on all fronts. But by and large, at least conceptually, these tests should save quite a bit of money in terms of diagnosis and treatment because they're just much more powerful and profound.

But these things are not a dollar test. They're the big ones. I'd keep my eye on them over the next five or 10 years because there's a lot of competition coming and a lot of venture money coming, which is always a good sign, a lot of venture money coming to this field.

DR. MILLER: Ralph, one more iteration on this, and I know we've got to stop. But one of the things I think that really scares us about this is it is like devices and things like that, these will come on the market, an administered price system will have no way of pricing it, and the information imbalance will be entirely held by whoever manufactured the test. And the system, like technology, they will be able to extract very large payments out of the administered price system.

So I think that's some of what makes us nervous

here is we don't have a good way of capturing that phenomenon.

MR. MULLER: I'd just say I'd spend more time then in trying to figure out how to price blood counts.

DR. SCANLON: I agree with you that the administrative pricing system is going to be challenged at that point but potentially competitive bidding is also going to have challenges. it's going to depend upon how the suppliers in that market develop. There is the potential that we would ultimately decide that we want to deal with these really small things where the marginal costs are very, very low versus these rarer things where it's very high in different systems.

And I agree with your point about the demo is going to happen. We should be positioning ourselves to learn the most we can from it.

MR. HACKBARTH: Okay, anybody else?

Okay, thank you very much.

We are now to the public comment period. We'll have a brief public comment period.

* MR. DOUGHERTY: Hi, I'm Bob Dougherty from the American College of Physicians. I will keep my comments

brief because I see everybody is about to head out.

I wanted to comment on the care coordination presentation earlier today and the discussion, which I thought was excellent. A lot of the concepts that were presented in terms of the physician role in care coordination, integrating that better with the disease management companies are things that the American College of Physicians have been talking about in our advanced medical home paper.

A few observations, though. The two options you put forward, the staff put forward, one was the more integrated large group. And the other was the care management organization plus the small physician practice.

The other way of looking at it is to create a model that works with the small physician practice epicenter, and that practice then may have arrangements with disease management companies and others to provide the full spectrum of services needed.

The concept we're looking at is the process where practices would qualify and be recognized as advanced medical homes and they would take on responsibility for full care coordination, not just kind of having that kind of

disease management function in place either internally or through an arrangement with one but, they would be responsible for the resources used. There would have patient-centric services like ease of access scheduling. They would use health information technology to measure and report quality. And they'd be accountable for the quality that they provide, the total cost of care they provide, and patient experience measures.

It's a different way of trying to transform small practices by using care coordination to get and provide the kind of care that we think patients want and will particularly be useful for patients with multiple chronic diseases, although it may work very well for patients with acute illnesses, as well.

So my suggestion is if you think of the continuum from the large integrated group to a model that puts the resources on the care management company and says plus the physician, that you think of a model where you can really change the reimbursement structure to enable practices to use technology and office redesign to coordinate and manage and arrange for the care of their patients where the physician has that responsibility.

And in terms of control over resources, the physician has control over a lot more resources, that was discussed earlier, than disease management companies have.

So again, it's a suggestion and it's something we'd like to talk further with you about.

MR. WATERS: Good afternoon. My name is Bob Waters. I have actually spoken to you previously on behalf of the American Association of Bioanalysts.

This afternoon I'm here on behalf of the Clinical Laboratory Coalition. The Coalition is comprised of laboratory groups that represent the full spectrum of health professionals and laboratory facilities that are involved with the nation's Medicare population. It includes the American Association of Bioanalysts, the American Association of Clinical Chemists, the American Clinical Laboratory Association, American Medical Technologists, the American Society for Clinical Laboratory Science, the American Society for Clinical Pathology, American Society for Microbiology, AVAMED, Clinical Laboratory Management Association and the College of American Pathologists.

We don't agree on a lot of things. But we do agree on several of the key issues that you actually, I

think, are going to have a major impact in terms of where Congress heads in this direction.

First of all, we believe it's important that MedPAC places this issue in its proper context. Laboratory tests, I think, are conceded by almost every medical professional I know as being a critical component for not only diagnosing and treating the patient. But they've been recognized as an important part of clinical practice guidelines. Laboratory services are also very cost-effective and provide enormous value to patient care.

In recent years Congress and health quality organizations have recognized the value of laboratory tests. In fact, 80 percent of the clinical evidence- based guidelines for the most costly disease conditions specified the necessity of ordering clinical laboratory testing.

Congress in a recent years has also taken some action to actually increase laboratory testing through expanded screening services such as a PSA, diabetes, colorectal cancer and cardiovascular health. Many of these were mandated as recently as the MMA, which was just passed.

It's actually in the public interest for Congress to provide more rather than less of these valuable

preventive screening tests.

Now as far as the issues you're considering, a couple issues that come at hand. One is what happens, why is the aggregate number -- what drives costs in the laboratory area?

Congress and the executive branch have done a very good job controlling price in the laboratory field. Unfortunately for us, laboratory payments have not remotely kept pace with inflation. Overall, Medicare fees for laboratory services have been reduced by 40 percent in real terms between 1984 and 2004. Our statutorily mandated inflation updates, we have not received them in 11 of the last 15 years. The national limitation amount that controls laboratory tests has actually been ratcheted it down.

So this is a reduction in real terms. This is not an anticipated growth or an increase that we thought we might have got that we would like to have. It's been an actual reduction in the amount we get.

We have succeeded in controlling the price for laboratory tests.

There are a few provider groups who have been asked to repeatedly absorb similar real reductions and we

appreciate that MedPAC has recognized that fact, and I think it's been pointed out in some of the briefing materials pretty well.

As you look at solutions to what should be done in the laboratory area, we would urge you to be wary of precipitously moving to any type of one-size-fits-all or type of solution that sounds good, is a nice sound bite, but has never really been tested.

This model is actually a radical departure, competitive bidding, from the current system. And it's a model that has not yet even been designed, much less tested. To move to the implementation phase, which is actually suggested in the President's budget this year, without designing it, testing it, could have serious and irrevocable consequences to this segment of the market.

Clinical laboratory services are just that. They are a test. They are a service. They are not a commodity. They are not a crutch tip. You can't measure them as they come across the assembly line to make sure they're all uniform and done in the same manner. These complex services require significant training and expertise to perform and interpret accurately. And the end goal of positively

affecting patient outcomes could be seriously jeopardized if the system is not designed correctly.

The Clinical Laboratory Coalition has grappled for a number of years with the issue of competitive bidding and we've become increasingly convinced there has not yet been designed a bidding model that could accurately take into account a number of the objectives that need to be dealt with. And that includes ensuring that the laboratory services are fees below the current reimbursement rate while simultaneously maintaining quality and access of care and keeping pace with improvements in diagnostic technology, and ensuring that all geographic settings and service delivery settings such as nursing homes continue to receive the range of highly qualified testing that's essential to caring for those patients.

In summary, laboratory testing plays an absolutely essential part in the delivery of health care quality. Laboratory tests provide physicians with objective data that they absolutely need to properly diagnose patients and by equipping physicians with critical information, laboratory tests ultimately will lives and reduce overall health care costs.

I'd like to thank you for the opportunity to speak on behalf of the Clinical Laboratory Coalition and I would like to add just a couple of comments with my AAB hat on, if I could, in response to some of the points that were recently raised.

I actually agree with some alarm the possibility that people might move down this path toward competitive bidding. And I would urge you that if you were speaking to or going to provide a report to the Administration in terms of things that they ought to look at, in terms of designing any model in this area, you ask them about five critical questions.

One, who's going to determine market share? What part of the government is going to decide how much my community laboratory gets? How much a large national laboratory gets? How much goes to the hospital? That has never been answered in 20 years of trying to design competitive bidding in this area.

Secondly, who's going to ensure that nursing homes get service? That they're not redlined, that they're too hard to serve and they're too costly.

Third, who's going to protect against low-

balling? Some people have deeper pockets than other and they can come into a market for three or four years and give you a low ball price just to clean out the competition.

Fourth, what will the new market look like when you're done? You may find out that the market looks radically different. What you've got is you've got only one or two labs to choose from. You still have to have a regulated price because now you've winnowed the competition.

And finally, please, please don't move precipitously to implementation in this area before you know what you're doing.

We appreciate your thoughtful consideration and your indulgence with my long comments.

Thank you very much.

MR. HACKBARTH: Okay, we reconvene at 9:00 a.m.

[Whereupon, at 5:12 p.m., the meeting was recessed, to reconvene at 9:00 a.m. on Friday, March 10, 2006.]

MEDICARE PAYMENT ADVISORY COMMISSION

PUBLIC MEETING

The Horizon Ballroom
Ronald Reagan Building
International Trade Center
1300 Pennsylvania Avenue, N.W.
Washington, D.C.

Friday, March 10, 2006
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.
NANCY-ANN DePARLE
JENNIE CHIN HANSEN
NANCY KANE, D.B.A.
ARNOLD MILSTEIN, M.D.
RALPH W. MULLER
ALAN R. NELSON, M.D.
WILLIAM J. SCANLON, Ph.D.
DAVID A. SMITH
NICHOLAS J. WOLTER, M.D.

AGENDA PAGE

P R O C E E D I N G S

MR. HACKBARTH: Good morning, everybody. We have three sessions today, one on cost effectiveness, and then two about trying to get prices more accurate for hospice and physicians, respectively.

Nancy, will you do introductions? Thank you.

* MS. RAY: Good morning. I'd like to welcome Dr. Neumann and Dr. Cohen. The three of us are going to be talking to you about cost effectiveness this morning.

Recall last spring we had an expert panel that included Dr. Neumann which discussed the use of cost effectiveness by Medicare. One of the issues that was raised was the lack of standardization of the methods and assumptions of cost-effectiveness studies, so we asked Dr. Neumann to look at the methods and assumptions of cost-effectiveness studies for two Medicare coverage services. They are going to go ahead and present their analysis and results. Following their presentation I'm just going to give you a couple of additional ideas for you to discuss.

Peter Neumann and Josh Cohen are with the Tufts University School of Medicine's new Center for the Evaluation of Value and Risk in Health. Dr. Neumann is the

director of the center.

DR. NEUMANN: Thank you very much, Nancy. Good morning, Mr. Chairman, and members of the Commission. As Nancy mentioned, my name is Peter Neumann from Tufts University New England Medical Center and I'm here with my colleague Dr. Josh Cohen, also of Tufts. We're pleased to be here today to talk to you about cost-effective analysis and the Medicare program.

As many of you may recall and as Nancy mentioned I was here about a year ago to talk about the challenges and opportunities in using cost effectiveness on a panel with Dr. David Eddy, and how cost-effectiveness analysis might be used to inform coverage and reimbursement decisions for the Medicare program.

As we highlighted last year, cost-effectiveness analysis offers a potentially valuable tool to help target resources more efficiently and to avoid paying for health care that offers little or no benefit for the dollars expanded. Medicare has chosen not to use this tool in the past for possible reasons we discussed last year; namely, that it is perceived as a tool for rationing health care in ways that may be politically unacceptable, but also because

of ongoing concerns about the methodology and the feasibility of using the approach.

One of the key challenges that Dr. Eddy and I discussed and that was mentioned by the commissioners last year pertained to the quality of the methodology of cost-effectiveness analysis and whether existing studies are robust enough to rely on for Medicare decisionmaking. Subsequent to that discussion, my colleagues and I were asked by MedPAC staff to review and analyze cost-effective analysis for selected Medicare services.

What we'd like to do this morning is discuss our methodology, our findings on the cost-effectiveness of two selected services, colorectal cancer screening and implantable cardioverter defibrillators, and then offer some concluding observations. We look forward to your feedback and questions.

Our objective was to evaluate the potential for use of cost-effectiveness analysis in the peer reviewed literature to characterize the cost-effectiveness of major services covered by Medicare. Our specific research questions were as follows. What methodologies and assumptions are used? To what extent is there concordance

of assumptions, methodologies, and results across studies? And three, can influential study assumptions be identified?

We emphasize at the outset that our purpose was not to delve into technical aspects of colorectal cancer screening or implantable defibrillators, per se, but rather to discuss more broadly what these case studies tell us about the cost-effectiveness methodology and how it might help Medicare decisionmaking in general.

Per the scope of work agreed upon with MedPAC staff we first identified four candidate Medicare-covered services for evaluation. Our intention was to identify examples that included pharmaceuticals, medical devices, surgical procedures, diagnostic procedures, and cognitive services covered by Medicare. The aim was to explore the completeness of the cost-effectiveness literature for these four services and then to select two for further exploration. In conjunction with staff we selected four: colorectal cancer screening, implantable cardioverter defibrillators, positron emission tomography, PET Alzheimer's disease, and erythropoietin for cancer patients undergoing chemotherapy.

Just to orient you a little bit to the cost-

effective analysis and the cost-effectiveness ratio, cost effectiveness is a word that's sometimes used loosely and generally. We'll use it specifically in this presentation. The cost-effectiveness ratio as we define it has, in the numerator, the incremental costs associated with a new technology versus existing technology or services, and in the denominator, life years gained or quality adjusted life years gained. There are other ways to measure cost effectiveness but we used this kind of standardized ratio in this talk today.

