THE NATIONAL QUALITY FORUM

NATIONAL VOLUNTARY CONSENSUS STANDARDS

FOR PATIENT OUTCOMES

CHILD HEALTH STEERING COMMITTEE

THURSDAY

MAY 6, 2010

The Steering Committee convened at

8:30 a.m. in Suite 600 North of the Homer Building, located at 601 13th Street, N.W., Washington, D.C., Charles Homer and Marina L. Weiss, Co-Chairs, presiding.
PRESENT:
CHARLES HOMER, MD, CO-CHAIR
MARINA L. WEISS, PhD, CO-CHAIR
DAVID R. CLARKE, MD, MEMBER
SHARRON DOCHERTY, PhD, CPNP (AC/PC), MEMBER
NANCY L. FISHER, MD, MPH, MEMBER
FAYE A. GARY, EdD, RN, FAAN, MEMBER
KATHY J. JENKINS, MD, MPH, MEMBER
PHILLIP KIBORT, MD, MBA, MEMBER
ALLAN LIEBERTHAL, MD, FAAP, MEMBER
THOMAS McINERNY, MD, MEMBER
LEE PARTRIDGE, MEMBER
DONNA PERSAUD, MD, MEMBER
GOUTHAM RAO, MD, MEMBER
ELLEN SCHWALENSTOCKER, PhD, MBA, MEMBER
BONNIE ZIMA, MD, MPH, MEMBER
MARK ANTMAN, DDS, MDA (via telephone)
LISA BERGERSEN, MD
JAY BERRY, MD, MPD
KERRI FEI, MSN (via telephone)
BARBARA FIVUSH, MD (via telephone)
KIMBERLEE GAUVREAU, ScD
CRAIG LILLEHEI, MD
ELLIOTT MAIN, MD (via telephone)
NINA RAUSCHER, MS, RN
SCOTT STUMBO (via telephone)

SONJA ZINIEL, MD

NQF STAFF MEMBERS PRESENT:

HEIDI BOSSLEY, MSN, MBA, NQF STAFF
HELEN BURSTIN, MD, MPH, NQF STAFF
NICOLE McELVEEN, MPH, NQF STAFF

ASHLEY MORSELL, NQF STAFF
NALINI PANDE, NQF STAFF
SUZANNE THEBERGE, NQF STAFF
REVA WINKLER, MD, MPH, NQF STAFF

NOT PRESENT:

JANE PERKINS, JD, MPH, MEMBER
C-O-N-T-E-N-T-S

Call to order and Welcome
   Co-Chair Charles Homer

Introductions

Recap of Previous Day's Deliberations
   Nicole McElveen

Measure OT3-27
   Ventriculoperitoneal shunt malfunction rate in children.
   Dr. Jay Berry

Discussion of Importance

Vote - Importance
Discussion of Scientific Acceptability
Vote - Scientific Acceptability
Discussion of Usability

Vote - Usability

Discussion of Feasibility

Vote - Feasibility

Vote - Overall on Measure OT3-27

Measure OT3-28
   Standard mortality ratio for neonates undergoing non-cardiac surgery
   Dr. Craig Lillehei

Ms. Kimberlee Gauvreau
C-O-N-T-E-N-T-S (Cont'd)
Discussion of Importance 68
Vote - Importance 85

Discussion of Scientific Acceptability 86
Vote - Scientific Acceptability 96

Discussion of Usability 96
Vote - Usability 107

Discussion of Feasibility 107
Vote - Feasibility 109
Vote - Overall on Measure OT3-28 109

Measure OT3-029 110
Standardized adverse event ratio for children and adults undergoing cardiac catheterization for congenital heart disease

Dr. Lisa Bergersen 110

Measure OT3-029 (Continued)
Discussion of Importance 112

Vote - Importance 121

Discussion of Scientific Acceptability 121
Vote - Scientific Acceptability 143

Discussion of Usability 144

Vote - Usability 145
C-O-N-T-E-N-T-S (Cont'd)

Discussion of Feasibility 145

Vote - Feasibility 145
Vote - Tabled 152
Measure OT3-031 153
Healthy term newborn

Dr. Elliott Main 154
Discussion of Importance 158
Vote - Importance 189
Discussion of Scientific Acceptability 189

Vote - Scientific Acceptability 190

Discussion of Usability 190

Measure OT3-031 (Continued)

Vote - Usability 192

Discussion of Feasibility 192
Vote - Feasibility 193
Vote - Overall on Measure OT3-29 193

Measure OT3-48 194
Plan of care for inadequate hemodialysis
Ms. Kerri Fei 195
Dr. Barbara Fivush 196

Discussion of Importance 198
C-O-N-T-E-N-T-S (Cont'd)

| Vote - Importance | 226 |
| Vote - Scientific Acceptability | 226 |
| Vote - Usability | 227 |
| Vote - Feasibility | 228 |
| Vote - Overall on Measure OT3-48 | 234 |

Measure OT3-46  
Validated family-centered survey questionnaire for parents' and patients' experiences during inpatient hospital stay  
Dr. Sonja Ziniel  
Discussion  
Vote - Importance | 288 |
| Vote - Scientific Acceptability | 290 |
| Vote - Usability | 291 |
| Discussion of Feasibility | 292 |
| Vote - Feasibility | 296 |
| Discussion | 297 |
| Vote - Overall on Measure OT3-28 | 305 |
| deferred until later |
Measure OT3-041  
Children who attend schools perceived as safe  
Measure OT3-042  
Children who receive the mental health care they need
C-O-N-T-E-N-T-S (Cont'd)

Measure OT3-044 321
Children who have inadequate insurance coverage for optimal health

Measure OT3-045 324, 342
Measure of medical home for children and adolescents

Measure OT3-050 327
Children who receive standardized developmental and behavioral screening

Measure OT3-044 328
Children who have inadequate insurance coverage for optimal health

Mr. Scott Stumbo 329

Vote - Importance 330

Vote - Scientific Acceptability 340

Vote - Usability 340

Vote - Feasibility 341

Vote - Overall on Measure OT3-044 342

Wrapup 348
CO-CHAIR HOMER: It is 8:30. It is a few minutes after, so I think we should get started because we have a lot more to cover today.

Good morning, everybody. Thank you, members of the Committee, for coming back after yesterday’s experience. That’s always a vote of confidence.

(Laughter.)

We do have a number of new members here, and we also have some new guests. So, should we just go around the room and everyone introduce themselves, first among the Committee members and then our guests and speakers and members of the public afterwards?

So, I will start. My name is Charlie Homer. I am CEO of the National Initiative for Children’s Healthcare Quality, and with Marina, always happy to co-chair the Committee.
DR. WINKLER: Hi, everybody. I'm Reva Winkler. I am NQF staff.

MEMBER PERSUAD: Donna Persaud, Parkland Health and Hospital System, Pediatrics, in Dallas.

MEMBER McINERNY: Tom McInerny, Golisano Children's Hospital, University of Rochester Medical Center.

MEMBER KIBORT: Phil Kibort, Vice President, Medical Affairs, Children's, Minnesota.

MEMBER FISHER: Nancy Fisher. I am the Chief Medical Officer at Washington State Health Care Authority.

DR. WINKLER: I will remind everybody to please use your microphones.

MEMBER CLARKE: David Clarke, pediatric cardiac surgeon, Denver Children's Hospital, retired.

MEMBER JENKINS: I am Kathy Jenkins. I am a pediatric cardiologist at the Children's Hospital in Boston and the Chief
Safety and Quality Officer.

And as I said yesterday, the Program for Patient Safety and Quality is a measure developer on the agenda for today. So, I will be recusing myself from that part of the discussion.

MEMBER PARTRIDGE: I am Lee Partridge, the Senior Health Policy Advisor at the National Partnership for Women and Families.

MEMBER GARY: I am Faye Gary, child psychiatric nurse, Case Western Reserve University, Cleveland, Ohio.

MEMBER ZIMA: I am Bonnie Zima, child psychiatry, UCLA.

MEMBER DOCHERTY: I am Sharron Docherty, the Duke University School of Nursing, and I am representing the National Association of Pediatric Nurse Practitioners.

MEMBER RAO: Goutham Rao from the University of Pittsburgh. I run the Pediatric Obesity Center at Children's Hospital,
MEMBER LIEBERTHAL: Allan Lieberthal, Kaiser Permanente, Panorama City, California.

MS. MORSELL: I am Ashley Morsell. I am NQF staff.

DR. BURSTIN: Hi. Helen Burstin, the Senior Vice President for Performance Measures at NQF.

Sorry I couldn't be with you yesterday. We had our board meeting. Kind of a hard thing to pass up.

MS. THEBERGE: Hi. I am Suzanne Theberge, NQF staff.

MS. BOSSLEY: Heidi Bossley, Senior Director, Performance Measures, NQF.

MS. McELVEEN: Good morning, everyone.

Nicole McElveen, NQF staff. We can now allow some of our guests to introduce themselves briefly.

DR. BERRY: Hi. I am Jay Berry, a
general pediatrician, Children's Hospital, Boston.

MS. GAUVREAU: Kim Gauvreau, also from Children's Hospital, Boston, a biostatistician.

DR. LILLEHEI: I am Craig Lillehei, a pediatric surgeon at Children's Hospital in Boston.

DR. BERGERSEN: Lisa Bergersen, a pediatric interventionalist at Children's Hospital, Boston.

DR. ZINIEL: Hi. My name is Sonja Ziniel. I am the Senior Survey Methodologist of the Program for Patient Safety, Quality, and Clinical Research Program at the Children's Hospital, Boston.

MS. RAUSCHER: And I have the privilege of serving as the steward for this group for Children's Hospital, Boston. I am Nina Rauscher, the Executive Director for the Program for Patient Safety and Quality.

MS. GALLAGHER: I am Rita Munley
Gallagher, Senior Policy Fellow in the National Center for Nursing Quality at the American Nurses Association. I have the privilege of supporting the work of the NQF nursing organizational members.

MS. McELVEEN: Operator, you can open up the conference line, and we can allow some of the participants who called in to also introduce themselves.

OPERATOR: All lines are open.

MS. McELVEEN: Do we have any Steering Committee members or audience members, measure developers, who have called in to listen to our meeting today?

DR. ANTMAN: Yes. Mark Antman from the AMA PCPI.

MS. McELVEEN: Anybody else?

(No response.)

Okay. I just wanted to quickly do a recap of our deliberations yesterday. I was looking through some of the measures to try to really capture how many we passed, how many we
would probably review on a future conference call, and how many the Committee just didn't feel were appropriate for endorsement.

There were about three measures which you did recommend for endorsement.

I am sorry, did someone call in?