This slide summarizes the results of a Medline search for cost-effectiveness analyses that report cost for life year, or cost per QALY, quality adjusted life year, for these four services. It shows you something about the availability of studies and the completeness of the literature for these four services.

As the table shows, we found 26 cost-effectiveness analysis for colorectal cancer screening, 14 for implantable cardioverter defibrillators, four for PET for Alzheimer's disease, and five for erythropoietin in cancer patients. Based on this information we selected colorectal cancer screening and implantable defibrillators as the two services

we would examine in detail.

We then examined key methodologies and assumptions used in each study to get a sense of the quality and completeness of this literature. Let me to briefly review each of these categories. The funding source. Most studies explicitly report the funding source with government funding the most common source. The type of model used in the analysis. For colorectal cancer, screening a type of model called Markov modeling was typically used. For implantable defibrillators, statistical comparisons were used.

In terms of identifying the software, typically software was specified for the type of simulations conducted, though this was less so for other types of models.

The perspective or viewpoint of the analysis. Almost all studies used a health care perspective focusing on the health care costs in the studies, although this is not always reported, and some investigators report a societal perspective in which total societal costs and benefits are noted. In terms of the costs used, almost all studies included only health care costs, excluding non-health care costs. In terms of discounting future events,

most studies discount future costs and benefits at a 3 percent discount rate.

In terms of the clinical outcomes, most studies measure outcomes in terms of net life years rather than net QALYs gained. In terms of uncertainty or sensitivity analysis, most studies include univariate analysis whereby one variable at a time is varied. Probabilistic methods whereby many variables are varied simultaneously are generally not used.

Finally, we examined the extent to which analyses reflected recommendations of the U.S. Panel on Cost Effectiveness in Health and Medicine. As the slide shows, most studies discounted costs and benefits. Most studies use life years rather than QALYs, as recommended by the panel, and most studies used the health care payer perspective rather than a societal perspective as recommended by the panel.

DR. COHEN: I'm going to step you through some of the findings that we had. First on the concordance of the methods used across studies, first for colorectal screening, we found that discounting approaches tended to be similar, the perspective was also similar. The health care payer was

typically used. And even if they used the term societal, usually the way we read it, the study, they were using the health care payer.

In terms of model structure, 11 of the studies that we looked at in colorectal cancer screening used Markov modeling, and that's a type of simulation that can be to extrapolate beyond empirical measure of results, such as in a randomized controlled trial. Fifteen others used other types of modeling, including statistical comparisons, which would be essentially more directly taking the results from the empirical studies.

Comparisons across studies were complicated by differences in the interventions, comparators, and populations analyzed, and the united used to quantify benefits. To compare different values from different studies you'd have to really match on all of those things and there's a lot of different combinations for colorectal cancer screening.

So there's two slides now which you can look at simultaneously in your handout and I'll talk a little bit about them. This slide shows results for the CRC screening. There are lots of different interventions evaluated, as I

said. Our analysis in terms of comparing the values from different studies, we limited that to those that compared screening to a no-screening alternative, just so we could get some numbers to compare. Without this constraint it's difficult to find multiple estimates that match both the intervention technology and the frequency, and the comparator technology and the frequency.

On this particular slide you can see that the alternatives examined, including colonoscopy, CAT scan colonoscopy, DCBE, FOBT, sigmoidoscopy, and combination therapies. The slide shows that the variation in frequency of screening examined and the variation in the results from $3,000 for some options to $26,000 in others.

This slide shows the concordance across studies. In the right column what you have there is a statistic called the coefficient of variation. All that is is the standard deviation of the values in that particular set divided by the mean. It gives you an idea of how much variation there is. The values you see there are relatively small given that you're taking these values, these cost-effectiveness ratios from different studies.

Even more importantly we think is that if you look

at the values qualitatively the concordance is pretty good. The values fall below typical benchmarks, even $50,000 per quality adjusted life year. They fall into that category. So if you look at them collectively, even though the numbers they may vary, they give you the same sort of policy result.

Which of the assumptions were most influential to the results? We found that studies tended to evaluate varying set of assumptions in the sensitivity analysis. Now that's not necessarily an indication of how much the methodology of the studies varied. But when they went and looked at how different assumptions influence the results it was more difficult to compare their sensitivity analyses.

Certain assumptions were evaluated in a relatively large number of studies and the general implication therefore is that they are important. For example, polyp dwell time and the diagnostic test sensitivity and specificity. Even reading the studies you can pick that up from the text even if it doesn't come out directly from the numbers.

Now turning to ICDs, we examined concordance in a similar fashion as we did for CRC screening and we found that discounting approaches were similar across studies.

Moreover, the perspective taken was similar, usually the health care payer. There was some variation in model structure. Five studies used Markov modeling and eight used statistical comparisons, and that was somewhat flipped from what we saw in the CRC screening. Here was there was more use of statistical comparisons and that's because there was a richer set of randomized controlled trials to draw on so it was easier to use the directly available effectiveness information.

This study summarizes some of the results. In the case of ICDs there were many fewer types of comparisons. So on this slide generally you had three types of comparisons, ICDs versus pharmaceuticals, ICDs versus no therapy, or pharmaceuticals versus no therapy. So you had a larger number of data points to directly compare.

Now the concordance here is somewhat less impressive than in the case of CRC screening. Note that the standard deviation in the first row is greater than the mean. However, when we look more closely at the values, if you look at the report that we submitted you'll see that that is driven in particular by a couple of data points. I seem to remember one was around $200,000 per quality

adjusted life year and another was on the order of $600,000.

If you even look more closely at what's driving these differences, it's really differences in the assumed effectiveness of the device. There was a recent paper by Sanders in 2005 that provides some indication of what gives rise to differences in ICD effectiveness estimates. Of the eight RCTs in the Sanders paper, six translated into cost-effectiveness ratios of $34,000 to $70,000 per quality adjusted life year, indicating reasonable value for the money. Two of them, on the other hand, the other two randomized controlled trials are associated with increased mortality risk and hence led hands led to the finding that ICDs were dominated by the control group.

Now those two particular randomized controlled trials, there are different interpretations as to why they gave such different results. The CABG trial, patients had undergone revascularization before implantation of the ICD and it's thought that that procedure may have achieved the available benefit to this population leaving no incremental benefit to the ICDs. The DINAMIT trial, the second one that yielded this increased mortality risk for ICDs, the implantation was done relatively soon after the event,

compared to other trials and for whatever reason, for example, the heart may not been sufficiently strong for the procedure, that timing may have affected the benefit.

The bottom line here is that just as in -- the cost-effectiveness numbers can vary a lot but that's because underlying those cost-effectiveness numbers, especially in the case of ICDs, are big differences in the effectiveness estimates. So of course we're all familiar with how hard it is to deal with that and that translates into this arena as well.

In terms of the influential assumptions, several assumptions were influential. Most of the variation is due to the assumed effect, just as I said, and because that's part of your cost-effectiveness ratio it translates into differences in the cost-effectiveness number that you get.

DR. NEUMANN: So we'd like to end with some observations and general conclusions about these particular services, but also more generally about the prospect for using cost-effectiveness analysis in the Medicare program. These kinds of analyses we believe show both opportunities and challenges in using cost effectiveness in Medicare.

On the one hand, it shows that for high profile

and potentially high cost Medicare reimbursed procedures there are numerous cost-effectiveness analyses in the literature and these studies provide a ballpark estimate of the costs and clinical consequences of using these services. The information could be used to help inform Medicare coverage and payment decisions. Other public payers, internationally certainly, have incorporated such information into their decisions.

The studies also reveal challenges in using the information. There's variation in the methods used, for example, in the costs considered. There's also variation in the populations, the comparators, and the clinical data underlying the analyses.

Despite this variation, the literature does provide a useful range of estimates. Also we believe that it's important to recognize that even clinical studies, even rigorously done randomized controlled trials of a particular technology or service often suffer from these same problems, that is, variation in the design, different populations studied, different comparator interventions and so forth.

Finally, for decision-making purposes, CMS, Medicare may want to undertake its own synthesis and review

of existing literature and its own analysis of that literature on a case-by-case basis. That is to say, simply taking numbers from the literature is likely not going to be enough.

Thank you very much. We look forward to your questions.

MS. DePARLE: Thanks, Peter and Joshua, for another really interesting presentation.

One thing that occurred to me given your last point, Peter, about CMS may want to do its own analysis is, at least for ICDs and maybe for the other things that you studied here, weren't they already covered by Medicare before the analysis was done? So I'm getting deep into operations here before we've figured out how to do this, but how would CMS -- it sort of a chicken and egg thing. Most manufacturers would argue, we've met the FDA standards, we've shown this is safe and effective. You need to let it be diffused now so we can see what it does with other populations. So how would you then go back and make these kind of analyses and change it?

DR. NEUMANN: It's a very good question and I think not only for ICDs but other technologies this often

happens, that a technology is approved for the marketplace by the FDA. The question then becomes for Medicare, how do cover it? In which populations do you cover it? How do you pay for it?

With ICDs the case was, it was covered but then additional clinical trials were done raising the possibility of covering this technology in expanded populations; namely, prophylactically for people in primary prevention. A series of clinical trials were done in 2000, 2001 and 2002 and then Medicare made a decision and decided to expand the indication, expand the populations for ICDs. They went back again after an additional trial was done in 2003 and 2004 and in January of 2005 expanded yet again.

I think it indicates that simply having the FDA approval is not enough for the specific Medicare decision in terms of populations, and also perhaps in terms of exactly how you cover it. Colorectal cancer screening is covered differently depending on the risk group, for example.

MS. DePARLE: I guess I'm a more raising a -- I agree and I assume you agree that the process they followed on ICDs was appropriate, the looking at the evidence. But what if they did the kind of analysis you're suggesting and

it turned out that something didn't meet the standard? Would you suggest then they should go back and withdraw coverage or change the -- because then you have doctors using it, people believing in it. To some extent, lung volume reduction surgery is an example of that, which you didn't look at.

DR. NEUMANN: Right, we didn't look at but we have looked at it in other studies. I think it raises all kinds of challenges for Medicare. What do they do? They could do coverage with evidence development as they're trying to think through now. They could try to go back and look ask for another clinical trial or have a clinical trial done as they did with lung volume reduction surgery.

That's an expensive proposition and a long term in terms of time decision, and one that may not be practical for a device like ICDs which is already out. But you're right, in terms of the physicians starting to use it and how Medicare makes decisions and tries to limit coverage or expand coverage I think is just an ongoing challenge that they need to decide on a case-by-case basis.

DR. MILLER: Can I also just add one thing to that? I just want to remind the commissioners that the other

thing we've said about this whole line of discussion is it doesn't have to always be about coverage. It could be about whether your payments are set differently.

Also Glenn has made the point in previous conversations that depending on research, let's say after it's been disseminated you get results that suggest it's not as effective, you could use that information for pay-for-performance purposes and pay differentially on who does and who doesn't use these types of things.

MS. DePARLE: I remember this and we had a discussion. I've often thought we spend a lot of energy now I think on new technologies. We spend very little on some of the things we're spending hundreds of millions of dollars on every day that we don't know if it works or perhaps we could know that it doesn't work that well. So payment would be a good way to influence that.

MR. HACKBARTH: Nancy, I apologize. I forgot that you had some additional comments, so let me go back to Nancy.

MS. RAY: That's okay. I just wanted to pick up on Peters's last point about CMS undertaking its own synthesis. As everyone is well aware, CMS does consider

clinical information when making national coverage decisions and the agency is increasingly linking those decisions to collecting clinical data in registries, for example. The agency does not explicitly consider cost information or cost-effectiveness analysis, and Peter and Josh's analysis I think raise some issues about the consistency of methods and assumptions across studies.

Peter also referred to the Panel on Cost Effectiveness in Health and Medicine. They made a series of recommendations back in 1993. It was a panel convened by the U.S. Public Health Service and I think there was 13 non-government scientists on the panel, and they made a recommendation about the use of a reference case, which is a set of standard assumptions and methodologies that studies should use. And they did so in order to improve the comparability of analyses.

Revisiting these standards by some public groups, including Medicare as well as private groups, is one option here to think about. Doing so might lead to even more improvements in the consistency of methods and assumptions across studies.

Your briefing paper also raises some issues to

think about if Medicare were to try to develop the infrastructure to consider both clinical and cost effectiveness. Considering effectiveness information could mean reviewing the information just like what Peter and Josh have done for us. It could also mean conducting a cost-effectiveness study when the literature does not provide a clear indication of the effectiveness of the service.

Your briefing material raises three issues to think about. There are clearly more to think about if Medicare were to move forward and develop the infrastructure. The first question would be who would sponsor the research? It could be Medicare or it could be Medicare with other public payers like the VA, as well as private groups, private employers, private purchasers, private payers.

Who would conduct the research? CMS is one possibility. They have some capability, but they do ask for assistance from AHRQ and from their coverage advisory committee when making national coverage decisions. AHRQ has developed some infrastructure. They have set up 13 evidence-based practice centers that they use when they are conducting their comparative effectiveness research under

the MMA 1013 program as well as their technology assessments.

A third option -- of course there's more here -- would be one or more independent groups to conduct the research.

The last issue that's raised in your briefing material is who would fund the research. Discretionary federal or private funding might be vulnerable to uncertainty. One researcher suggested a method that is not linked to either annual federal appropriations or discretionary private funding.

So that concludes my other additional topics you may want to discuss.

DR. REISCHAUER: Thank you for the paper. I thought it was really excellent and the presentation was a very way of concise way of summarizing what took a lot longer to read.

The question that I had when I was reading this last night was one of age. Are these studies all done of people who are 65 and older or not? For a lot of this type of analysis, these ratios, I would think would change dramatically as one aged. That's question number one.

Question number two is, would I be wrong in assuming that the failure to use social cost is less of a problem when we're talking about Medicare than if we were talking about the working age population, simply because the major component that's left out is foregone income and out-of-pocket spending, which is also probably less for the Medicare folks than for others.

DR. COHEN: In terms of your first question on the age and whether we limited somehow the inclusion criteria for studies, we did not. For ICDs, that really wasn't even a criterion that was specified in studies. For colorectal cancer screening it's a bigger deal.

There were some studies, not a huge number, that looked at how differing the age affected the cost-effectiveness ratio. So there wasn't a lot of information where you could say, all right, here's five different ratios that compare starting screening at age 60 versus age 50. So that did not show up as something that we were able to tease out and do some sort of analysis on.

In terms of the social costs, your second question, I thought about that myself. I guess, not that I'm someone who knows a lot about these conditions. The one

place where I think that including social costs could make a difference is if you do have an appreciable number of people who have heart attacks before they retire there could be a productivity loss that does affect the answer in some appreciable way. I don't know whether that's with really true. I was just guessing about it in my mind. It's something that would have to be resolved by analysis.

DR. NEUMANN: I would just add one thing because I think your observation that perhaps because it's an elderly population foregone income is not important is a good one. We can speculate that that may be a lot more important in younger populations.

This issue of societal versus health care costs comes up a lot and sometimes it's argued that for the Medicare program, funded by general revenues as well as other sources and it's a large social programs, should take a societal perspective. Even if we don't include the foregone income there could be other costs that are not included in the analysis that might make a difference for an elderly population.

For example, do you include nursing home costs and custodial care and such as an issue? It could make a

difference whether you're taking a societal perspective and health care perspective.

DR. REISCHAUER: Just as a practical observation, even if you found that age was terribly important I can't imagine the political circumstances that would allow us to say, implantable defibrillator if you're under 64 but not if you're above. So it's interesting for analysts but does not have any practical policy ramifications.

DR. MILSTEIN: I have a couple of questions. First, there are many people who have expressed skepticism as to whether or not there is the funding availability or political will to force the kind of degree and granularity of randomized controlled studies that you'd need to begin to really map all of the -- even a fair percentage of the treatment-treatment indication diads for which one would want information if I were putting together a careful policy.

In response to that some have said, can we take advantage of the fact that in America there is such wide variation in the rate of uptake of new treatments, new technologies and use the Medicare database, perhaps with some expansions, as our pretty good database for purposes of

conducting non-randomized control studies? I wonder what your thoughts are about that? Specifically, has anybody given any thought as to the incremental data elements one would have to routinely collect as part of the Medicare billing data set in order to give us a pretty good start on that kind of a course?

I'll hold my second question because my first question is complicated enough.

DR. NEUMANN: There is a lot of discussion about the fact that you don't have the kind of treatment-treatment randomized trials that you would ideally like to make decisions and what do you do about it? Whether existing Medicare databases are sufficient is a big question but I think certainly it's true that there's a lot of activity that's going on to try to tease out of all kinds of non-randomized data evidence, treatment effects out of observational databases and so forth.

There is an effort that is the drug effectiveness review product that's an alliance of 16 states, mostly Medicaid programs, and a couple of non-profit organizations that have banded together and are looking very hard at this science of reviewing the totality of evidence including

randomized controlled trials and all their non-randomized evidence and trying to put a lot more rigor behind it. So I think that activity and others like it are certainly shedding light on areas where you don't and probably won't ever have head-to-head comparisons. Some of the Section 1013 work that Nancy mentioned is also getting at non-randomized evidence.

It's certainly an area that needs a lot more activity and I think Medicare databases are great resources to try to exploit for those kinds of analyses with all of their limitations and selection effects of everything else.

DR. COHEN: I'll just add one thing to that and that is that you can get some information just from modeling, extrapolating, simulation to extrapolate beyond trial results. Now obviously you'd much rather have empirical information, but sometimes you can do modeling and if you do your uncertainty or sensitivity analysis correctly you can establish that even though you don't have the empirical information you can be pretty sure that the result is in some range that is either this is a no-go or it's really good value for the money. So you can qualitatively get your answer even if you haven't been able to measure it

directly.

DR. MILSTEIN: I'd like to follow up on -- it's actually a combination of both Bob and Mark's comments, appreciating the fact that, particularly in view of the politics and the information imperfection it might be difficult to move ahead with cover/non-cover. I wondered if you could comment on the option that Mark touched on, which is to vary the amount paid either for the technology or the professional services associated with service types that are more rather than less cost effective or vice versa? Are you aware of other countries that have successfully faced the political challenges associated with that kind of a policy?

DR. NEUMANN: I think you're right in the sense that the cost-effectiveness analysis is often framed as a tool to say we cover or we don't cover, when in fact I think the way it's been used is much more nuanced. It's a tool to figure out where to cover. So we cover this technology and we use the cost-effectiveness analysis to figure out that the really good value for money or the cost effectiveness is in this population, defined by clinical characteristics or age or even other dimensions.

I think that's been the case, for example, in the

U.K. with NIHCE, National Institute of Health and Clinical Excellence. Even in the U.S. I think there's some emerging evidence that cost-effectiveness information and related information is being used in terms of informing formulary placement decisions. So we cover this drug but we put it on this tier, or we cover this drug but only as second and third line therapy and after you use the cheaper job drug first.

It could be used, as you said and Dr. Miller said, to inform payment policy, either which DRG it goes into or perhaps whether it warrants add-on payments and a number of other options.

MR. BERTKO: Just a quick to follow up to Arnie's first question, actually two parts. But the first is, I noted that you chose for AHRQ, the work here for MedPAC, to pick two of the four study procedures because they had more studies. The question is, how many are good enough? Would one really big study be sufficient or would you want to have multiple studies? Any thoughts about what's the threshold for that?

DR. COHEN: I think it really has to be looked at on a case-by-case basis. If there is an important decision

that has to be made, and say there's one study and it's really good. I think you have to look at it and you have to say, is this really answering the question that we want answered? Is it using the assumptions that we think are valid assumptions? Can we even tell what it's doing? That's an important thing.

If you don't have that then you have to think about what is needed next. So there isn't a hard threshold, of course. I think it's just a number of considerations that you have to think about.

MR. BERTKO: The second, to more directly connect with Arnie's first question is, the Medicare database is very rich in claims data. It allows you to identify diagnoses, multiple comorbidities, but not medical records. I know that people like RAND have gone back and done medical record review, which is very expensive, but given the difficulties of finding randomized controlled trials that expense might be minor compared to the time and effort going forward. I just wondered if you had any thoughts about that idea.

DR. NEUMANN: I think you're right. Claims data have a lot of advantages. One of the limitations is they

tend not to have rich descriptive clinical information that you can get out of a medical record. So certainly there are advantages to going to records and it is costly, to be sure.

But again, I think to get back to the comment earlier, we need to push on the science of assembling non-randomized information, and our databases are getting better and better, and our ability to link claims data to clinical data is getting better and better. So I think those approaches are very valuable and will be even more so in the future.

DR. KANE: I have two questions that are totally unrelated but I'll ask them both at the same time. One is, do you expand the group of clinical trials that you look at and cost-benefit analyses you look at to international settings? If you do, do they give you roughly the same types of answers or greatly different, and would you have any notion of why? That's question one.

The other question is, apart from using this for coverage decisions or payment decisions, can you see it having any use in influencing consumer choices?

DR. NEUMANN: First to the international question. The convention is internationally patients' biology doesn't

change but health systems do. So the clinical trials might be relevant. There may be exceptions in certain populations even on clinical data but by and large clinical data, if the trials are well done, and randomized, are generalizable.

Health systems often are not, so that a hospital length of stay might be much longer in Japan or Germany. To use the economic data from an international trial, from an international cost-effectiveness analysis, may not be as generalizable. But again I think you need to look at that on a case-by-case basis. Even though the study is done in Sweden it still may yield some important information in terms of the cost-effectiveness analysis.

If, as does Josh was saying earlier, if it's very, very clear that it's cost effective and you can see why in Sweden it may well be that you for comfortable enough that conditions are generally holding elsewhere in the U.S.

DR. KANE: That suggests that how the practice around that technology is delivered as a big effect on your result. So if you're in Sweden you may have a very cost-effective treatment but in the U.S. you may not because of the differences in practice.

DR. NEUMANN: It's certainly possible.

DR. MILLER: Consumer choice?

DR. NEUMANN: How the cost-effectiveness information is dealt with for consumers and some of the consumer issues is a question people debate. Often these studies are targeted at managers and policymakers and they are done for people who are thinking about broader societal resource allocation decisions. To try to interpret a cost per QALY ratio as a consumer is difficult.

Nonetheless, I think there is an attempt to try to in some way marry cost-effectiveness information with the kind of burgeoning consumer movement. Maybe it means using the cost-effectiveness information to influence what tier or how the cost sharing is done. So the idea is to try to match value with giving incentives for consumers to do good things.

So even if an expensive new technology in terms of its price shows to be very good value for money from a longer-term perspective, you don't want the patient have high cost sharing on that because you don't want to offer a disincentive for the consumer to use that.

MR. HACKBARTH: Can I just follow up on that, Peter? For patients, even if, because of insurance

coverage, cost-effectiveness research has less significance for them personally, certainly risk benefit analysis can be very important. There are those who feel that that's a major opportunity for us to improve the health care system systematically, in a user-friendly way provide information to patients about risks and benefits of alternative therapies.

To what extent does the database necessary to feed the cost-effectiveness research for payers also support risk benefit information for patients?

DR. NEUMANN: In the cost-effectiveness framework we're dealing with cost per unit of health, cost per life or cost per QALY. The QALY has some strong assumptions attached to it which may or may not incorporate some of the risk benefit information that you've like to convey to the patients. It may well be that technologies are associated with risk trade-offs, in fact they undoubtedly are associated with risk benefit trade-offs that you'd like to convey to patients and have them much better informed about the risks and the benefits that they themselves face.

It may be that the cost-effectiveness ratio, some of the ones we presented here, obscure or mask some of those

risk benefit trade-offs and it may be you need to do a separate analysis or present both of those pieces of information to the decision-makers. It could be that you provide it to the payer, the physician and the patient and they use all of that information in making their decision.

MS. RAY: Can I just point out, I think on everybody's chairs there was an article, I think it was from the New York Times. I think it showed a really good example of how the risk benefit information trickled down to physicians for lung reduction surgery and that in turn help physicians talk to patients about going ahead with the surgery, or at least the indication of the article suggests not going ahead with the surgery. I think that's one example to follow up on your point where this information has been put together and used.

MR. MULLER: Just a technical question. One of the new biologic drugs is likely to be very effective for colon cancer for a certain subset of the population but not for another subset based on genetics. How does the ratio change? If you pick the population which we can test now where it's likely not to be effective versus the one -- how does the calculation change? If it's, current evidence,

clearly ineffective versus highly effective.

DR. COHEN: It depends on some of the specifics of the drug. For example, -- I don't know anything about this, but if this is the type of drug that is effective if you can catch the disease before it manifests itself in an obvious way then clearly that's going to make screening more cost effective because it means that when you catch something early you're going to get a greater benefit. If this were some sort of drug that we're able to knock out the cancer later on in the process than the incremental benefit of catching it really is not going to be as great.

Then there is the issue of whether you can identify this specific population for which the drug is beneficial. If it's 5 percent of the population but you need some sort of expensive test to identify who that 5 percent is, then it sort of becomes a moot point. You may as well just go on with your general screening.

MR. MULLER: But does the calculation, would it also take into account the treatments you could avoid by identifying the set of the population for whom the therapy would not be effective? Does that go into the calculation as well? Almost like a cost avoidance type, is that --

DR. COHEN: Sure, in principal.

MR. MULLER: I think in many ways right now many of the cancer drugs are, as you know, applied to whole populations. Part of the promise of the new biologics is that they could be used for the 30 or 50 percent of the population for which they're most effective. In that sense, cancer therapy can be quite expensive, $50,000 or $100,000, et cetera. So to the extent to which one can cost avoid, if that's a verb, that $50,000 or $100,000 treatment, in that sense the test I would think becomes very cost effective. I'm just wondering whether the calculations take those kind of considerations into account.