(No response.)

There were actually three measures that we did review and move forward with endorsement on. That was the number of school days missed due to illness for children; children who have no problems obtaining referrals when needed, and, also, children who live in communities perceived as safe. Those are the three I have on my list.

We also tabled a few measures. Some were due to allow the measure developer to provide some further clarifications on a measure, and others were the larger-serving measures submitted by the CAHMI developer, and where NQF staff is going to work with CAHMI to, hopefully --
CO-CHAIR HOMER: If we could ask
the person on the phone who is calling in to
put his phone on mute? We are hearing a good
deal of static which is broadcast over our
speaker system.

Thank you.

MS. McELVEEN: So, we will look
into those larger-serving measures and gather
the questions and some of the additional
materials that you will need to fully evaluate
those.

There were about three measures
that were out of scope or either considered to
be a process measure, which again we discussed
yesterday possibly taking some of those
measures and moving them on to the second
phase of the Child Health Project.

And it looks like there was one
measure that the Committee agreed was not
appropriate for endorsement. That was the
children living with illness and the effects
of that condition on their daily life.
So, again, we will summarize all this information and get this out to the group in a meeting summary, but I just wanted to do a quick recap before we begin today.

Helen, do you have any comments?

DR. BURSTIN: No.

MS. McELVEEN: Okay. So, we are going to start with some of our more clinical measures, which will be a little bit of a change from yesterday.

We are in Work Group 1. So, if you all have the materials, either on your computer or printed, you can go ahead and pull up the table that we have compiled of the Committee reviewers, their initial comments on this particular measure.

The first measure we are taking up is No. 27. We do have our measure developers and a lovely team of folks back there who have worked on these measures.

Did you all want to take some time?
CO-CHAIR HOMER: Barry, would you like to present the measure?

DR. BERRY: Sure.

CO-CHAIR HOMER: That would be great.

DR. BERRY: Thanks very much for having me today. It has been a great opportunity to develop this measure with our pediatric neurosurgeons at Children's.

This measure reflects sort of bread-and-butter procedure by the pediatric neurosurgeons there. It is also very important to me. I have a clinic that is full of children with special healthcare needs, especially those who are technology-dependent. We are seeing a lot of readmission rates around these children, especially with malfunctions. So, that is why I was brought to the table to help these guys.

It has been fun developing the measure. In terms of the neurosurgeons' acceptance of it, it seems that most
neurosurgeons across the country feel that shunt malfunction is on their radar and it is something that they consider an outcome already.

So, the challenge for us was how to take that measure clinically and plug it into administrative data in order to pull out a valid measure. So, we spent most of our time searching through the codes and figuring out the best way to do that, and then, also, looking at populations that might be at risk and the case-mix adjustment issues in trying to figure out how to risk-adjust those things or whether to exclude them in the end.

So, there were a number of parts of the measure that we actually had built in initially as risk-adjustment, and then we ended up excluding them to try to homogenize the measure a little bit. That is why we excluded the population with spina bifida and also with other types of shunts that could be placed, that go not into the abdomen, but into
other places.

We have been using the measure for a while at our hospital. It has been accepted, and the neurosurgeons feel like it has helped change their care and their approach to the operation. So, we are proud of it.

CO-CHAIR HOMER: I would like to invite the other members of Work Group 1 initially to either make comments or ask the developers in areas. I would suggest we go through the sequence of the areas, the criteria that we need to do in order to approve, the first one being an indicator of the importance.

MEMBER RAO: Dr. Berry, just a couple of questions. I mean one of the questions that came up for me is, not being familiar with this area clinically, is, how common is shunt malfunction?

The other more important question from my standpoint is, how much of shunt
malfunction is actually due to procedural
issues as opposed to something that had just
happened spontaneously?

If you could address those two?

DR. BERRY: Right. So, in terms
of the commonality of it, we think that
probably you are looking at an overall average
of around 10 percent. So, 1 in 10 shunts are
malfunctioning within 30 days of being placed.

In terms of the variability of
that among hospitals, it seems that there is
around a four- to fivefold difference in the
variability. So, you can look at rates that
are going between like 3 to 25 percent. If
you expand out beyond 30 days, we see rates
that climb up much higher than that.

In terms of the quality of the
operation and how that can affect the
outcomes, the surgeons feel strongly that one
of the largest indicators of the shunt
survival is due to the actual placement. I
mean it actually is the angle and the
insertion of the shunt into the brain and also into the abdomen, and the way that the shunt is routed to make sure that it is not at risk for being kinked or broken, or that somehow it is being placed that would impede the flow of cerebral spinal fluid.

They also believe that there are a fair number of malfunctions that are due to infection. So, in the operating room, trying to increase the efficiency of the operation being performed, double-gloving, antibiotics at the procedure, et cetera, are all process measures that they feel relate to the outcome. So, they do feel that there is a strong bit of clinical happenings that are associated with the malfunction rates.

CO-CHAIR HOMER: Please, Faye. Please use your microphone. Thank you.

MEMBER GARY: Would you just say a bit more about infection? Could you just make one or two additional statements about the rates of infection and what kinds of
complications that might cause? And the other
issue is, what are the professional healthcare
providers that help take care of these
children, and did you get any feedback from
any of them?

DR. BERRY: Sure. So, in terms of
infection, the prevalence rates of infection
within the malfunction rates, you are probably
going to have around a third to a quarter of
these that will be associated with infection
in terms of the ones that are malfunctioning.

What infection means is that you
likely have bacteria that are getting into the
shunt. If the bacteria are inside of the
shunt, that is a direct route into the brain.
So, essentially, when you are talking about an
infected shunt, you are talking about treating
a child with suspected meningitis.

It is a problem. The shunt has to
be taken out. You are looking at maybe a 14-
to 21-day course of antibiotics, externalizing
the shunt. You still have got to deal with
the pressure when the shunt is removed to make sure the kid is safe, and then you have got to put another shunt back in. So, infection is a big deal, and they take it very seriously.

In terms of the co-management, the other operating staff, in addition to the neurosurgeons, feel like they play a heavy role into the process. Again, they try to really streamline as much as they can in the operating room the time of procedure and time to completion. So, having the surgical assistants there and everyone else onboard with exactly what is going on and making sure that they are comfortable with the procedure makes a difference.

When a child is out of the operating room, then at our hospital there is a good bit of co-management that goes on between some of the general and developmental pediatricians and the surgeons to help manage these children afterwards. Sometimes it is harder than you would think to determine
whether a child actually has a shunt malfunctioning or not. So, when a child has symptoms that are suggestive of it, oftentimes the surgeons will consult with us, if we know the children very well, to determine if we are highly suspicious of that happening or not.

CO-CHAIR HOMER: Dr. McInerny, Tom?

MEMBER McINERNY: Yes, I think this is a terrific idea. It reminds me a little bit of what we have been able to do with central line infections. You know, we used to consider them, well, that is just part of putting central lines in, and now we know that if you do things correctly, you can avoid that.

A couple of questions. I am wondering, in Boston are they using checklists when they are doing these?

DR. BERRY: Surgical checklists, I am not sure if they are using the checklists or not.
DR. LILLEHEI: Yes, as one of the surgeons in the operating room in Boston, yes, checklists have become a part, a required part.

MEMBER McINERNY: Okay. So, that should help.

And two, essentially, you are sort of providing a 30-day guarantee. I am wondering why you pick 30 and not, say, 60, 90, or 365 days. Any evidence to suggest -- because my experience has been 30 days, you know, you may get some, but another month or two or three later you are going to get more. So, where do you draw the line? Can you perhaps extend it to more than just 30 days?

DR. BERRY: It is a great question. We really argued about this for a while.

So, it seems that the majority of shunt malfunctions are occurring closer to the operation than later out. Now, if you do expand out to 60 or 90 days, you are going to
pick up more signal.

However, it was a little bit of a
dance with the neurosurgeons in terms of how
the quality of the operation was related to
the outcome. So, they sort of felt like, yes,
well, the further you are going out, the less
likely it was associated with a previous
operation. So, in that regard, we sort of
negotiated and ended up on 30 days.

However, I would say that I think
that we are minimally considering going out
further, if that is important to the group.

CO-CHAIR HOMER: Phil? If I could
also ask the questions right now, I would like
them focused on the importance question
particularly. We can deal with some of the
other issues as we go through, but go ahead.

MEMBER KIBORT: All right. So,
from my perspective, and I will concur that
there is importance there. I think most
active children's hospitals believe that this
is a major problem. I think there are data
about anywhere from 3 to 20 percent or 25 percent is true. So, for me, it is an important operation.

And in some hospitals, the hospitalists also take care of the patients post-op, as do our neonatal or our pediatric nurse practitioner hospitalists. So, it crosses different aspects, different professionals.

CO-CHAIR HOMER: David?

MEMBER CLARKE: Just one issue that I am not sure that the Committee is really aware of related to the importance of this measure is, what are the implications of shunt failure, particularly acute shunt failure, from the standpoint of morbidity/mortality of the patient, and also the cost? My impression is most of these are emergencies, particularly when they occlude. Would you comment?

DR. BERRY: Thank you.

So, they are considered
emergencies, and if not treated promptly, there is a high risk of death. If death does not occur, then you are looking at essentially a lot of permanent neurologic sequelae from pressure on the brain.

In terms of the economic impact, we were able to go back and look at some of the HCUP data from AHRQ that has been published on this. And it is estimated that there are probably around 10,000 admissions a year associated with shunt malfunction in children, and the mean cost of those admissions is around $17,000 to $20,000. So, you are looking, I think, at around $200 million annually just in shunt malfunction admissions.

CO-CHAIR HOMER: That is very helpful, David. That last point is the kind of data that I was looking for in figuring out the importance.

I understand the clinical importance and the frequency of shunts that
are put in that fail. One thing in the measure specifications, in your description, though, that concerns me is if it requires three-year averages, three-year running averages, to come up with stable rates sufficient for conducting analysis and benchmarking, what are the implications of that in terms of really the frequency and our ability to use it to actually track changes?

DR. BERRY: So, I was thinking more of the three-year running average less in order to collect the numbers --

CO-CHAIR HOMER: Okay.

DR. BERRY: -- but more to stabilize the confidence intervals of that, and, also, so that you are not trying to change or do not change the quality of care that you are doing for these things just because of a quarter where you may have looked bad or maybe a year. So, we thought that it stabilized the measure to median in terms of more of the variance than it did the actual
signal.

CO-CHAIR HOMER: Okay.

DR. BERRY: And that was sort of my approach to it.