DR. NEUMANN: It should. The methodology is certainly flexible enough to accommodate assumptions. To the extent you have clinical information it should be in there.

MR. MULLER: That obviously is -- we had part of this discussion yesterday but there's a lot of cost avoidance by targeting these drugs much more effectively and avoiding $50,000, $100,000 therapies where they're not effective and targeting them on the people where they are.

DR. CROSSON: Thank you. One of the questions I

think for the Commission to consider is the cost effectiveness of recommending that CMS think about using cost effectiveness in anything, coverage or payment determinations. So as we think about that I'm thinking that a lot of this, particularly if we're going to be looking at payment mechanisms, we're really talking about using cost-effectiveness analysis in a comparative effectiveness way because most of the time -- not always but most of the time you're talking about doing this versus doing this other thing which is standard practice.

It seems to me if that's the case most of the time then you have sort of a two-by-two table in your head where over here you've got small or large differences in quality and here small or large differences in cost. Then if you apply that two-by-two to Nancy's universe of not just what's new but what Medicare is paying for you've got four different cells. And the cell that seems to be the most attractive is the cell that has small or no differences in quality but large differences in cost. That's the cell -- just to say, the cell that has big differences in quality and small differences of cost, that's kind of a slam dunk.

If there are big differences in quality and

doesn't cost much then you expect that things are going to go in that direction. If you've got big differences in quality and big differences in costs then that takes us into the hard analysis like the British are doing because you end up basically saying yes or no to something which does make a difference but it costs a lot. That's the political minefield, I think. If you've got small differences in quality and small differences in cost, who cares?

So the box we're really talking about is the box theoretically where there are small or no differences in quality but big differences in cost.

So the question is, do you have any intuitive sense of whether that in fact is a big box or a little box?

[Laughter.]

DR. NEUMANN: There's an awful lot of new drugs, technologies, procedures out there that are expensive and probably have positive benefit. So the box is probably pretty large.

There also, as Nancy-Ann was saying earlier, a lot of existing things that we do you that really haven't been subject to some of the scrutiny that also may well have positive benefit but also positive cost. These are

empirical questions. But there's an awful lot that one might look at.

One of the challenges is to figure how to prioritize the big-ticket items. Often they're the ones with the biggest budget impact. That I think is a key question.

MS. BURKE: I wonder if I could ask just a follow-up question for either Nancy or for either of you. That is, in looking at Great Britain's process and the NIHCE process I wonder -- there's a brief description in the materials but I wonder as to how they make the decision as to the procedures they refer to the advisory process. There's a reference to a base amount in terms of cost. But I wonder if you could give us just a two-minute -- to your point, Peter, which is, there is an enormous universe out there and query, given limited resources how best to target your resources were you to begin to get into this business.

Can you give me just a couple of minutes so I fully understand how they go about making that decision? Have they gone backwards or are they only going forwards in terms of new applications as compared to existing? And how are they discriminating among all the things that come on

the market as to which of them they will refer to the advisory process?

DR. NEUMANN: I'll try to take a stab at it. I must say I know something about the NIHCE system but there are people who know a lot more about it than I do. But I think they have tried to think hard about how to identify those procedures to look at in the first place. They have what I think they call a horizon scanning group and process where they have people who are looking for things coming, and also existing, that are costing a lot of money, that are areas of uncertainty, that are areas of perhaps some clinical controversy, and certainly good candidates for scrutiny of clinical and cost-effectiveness evidence.

There are also people who -- and I think in the U.K. and at NIHCE as well in particular perhaps, who are trying to develop some formal methodology for doing this with value of information analysis, where they're trying to actually formally estimate the costs and the benefits of collecting information in the first place. It's a methodology that's fraught with its own challenges and uncertainties but I think that's how they do it.

I don't know if I answered all of your questions.

MS. BURKE: That's certainly helpful. I think, Glenn, one of the issues for us as we go forward and as we begin to think -- I think Nancy's done a nice job of identifying three of the key issues. I would add to that issue not only these sort of functional questions of who sponsors it, who conducts it and who funds it, but also the fundamental question is how one makes a decision as to where one prioritizes one's efforts. To the extent that we go in this direction, the universe -- the box, to answer the question that was asked, the box is potentially quite large or not.

It was interesting to watch the British decision on the Alzheimer drugs which has met with some anxiety, not surprisingly.

But again, I think one of our questions, were we to go forward and I think it's something that the Commission would need to look at it is, all things being equal how would you even begin to approach the process? How would you begin to set priorities? Because that issue -- I mean, I think back to some of the OTA issues and I think back to some of the AHRQ problems that arose, it was about what you chose to do, what was the reason. Was it solely based on a

set of criteria that were clearly established, that had to do with either clinical effectiveness or controversy over its use or its cost? I think clearly delineated that and creating a transparency will be critical to making this a process that people are comfortable putting into play.

MS. HANSEN: Actually I'd like to follow up with what Sheila just brought up about the decision-making process. I know that AHRQ in its process now of looking at comparative effectiveness is making this a very public, transparent process where it's on a web site. They have stakeholder groups that really talk about this, and they are actually trying to, from an AHRQ perspective, choose different methods and have the criteria definitely transparent.

I think they were really stung by the previous process of just all the different special interests. But now it is publicly there on a web site and available out publicly.

MS. BURKE: That's probably, Nancy, worth our getting a hold of. As I recall the last big issue for them was back surgery, was it not, orthopedic stuff? But if they've moved in this direction it would be interesting

going forward for us to understand and to get further information on NIHCE. But also if AHRQ is going in this direction it would be nice to how are they in fact establishing their criteria and what's the basis.

MS. HANSEN: They've chosen two topics that are public right now. One is GERD, the gastroesophageal reflux disease as well as positive mammograms and what are some of the follow-ups. But separately, this segues to a question or comment that I don't know whether that was an underpinning of Ralph's comment about effectiveness for certain groups.

My question is relatively broader and that is, in terms of looking at all of these studies whether or not in terms of coverage decisions or preventive, kind of encouragement of taking on services, whether considerations are differentially done for populations that may be racial groups that are more predisposed as well as people who might be predisposed say for breast cancer. Are there ways to take a look at some of these studies with any kind of volume that looks at populations a little bit more differentially for this?

DR. NEUMANN: Many of the studies do stratify on

lots of risk factors. It's all a matter of what the particular investigators happen to do, but they often will take into account family history, perhaps race or ethnicity, certainly clinical risk factors. One can certainly do analysis, clinical analysis, cost-effectiveness analysis, on those dimensions. Then whether or not you want to make actual coverage decisions on those dimensions that's a question for the decision-makers.

MR. HACKBARTH: Between this session and the one a year ago we've heard that there are a series of challenges that must be addressed to expand use of cost-effectiveness analysis. We need better funding for clinical studies so we have a better idea of what works and what doesn't. The number and quality of the studies themselves varies. We need standards on cost-effectiveness analysis. There's a rationing reticence, shall we say, both in the public and private sectors. So there are a lot of fronts where we need to do work.

What I wonder is whether this is an area uniquely or almost uniquely calling for public/private collaboration if we're going to make headway in addressing these multiple challenges.

I think, focusing on the rationing reticence for a second, I think there's a dynamic whereby Medicare is reluctant to be seen as more restrictive than private payers, that's a politically untenable position. And private payers are often looking to Medicare for leadership for a variety of reasons. And so everybody is saying, you go first and we're not going anywhere as a result.

In some ways this seems analogous to me to maybe the pay for performance area where if the public and private sectors together come to build infrastructure and invest in research and development of standards, that both would be significantly better off and we'd have a much greater likelihood of advancing the cause.

Any thoughts either from the panel or commissioners about that?

DR. REISCHAUER: The problem is this is a public good. It's really an international public good so there's no incentive for the private sector to invest in it. It's the kind of thing that even if you do produce the public good, politically you can't do it unless everybody else does it, in which case whatever benefit you might have has disappeared completely.

I'm an advocate of this being a public/private, not voluntary kind of thing, because I think you have to coerce the private sector into contributing.

MS. BURKE: Bob, I'm not sure I would agree necessarily that it's not in their interests. They, arguably have, certainly with the bigger plans, with United or Wellpoint or Aetna, they've got 15 million, 20 million lives on the line as well. So the value, if you assume that there's a quantitative value in not doing things that aren't cost effective, would be to their advantage as well I would think. It's not just a public good.

DR. REISCHAUER: But they can't capture the benefits from it.

MS. BURKE: You mean for certain age cohorts or generally?

DR. REISCHAUER: Just in general. They discover treatment A is not cost effective. Let's say it's in existence already and everybody is providing it and they're going to clamp down on it. They're going to get into some political problems just as Medicare has --

MS. BURKE: But they don't seem to be reluctant --

DR. REISCHAUER: -- in denying it to their people.

If they've done all the research and it comes out and they can impose it, then Humana can glom onto that knowledge for free, and these things are not cheap to do.

MS. BURKE: I would think it would depend on the makeup of their population. All things being equal you're right, anything they do that gains knowledge advantages everybody. But I've got to believe there's some advantage to them.

MR. BERTKO: Bob, some of these are clear-cut and easy, others are much more difficult. I'll give you one example that's simplistic.

There is apparent, in some places, overuse of human growth hormone off label. It's an approved drug. It's useful for certain people, and some of us have seen some spikes that are clearly inappropriate so, ping, it goes away. It's to our benefit, it's to our customers' benefit, the large employers, and in fact arguably it's a good idea for our members because it was a bad idea to treat just short but normal kids with human growth hormone.

DR. REISCHAUER: So everybody does the same thing you do but you've done the research to show --

MR. BERTKO: But it was sequential. One company

found it, other companies heard about it. Horizon scanning is an appropriate term here. We didn't even know it was happening until we heard via, I'll call it the industry gossip line, rumor that it was happening and then we found some happening in some geographic locations.

DR. NEUMANN: I guess I would just say, I do think there's a process here. Maybe it should indeed involve public and private groups. But I think there's a process that has gone on in other places that have used it, in the U.K. for example, that involves input from all kinds of stakeholders, and the public, and feedback and so forth.

I think part of the answer may lie in how this is done, not only in terms of the process but in terms of how it's framed. That is, I think cost-effectiveness analysis is often framed as a tool, as we discussed earlier, to deny coverage, to ration, when indeed it should be seen as a tool to improve the value of the care delivered and to try to inform the types of coverage and payment decisions that are made. I've argued in the past, perhaps the term itself, cost effectiveness, has become a kind of pejorative because it sounds like it's about cost containment. So maybe we need a new term as well.

But regardless of what term we use I think we do need to frame it in the right way for people.

DR. MILSTEIN: Reflecting on Glenn's question, it seems to me than Congress has in some ways already taken the first step forward with respect to implementation of what could be framed loosely as cost-effectiveness analysis, both in MMA and the Deficit Reduction Act. In MMA they basically said, we want payment levels to providers geared to what efficient providers need. They didn't tie the definition of efficiency to cost effectiveness. Then in the Deficit Reduction Act they signaled an interest in provider pay for performance to take into account both effectiveness as well as efficiency. And I think efficiency translates into resource use, and the ratio between resource use in an administered-price environment, resource use and effectiveness becomes a proxy.

What is left a little vague is the time frame of reference. Cost effectiveness is a lifetime assessment, whereas some of the other time units on which efficiency and effectiveness can be calculated are much shorter than that. But that really remains an unspecified aspect of the policy that's already been laid out.

So essentially we already have a signal from Congress that with respect to both effectiveness and efficiency we should move forward in our payment system to gear it to both of those characteristics.

I want to go back to Mark's question and ask whether there are any countries globally that have taken into account cost effectiveness, efficiency, effectiveness, any of the above, in setting the service payments levels? I think, for example, in the last Commission meeting or the one before, the CDC when they rank preventive interventions smoking cessation counseling keeps coming out number one off the charts in terms of not only cost effectiveness but it's actually a cost savings intervention. Yet there's been no effort in the Medicare program or many other programs to more favorably reimburse that very high yield service.

Have any countries moved forward on, I'll call it pay for performance but in which the unit of analysis is the service rather than the providers' practice writ large?

DR. NEUMANN: There are many countries I think that are trying to use cost-effectiveness information, sometimes an indirect ways perhaps, to negotiate prices down. Now for example, your drug is very expensive, it does

offer some benefits but at the price you're offering it it's not cost effective. It becomes cost effective at a much lower price. Now sometimes the rules don't allow them to do that explicitly but that seems to be the outcome of the process.

The other part, there's experimentation in trying to use cost-effectiveness analysis in doing risk sharing arrangements. The famous example is the MS drugs in the U.K. There was a lot of uncertainty about whether they were cost effective or not. It depended on whether you believed assumptions about long-term effective based on short-term trials. The decision that was made was, we'll cover your drugs and we'll actually see. We'll look and see if they work overtime. If it turns out that they do, we'll give you the higher price. If they don't then you get a lower price, so you have to pay us back, in that sense.