MEMBER LIEBERTHAL: I would like to ask members of the Group 1 why they chose partially rather than completely as far as impact. We went from yesterday these very broad measures that had value as far as populations and government to now a very operational small volume, but to this specialty very important measure that is a true outcome measure. I wanted to know why people considered it only partially meeting the impact, that one.

CO-CHAIR HOMER: For me, it was basically we didn't have population prevalence data. We didn't have the financial data. The three-year average concerned me. I think this was the measure with the lack of improvability over time. Well, there was the variation -- I may be confusing with a different one where
it had been tracked, but there hadn't been changes. If I am confusing it, please tell me.

But those were why I put it only in the partially rather than the completely area. So, even if it was clinically -- again, I am operating a little on the assumption that NQF already has 600 measures, many of which are -- 450 -- many of which are clinically-accurate, but relatively low-prevalence conditions and so aren't going to have broad impact on changing.

So, that is why I wanted to make sure this was something that was not only sort of valid and clinically important for a very small subset, but actually was worthy of investing the resources in maintaining and continued for a significant impact. That was my personal reason for only putting it at partially rather than completely.

DR. BURSTIN: Just one comment on the criteria, and the way we read this
specifically was that impact could either be
in terms of broad impact, broad population,
big numbers, or a small population with a
significant impact on them. So, I think part
of what you have heard is the significant
impact on a small population is quite
reasonable as well.

MEMBER DOCHERTY: Yes, and I would
say that that was really my learning curve
over the past two days, is that it had to do
with the broad population. I was thinking
more of the impact broadly and now realizing
that this is a very strong measure of a
specific group.

MEMBER RAO: My concern is, and I
think Dr. Berry has addressed this to some
degree, I was under the assumption that the
vast majority of shunts are placed in a
handful of hospitals, and therefore, it would
be harder to pick up variation. But it seems
like there's a lot of different places where
they are performing the procedure.
DR. BERRY: That is right.

So, going back into the HCUP data, not in the nationally-weighted data, but in their actual sample from 38 states in 2003, there were over 300 pediatric hospitals that were performing these across the country. About 70 percent of those are considered by NACHRI to be teaching hospitals in some way. Thirty percent are community hospitals. So, we think that there is a lot more bandwidth out there for this than we initially thought.

CO-CHAIR HOMER: So, Tom?

MEMBER McINERNY: Yes, just a quick question. Is this somewhat similar to pediatric cardiac surgery in that, the more you do, the better you are, the less you are going to have some failures?

DR. BERRY: So, that is emerging. You know, the volume/outcome relationship for this over time is emerging as in cardiac surgery, yes.

CO-CHAIR HOMER: So, just
following our procedure, the first thing we need to do is vote. Because importance is a threshold we have to vote on the importance first, and then we can go on to discuss the other elements of the measure.

So, I would like to have a vote from the Committee on whether this measure meets the threshold criteria for importance.

So, all in favor raise your hand.

DR. WINKLER: Marlene, are you on the phone?

(No response.)

No.

CO-CHAIR HOMER: Terrific. So, it meets that criteria. So, let's move on to discussion of scientific acceptability.

Any members of the Committee, the Work Group, want to make any comments?

MEMBER CLARKE: I have a couple of points that I would like to ask about. Going back just tangentially to the 30-day issue, one of your data-gathering points was
reoperation for ventriculoperitoneal shunt
during the same hospital admission. I am
wondering, does the 30-day rule still apply in
that situation?

And I would also like for you to
comment on the exclusion of the children under
one-month of age or children with spina
bifida, which are known risks for shunt
failure that are, I guess, at this time
considered non-preventable. But my concern
about that is that, if you don't measure those
kinds of things, they never become preventable
because they are not identified.

One of the things that has been
applied, for example, in the STS database is
that the data is harvested, but these things
are initially excluded from analysis until
they determine exactly what their relationship
is to the overall measure.

Could you comment on those points?

DR. BERRY: Sure. So, if a child
receives an initial shunt, remains in the
hospital past 30 days, and has a shunt malfunction at 30 or greater days, then they are not counted.

In terms of the age less than one month, we are understanding now that there is a lot of treatment variability going on across the country in how to manage hydrocephalus in those kids with modalities that lie in addition to VP shunts, which makes it more complicated to study.

There is an endoscopic third ventriculostomy and a reservoir of things the neurosurgeons know much more about than I do, but they felt like it was better to pursue what is actually going on among the treatment modalities for those kids than to single out shunts in those kids less than 30 days for the measure. They thought that it made it more homogenous across hospitals to exclude them.

In terms of the spina bifida, I think you have a very valid point. So, when we initially created the measure, we included
spina bifida within the cohort and we risk-adjusted for it. Then, after some discussion, we thought it was best, again, to draw a nice circle around the measure and exclude the children with spina bifida because they weren't exactly sure what is going on and why their malfunction rates are so high.

We did have discussion yesterday about whether it would be appropriate to have a subdomain measure for those kids. I feel that that would be important for something for us to do as we test the measure and move forward. I agree with you, if there is a lot of signal without that group, we should not exclude it.

MEMBER PERSUAD: What percentage do you know of children who have shunts placed have shunts placed for spina bifida?

DR. BERRY: Total?

MEMBER PERSUAD: Yes.

DR. BERRY: I think you are looking at around 10 percent, 10 to 15 percent
at the most.

CO-CHAIR HOMER: So, Jay, could you talk a little about the validity assessment? Have you compared, for example, the chart review data with the PHIS data, et cetera? So, just technical aspects of the measure.

DR. BERRY: That was our first step. The neurosurgeons at first did not trust the administrative data whatsoever, which was a great process.

(Laughter.)

Luckily, they were collecting their own. They had their own registry, which was nice. So, they had all of their shunt patients lined up. Then, we went through and correlated that with the administrative data from our hospital first. That made them much more comfortable when they saw that the same patients were showing up.

I don't have specific specificity or sensitivity data for you, but there was a
litmus test of this work and the neurosurgeons bought it, which made me feel good. It seems that the codes are okay.

I mean to have a code that is specifically for ventriculoperitoneal shunt is very well-circumscribed. There is not a lot of noise in that code from other things that can be thrown in. And they had specific codes for shunt removal, shunt revision, et cetera. I think they have done a nice job upfront of sort of thinking about these codes. So, we like the face validity of the codes as they are.

CO-CHAIR HOMER: Any other questions about different aspects of that? Ellen, please.

MEMBER SCHWALENSTOCKER: This may actually be a feasibility question, but the one question I had is about the 30-day and whether they always come back to the same hospital, or how can you capture a 30-day rate if a child is admitted to a different hospital?
DR. BERRY: That is a very, very important question. So, the data that you see before you captures only kids who come back to the same hospital. We polled a number of neurosurgeons and did a few key informant interviews to just try to get a sense of could these kids go elsewhere. Because you could imagine if a kid lives in a more rural area, say, they have shunt malfunction; they may not have time to come back to the tertiary care center where they were operated first. However, the surgeons feel strongly that the vast, vast majority of the time the kids are coming back to the same hospital.

CO-CHAIR HOMER: Jay, did these data come from the PHIS database?

DR. BERRY: That's right.

CO-CHAIR HOMER: Describe how you have applied this or have you tested this with other discharge data for non -- since you said only 30, you said a very significant number of children have these procedures that are not in
tertiary children's hospitals, so would not be in the PHIS database.

DR. BERRY: That's right. That's right. So, beyond PHIS, we know that using the HCUP and AHRQ data, that the codes are being used across the country nationally. Now the problem with the AHRQ data is that we do not have the ability to link patients across hospitalizations at the moment.

However, there have been a few states that have been released in the last few months where they have their patient linker, which is allowing that process to occur, with Claudia Steiner from AHRQ, I think with the ultimate hope that they will be expanding out for longitudinal data as it grows over the next few years.

So, our next, I think, is to look into that small sample of AHRQ data, have some of the community hospital cohort included, and start to test the measure there to see if the codes are lining up appropriately.
CO-CHAIR HOMER: And somewhere in the specifications it says the measure hasn't been tested, but, in fact, you are using it. It sounds like maybe you are doing more than you gave yourself credit for.

DR. BERRY: That is a good question, Charlie. I wasn't sure exactly what the testing meant. I mean, in terms of what we have done at our hospital, we have done the chart review. It has been plugged into PHIS. We have looked at the rates and benchmarked and targeted against other hospitals, and we have acted on the data.

So, to a certain extent, I mean we are using it, but the gold standard to me, if you are really going to test it, I think, would be to go out and do a multi-institutional chart review and validation process to make sure that there is not a lot of coding variability, et cetera. That hasn't been performed.

CO-CHAIR HOMER: It really was
just performed at your institute? When you said you compared to registry data, that was really only within your own institution?

DR. BERRY: Exactly. Right.

CO-CHAIR HOMER: Okay.

MEMBER CLARKE: I have a question relative to your three-year rolling data plan. Does that apply only to single institutions or, if you are implementing it more broadly -- in other words, if the entire group of 70 institutions, academic institutions, are participating in the measurement, does that modify your need to do that?

DR. BERRY: This is a great question. I think it depends on what you are going to use the data for, what's the actionability of the data. I think if you are looking on a population level of are we getting better with shunt care, are we decreasing the malfunction rates, I don't think you need the three-year rolling average. I think you can do that on the population data
with a year of data and be fine.

I think, however, if you are trying to look at yourself and say, within this cohort of hospitals, are we doing better; are we doing worse; are we in the middle; do we need to think about changing or not changing our care, then I do like to incorporate the variance surrounding that measure and making sure that, before I say that my hospital has worse malfunction rates than Hospital B, that I sort of look around the noise of that signal and make sure that it is not due to the noise, that it is the signal. So, I would say go for the three-year if you are doing that.

CO-CHAIR HOMER: Just one more technical question on this, and then I think we can probably on to vote on this. Have you looked at disparities issues? Have you looked at variation in rates across different populations?

DR. BERRY: We did, and this is
what we found. So, in our bivariate analyses we found that non-Hispanic Blacks have higher rates of shunt malfunction compared to Whites. Now when we threw that into a multivariate model, controlling for other things, the effect went away. So, I am not sure if there is something there or not, and at this point I felt that it was best not to cull it and put it into the measure.

However, if people are thinking that it is important to present within the measure subdomains rates within different race/ethnicity groups, we are certainly amenable to doing that.

CO-CHAIR HOMER: I mean the NQF process is to stratify results by different populations rather than adjust. So, okay, the point is, again, you looked at it and that it is feasible to look at within the dataset.

DR. BERRY: Yes.

CO-CHAIR HOMER: I think that is what is important from the NQF perspective.
Members of the --

MS. BOSSLEY: This is Heidi. Can I just jump in?