DR. MILSTEIN: First of all, that was very informative. My question pertained to the level of professional payment. In other words, are there any countries that are moving ahead, for example, with paying for smoking cessation at a substantially higher -- or beginning to vary that based on demonstrated high levels of

cost effectiveness? It's taking the other, in some ways the less politically challenging cell of Jay's four cells, which is the services that are off the charts in terms of favorable cost-effectiveness rating and pay them more generously. In an overall constrained environment it has the effect of paying less generously those things that aren't in that favored cell.

DR. NEUMANN: I get it. So to use the information to give incentives to do good things, and maybe pay people more. I can't think of any offhand. Maybe there are. It certainly seems reasonable to do and I've certainly heard people mention, for example, tying it to pay for performance in a way that you suggest. But I don't know of any actual cases.

MS. RAY: But that is an issue that in the future we could explore it. We could look and see what's going on in Canada and the U.K. and Germany and Australia and New Zealand, for example, to see if there are any cases That could be on our future work agenda certainly.

MR. HACKBARTH: Unfortunately, we're going to have to bring this to a conclusion and move on. Thank you very much. Very well done, Peter and Josh.

Okay, next up is payment for hospice services.

* MS. LINEHAN: Good morning. This session is about Medicare's hospice benefit. The hospice payment rates were developed 25 years ago and since then the use of hospice has grown and the provision of hospice has changed. These changes to the use and provision of hospice care that I'm going to review motivated us to assess whether payment could be adjusted using patient characteristics to improve the accuracy of the payment system.

I'm going to present some background and that is going to set the stage for why we contracted with Dr. Melinda Beeuwkes Buntin at RAND and her colleagues to undertake an analysis of possible payment system refinements. Melinda is a health economist and co-director of the Center for Health Care Organization, Economics and Finance at RAND. After I review the background she's going to discuss their results based on an analysis of data from one large hospice chain provider.

The CMS office of the Actuary estimates that Medicare spending on hospice will be $9.8 billion in 2006. Spending on hospice services is projected to increase at an average rate of 9 percent per year from 2004 to 2015. This

growth rate is more than the rate for hospitals, physicians, SNFs and home health services. Medicare is by far the dominant payer of hospice care. The National Hospice and Palliative Care Organization reports that Medicare paid for 88 percent of total days in 2004 in the every facility.

Hospice was added as a Medicare benefit in 1983. The benefit covers palliative and support services for beneficiaries who have a life expectancy of six months or less and who agree to forgo Medicare coverage of curative treatment for their terminal condition. Covered services under the benefit include skilled nursing care, drugs and biologicals for pain control and symptom management, physical, occupational and speech therapies, counseling, home health aide and homemaker services, short-term inpatient care and other services necessary for the palliation and management of the terminal condition.

Hospice care is and always has been carved out of Medicare's managed care benefit. Beneficiaries do not have to disenroll from their MA plan but they may choose to do so.

The payment methodology and Medicare's four daily payment rates were developed using cost data from 26

hospices providing care to Medicare patients with terminal cancer under a HCFA demonstration project between 1980 and 1982. The base rates have been updated for inflation but the payment methodology and the base rates haven't been changed since the initiation of the benefit.

The four categories of care that are shown on the screen are distinguished by where they are provided and the intensity of the service, and the dollar amount following the names of the days of care are the 2006 daily rates. The vast majority of hospice days are routine home care days, that first category listed. In 2003 they were 93 percent of days billed. Routine home care is the default payment category that hospices are paid if one of the other types of care aren't provided.

There's another feature of this payment system that bears noting. There are two types of caps. There is a cap that 20 percent of the total agency's days can't be, or no more than 20 percent of days can be for inpatient types of care. The other type of cap is an aggregate annual spending cap. In 2005 it was around $19,000. That amount is multiplied by the number of Medicare patients seen by the agency. If total payments to the agency exceed that amount,

they have to pay that amount back to the program.

Hospice services are characterized by growth. Hospice has become much more widely used as the visibility and acceptance of hospice care has increased. This share of Medicare fee-for-service decedents electing hospice grew from 22 percent in 2000 to 31 percent in 2004. Between 2000 and 2004 the number of Medicare hospice users increased almost 50 percent, the days of care doubled, and payments increased 130 percent. As this shows, the number of days increased more than the number of users. When we look at data on changes in the length of enrollment we see that between 2000 and 2004 the median length of enrollment remained at about two weeks but the mean length of enrollment for a beneficiary in hospice increased from 51 to 67 days.

The mean length of enrollment was driven up by the upper end of the distribution having increasingly longer lengths of stay before they died. 25 percent of beneficiaries dying in hospice were enrolled for less than a week and that persisted over time. That might be suboptimal because the patient and family may have benefitted from a longer hospice enrollment.

This distribution likely reflects several factors related to the structure of the benefit: the difficulty of estimating the amount of time a patient has to live, and the election of hospice only when death appears imminent. It also reflects that the benefit has expanded beyond cancer patients to patients with other terminal conditions such as neurodegenerative conditions and cardiovascular disease. In 2003 more than half of all Medicare hospice patients had a non-cancer terminal diagnosis. On average non-cancer patients tend to have longer lengths of enrollment.

Another change since the implementation of the hospice benefit, and even in the past five years, is in the composition of hospice provider types. As you can see in this chart, between 2001 and 2006 the number of hospices increased and that increase is attributable to the growth in freestanding, here labeled non-provider affiliated to make the point that they're not necessarily a freestanding building somewhere out there but they're not affiliated with a home health agency or hospital or a SNF.

Not shown on this chart but noted in your paper is that new hospices are nearly all for-profit. As of February 2006 47 percent of hospices were for-profit and that's

compared to 31 percent in 2001.

MedPAC has not done a formal payment adequacy analysis of the hospice sector, including look at hospice agency margins, like we do with other providers. There is, however, some information on agency margins from other sources. But these are not necessarily representative of the entire industry and given the recent changes may not even reflect the current state of the sector.

GAO found that freestanding hospices had Medicare margins of over 10 percent in 2001, but margins vary by the type of day of care, suggesting that the relative values of Medicare rates for different payment categories may need to be recalibrated. They also found that smaller hospices had higher costs. NHPCO, the National Hospice and Palliative Care Organization data on margins from 2004 showed margins of 11 to 19 percent, again varying by the size of the agency. These were voluntarily reported and not necessarily representative of the entire industry.

An analysis of margins using freestanding Medicare cost reports that was published in the Journal of Palliative Medicine found margins varied by the size and for-profit/non-profit status with the median for a large for-

profit agency at 18 percent but the median for a large non-profit at just 2 percent. In addition, SEC filings from publicly-traded hospices report that they are acquiring and opening new hospices and have growing average daily censuses which is consistent with the increasing use.

I'm going to turn now to Melinda and she's going to discuss the results from their analysis of patient level costs using one chain provider's data.

DR. BUNTIN: Thank you, Kathryn, and thanks to the Commission for having me. It's always a privilege to be here. Kathryn has really summed up the motivation for my empirical work that I'm going to be presenting; namely, that we have a per diem system currently that's based on four categories, but that it was implemented about 25 years ago and since then there's been a large change in both the types of patients seen by hospices and the providers serving them.

This led us to three specific questions we wanted to investigate. First, how well does the current per diem system reflect current hospice costs? Second, should case mix adjustment be considered, specifically case mix adjustment using, for example, patient diagnoses as has been done with other prospective payment systems? And third, are

the beginnings and ends of hospice stays more intensive? These are all questions that have been raised in prior literature by the GAO reports and in the Commission's June 2004 report.

In order to address this we needed to have data in addition to Medicare claims data which are very limited in the hospice area. So as Kathryn said, we arranged to obtain data from a large for-profit hospice chain. These data contained information on the frequency, timing and duration of visits to hospice patients and on the type of staff providing those visits. It also contained rich patient level data on things, for example, like marital status, nursing home residence, and discharge status.

That probably requires little explanation. Most patients, 90 percent or so, die while in hospice, but there are some who are discharged either to move to another area, to go to another hospice because their prognosis is extended or because they decide to see curative treatment and that affects their costs of the pattern of care they receive.

There were, however, a number of limitations to using these data. First, it only covered about 6 percent of the hospice population and during the time period we

examined only encompassed one chain provider and about two dozen sites.

This provider saw a slightly different patient mix than the Medicare hospice population as a whole, particularly it saw a little less lung cancer and debility patients, more of the chronic diseases like cardiovascular, cerebrovascular, neurodegenerative disorders, and had more patients who were in the oldest age category. The hospice also had slightly different practice patterns. It used inpatient care to a greater degree, it did not use respite care and it had very favorable negotiated pharmacy and supply rates. I should also note that they had higher mean but lower median lengths of stay than the industry average.

So even though we had these very rich data we also still had to impute costs for direct patient care. We did this using the information I describe on the visits and BLS wage data on relative wages and different labor categories. So again just to reemphasize this, we're not including in here drug costs, supplies, overhead, things like that. We're just looking at the direct costs of patient care visits.

When we did this, however, in response to our

first question we did find that the per diem system is reflected very well in current visit and visit cost patterns. This bar chart shows R-squared so the proportion of variation explained simply by using the number and type of visits that a patient received. In other words, the variation in cost was really explained by the patient's variation in length of stay or days of care.

Now there are two possible reasons for this. One is that this provider may have responded extremely well to the current payment system. The other is that the needs of dying patients could be relatively, clinically similar. When we spoke to the clinical advisors on our project they were actually not terribly surprised to see this. They did feel like on a daily basis the needs of dying patients were relatively similar.

We did, however, go ahead and look at whether additional variance could be explained using this rich set of demographic and diagnostic information. The green bars on this chart are the same as you saw on the prior chart. The much smaller blue bars are the portion of variation that we're able to explain using that rich set, again, of demographics and diagnoses. When we combine all three

categories of information into a model we're able to actually, in a statistical sense, explain a little less of the variation given that we're adding so many co-variates to the model. So really the per diem system does seem to be reflecting costs well.

Another way to look at this is shown on the next chart where you can compare the predicted total visit cost using just the days of care model versus the model that includes the types of days of care, demographics and diagnoses. The takeaway point here is that these bars are very similar.

Finally, we did find some evidence that more intensive care is delivered at the beginning and end of hospice stays. Here the green bar shows the average visit labor cost or number of visits across an entire stay. The red bar shows the average in the first three days, the pink in the days that are neither the first or the last three days of a stay, and the blue, the last three days of a hospice stay. You can see that in particular resource use is more intensive during those last three days of a patient's stay in hospice when they're actively dying and need a lot of services.

So to conclude, the current per diem system reflects resource utilization in this particular hospice chain well. Again, perhaps because the chain has adapted its practices to the payment system parameters or perhaps because the clinical needs of these patients are relatively similar on a daily basis. Potential case mix adjusters really added little explanatory power conditional on days of care.

I'm going to pause here and anticipate a question about selection that the Commission may have. Similar to what I said about adapting to practice patterns, we are looking at the data for the patients who are actually enrolled in this hospice. So if it's the case that patients are adversely selected against, for example, certain category of very expensive cardiovascular patient just isn't admitted to hospice, then it won't be reflected in our data. That said, there is evidence in the literature that these large chain hospices don't have the type of explicit admission criteria that some of the smaller hospices say that they're forced to.

Again, in response to our third question we did find that greater compensation for the first and last days

of hospice care could be warranted. But I would caution that these results should be validated with a more representative dataset and with complete patient level costs.

I think Kathryn is going to wrap up with some implications.

MS. LINEHAN: I'm just going to review some possible directions for analysis in the hospice sector. There's something that we could clearly do here which is to analyze payments and costs at the facility level like we do for other types of providers. Based on the evidence available payment levels are generally favorable, but a deeper exploration could show whether there's variation in costs and financial performance by agency size, geography and other characteristics of the facility. This could help us assess the adequacy of the base rates.

The second thing that we could think about, although it would require CMS or someone collecting additional data, we undertook the case mix analysis with the proprietary data to determine whether it would suggest the viability of adding case mix adjusters to the payment system. The results, as Melinda said, of RAND's analysis

don't make a compelling case that case mix adjusters based on patient characteristics would improve the accuracy of the payment system. However, depending on how you feel about the limitations of the data, doing an additional analysis on a more representative population with a more fully defined dependent variable could lead to different results. But like I said, the data don't currently exist to do this kind of analysis.

Also suggested by RAND's work, that redistributing payments from the middle days to the first and especially last days of the stay would more accurately reflect the costs incurred at these stages of the hospice stay. There's evidence of two distinct populations of patients in hospice, a persistent share of patients with short stays and those with increasingly long stays at the upper end of the distribution. Paying more at the beginning and end of the stay would raise the average payment per day over the entire stay for shorter stays but lower the average payment per day for longer stays. But again, testing on a larger population would still be required to know whether we'd see these same patterns in a more representative sample of hospice agencies.

Finally, we could consider other policy issues such as whether, in the interest of coordinating care for Medicare beneficiaries, that hospice should, like other Medicare covered services, be included in the managed care benefit given that hospice is no longer a new benefit and is covered by commercial insurers for non-Medicare populations.

Now I'm done and I'll take any questions, and Melinda as well.

DR. REISCHAUER: I found this stuff really interesting and I want to ask Melinda some questions. I guess, as you suggested, it's not surprising that when you give somebody $500 they spend $500. The real question is, is this the optimal or best level of care? And how do we tease that out? You've gone through different types of people and shown that the costs are close to what they're paid in those situations. But I wondered. can you break it between for-profit and non-profit and see if there's any variation that way?

And then the other way of trying to answer a question like this would be to say, what kind of services for similar types of people are provided in hospices in other countries where maybe the levels of payment are

different from ours?

DR. BUNTIN: Unfortunately, we only have patient level data from one hospice chain which is a for-profit chain so I can't compare the practice patterns in a for-profit versus a not-for-profit, but I do think that that would be interesting, in particular because there are reports about different margin levels across the two types of providers. So that would be interesting to do but I'm sorry I can't answer that question.

MR. HACKBARTH: The for-profit would presumably be less likely to spend $500 because they have $500.

DR. REISCHAUER: They didn't. They spent 75 percent of $500, right? There was a margin there.

DR. BUNTIN: Actually, I would like to clarify. What we looked at here was the variation in days of care, explained variation in these visit costs, not that they spent the entire $500. So we didn't look at that payment adequacy issue. But given what Kathryn said, given what this provider publishes in its industry reports it is clear that they are making overall a healthy profit margin.

DR. SCANLON: When you say days of care. is this days in the episode or days that they actually received a

service in the episode?

DR. BUNTIN: This is days in the episode. Remember, there's a default category here so if you don't receive any -- you only receive the inpatient respite or continuous care rate if you're getting one of those three types of services. Otherwise the default category is that you receive the routine home care rate.

MS. LINEHAN: You did look though, I think, at the number of days in the episode and the number of days that they actually had a visit and didn't you find an average rate of --

DR. BUNTIN: There's an average of about 1.5 visits per day. So people actually get services from a variety of different disciplines. That doesn't mean that everybody is getting a service on every day. We didn't actually look at it that way. We could. But people are often getting a visit from a home health aide and a nurse, or a therapist and a daily home health aide, for example.

DR. REISCHAUER: Do you have any way of saying if a mix of folks is different between for-profit and not-for-profit?

MS. LINEHAN: We could look at some

characteristics using claims but we could know their diagnosis, we could know their age. We couldn't know things very easily like whether they were in a nursing home. And we don't know whether patients actually got a service on any given day using the Medicare claims data.

DR. MILLER: Let me just ask one other way of asking that question. So there's four different payments that a person can get. From the claims data can you say between for-profit and non-profit, tell the mix of that? Which I know is pretty gross but still it's --

MS. LINEHAN: Yes, you could.

DR. BUNTIN: Yes, and we compared this hospice chain to all freestanding hospices and to the Medicare hospice population as a whole. We didn't break it down for-profit versus not-for-profit, but we could do that given the information we have.

MR. SMITH: On that point, do we know anything about patient characteristics and point of service? Is there something that explains at the patient level when they're in a hospital. when they're in respite care, when they're at home? Is that likely to be situational? Is it likely to be patient characteristics? Is it associated when

in the episode -- are you more likely in a hospital at the end of the episode or more likely to be at home?

MS. LINEHAN: We didn't look at that. I don't think you looked at that either, Melinda. This particular chain didn't have any days of inpatient respite care. They would have the other inpatient category. We could look at where in the episode they used different types of days of care, and we could look by patient characteristics whether that varied. But the vast majority of days are the routine home care days, like 95 percent of the days. So just at the median patients don't have any other types of days.

Sorry that was unclear. They don't have use of any other types of days except routine home care days.

MS. BURKE: This was really quite helpful. I am, again, at the risk of asking to have more work done, I'm not pushing that but I am actually interested in some of the statements that were made and some of the points that were made. And that is understanding -- I mean, I remember only too well why we created this benefit and how we structured it at that time.

I am interested in understanding the comment that to a certain extent the nature of the patient has changed.

I noticed in the materials that the greatest lengths of stay tended to be around patients who had neurological conditions; not terribly surprising, largely I suspect Alzheimer's patients. Interesting. It is a good thing that people are beginning to use this benefit with somewhat different diagnosis than what we originally expected, which were largely cancer patients. But it would be interesting to understand how in fact the benefit has changed and the nature of the patient. The distribution across diagnoses. Who's now using it.

I was also struck -- and there is, as I recall from the materials the average was about 50 days for that particular population. We were constrained at the time of the creation by the fear that people -- one, the requirement that people chose between essentially curative services and palliative services was a conscious one because we wanted to be sure that people were making an actual decision. So it is interesting that you are now seeing more people make those decisions.

But I would be interested in understanding whether because of that the nature of the benefit has also begun to change. For example, your comment that you didn't provide -

- we didn't ask you to but there's no understanding as a result of this work as to what has occurred with respect to the use of pharmaceuticals. That, of course, was one of the fundamental issues at the time was the flexibility in the use of pharmaceuticals that was not prevalent in the more curative services; that there was more freedom. It would be interesting to understand whether that's still very much a part of what that occurs and whether that has changed.

There's also the reference to the fact that this particular chain had no respite, I found stunning, since that is a fundamental piece of the presumption that people in fact are staying in a home-based setting and then essentially you're relieving -- so I'm interested that they had a greater inpatient use, is what I understand, but no inpatient respite use.

There was also a desire to essentially keep people out of institutional settings. So if we're suddenly moving to more inpatient use I'm interested in understanding whether that's stabilizing someone, whether that's determining whether their pharmaceutical needs have changed and they have to alter it. What exactly is leading to that?

Again, some fundamental understanding of how this

benefit has in fact changed over time, and how the patient has chosen. And our capacity to manage these patients has also changed. We can now manage people in a home-based setting that I couldn't have cared for in an acute care unit in 1922 when I was trained.

[Laughter.]

MS. BURKE: But it would be interesting to understand whether this benefit has fundamentally begun to change. And if so, what is it, in addition to the payment system, do we need to think differently about the patients that are being seen, why they're being seen, how they're being cared for, and by whom? There's just interesting little facts that came out of this that lead me to wonder what in fact is going on here is.

MS. LINEHAN: I can answer some of that. With respect to the use of no respite care, we asked about that and heard that -- we asked two things. One, why people don't use it, and the answer was caregivers may be reluctant to actually put a family member in an institutional setting when they want to care for them at all. So even though it's available they might not want to use it.

Then I also asked whether family members actually

know that that's available to them, and maybe there's some question about whether when they hear what they can receive under this benefit that they hear that that's a component of it.

With respect to how this has changed over time, we were limited. Drugs are obviously a big piece of this. GAO found when they did their study that in the routine home care day payment category that the mix of services changed over time. There was home health aides, supplies, outpatient service costs as a share of the cost of day, that declined. Then the cost of nursing, drugs --

MS. BURKE: I'm sorry, home health aide visits declined?

MS. LINEHAN: The cost as a share of the total cost of the day declined.

MS. BURKE: What increased?

MS. LINEHAN: Drugs, nursing, social services, DME. We don't have data at the patient level. There are no data at the patient level to look at this issue. So we are kind of limited with what we can do with administrative data.

MS. BURKE: That make some logical sense. If you

assume that the acuity of the patient has increased, the application of a greater range of highly skilled services, the presence of DME, the presence of pharmaceuticals would suggest it's not simply a nursing home check-in. So that logically make sense to me.

Query what that tells us about the management of who these patients are.

Interesting that the caps don't seem to be -- which is about $19,000 on average-- that the caps don't seem to be being preached to any great degree, which is interesting if in fact the acuity of the patient has increased and the skill set required has increased, where the trade-offs are, particularly since you've got longer lengths of stay.

DR. BUNTIN: Actually, I would like to pick up on this issue of the acuity of the patients and I think it's related to the cap issue. As Kathryn was saying, there's a little bit of the bifurcation in this population going on. We might have a predominantly cancer population that maybe is entering hospice later than they used to, this 25 percent that has a stay of a week or less. They are, arguably, higher acuity than they used to be, perhaps because of

advances in medical technology that bring them -- preserve hope on until that point. But they're being balanced by this larger population of non-cancer patients many of whom have very long lengths of stay. If you are a savvy hospice you can balance these two populations and not hit the cap.

Also a little side point on the respite care. It was interesting when we asked questions about this, when you're talking about a patient who's primarily cared for at home an argument clinically was made to us that it doesn't make a lot of sense to take that patient out of the home, put them someplace else to give the family member respite. It may make more sense to bring people into that home to give the family support. So it's not clear whether not using respite is actually better for the patient.

On another point on that is --

MS. BURKE: But that assumes you define respite as only involving inpatient. In fact respite can involve a home-based service. So in this case did they provide no inpatient respite but were they providing backup services at home that allowed people to stay home and provide backup?

DR. REISCHAUER: Is it continuous home care?

MS. LINEHAN: No, that's not the same thing.

DR. REISCHAUER: So the margin, I would think, of inpatient respite care which is paid 131 versus routine home care must be hugely different. I suspect the former has a negative margin.

DR. BUNTIN: We also saw some evidence that there were patients seen by this particular hospice who were admitted to the inpatient unit but because there wasn't clinical justification for them being in inpatient care Medicare was actually only being charged the routine home care rate. So that may be another way that they're dealing with patients who aren't able to be supported at home. Again, that's anecdotal evidence.

MR. HACKBARTH: We need to do a time check here. We're running behind schedule so we can spend I think roughly another 10 minutes or so on hospice, because I know people have planes to catch today.

I have on my current list, John, Arnie, Ralph, Jennie and Bill, and we can make it through that list if people ask very focused questions.

MR. BERTKO: I'll try to be focused here. I guess what I'd want to do would be to separate out the margin and how much we pay from the methodology. Your slides appeared

to me to say that the methodology worked reasonably well. Arguably, you could redistribute. But if it ain't broke we shouldn't fix it. I had a reason personal experience with hospice care. It was a good experience. It was useful. But I can't imagine people tweaking the system once it starts for that. This was a non-profit.

So I guess I just wonder, should we continue to say -- we could make it more complex but would it help very much? Knowing more about what we pay for is good but we do need to change?

DR. MILLER: Maybe I could say something about that. When I was listening to all of this I had something -- I'm trying to explore sort of case mix. Given the lack of the data and given the lack of a relationship here maybe you don't want to go. But maybe there are more broad structural things, if you are seeing the intensity at the beginning and the end of the day. I think that's the level of adjustment that we're thinking of as opposed to something much more complex than that. You guys may have a --

MS. LINEHAN: I think that's what we were thinking. We undertook this because we wanted to explore whether this seemed viable, but we don't have any compelling

evidence that it does. So it's really hard to say, yes, go out and collect all these data to redo this payment system in a way that won't necessarily improve it a whole lot.

MR. SMITH: Mark, I'm not at all sure why we want to tweak this. It's interesting, more intensity at the beginning and the end. But there's not and end if there's not a middle. It seems to me that this system, assuming we still don't know about adequacy, but that this system has the interesting virtue of being a per diem system that approximates an episode system. We don't exactly know what's going to go on in the second day so why should -- if we figure out the episode payment is adequate who cares how it's distributed? And why make it more complex and create a set of medically unuseful incentives to prefer one behavior on day three and a different behavior on day seven?

DR. MILLER: The only thing I would say to that is if you get inside and find -- think of some of Bob's questions about is there differences in different types of facilities, for example, or different kinds of patients, that you may want some torqueing of the payment system if there are some differences. We just haven't gotten deep enough to know. I think your question is fair, but to know

whether maybe in a more adequately set payment, if there are underlying differences, that you would want to torque the payments a bit.

DR. BUNTIN: If I could just add a little bit to that. I think in all the administered pricing systems that the Commission looks at the desire is to equate marginal costs, to approximate marginal cost pricing. So if we do have beginnings and ends of stay that are more expensive and that is driving hospices to seek longer stay patients, then you're skewing them towards a certain type of patient. What we want to do is make the system neutral across all patients so clinically people get what they should.

DR. MILSTEIN: We tested some case mix adjusters and we found that they didn't account for any kind of variance. As I looked at these adjusters I'm not sure I would have predicted they would've accounted for a lot of variance. Maybe Jennie when she comments can amplify on this, but these would not have been the variables I would have tested. I would have focused more on variables that researchers like Judy Hibbard have now begun to develop where you're essentially coming up with quantified indices of patient and caregiver confidence and self-confidence in

self-managing their part of the bargain. I wondered is there prior research testing those variables rather than some of these demographic variables, because I would guess they would account for more variance?

DR. BUNTIN: It's a good question. I didn't emphasize it but one variable that we had that prior researchers have not looked at to my knowledge was we were able to look at marital status. We actually thought that single patients would need a higher number of visits to be able to be maintained at home. We actually did not find that was the case. If anything, the patients who were married were living with someone seem to get more visits. We don't have a good explanation for that except perhaps, I could hypothesize that they had a better advocate, but that was not borne out by the data from this particular provider. I think it is an interesting question though and it is not something that I know of anyone being able to look at.