Going back to the evaluation criteria, which all of you have been working off in rating all these measures, the key piece on testing, I want to make sure you all understand why staff rated this as not tested is it hasn't had reliability testing, the test/retest or some type of look, and it hasn't gone through any validity testing as well, which is something that you all can decide is okay for this measure, but we would really feel that it needs to have a time-limited endorsement, which means they have 12 months, or we will negotiate with them -- I think sometimes it takes a little longer -- to come back and provide that information.

I think the key piece that we always want to make sure is any measure you put out there for public reporting, anyone else who goes and does the same thing with the
specifications that they provide can be replicated to the greatest extent possible. We don't know that yet, that you can with the way this measure is specified. So, that is really, I think, why we had it labeled as needing time-limited endorsements.

Does that make sense to everyone?

CO-CHAIR HOMER: I think NQF, quite appropriately, is tightening its criteria. Certainly, this is more tested than a number of measures I know when I was on the Ambulatory Steering Committee -- (laughter) -- which was we sort of kind of think this is a good idea, and we could actually pull the data. That was viewed as testing.

This one has been validated in one site, but not in multiple sites. And test/retest in this seems like that, you know, with administrative data, I am not sure that concept is really quite applicability, but that is probably getting too deep into the weeds.
DR. BURSTIN: I'm sorry, we can also take just a closer look at the testing and get back to Children's as well, just to be sure.

MEMBER DOCHERTY: Yes, I was just going to say that I was less worried about the validity than the reliability, and that there should be some formal measure that across sites people are --

CO-CHAIR HOMER: Well, I guess we go through all the criteria and then we vote. We will go through.

So, let's move on to -- and we may have already addressed this -- the usability? Do you want to vote on each section? Okay. I forgot.

So, then, to vote on the scientific acceptability, how many feel it is completely meets criteria?

And how many feel it partially meets criteria?

Okay. And does that get everybody
or are we down to minimally -- okay, good.

All right. So, moving on to usability, that is, is it understandable? Is it harmonized? Are there any other measures out there? And does it provide added value?

Any comments from the Work Group on that?

MEMBER PERSUAD: I think I would like to ask, you said you are using it at your institution. So, what I want to know is, what has happened since you started the measure? What has happened to your rates and what have you done?

DR. BERRY: Since starting to measure, I think the first thing that happened there was a little of a Hawthorne effect going on, which was fantastic. I think it just got people thinking about malfunction.

And it also had non-neurosurgeons thinking about malfunction as well. I mean hospitalists, other people when they were admitting to our services said, "Jay, we had
another kid that was readmitted with a shunt malfunction," you know, blah, blah, blah. So, it created a lot of buzz.

The second thing that happened is that the neurosurgeons really felt like they needed to streamline the amount of time it was taking to perform these operations initially in the OR, and that they really needed to have a core competency within a small group of staff in the OR to make sure the operation was right.

So, they have actually tried to decrease the number of personnel that are physically in the room during the operation because they feel like the more people that are there, strictly adding another person may increase the risk of the child having an infection. So, they really are trying to make a difference.

We have seen some small decreases in our rates. Now, if you look at the confidence intervals around that, they haven't
changed significantly, but we have seen a little decrease in our signal since the measure was put onboard.

CO-CHAIR HOMER: Any further questions about usability?

(No response.)

All right. So, why don't we call a vote on how many feel this completely meets usability criteria?

That's everybody, right? No?

DR. WINKLER: You're a partial?

Many people are looking at Ellen back there. Partial? Okay, good. Okay.

MEMBER PERSUAD: Ellen, are you a complete or partial?

MEMBER SCHWALENSTOCKER: Partial.

CO-CHAIR HOMER: Okay. All right. So, let's move on to the feasibility. I think you have already addressed many of the questions there, which is that it is feasible within the PHIS database, may be feasible in the other ones, but hasn't been, because of
the idea of whether you can actually track individuals over time, hasn't yet been applied. Is that correct?

DR. BERRY: That is correct. I think that is the data that you are going to need to really establish your targeting and benchmarking.

I mean I would hope that most hospitals across the country have enough admin data in terms of for every admission they have the procedures and the diagnoses that occur in order to bill for them, that they have an internal structure which from their admin data they can pull their own rates.

So, I think you are looking at more the national databases, then, to determine, okay, how well are we doing compared to other hospitals?

MEMBER LIEBERTHAL: What exactly is the PHIS database?

DR. BERRY: So, the PHIS database is a database of inpatient hospitalizations
for 42 freestanding children's hospitals across the country. It is unique in that the patients are linked across multiple encounters. So, you can track a patient over time to see the number of times they are hospitalized, and for each admission you have the diagnoses and procedures that occurred, demographics, et cetera, to allow you to pull data such as this.

MEMBER LIEBERTHAL: So, for patients who do not receive the procedure at one of these 42 hospitals, their own administrative data would have to be used, is that correct?

DR. BERRY: That is correct.

MEMBER LIEBERTHAL: And you are basing this on assumptions that they have accurate databases that they can pull data from?

DR. BERRY: Yes, some type of administrative billing database that the hospital would use for their coding, which
would be the same that PHIS is pulling from our hospital. It is the same sort of coordinated set. But the assumption would be that they have that same similar dataset.

MEMBER CLARKE: I would like to ask if there exists a Neurosurgical Society-based database that would cover this issue, and would that be useful?

DR. BERRY: Yes. So, one product that has emerged from this work to start is the creation of a multi-institutional Hydrocephalus Collaborative, which is now being started up by John Kestle in Utah, and one of the collaborators in some of our work, Tamara Simon.

It is really good stuff, prospective data collection, looking at very, very specific variables around quality of care around the shunt procedure. Hopefully, we will see data from them in the next year or two.

CO-CHAIR HOMER: That's very
exciting.

Nancy?

MEMBER FISHER: I would like to make a comment about this. In the State of Washington, we have been doing collaboratives like this. We have done it around cardiovascular surgery. We now are including some things in cardiology and PCI. We have done it around surgical procedures that we thought, like for appendectomies you ought to be able to do an appendectomy. And we were quite surprised to see the variation.

One of the things is that we have used the three-year rolling average. It does eliminate problems when people think that they are being unfairly targeted for something that it was just it happened.

The core thing that we found that was going on when you started looking at administrative data was one is the people that were extracting the data. And we even get asked by the hospitals to send out people to
make sure that we could look at this and do validation on it.

The other thing is, when you first start out, whether people take it, if you are collecting the data, they say they will, if they take it seriously. They have been pinged and looked bad because they did sloppy data collection. But all you have to do is be pinged and you put yourself back together.

The other thing that I found good about what he was saying is, if you want to go into different hospitals, what we found out is the key is to get a physician in that specialty to be your champion. That is the way to get in. This is one way -- I mean I am very glad that you realize about the data and stuff because the first thing we had to do was get this data. People got sick of hearing about it. So that we answered everybody's questions about the data, so they could believe, yes, maybe you do have a problem.

I am really happy to hear that you
are going to do a collaborative about that.

CO-CHAIR HOMER: Tom?

MEMBER McINERNY: Just a quick

question, sort of suggestion. As more and

more hospitals migrate to electronic medical

records for both their inpatients and their

outpatients, would you foresee that maybe

sometime in the future you would be able to

use that data and get rid of the

administrative data?

DR. BERRY: Oh, boy, that would be

absolutely fantastic. I mean to move beyond

codes, to move into a lot of clinical detail,

the size of the shunt that is placed, the time

in the operating room, you know, very, very

specific clinical details going into it will

trump this stuff like no tomorrow. So, I

can't wait for that day.

CO-CHAIR HOMER: All right. So,

in terms of feasibility, I suggest that we

call a vote.

How many feel this completely
meets criteria for feasibility?

        And partially?

        Okay, good.

        All right. So, now it is time to call the vote on the overall measure. I think the recommendation we are hearing from staff would be that this should be recommended for time-limited approval, pending additional testing. I think particularly the idea of looking at validity across multiple institutions and potentially expansion beyond the CHCA dataset seem to be the two areas we would like to see additional testing on.

        CO-CHAIR WEISS: With no specificity about the time limit, right? They would work that out?

        CO-CHAIR HOMER: It's 12 months, generally?

        DR. BURSTIN: It's generally 12 months, but if there's a little wiggle room, we can do it.

        CO-CHAIR HOMER: Okay. So, all
those in favor of conditional approval -- I'm sorry -- time-limited approval? Thank you.

There you have it. All right.

Thank you very much.

DR. BERRY: Thanks for your time.

Thanks for everything.

MEMBER PERSUAD: Charlie, I have just two final comments about that measure before we pass it.

CO-CHAIR HOMER: Yes, please, Donna. We did pass it, but before we move on it.

MEMBER PERSUAD: Well, before we move on it.

CO-CHAIR HOMER: Okay.

MEMBER PERSUAD: One is I may have just blanked out over the discussion regarding when children get readmitted from different institutions, and if it's possible to work that out in the follow-up period through the PHIS. I don't remember what he said. There is a way to do it with the PHIS database, but
getting it cleaner to where not only from the institution where you did the surgery, if you readmit to another hospital for shunt malfunction, if you can get that into the data? It may not be doable just yet, but --

CO-CHAIR HOMER: If they do that collaborative, it would be.

MEMBER PERSUAD: Yes, if they did a collaborative, I guess they could sort that out there.

CO-CHAIR HOMER: That's a good question. I guess that is the challenge of not having a Medicare database, that you can't track individuals across institutions. But, okay, so something during the test period to encourage them to look at. That is a great suggestion.

MEMBER PERSUAD: And, then, speaking for Marlene in her absence over the toolkit issue, since this group has a checklist already, where the measure is published, the checklist could become
available or I guess the collaborative would probably come up with tools for having better rates.

CO-CHAIR HOMER: So, that is a question I think for Helen, which was Marlene --

MS. RAUSCHER: Could I just ask the measure developer to come in?

CO-CHAIR HOMER: Sure.

But just sort of more as a policy or process, Marlene Miller suggested yesterday that, when we approve or consider a measure, the idea of linking that to a quality improvement toolkit. I didn't know whether NQF had considered as part of its process making those available together with their measures.

DR. BURSTIN: We haven't done that to date, but we are moving towards trying to create this relational database. We are calling it MAPS, Measures and Practices. It will try to package everything together saying
here's the measure; here's the practice;
here's related information. It is all sort of
developmental, but that is something we can
work on as well.

CO-CHAIR HOMER: So, Jay, there
were really two sets of questions that were
raised. One was the idea of linking across
institutions. So, a child gets operated on at
Boston Children's and shows up at some other
institution in town, for example. Is there
the capability or at least can you look during
the testing period at that potential to look
at? That was one of the questions.