DR. MILSTEIN: Maybe next time we can test state of the marriage rather than marital status.

MR. MULLER: Even with the growth of utilization we know from our reports a year or so ago that there are still some considerable underuse of hospice. For example, I

cited earlier in terms of many cancer patients in between the patient, the family and the provider community, people are still pretty awkward in going to the hospice decision. So in terms of the hospice still being a very reasonable cost alternative to the inpatient and nursing home stay it's appropriate to consider that alternative as well. It still relatively cost-effective.

So I would say there -- and one can start seeing this especially with more and more people with neurodegenerative disease in the hospice, that there's also going to be a fairly cost-effective alternative to that as well. So I think one could see that even with this growth there are disease categories in which one can and should anticipate more hospice use as people become more familiar -- if not comfortable, more familiar with how to make that decision.

MS. LINEHAN: Since you asked I would just mention one recent study on the issue of cost to Medicare of hospice and cost savings done by Diane Campbell. They found that young patients and cancer patients, the use of hospice in young and patients with cancer diagnosis saved Medicare money, but actually cost Medicare money for older and non-

cancer patients. Just to wrap some numbers around that.

MR. MULLER: Why is that, on the latter category?

MS. LINEHAN: I think part of their explanation was that cancer patients have -- their trajectory is a shorter period of obvious decline. So they get into cancer -- they're not long stay patients. It's driven by longer lengths of stay for the non-cancer cohort.

MR. MULLER: The older cohort of cancer patients you said it was not cost effective? Did I misunderstand you?

MS. LINEHAN: Older and non-cancer, yes.

DR. BUNTIN: I guess their rationale was that for the cancer patients, again, they had a shorter length of stay and there was more potential for avoiding a costly hospitalization during that period right at the end of their life. With the longer stay, non-cancer patients it was less certain that they would be avoiding that stay and they were receiving more supportive services at home than they would have in the absence of the hospice benefit.

Now whether or not the patients received benefit for this that's completely --

MS. LINEHAN: That wasn't included in the

calculation of the benefit. It was the benefit to the family, the quality improvement to their life. That wasn't factored into this. It was Medicare costs.

MR. MULLER: I think certainly when one has a service that one is not getting then the comparison is it's going to cost more. My point was that when this is an alternative to more expensive institutional, and as we noted either last year or the year before, still in many cancer cases the decision to go to hospice is not made for the reasons we've discussed. To the extent that people become more comfortable making that choice then it truly is a cost-saving alternative. I'm not arguing that if it provides services in patient subsets that have not been receiving before then obviously by definition it costs more.

MS. HANSEN: It's interesting, I just have a hypothetical thought about why the non-cancer elderly people might cost more. I think that if many of these people are end stage people with dementia as well as other comorbidities, typically dementia itself is not a payable diagnosis in Medicare. Whether or not this is actually in some ways a new resource for that end of life, because this is not something that nursing homes would normally cover.

So just a thought as the question was being raised about the dementia factor.

The question I was going to have is I was really intrigued by page 21 where you give the demographics of the hospice participants for the chain vis-à-vis the all Medicare. Something striking to me that I found unusual and it may be relative to this chain, but the use of hospice by racial minorities was extraordinary given the population of Latino-Hispanic population in the general Medicare being 1.3 and that the chain's percentage is 11.3. Any explanation?

DR. BUNTIN: Yes, that is actually an artifact of where this chain is located I think more than anything else. In the hospice population as a whole we do see lower rates of use of hospice among most ethnic minorities than their proportion in the Medicare population. So it's an artifact of where they're located.

DR. SCANLON: I guess this may be multiple questions. It's about the issue about how the hospice benefit differs for persons who are residents of nursing home, not Medicare covered but Medicaid or private pay. Given that the nursing home provides all the supportive services that an individual needs and some of those

supportive services for someone living at home are coming from the hospice the question is, is this a characteristic of a hospice patient that should be used to distinguish payment in some kind of a system?

Also I guess from the perspective of what's happened over time that we've seen this significant growth, is there a disproportionate concentration of that growth among nursing home residents? And how might that relate to the type of agency that's actually providing the services, since we've also seen a change in the composition of the industry in a relatively short period of time?

I think this may also relate to the issue of nursing homes and how the hospice benefit changes relates to Bob's comment about international comparisons, because residential settings of the elderly are often very different in the international settings with respect to the kinds of services that come with your residence as opposed to what happens to people at home here in the U.S.

The last thing I guess is a caution about the idea of what we know about the needs of patients being met, which is we don't know virtually anything at all, because what we're talking about here is hospices provide some supportive

services of which we have no sense of what share of supportive services that someone is getting. It's the same problem we have with the home health benefit which is that we don't know, in terms of how people's needs are actually being met by just looking at the services that they're receiving because we don't actually go out and measure any -- we have no metric of what unmet supportive services needs there might be.

DR. REISCHAUER: Let me just have a final comment, playing off of Dave and John's notion that if it ain't broke don't try and fix it. I'd be a lot more agnostic about whether it's broke or not. What we've done is looked at some information from one for-profit chain and drawn a conclusion that that ain't broke. But who knows. We don't hear a lot of complaints, I don't think, out there. But when you have a 15 percent to 19 percent margin that covers up a whole lot of complaints. Everybody can be happy.

The question is, what if the margin were 5 percent, what would the situation look like? So let's keep an open mind.

MR. SMITH: Bob, I think you're right but I think a traditional MedPAC adequacy analysis can get at that

question independent of the differential payment for site of treatment. Just a personal footnote, John and I both have recently come off hospice experience with not-for-profit hospices. Nothing at all systematic but surely it forms part of our reaction.

MR. HACKBARTH: Thank you very much. Good job.

We are now to our last session on physician practice expense.

* MS. RAY: Good morning again. I presented a work plan in November to look at issues about the data sources and methods used to calculate practice expense payments. Ariel and I are back here to follow up on that. Our work today fits into our broad agenda to examine physician payment issues, including the SGR and the unit of payment.

Recall that in our March 2006 report commissioners made a series of recommendations to improve CMS's process for reviewing work RVUs. These recommendations addressed the concern about the mispricing of services in the physician fee schedule. The Commission and others have argued that inaccurate pricing may be leading to increased volume in areas such as imaging.

We are now turning our attention to the other

major component of the physician fee schedule, practice expense. Our analysis of practice expense also addresses this pricing issue. In today's session we are asking you about ways to improve two key data sources CMS uses to calculate practice expense payments. Today's discussion is particularly relevant. We may be on the threshold of a major change. CMS has given a strong indication that it is interested in changing the way it uses to calculate practice expense payments. These changes may be out in this summer's proposed Part B rule. Thus, today's discussion may provide input into the agency's deliberations.

Practice expense payments are important. They account for about half of the payments to physicians. Given the magnitude of dollars involved, inaccurate payments can boost volume for services inappropriately and undermine access to care. Some of you have expressed concern that inaccurate payments can make some specialties more financially attractive than others. These are points that you just made in our March 2006 report.

CMS divides practice expense into two categories, direct and indirect. Indirect account for at least 60 percent of practice costs for most specialties. So like I

said, CMS uses two sources. The first source gives information about total and hourly practice costs for each specialty. The second data source provides estimates of the direct resources used to provide each service.

Very, very briefly, CMS currently calculates direct and indirect practice expense payments by taking total costs per specialty and allocating those costs to individual services based on resource estimates. This is called the top-down approach. CMS is considering changing how it calculates direct practice expenses by going to a bottom-up approach, or simply summing the resource estimates for each of the 7,000 or so services in the physician fee schedule.

So the first data source that CMS uses is called the SMS survey. It's a multi-specialty survey. It was last conducted by the AMA in 1999. So needless to say it is old and it probably does not do a great job at capturing current practice patterns, medical equipment and medical costs. It also does not include information for all specialties paid for under the physician fee schedule, particularly non-physician providers.

As a way to update the data, specialties could

submit to CMS updated total practice cost data and CMS allows specialties to do so through March of 2005. Few specialties have done so. To date CMS has accepted data from 13 groups and the fee schedule is currently, of those 13 groups, from six groups.

Under a voluntary updating process the fee schedule may no longer accurately reflect the relative resources required to provide a service because CMS incorporates these changes budget neutral. Therefore payments may shift from specialties without updated data to those specialties with updated data.

Medicare needs current data for all specialties to determine if the relative costs of operating a practice has changed across specialties. We would like the Commission to discuss different ways that Medicare could obtain more current information. One way is for a private sponsor, say a consortium of physician and non-physician groups, could collect the data and CMS could purchase the data from the private group. CMS staff have expressed an interest in this approach.

Of concern is whether all specialties would participate, particularly the 13 specialties with more

recent practice data accepted by CMS. If history is any guide, a voluntary effort, whether it's public or privately sponsored, will have a low response rate. Even when specialties collect their own data the response rate is low, about 20 percent.

A non-voluntary public effort may not be too popular with providers. One overarching issue CMS would need additional resources to obtain new data.

Moving to the second data source CMS uses to derive practice expense payments, it's called the direct resource database. You may have heard it called the CPEP database. Is it essentially a micro-costing database of the non-physician clinical labor, medical equipment and medical supplies required to provide nearly all of the services in the fee schedule. Here's an example of the direct resource for one urology service. You'll see here estimates for the clinical staff needed before and during the procedure, medical equipment, and medical supplies. So you multiply this by about 7,000 and that's the CPEP database.

CMS assigns a separate price to each of these direct resource estimates to estimate the total direct costs of a service.

Are the direct resource data accurate and complete? Getting the data accurate is especially important if CMS goes to a bottom-up method. We have found that there are certain challenges in maintaining the direct resource database. There are a lot of values here. Some of our initial concerns surround the accuracy of the database.

An AMA subcommittee called the PEAC, the practice expense advisory committee, went about between 1999 and 2004 and refined the values that were originated in the mid-1990s by the CPEP panel. The PEAC made assumptions about the use of labor, equipment, and supplies and applied these assumptions to similar codes called families of codes. It is unknown whether these assumptions have been applied consistently to all related services, particularly those services that the PEAC refined early in the process.

Having a continuing review process here may be worthwhile. Indeed, the agency has stated that there needs to be such a process but has not proposed any specific plan for doing so for both inputs and prices.

With a discussion about updating data you might have a question about the five-year review for practice expense. The statute requires the Secretary to review that

make adjustments to the relative values for all physician fee schedule services at least every five years. CMS has not yet proposed a five-year review of practice expense RVUs. The resource-based practice expense RVUs became fully implemented in 2002.

Ariel is now going to discuss some of the challenges in keeping the prices assigned to CMS to each direct resource up-to-date.

MR. WINTER: Before we discuss the options for keeping the input prices up to date there are some challenges to keep in mind. First, there are over 1,000 unique supplies and over 500 equipment items in the database so we need to be aware of CMS's administrative burden. Also, specialties have a weak incentive to request a review of overvalued input prices.

With that in mind, these are some options we're going to talk about for CMS to consider pursuing. One is to set a reasonable schedule for updating clinical staff wages, and supply and equipment prices. Second is reviewing the prices of new, expensive supplies and equipment more frequently. And finally, revisiting the assumption that all equipment is used at 50 percent of capacity, which is part

of the formula for determining equipment prices per service.

CMS last updated clinical staff wages for the 2002 fee schedule and has not indicated when the next update will occur. Wage growth for different types of staff varies. At the lower end, wages for lab technicians increased by 14 percent between 1998 and 2001. By contrast, wages for medical assistants grew by 63 percent. If wages are not updated regularly, services could become misvalued over time. Although an annual review of wages would be probably too burdensome for CMS, it is perhaps feasible to review them every three to five years.

As procedures shift from hospitals to physician offices, supplies and equipment become a more important part of practice expense. Supply and equipment prices were updated between 2004 and 2006. To update the prices CMS examined vendor catalogues and web sites and asked specialty societies for invoices. Manufacturers and specialties can ask CMS to change a price they believe to be incorrect. These groups have a stronger incentive to identify undervalued items than overvalued items. This is particularly a problem with regards to new, expensive supplies and equipment which can account for a large share

of a service's practice expense.

Prices for new items are likely to drop over time as they diffuse into the market and as other companies begin to produce them. Thus CMS should probably review expensive new items more frequently than older items, perhaps every year or two.

In fact, the AMA's relative value scale update committee, or RUC, recently requested that CMS re-price new high cost supplies annually. Because it would be too burdensome for CMS to review all of the remaining older items at the same time it could periodically review a sample of these items. The concept of re-pricing new items to reflect cost changes is similar to a recommendation you made in the March report, that the work RVUs of new services likely to experience reductions in value should be reviewed in a timely way.

Unlike supplies which are used only once, equipment is used repeatedly so CMS has to spread the cost of equipment over many uses. To derive the cost of a unit of equipment per service, CMS multiplies the number of minutes it's used for that service by the cost per minute. The cost per minute is based on the equipment's purchase

price, how frequently it's used, the cost of capital and other factors.