DR. BERRY: I think we should
explore it. I am wondering, I think that AHRQ
may actually have more data that is just not
publicly available yet.

CO-CHAIR HOMER: Okay.

DR. BERRY: And I feel comfortable
talking with them and asking them if we could
do something like that through the state
inpatient databases and merging them together.
CO-CHAIR HOMER: And the other question was, building on conversation we had yesterday, was the desirability of linking measures with quality improvement-related toolkits. NQF is in the process of putting together a database that you could just cue up your issue and you would link a variety of things. So, I think more expression of interest in that toolkit being made broadly available as the collaborative moves forward.

DR. BERRY: Sounds great.

CO-CHAIR HOMER: Did I capture that, Donna?

MEMBER PERSUAD: Yes. Thank you.

CO-CHAIR HOMER: Good. All right, thank you very much. That is really great.

MS. McELVEEN: Okay. Moving on to our next measure, Measure 28, is the standardized mortality ratio for neonates undergoing non-cardiac surgery. This is the ratio of observed-to-expected rate, observed to -- yes, ratio of observed-to-expected rate
of in-hospital mortality following non-cardiac surgery among infants less than 30 days of age and risk-adjusted.

So, this is, again, under the same group. We will open it up for importance.

CO-CHAIR HOMER: Or should we ask the presenters --

MS. McELVEEN: Oh, sure. Yes.

Absolutely.

CO-CHAIR HOMER: -- to briefly describe the measure?

Maybe also, Dr. Lillehei, having heard the conversation before, maybe sort of focusing some of your comments on some sort of sequentially thinking about the importance of the measure and then its scientific credibility, et cetera, that would be great.

DR. LILLEHEI: Certainly. I can try to do that.

CO-CHAIR HOMER: Thanks.

DR. LILLEHEI: Together with Kim Gauvreau, the statistician, I am a pediatric
surgeon, and together developed this model.

One of the problems that we face in surgery is, obviously, an ability to risk-adjust, and that is particularly a problem in pediatrics and pediatric surgery, where we have a wide variety of different sort of problems that present in children.

What we have done especially is for a variety of different diseases, we will pick a disease and then study that either in a particular institution or across institutions. But, again, it is limited to that. Whether it is gastroschisis or omphalocele or diaphragmatic hernia, necrotizing enterocolitis, a variety of those studies have been done.

We were looking for a broader measure, and specifically a broader measure to look at neonatal surgery. So that, as you see in our measure, we are focusing on operations that occur within the first 30 days of life. Those are primarily congenital lesions, but,
again, they are infrequent. Surgery in 
children is, fortunately, rare. In order to 
develop a risk-adjustment method, we used a 
strategy where we could combine procedures of 
similar risk.

Now, in order to do that reliably, 
first of all, we chose to use a large national 
database. So, this is based on the KIDS 2000, 
where we have developed the model. And within 
that, we focused on procedures that had at 
least 20 cases. In those circumstances where 
there were just a handful of cases, we didn't 
feel that that gave us a reliable tool for 
assessing risk. So, they are limited to cases 
that are greater than 20 cases.

Of those, there are 63 different 
procedures that are included out of a total of 
some 570, something like that. But, in fact, 
those 63 cases account for almost 85 percent 
of all the procedures done. So, we think it 
is based on a large sample from that KIDS 
database.
With that, then, to take you through the rest of the model, we developed risk categories based on that. Then, as a measure, we thought the most reliable measure and, in fact, in many ways the most important measure for us was in-hospital mortality. So, that was something that we could measure that we felt was reliably reported, even within an administrative database, and would allow us to make assessments from institution to institution and risk-adjust appropriately.

CO-CHAIR HOMER: Kim?

MS. GAUVREAU: I don't have anything more to add at this time.

CO-CHAIR HOMER: Members of the Work Group, questions? Again, probably start with maybe first your observations and then questions, starting on the importance first.

MEMBER RAO: Sure. I think that my concerns and observations were actually the second lengthy comment up there. It is just, essentially, I am not sure how many children
actually undergo surgery in the neonatal period. The incredible variety of procedures, even 63 procedures, seems like it is too heterogeneous to mix together.

And finally, just looking at the data that you had, only one hospital had a significantly different rate of neonatal mortality listed there.

So, if you could address some of those points?

DR. LILLEHEI: There's no question that the heterogeneity of cases is a challenge in making that assessment of neonatal surgery, newborn surgery. But I think the most telling thing about that is that, in fact, when we made this analysis, we derived it in the KIDS 2000 base and then validated it in the 2003, that our ROC curve, that it was actually quite reliable, that 90, 92.92 was the -- they were under the curve with the ROC curve. So, that it showed that we really were able to reliably risk-adjust in this population.
So, to be sure, it is a challenge, and there is a considerable heterogeneity, but, in fact, the results seem to underscore that.

CO-CHAIR HOMER: When it comes to the second question of the lack of variety, can you describe the variety across sites?

DR. LILLEHEI: Yes. We presented the table where you saw 15 different institutions. I think what you could see within that table is there was a considerable variability from institution to institution. However, you are quite correct that it was only one in which that achieved statistical significance. So, it may be that, by virtue of the fact that we are dealing with relatively small numbers, that we would need a larger time period to accrue greater numbers and highlight some of those differences from institution to institution.

MS. GAUVREAU: Right. We were only looking at one year of data in that case.
So, maybe something like we were talking about with the previous measure, maybe it needs to be a two-year measure or a three-year measure.

We have started looking at this method a little bit in the PHIS database, and in that case we were using three years of data to look. The confidence intervals are narrower in that case.

CO-CHAIR HOMER: Faye, please.

MEMBER GARY: I was wondering, across the 15 different institutions and collectively, do you have data about the mortality for subgroups of populations, African-Americans, American Indian, Hispanic, et cetera? Do you have those data? And if you do, could you share those with us collectively? And if you see a variability across the 15 different institutions?

DR. LILLEHEI: No, you've cut to the core, but the exciting thing about this method for us is the ability to look at that administrative database, which will have
access to things like race, insurance, type of hospital, those sorts of data. But I don't have that data for you today, no.

MEMBER GARY: Those data are forthcoming?

DR. LILLEHEI: Well, that is something that we are working on right now. We developed that model for just that reason, to be able to look at those sorts of issues. You bet.

CO-CHAIR HOMER: Tom?

MEMBER McINERNY: I think we all know that morbidity, I mean mortality in children, even in neonates, is so rare that it makes it difficult sometimes to come up with significant differences. I wondered if you have looked at, if you could take a combination of morbidity and mortality, such as needing a second operation or getting an infection, and if you might, then, be able to get a bit more variability in the data by adding in morbidity to mortality?
DR. LILLEHEI: To be sure, the more we know and understand of how that surgery impacts on children, the more valuable that becomes. The problem for us or the challenge for us is that in administrative databases they have certain limitations, of which you are all quite aware. We really felt that what we wanted to be able to generate is a very reliable method for assessing that risk. We thought that, as such, mortality is the most reliable measure to do that.

CO-CHAIR HOMER: David?

MEMBER CLARKE: Yes, I certainly agree with that. I have been there and tried to do that, and it is tough.

The one thing on my first run-through of the application that I came across that I think is a fatal flaw, but it apparently is not, potentially a fatal flaw, is basing the total project on operative cases. Cases don't die; patients do.

When you have patients who have
multiple procedures, particularly, this becomes an extremely unreliable method of measurement. But I was talking to Kim a little earlier, and apparently this is not exactly true.

DR. LILLEHEI: Yes, Dr. Clarke, we perhaps didn't make that clear enough in our application. But, in fact, no, it does refer to specific patients. When a patient has more than one procedure, which obviously is a not uncommon occurrence, they are assigned to the highest risk category associated with those procedures.

MEMBER CLARKE: So, the mortality is only attributed to one procedure? Because, otherwise, you wouldn't know what you were talking about.

DR. LILLEHEI: Correct.

MEMBER CLARKE: The other question that I had was the ordinary criteria for operative morality is either death during the same hospital admission or death prior to 30
days. I am wondering whether your limitation
on data is the reason that that is not
included or whether it would just be so rare
in the neonatal population that a child would
die within 30 days of operation but after
hospital discharge, that it is probably
unnecessary to look at that. Any comment?

DR. LILLEHEI: Well, fortunately,
it is rare, but to be sure, in-hospital
mortality was once again what we felt was the
most reliable piece of data that we could gain
from that administrative database. So, that
is why it was chosen.

MS. GAUVREAU: And also because
the database we were using didn't allow us to
link multiple admissions on the same patient.
So, if a patient was discharged, was
readmitted, and died subsequently, we wouldn't
know that.

MEMBER DOCHERTY: My concerns were
similar to David's in that I was trying to
make sure I understood that the mortality
associated with these infants could be directly related to the surgery and not other things that happen, the other morbidity that is associated with this age group being hospitalized. But you are pretty certain that you will be able to, that the database will be able to link that mortality specifically to their surgical outcomes and not iatrogenic things for infants in hospitals?

DR. LILLEHEI: Yes, I think that is a good question. A couple of different things.

No. 1 was that we looked at a variety of different clinical variables that might impact on outcome, mortality in this case, and there were only two that showed up in that. That was serious respiratory diseases and necrotizing enterocolitis. They are included, those two clinical variables, and only those two, are the ones that are included in our model and are part of this risk-adjustment method.
CO-CHAIR HOMER: So, just to focus the conversation, this has been a fantastic conversation, but, again, the first question we want to ask is, is this important enough? That is, either prevalent enough or for a certain population important enough for us to feel that it is worth proceeding with the review of the other detailed attributes of the measure.

MEMBER RAO: As the measure is currently structured?

CO-CHAIR HOMER: As the measure is currently structured, right.

MEMBER CLARKE: I have a question.

CO-CHAIR HOMER: David?

MEMBER CLARKE: Do we have to rely on what's present in the application or do their comments contribute? Because it changes everything.

CO-CHAIR HOMER: The comments count. I mean that is why they are here. So, your main concern was this
issue of if there were multiple surgical procedures done?

MEMBER CLARKE: Right, and using procedures as your basis, and then you have multiple procedures per patient. Then, you try to attribute mortality to what procedure. It was not clear how that was going to be handled. But the way that it is being handled I think is perfectly appropriate.

CO-CHAIR HOMER: So, going back to this question of in-hospital mortality versus a 30-day kind of mortality figure, have either you or anyone else looked at basically a survival curve post-surgery of when, for children who do die post-surgical, when that actually happens and how many or what proportion of deaths might be the child is home and an untoward event happened, and they ended up being rehospitalized? Do you have any sense, either from this analysis or from Medicaid claims or other cohort studies, or anything like that?
DR. LILLEHEI: No, Mr. Chairman, I don't think we have it beyond my own experience as a pediatric surgeon that, yes, when children die of neonatal surgery, that is typically they don't make it home, yes. Yes.

CO-CHAIR HOMER: That makes sense.

Nancy, please.

MEMBER FISHER: I think I just need a little bit more clarification because what I am trying to see is, first of all, my understanding is that there is a small percentage of kids that die from surgeries in this age in pediatric hospitals, that we are looking at a small number of people. Then, you start talking about things like gastroschisis, operating on neck, operating on -- my question is, you have to operate on those kids with gastroschisis or they will die. So, my question is, what exactly are you trying to get at to improve?

I am just sort of confused about it because it is not like you have a choice.
I mean, have you run some data, and what you are talking about is the kids that died from the operation because it was the really the infection; it wasn't the procedure? If you don't have a choice about operating on somebody, you have to operate.

DR. LILLEHEI: I think part of -- Kim, did you want to respond?

MS. GAUVREAU: I was just going to add the piece of information that in the KID database that we looked at to develop the model, there were about more than 5,000 neonatal surgeries, and that does not encompass all states. That is only an 80 percent sample of cases. So, we were able to extrapolate that probably somewhere between 9 and 10 thousand surgeries happen in the United States each year, just to put it into context.

CO-CHAIR HOMER: Nancy, if I could answer just for a second?

MEMBER FISHER: Yes.

CO-CHAIR HOMER: I mean, if we
I think of the last presentation on shunts, I mean those children need their shunts, too. So, the question is, is there variability? Children will need surgery for their gastroschisis or the other conditions. And the question is, is there variability in mortality rates across institutions that, presumably, is attributable to some element of the care that they are receiving in those institutions? I think that fundamentally is, is there improvability based on variability across institutions for kids who need surgical procedures?

MEMBER FISHER: I see this a little bit different from the last one. I just can't get my hands around -- you know, to me, I don't know, I just can't seem to get my hands around how you are going to improve it. Maybe it is because the other ones had said they had looked at it, and they said that, when they did the procedure, it was the angle at which they put the shunt in. I can see
that it is a device that you are putting in someone, and the device could malfunction or you could do something to the device. I am not seeing that with these procedures. I guess that is the difference.

        CO-CHAIR HOMER: Okay. Dr. Lillehei?

        DR. LILLEHEI: Well, I think that your point is well-taken that this is not a specific surgical, telling an individual surgeon or identifying even an individual surgeon as to their specific outcomes, but we are really looking at a broader level at an institution, all of those things that impact on successful surgery in neonates.

        And what we anticipate is that, in fact, we will see variability, whether it is from institution to institution, whether it is socioeconomic groups, whether it is regions. To understand by identifying that variability, then, hopefully, we can move to the next, which is to say, how can we impact that; how
can we change that, whether it is access to
care, whatever that might be?

But you are right, what we are
looking at is for what that variability is and
then how we might use that to change practice.

CO-CHAIR HOMER: Thank you.

Faye?

MEMBER GARY: I just have a quick
question, please. That is, I know the study
involves 15 hospitals. Is that correct?

DR. LILLEHEI: No, no. Actually,
the study, we just gave you data, a table of
15 hospitals for the evaluation form. In
fact, we have applied it to the entire KIDS
database, which is surgery in -- what? -- 37
states that provide data to the KIDS. We have
actually applied it to the PHIS as well. So,
no, our intent, this is population-based.

MEMBER GARY: Well, then, I would
just amend my question just a little bit. I
was wondering if you also have any data, or
will have any data, about the basic
characteristics of the hospitals or the
populations that these hospitals serve.
Because I think that, based on the populations
that they serve, you will probably see some
influence with the outcome. How will you
address that issue?

DR. LILLEHEI: Absolutely. Those
are elements that are available in the KIDS
database. So, the nice thing about using this
database is it will allow us to interrogate
just those sorts of questions about what other
factors related to those patients.

CO-CHAIR HOMER: I have just one
more observation, and maybe the NQF staff can
help me on this. I see this as analogous to
the hospitalized standardized mortality rate
measure in adults. It is not precisely the
same because that is overall hospital.

I mean my understanding is you
probably do have on the adult side some
indicator of hospital-specific mortality rate
or standardized mortality rate that you use as
an overall performance measure, or not?

DR. WINKLER: Well, if you are
talking about a multiple-surgery, across-the-
board kind of measure -- is that what you are
talking about for adults in a surgical
measure?

CO-CHAIR HOMER: Well, I know,
again, at the IHI they certainly use as
quality improvement --

DR. BURSTIN: We have not brought
on any of the HSMR measures. I'm sorry. None
of the HSMR measures have come to NQF yet.
All of our mortality measures tend to be
procedure- or condition-specific, although we
do have composites of selected mortality for
certain procedures.

DR. WINKLER: In another part of
the project, the main Steering Committee is
reviewing surgical complications, which
includes mortality, but also other serious
morbidities for all of the age 65 patients,
but it encompasses a wide variety of
surgeries. So, we are getting there.

MEMBER McINERNY: And, Charlie,

there is the NSQIP, the National Surgery Quality Improvement Program.

CO-CHAIR HOMER: Right.

MEMBER McINERNY: Which is for adult surgeries. But there is now also a pediatric NSQIP that is up and running. I don't know if they reported any data yet. I think they are still collecting it. So, this would be similar.

CO-CHAIR HOMER: So, then, why don't we vote on, if there are no questions further, let's vote on the importance question.

So, all those who feel -- again, this is a dichotomous, yes, this is sufficiently important based either on its prevalence or within a narrower clinical area, which this, presumably, isn't, but within a narrow clinical area, whether it is an important measure.
So, all those who feel this meets the importance criteria raise your hand.

Good. Okay.

And all those who do not feel this meets the criteria?

Good. Thanks.

I think we have already had a good bit of discussion, but let's move on with any additional questions or observations related to the scientific acceptability of the measure.

Are there further either observations from the Work Group or questions from any members of the Committee around the scientific acceptability?

MEMBER LIEBERTHAL: Yes, I am still having trouble with the risk-adjustment model. There is such a broad variety and there are so many underlying conditions that affect it, that I don't know that the risk adjustment is adequate to give good information.
CO-CHAIR HOMER: So, Kim, could you tell us about, first of all, how the risk adjustment, how the panel created the risk-adjustment measure, and then perhaps, for some of us whose statistics are a little rusty, describe what an ROC curve and the area under the ROC curve means, and things like that?

MS. GAUVREAU: So, the risk categories were derived primarily empirically in that we looked in the dataset, the KID 2000 dataset, and looked at, for our procedures that occurred at least 20 times in that dataset, we looked at the in-hospital mortality rates.

Then, we grouped procedures by those rates. We started out with more than four categories, so looking at lots of possible different splits in the data, but knowing that we somehow wanted to group procedures by mortality rates.

We, then, sort of worked backwards and looked at adjacent categories. If there
was not much overlap between them, we would
collapse them. We did that both looking at
actual mortality rates and by putting things
into logistic regression models and looking at
odds of in-hospital mortality relative to what
we considered to be the baseline group.

And for the baseline group, from
the very beginning, we said that we were going
to include procedures with a less than 2
percent mortality rate. There were a lot of
procedures that didn't have any mortality in
the database at all, but we would not be able
to fit our regression model if we had a
category with no deaths. So, we wanted to be
sure we would have at least some deaths, even
in our baseline risk category.

So, then, we looked at the various
cutpoints and found the one with the best
discrimination. That was measured by the C-
statistic or area under the ROC curve, which
basically is a measure of how well the model
is able to predict who dies and who doesn't
die.

Just with the four risk categories that we ended up with alone, the area under the ROC curve was, I believe, .87. In general, anything over .8 is considered very good. So, the model, the risk categories were very highly discriminative about predicting in-hospital mortality in this case.

CO-CHAIR HOMER: And then you retest your validation?

MS. GAUVREAU: We validated that in a second dataset, the KID 2003, and found the area under the ROC curve to be, I believe in that case it was .85 or .86. I mean just only a little bit less.

I guess we haven't mentioned it. On top of that, so in addition to the risk categories, we did look at these other clinical factors that might contribute and help us to predict risk of in-hospital mortality, even beyond the risk categories. That was the necrotizing enterocolitis and
serious respiratory conditions. We considered
a list of about 10 or 15 other variables that
might help contribute to risk and found only
those two to be statistically-significant.

DR. BURSTIN: Just a quick
question, just because untested outcome
measures make people a little anxious. How
does this relate -- there's a reference you've
got at the bottom of 2a to an article that is
the Annals of Surgery in press. Is that the
risk model you are talking about with the
validation?

MS. GAUVREAU: Yes, it is.

DR. BURSTIN: Okay. Because I
think there is at least a comfort zone to know
the risk model used in this particular
measure --

MS. GAUVREAU: Yes.

DR. BURSTIN: -- has been
validated.

MS. GAUVREAU: Yes.

DR. LILLEHEI: Yes, it was just
published last month in the Annals of Surgery.

Yes. Sorry.

CO-CHAIR HOMER: Other questions about the scientific acceptability of it or comments from the Work Group members?

(No response.)

So, it sounds like some concern about the risk-adjustment issue. Coding issues, probably any sense of the validity of coding? Any concerns? I mean death seems like it is pretty reliably coded, and the procedure itself seems like it is going to be pretty reliably coded. That is basically all you need or?

CO-CHAIR WEISS: I actually have a question about that. I just wondered, aside from factoring out the cardiac procedures, is everything else in that bucket or did you make some selections about what you included in the non-cardiac surgery compendium or inventory?

DR. LILLEHEI: Yes, I don't know that we have provided that table, but, yes,
there were procedures that we viewed as closed procedures. So, I can list sort of a sample of that for you here. But the excluded procedures were any closed biopsies, closed reductions, superficial lacerations, catheterization, delutations, injections, aspirations, radiologic procedures, dental extractions, circumcision, and other incidental procedures.

So, yes, you're right. Of those procedures, there were certain ones that we excluded because we didn't think that they would really fit under the umbrella of neonatal surgery as we were trying to understand risk and mortality.

MS. RAUSCHER: We can provide a copy of the article for the Committee, if you would like that.

CO-CHAIR HOMER: That would be great.

MS. RAUSCHER: Okay.

MEMBER LIEBERTHAL: I would like
to ask, since you excluded a number of very minor procedures, why you included lingual frenectomy as a significant condition.