The frequency of use assumption is very important. If equipment is used at full capacity, the cost to spread across many services and the cost per service is lower. By full capacity we mean that is used during all the hours the practice is open for business. If equipment is used at lower capacity the cost is spread across fewer services and the cost per service is higher. Since CMS began using resource-based practice expenses is has assumed that all equipment is used 50 percent of the time.

Some equipment may be used less than half the time. This equipment would therefore be undervalued. And other equipment may be used more than half the time and would therefore be overvalued. The rapid growth of imaging services suggests that imaging equipment is used more frequently. Medicare spending for imaging grew by 60 percent between 1999 and 2003 to over $9 billion. This growth could be explained by new imaging providers entering the market, existing providers increasing volume per machine, or existing providers adding new machines.

We think that higher volume per machine probably

explains at least some of the spending growth because providers have a financial incentive to boost the use of expensive equipment. This is because a large share of the direct costs of imaging services are related to the equipment which is a fixed cost. Once imaging providers cover their fixed costs the marginal profit from each additional service is significantly higher.

This table illustrates the impact of changing the assumption of equipment use. Let's say a piece of equipment currently costs $100 per service using CMS's 50 percent assumption. If we instead assume that this equipment is used 75 percent of the time the price falls to $66.70, a 33 percent drop. This is because the cost is spread over more services. If we assume that the equipment is used 90 percent of the time the price falls to $55.60, a 44 percent drop.

It's important to note that the technical components of most imaging services are not currently valued using direct inputs such as equipment costs. Instead they are based on historical charges. Thus the impacts you see here would not apply to imaging under CMS's current methodology. However, CMS has given a strong indication

that it will eliminate the charge-based approach and instead use direct inputs to value imaging services. When this happens these impacts would apply to imaging equipment.

Here are some options for CMS to change its 50 percent equipment use assumption. First, it could develop a range of assumptions for different kinds of equipment. For example, rarely used equipment could be assigned to a 25 percent category, average use equipment could stay 50 percent, and frequently used items could be assigned to 75 percent.

One question to keep in mind is whether Medicare should pay for the higher cost of equipment that's really used. On the one hand, we have a principle that Medicare should pay for costs incurred by efficient providers. On the other hand, to not pay more could create access problems in rural areas or for services that are delivered infrequently.

A second option to improve this assumption would be for CMS to focus on expensive equipment which has the biggest impact on RVUs. Under either approach CMS would need to collect data on equipment use.

One option is to survey providers on their use of

equipment, perhaps as part of the practice cost survey that Nancy discussed earlier. Another option would be to analyze volume data from Medicare claims to see how frequently equipment is used.

We are testing the feasibility of both of these approaches which regards to two types of imaging equipment: MRI and CT machines. We are focusing on these machines because of the rapid growth of imaging procedures and the importance of pricing them accurately, especially because CMS has expressed a strong interest in using direct cost inputs to value imaging services.

First, we are fielding a survey of providers that have billed Medicare for performing MRI and CT scans. The survey includes physicians in freestanding imaging centers in the six markets listed on this slide. We chose these markets because they represent a range of geographic areas and a range of per capita Medicare spending. In addition, we have 100 percent Part B claims data for these areas. So in combination with the survey we will examine claims data on the volume of MRI and CT services performed by providers in these markets.

To sum up our presentation, we've highlighted some

concerns with the data used to determine practice expense RVUs, both the total practice cost data and the direct cost inputs. We've also laid out some options to improve the data. We're interested in getting your feedback on the issues we've raised.

Thank you.

MR. HACKBARTH: We have 20 minutes before we our scheduled adjournment and I'd like to allow at least a little time for the public comment period, so we've got maybe 15 minutes for commissioner questions and comments.

DR. REISCHAUER: Ariel, in this analysis you're going to do how do you know how many machines an imaging center has, number one? And of course, Medicare is not the only buyer of services. There are all the other folks.

MR. WINTER: Good questions. In terms of the number of machines per provider, we're hoping to get data on this from the survey. One of the questions we're asking is both how frequently do you use machines and how many machines do you have. So we can take an assumption from the survey. Another source of data is state certificate of need agencies for states that have these laws that approve the purchase and use of MRI machines. It's true for two of the

states in our sample, South Carolina and Massachusetts have these data at the provider level.

The other question is about what share of total services, total service volume is accounted for by Medicare. What we plan to do here is use an assumption derived from the National Ambulatory Medical Care Survey which is an NCHS survey. It has data on office visits by type of visit, so you can look at visits in which a radiology service was ordered or performed as well as data on the age of the patients. So we can take for all visits that involved a radiology service, what share were for elderly patients, and that would be our assumption for what share of the volume is for Medicare for these imaging providers.

DR. SCANLON: Two different comments. One first about the switch to the bottom-up method which is actually -- HCFA at the very beginning when the first rule for practice expense was proposed but never implemented there was up a bottom-up method too but it was different. I think I would you characterize the differences between what's being proposed now, if I understand what's being proposed now and what was done in the past is that in the past the practice expense values were a combination of the SMS data

and the CPEP, or PEAC-improved CPEP information and that right now CMS is proposing to use the CPEP information with price data to calculate the practice expense values for the direct components of these values, if that's right.

What it involves is an assumption that your data are good enough through the PEAC and your prices that you're going to get accurate estimates. The SMS data were used in the past to provide you a check on that. Now what we have to admit is that the SMS data are now six years old and for next year they will be seven years old so it's -- at best, some of it actually goes back to '95. So there's a question of how good of a check is that? So the assumption that maybe the PEAC data are better is potentially plausible but not verified.

This gets me to the second point which is the issue of the SMS data and what are we going to do into the future. I guess I've often or long felt that maybe what we need to really think about is mandatory reporting of this kind of information by a sample of providers whom we might compensate because they were unfortunate enough to be in our sample. We do not need the universe for this purpose but we do need the information.

What we should do in thinking about that is to think about the Medicare fee schedule as a public good because so many of the private plans use it. It's even of value to the physicians themselves in terms of understanding the differences in cost among services that they are providing. So without hearing the objections to that approach, that's where I've been leaning for a long time.

DR. KANE: On the input price piece, don't we have proxy inflation indices for all the other, like the hospital index? Why wouldn't we want to have a similar proxy inflation index rather than a direct measurement of wage increases for specific classes of labor in the office? Is there some reason we don't -- that's one of my questions.

The other is the 50 percent capacity assumption. I just don't understand how is capacity define? And then why would CMS want a 50 percent capacity rule rather a 75 percent? How did they come up with 50 percent? Was that just a political compromise or was it some balance of not wanting to over-incentivize excess volume versus -- they're incentivizing people to buy equipment that they're not going to use efficiently and I'm just wondering what's the root of the 50 percent? And also, why would we want direct price

measurement rather than proxy on an inflation index?

MR. WINTER: Good questions. On the first one, if you assume that all the inputs increased at the same rates, if all services had equal numbers of those inputs then the relatives wouldn't change. But it makes a difference where you have one service that uses a type of staff where the wages increase significantly, like medical assistants, and if you assume the average increase for that then their costs would be undervalued. If you took a service that used a type of staff where the wages increase slowly like lab technicians, at least based on the previous years they looked at, and you assumed the average wage increase, those services would be overvalued. They're overcompensating them.

DR. KANE: But it seems that we do that with hospitals and they have the same skill mix issues. I'm not sure I understand why we wouldn't do the same thing for physician practices.

MR. WINTER: With hospitals, that's used really to determine the market basket and here you're talking about estimating a resource for each specific service. You're trying to get the relatives right so the value of one

service reflects its true cost.

DR. KANE: You would know the mix of hours by skill mix and the input price would be reflecting a proxy for that particular skill mix. In other words, you would have the mix right but you wouldn't have the input. The input would be a proxy rather than a specific measure of the wage increase.

DR. MILLER: I'm not necessarily following this myself but let me bring some clarity to it for me. This is just for me. Aren't we talking about two different things here? You're talking about how to take a mix of services, increase them over time. I think part of what we're talking about here is because it's very service specific it's getting the mix right, because unlike in a hospital setting where you've got a large unit, if you've got the mix a little bit wrong and you're inflating it overtime there's a larger unit to put it over.

But here, the practice expense for an oncologist is extremely different, or the labor inputs for an oncologist are extremely different than some other and you're paying service by service -- I'm just wondering if we're talking past each other.

DR. KANE: I'm just talking about the input prices, not the skill mix.

DR. SCANLON: But I think what Mark is saying is what they have from the PEAC are real resource units but their heterogeneous units. They're hours of this type of labor versus that type of labor. In order to create the common measure, which is the estimated overall cost, they've got to start with the prices of those very specific things first. Then over time you could inflate things with an index until you decided that the mix of inputs had changed enough that you needed to go back to this first step, which is where they are now. They're basically at this first step trying to translate real things into monetary values.

DR. MILLER: To say it different way, we're trying to build the base that you would then inflate.

MR. WINTER: That's right.

DR. KANE: I'm just referring to page 10 where you're talking about updating input prices. I guess that's what I'm getting at.

DR. WOLTER: Just one comment.

MR. WINTER: Can I answer Nancy's second question about the equipment use assumption? This is actually sort

of a black box to us. In the 1997 proposed rule when they were developing the practice expense RVU system they said they hired a contractor, Abt, which recommended a 70 percent assumption. They did not cite any data in support of that. Then they said, based on comments we've received we've decided to go to a 50 percent assumption. It seems like that was a default that they went to, because they weren't able to get specific data on the use of different kinds of equipment across all payers and procedures. So because they weren't able to get this data they defaulted to 50 percent. That's the best we can make of it. As far as we can tell they've not revisited that since the decision was made in 1997.

DR. REISCHAUER: But this is 50 percent for use by everybody.

MR. WINTER: That's right, by everybody. Full capacity would mean if it were used during all the hours the practice operates, is open for business. So if you assume the average practice is open 50 hours per week, 50 percent capacity means it's used for 25 of those hours.

DR. WOLTER: I have a comment and then a question. The comment is really raised by the fact that you chose

imaging as one of the things to illustrate here. There is an interesting conundrum I think developing, and that is if you look at the issue of as the utilization increases and it's used 70 percent of the time or whatever, that would be a goal perhaps and it might drive how we looked at the resource use. As we look at these MRIs and CTs moving into small physician offices we're almost talking about creating an incentive for increased utilization of procedures which may or may not be always appropriate.

I'm very worried about what I'm seeing out there now in terms of the acquisition of this expensive technology in very small offices. I think that obviously doesn't get addressed here and wasn't intended to be addressed here but it strikes me as a paradox that we would try to price appropriately for 70 percent or 80 percent use in settings like that. So maybe we'll come back to that at another time.

My question is, issues around the geographic adjustment of practice expense have been raised in the past and I believe legislation about two years ago created a floor of some kind, 1.0 or something on practice expense. Am I remembering that right?

MR. WINTER: The floor was for the work, not for the practice expense.

DR. WOLTER: Thank you.

MR. WINTER: We do have on our work plan to look at whether -- the GPCI right now currently reflects an average use of supplies and equipment across all services which for services that use a lot of equipment and supplies like imaging could overstate on the geographic variations. So you might be overpaying in a high GPCI area and underpaying in a low GPCI area. So it's something we want to look at in the future is whether the GPCI could be changed to better reflect the mix of inputs where the prices actually vary geographically.

MR. HACKBARTH: Could we go back to Nick's first point because I was pondering the same thing, Nick, whether changing this assumption and making it more aggressive would encourage more inappropriate use. I guess the conclusion I came to is that if you look at it from the incentive facing the practice they profit, regardless of where this assumption is set, their incentive is to use the equipment more. They're going to move down their cost curve. They're going to increase both their total profit and their profit

per unit of service provided regardless where this assumption is set. I think that's right.

So I guess I persuaded myself that it was in the interest of the program to say that we should have a payment level that reflects a more efficient level of utilization of this service. It's not going to affect their incentive.

DR. WOLTER: I agree with that. I think that if the ability to profit on five imaging procedures per week as opposed to 100 that in fact there is a little different incentive there in terms of whether or not you want to put in the fixed cost of acquiring that equipment.

MR. HACKBARTH: That would argue for making a more aggressive assumption.

DR. WOLTER: Right. But I think my real issue is just at what point do we address the fundamental question about what settings is it appropriate for this equipment to be in, and self-referral and conflict of interest and that sort of thing.

DR. REISCHAUER: But once you have the machine, if your profit margin is $100 per use versus $5 per use your incentive to do more of it is greater when you're making $100 off it per unit.

MR. HACKBARTH: So again that would argue in favor of increasing the utilization assumption.

DR. WOLTER: There's two things. There's the barrier to entry and then there's if you think you can enter.

MR. HACKBARTH: Okay, we're going to have to conclude for today. Thank you very much.

We'll now have a brief public comment period. We only have about five minutes.

[Pause.]

MR. HACKBARTH: Thank you very much. Less than five minutes.

[Whereupon, at 11:39 a.m., the meeting was adjourned.]