DR. LILLEHEI: Fair enough. We can talk through one of the others, but, in fact, most of the time those procedures now are actually done, unlike perhaps an earlier era when they were done in the pediatrician's office, kind of a clipping at the bedside, now most often, in fact, they are done in the hospital and usually in an ambulatory setting with some element of anesthesia. Anesthesia and a surgical team, we felt that was kind of the bar that put us into this category, but I am open to --

CO-CHAIR HOMER: I'm sorry, I thought you said there actually did need to be some deaths for it to be included in your risk? No? I misheard that?

MS. GAUVREAU: So, we wanted our lowest baseline risk category to at least have some deaths, so that we could compare the
other categories to that one. In the end, the
mortality rate in that category 1 was .2
percent.

DR. LILLEHEI: No, but in answer
to your question --

CO-CHAIR HOMER: In individual
procedures --

DR. LILLEHEI: Yes, individual
procedures didn't need to have deaths, no.

MS. GAUVREAU: That's right, just
in the categories.

CO-CHAIR HOMER: Just in the
category?

DR. LILLEHEI: Yes.

CO-CHAIR HOMER: Okay. And again,
this issue of meaningful differences, so your
sense simply is, because you have only done it
one year and don't have multiple-year
averages, that your confidence intervals are
sufficiently --

MS. GAUVREAU: Are fairly wide.

CO-CHAIR HOMER: Are wide? So, if
this were trended over time, perhaps that
would result in significant narrowing.

All right. My next question will
be towards usability.

Any other questions on scientific
acceptability? Or we could move ahead and
vote on that. Or comments? Kathy, did you --

MEMBER JENKINS: Yes, may I make
one comment? Even in cardiac surgery where
there is twentyfold differences in the
country, it is extremely unusual in pediatric
sample sizes to find statistical differences.
I actually think it is extremely amazing that
in a single year of data we did find any
institution that was statistically different
with an area under the ROC curve of .9.

So, I think perhaps having looked
at rare pediatric procedures and outcomes more
than maybe people who do more work on more
common procedures, people may not be aware of
that.

MEMBER CLARKE: If I might
comment, I think that the kind of AUC numbers that they are giving us are absolutely incredible with the kind of broad categorization that we see here. I am very, very surprised at that, but that is really excellent.

CO-CHAIR HOMER: All right. So, why don't we go ahead and vote on the scientific acceptability?

Those who would consider the criteria completely met raise your hands.

DR. WINKLER: One.

CO-CHAIR HOMER: One. Okay.

Those who feel they are partially met?


CO-CHAIR HOMER: And minimally?


Usability. So, I think there one question I would ask is, has there been interest in the surgical community around this measure? And how does it seem to be received?
And how understandable have there been —
again, comparing a little with the previous
conversation, as you have vetted this with
your colleagues at the hospital, as you
discussed it at CHCA meetings or NACHRI
meetings or Surgical Society meetings, what
kind of interest is there in this? What kind
of reactions are you getting? Just to get
some flavor for the usability of the measure.

DR. LILLEHEI: Well, in fairness
to you, we are pretty early on in that. In
fact, the surgical community, at least at
large, are only those who have read the last
month's Annals thus far perhaps.

(Laughter.)

But, in fact, within the surgical
community, and specifically the pediatric
surgical community, we are, as a group, very
interested in understanding reliable ways of
risk adjusting for what we do, and then making
things better accordingly.

Now, to date, we have done that
with specific diseases, like congenital
diaphragmatic hernia registries or necrotizing
enterocolitis working group, but there are
certainly limitations to that analysis and
questions that really don't lend themselves to
answering in that context.

There is the pediatric NSQIP that
was alluded to earlier that is being developed
in an effort to allow us to better understand
what we do and how we might change things
accordingly.

But, in fact, specifically, as
regards our measure, no, we are just in the
process of springing it on them, if you will,
yes.

CO-CHAIR HOMER: Ellen? Or you
weren't on the Work Group. Any other members
of the Work Group? No, go ahead, Lee.

MEMBER PARTRIDGE: If I am
understanding this correctly, and I am looking
at it from the consumer perspective primarily,
this has the potential to be a very powerful
measure in the sense that I believe what you are trying to do is develop an overall measure of predictability of mortality of children prior to 30 days of age from non-cardiac surgery.

It is perhaps one of the two or three most powerful kinds of measures that anybody wants to know about a hospital. I guess it would be helpful to me if you would just take a minute and explain how you would explain this to a patient or a purchaser as a good predictor of whether I should hospitalize my child or have that hospital in my network.

Is that too tall an order?

DR. LILLEHEI: We do think that this has the potential for a considerable impact. Obviously, results of neonatal surgery, for those of us who look after neonates or have neonates, or whatever, is exceedingly important.

We do think that with this tool we will be able to dissect out whether it is
institution-specific, whether it is types of institution, whether freestanding pediatric hospitals do better or not, or whether a children's hospital within a large general hospital, a children's unit within that neonatal surgery is the same, whether anesthesiologists that have the ability to have access to pediatric anesthesiologists makes a difference versus not. I don't know the answer to those questions.

We may be surprised by some of the results, to be sure, but, yes, I think that is why we are excited about this particular study.

MEMBER McINERNY: Well, of course, you know, for adult cardiac surgery, this kind of reporting has been going on for at least a decade. But my impression -- and correct me if I am wrong -- is that, as far as either purchasers or consumers are concerned, whether they even know the data, No. 1, and whether it makes any difference where they decide to go,
No. 2, I haven't seen a lot of evidence that it has influenced that. Maybe I am wrong.

CO-CHAIR HOMER: The data on these kinds of things tends to influence providers a great deal. Because even though the expectation, of course, was that consumers would use it to drive, what frequently happens is we, hospitals and physicians, are very driven by comparative data. We all didn't want to be in the bottom of our classes, and therefore, we look at these data and it tends to drive improvement through that strategy more than --

MEMBER PARTRIDGE: I think we all know that. Of course, you are in the State that has been the leader in cardiac surgery reporting. But I think, in fact, it has impacted some of the purchasing patterns in your State.

So, in my understanding also, we haven't actually done a lot of testing of this yet. So, we would be, presumably, talking
about this as a time-limited, yes, because I
would want to get a little better sense of how
much variability we are really turning up. I
know you have been working at that.

CO-CHAIR HOMER: Allan, did you have a question?

MEMBER LIEBERTHAL: Yes, I did. I am looking for the exact wording. But you said something about, to be included, you would have to have more than 20 procedures, is that correct?

MS. GAUVREAU: That's right.

MEMBER LIEBERTHAL: Is that out of 20 individual procedures or 20 in a risk category?

MS. GAUVREAU: Twenty in an individual surgical procedure.

MEMBER LIEBERTHAL: Okay. One of my concerns about usability is that, if there are institutions that are doing fewer than 20, and they are excluded from the data --

DR. LILLEHEI: No, let me clarify
because I think we are leading you in the
wrong direction. When we developed the model,
in order to develop the model, what we used
were procedures in the KIDS database that were
20 or more. Okay? Procedures in the KIDS
database that did not have 20 procedures, we
did not include in our risk category. We
didn't put them into a specific risk category.

So, do we, by that, do we omit
certain rare procedures? Indeed, we do, but
the fact of the matter is our analysis
encompassed about 85 percent of the types of
procedures being done. So, that was just to
develop the model.

MEMBER LIEBERTHAL: So, hospitals
that are doing fewer than 20 procedures still
would be part of this?

DR. LILLEHEI: Oh, absolutely.

Absolutely. Yes.

MS. GAUVREAU: Yes.

MEMBER LIEBERTHAL: That answers
my question then.
DR. LILLEHEI: Thank you.

CO-CHAIR HOMER: Faye?

MEMBER GARY: I just had a quick question, and it is related to Lee's question. Lee asked about how the data will impact decisions, administrative decisions, et cetera, et cetera.

And my question is, have you all given any thought to how our outcome data can be used or will be used when communicating with consumers, i.e., parents and family members? Could you just talk about the usability of this data as it relates to I think one of the most important groups, and that is the parents of the child? I am trying to get translation here to how it appears in the clinical setting with the variety of different kinds of parents and families who might have this experience.

DR. LILLEHEI: Your point is well-taken. Obviously, that is a decision that, as a parent looking to a surgical procedure, that
is paramount in your mind.

This data is really, by virtue of

the fact that we are including a lot of
different types of procedures and combining
those risk categories, we are really not
looking at individual operations. So, I don't
think this is a tool for that individual
parent to decide whether I am going to have
Dr. Lillehei do my hernia repair or not, based
on that.

It is talking about institutions
as a whole, whether we think that institution
specifically or that type of institution or
that region of the country -- those are the
sorts of questions that we will be able to
answer about this, and not the specific one,
you know, where do I get my hernia fixed or
with whom?

So, I just wanted to underscore
that limitation of what we are going to be
able to answer.

MEMBER DOCHERTY: Yes, I think
what I like about it is that it reflects more than just the procedure itself, but the postoperative care that is given in a hospital. So, it is nursing care. It is all the post-surgical care, anesthesia, all of those things.

DR. LILLEHEI: Absolutely.

CO-CHAIR HOMER: Nancy?

MEMBER FISHER: I just wanted to make a comment in response to Tom's question. Large purchasers are looking at things like this. There is a large company that looks at what you have done with Leapfrog, and they set up their payment for you, so that if you are at a hospital that they approve of, you get 90 percent reimbursement; if it is not, it is 80 percent reimbursement. There has been talk about tiering hospitals, so putting them into three different tiers and then shifting it over.

So, there is all of this stuff going on sort of coming out of value-based
purchasing that is looking at all of this.

So, yes, this is very important.

CO-CHAIR HOMER: So, I think we are ready to vote on the usability criteria.

So, those who feel it fully meets the usability, completely meets the usability criteria?

All right. Those who feel it partially meets the usability criteria?

DR. WINKLER: Ten.

CO-CHAIR HOMER: And those who feel it minimally meets the usability criteria?

All right, good.

And then, moving on to feasibility, so that is how easy it is to collect, report on issues, concerns about exclusions, inaccuracies, and implementation issues.

Seems like it is pretty straightforward in that it comes from administrative databases. That is where you
have validated it. It is not a narrow dataset, that it is not PHIS, or it is something that really can be used on any standard set of discharge data. So, it seems pretty straightforward.

Tom?

MEMBER McINERNY: Do you have any idea, I mean, how much of the data-based person time does it take to collect and sort of analyze and report the data? Is this a 1.0 FTE for a full year or a .2, or do we know?

MS. GAUVREAU: Well, assuming the data is coming from an electronic database or an administrative dataset, it doesn't take very long at all. So, it is not based on primary data collection.

MEMBER McINERNY: So, you could probably somehow program it once --

MS. GAUVREAU: Yes.

MEMBER McINERNY: -- and then it becomes automatic?

MS. GAUVREAU: Yes. Right.
Exactly.

MEMBER McINERNY: Okay.

MS. GAUVREAU: And we have the program and documentation to do all that.

MEMBER McINERNY: Okay. Thank you.

CO-CHAIR HOMER: All right. So, I think we could vote on the issue of feasibility.

Those who feel it completely meets the criteria for feasibility?

DR. WINKLER: I think it is everybody.

CO-CHAIR HOMER: Very good.

All right. Then, we move on to voting overall for the measure. Again, I think because this measure has not been in general use, this would be a time-limited endorsement, presumably with the request for specific testing about applicability and usability, I think looking at the potential for narrowing the confidence interval by
extending the data over periods of time, et cetera.

So, those in favor of a time-limited endorsement of the measure raise your hand.

DR. WINKLER: I have got everybody.

CO-CHAIR HOMER: Good. Okay. We're done. Thank you very much.

All right. I think we will do one more measure before our break on my 10:30 break time rule.

MS. McELVEEN: Okay. Our next measure is Measure 29. This is the standardized adverse event for children and adults undergoing cardiac catheterization for congenital heart disease. This is the ratio of observed-to-expected clinically-important, preventable, and possibly preventable adverse events risk-adjusted.

DR. BERGERSEN: Hi. My name is Lisa Bergersen again.
I want to thank you for the opportunity to be at the table today from a smaller community of physicians who perform cardiac catheterization procedures on both children and adults for congenital heart disease.

Over the past 15 years, cardiac catheterization has evolved from diagnostic studies to primarily interventional studies with not an insignificant amount of morbidity associated with those procedures. So, as a field, it has become very important for us to understand the outcomes for these children that can eventually have bad outcomes in the catheterization lab.

We look at a lot of different measures: overall event rates, clinically-important event rates. But I chose this outcome to share with you because it has some face validity in its importance to us as a community as a measure to track.

That being clinically-important
events, those events that are life-threatening or potentially life-threatening to the child and with a potential opportunity for improvement in care. So, those that are either possibly preventable or preventable. As a community, we think that this is an important measure and outcome for the children.

CO-CHAIR HOMER: Terrific. So, why don't I ask members of the Work Group to comment and specifically, again, initially on the area of importance? A request to raise questions. Any comments from the Work Group? And then open to comments and questions from anybody else on the Committee.

David?

MEMBER CLARKE: Well, I think she has adequately addressed the importance of this. This is becoming a very strong interventional type of a specialty as opposed to diagnostic, as it was historically. This brings on whole new implications.
Basically, now they are doing a lot of cardiac surgery with a little, tiny tube. So, I think looking at adverse events and the things that can go wrong with that, which are not inconsequential, is extremely important as this specialty evolves.

CO-CHAIR HOMER: Any other questions from the Work Group members or the members of the Committee about the importance of this?

DR. WINKLER: I just want to follow up on one of the Committee's comments about why we are including adults as part of the measure. Can you give us some numbers? Because I am assuming you are defining the adults are probably young adults as opposed to I don't see a lot of 65-year-olds, but maybe.

DR. BERGERSEN: Well, it is congenital heart disease.

DR. WINKLER: Exactly. That is my point.

DR. BERGERSEN: So, the measure
was developed to capture our entire
population. Depending on the physician, their
case mix can have a varied amount of adults.
So, it was developed to capture the entire
case mix for institutions performing these
procedures.

We could limit this outcome
measure to children less than 21 years of age
without losing validity, we believe.

CO-CHAIR HOMER: So, for example,
an adult with an anomalous coronary -- now I
am showing my clinical ignorance -- but, you
know, anomalous coronary artery, or something
like that, which is presumably a congenital
problem, you are not talking about that?

DR. BERGERSEN: Right. So, among
our six physicians who primarily do neonates,
the percentage of adults that we do ranges
from zero to 15 percent. So, it is a small
percentage.

MEMBER RAO: My concern in raising
that question was that, is there a significant
proportion of adults who are going through revision surgeries, second or third surgeries for their congenital heart disease, as opposed to their first surgery? Because I assume if you are 25 years old, somebody has picked up on this at this point.

DR. BERGERSEN: Right. Exactly.

CO-CHAIR HOMER: But you still might be catheterized, presumably, in your pre-operative revision --

DR. BERGERSEN: Some of the adults that we catheterize will be -- and again, this depends on the case mix of the particular interventionalist -- it may include late-presentation ASDs, late-presentation PDAs, pulmonary hypertension, or redo operations, conduit revisions requiring human NMX pre-operatively.

But, like I said, we could limit this measure to less than 21 years of age.

MEMBER RAO: Yes, I just thought that maybe somebody going through
catheterization as an adult probably had a
different morbidity or risk than somebody who
is younger.

DR. BURSTIN: Having spent a
decade practicing at the Brigham and taking
care of a lot of these adult cardiac
surgeries, I mean, literally, you have 30- and
40-year-olds who had tetrology procedures a
decade ago who still go to Children's for cath
because they know that better than anybody
else does. So, the question would be, are
they really the same group or should you
really segregate it?

CO-CHAIR HOMER: Have you looked
at --

DR. BERGERSEN: We haven't looked
at these outcomes -- well, the risk-adjustment
method for this outcome was developed using it
with the adult population in it. We haven't
excluded them and gone through the same steps
to look at how the model performs, but I would
expect that it would perform at least equally
as well.

MS. GAUVREAU: But we also didn't find age to be a statistically-significant predictor of adverse events in our model.

DR. BERGERSEN: That's correct.

CO-CHAIR HOMER: Allan?

MEMBER LIEBERTHAL: Your specification for numerator specifies a pediatric cardiac catheterization lab. The majority of hospitals that are not children's hospitals use one cardiac catheterization lab for both children and adults. Using your definition, you would confine it to only specialized pediatric cardiac catheterization labs, which excludes a significant number of cardiac caths.

Then, when you start including stint placements in adults, you haven't really specified congenital heart disease. So, I think I understand what your intent is, but your wording can lead to the measure being applied differently than you intended.
DR. BERGERSEN: That is a good point. I think, to be more precise, it would be procedures done for congenital or acquired heart disease, congenital heart disease in the adult or child.

So, if you were doing, let's say you were an institution that was not a pediatric institution, but you were doing procedures for congenital heart disease. You could apply this measure.

Does that answer your question or address it?

MEMBER LIEBERTHAL: Yes, it does. I would just ask that, if we approve this, that the conditions be a change in the wording.

CO-CHAIR HOMER: Again, prevalence is not an absolute requirement, but do you have any sense of the number of procedures done in a year that this would apply to?

DR. BERGERSEN: Yes. We estimate that there's about 100 institutions across the
country that do regular cath procedures on both adults and children with congenital heart disease.

CO-CHAIR HOMER: And how many --

DR. BERGERSEN: And the volume there, our institution probably performs more than most at about 1200 a year.

CO-CHAIR HOMER: Yes.

DR. BERGERSEN: And then the other institutions, about between 300 and 600, some a little less. So, let's see --

CO-CHAIR HOMER: So, if you are thinking about 500 per institution, then you've got -- yes, okay, then you have about 1,000. So, 500; you said how many institutions, 100?

DR. BERGERSEN: About 100, yes.

CO-CHAIR HOMER: So, 50,000.

Okay, good.

DR. BERGERSEN: Fifty thousand.

MEMBER McINERNY: I wonder if we could change the numerator to catheterization
cases performed by a pediatric interventional cardiologist instead in a pediatric cardiac cath lab. Because I know our institution, we have an interventional cardiologist, but he is, as Al describes, he does his in a general cardiac cath lab, but there is a specialized, sort of a specialized room where he does it, but it is still considered an adult cardiac cath lab.

DR. BERGERSEN: That would be clearer, specifying it by the physician.

MEMBER LIEBERTHAL: I actually disagree with the wording. I would just say cardiac cath procedures done on congenital heart disease. Because what worries me is adult cardiologists who are doing procedures on adults with congenital, oh, yes, congenital heart disease. I think that quality measures may, hopefully, put an end to that.

(Laughter.)

CO-CHAIR HOMER: Okay. So, on our first criteria of importance -- and this is a
threshold criteria, so it would be either yes or no -- the question is, how many believe this meets the threshold criteria for importance? Show of hands.

Okay. So, now we can move on, which we have already been delving into, but we can move on to the issues of the scientific acceptability of the measure. So, why don't delve more deeply into that?

Can you talk a little bit more about how adverse events are defined, how reliable, adverse events, preventable adverse events, how reliable the identification of those are? You had some data in the report.

DR. BERGERSEN: Yes.

CO-CHAIR HOMER: But if you could talk more about that?

DR. BERGERSEN: Well, in 2003, 2004, reviewing the previous literature on cardiac catheterization and how people were reporting outcomes, most institutions would report them as minor or major. We thought
that didn't lend itself to -- could have
greater clarity by separating them out into
five categories.

So, we started collecting adverse
events at our institution using these five
categories of severity, one being an event
that happened, but there was really no
clinical consequence; two being a minor event;
three being something that was potentially
life-threatening, like a supraventricular
tachycardia that you had to cardio vert; four
being something that was clearly life-
threatening. You had to do CPR on a patient
because of an arrhythmia. And five being
death.

So, we started collecting our
adverse events using these definitions. Our
hospital later adopted them in other areas
across the entire hospital.

As a field, recently, we have
gotten together as a group and started to talk
about nomenclature and how we are going to
define both the procedures that we do and complications, because the nomenclature just didn't exist previously.

So, in the next year, our definitions for severity will be adopted by the International Pediatric Cardiac Code and will be available to the community to use. So, we will be publishing a complications list that has qualifiers for severity and definitions, as we have been collecting events over the past six years.

Currently, there are eight other institutions who are collecting data in a similar fashion and coding their events using our severity classifications. And I referred to them in the background material.

This is the Congenital Cardiac Catheterization outcomes Project. This group of institutions started collecting data in 2007, and we now have a dataset using uniform definitions since 2007. It includes now 13,000 cases.
CO-CHAIR HOMER: So, as I hear these, I guess the reason I was asking about it, but I think you have started to answer it beautifully, so we know that voluntary reporting of adverse events in hospitals is miserable, to put it bluntly. And we know that when people do audits using something like a trigger tool, they vastly increase by factors of five or ten the number of adverse events that are identified.

But this seems different. This you have got specific events that are defined and typically recorded in a procedure anyway, and you are capturing more routinely. So, again, it is different than routine hospital collection of adverse event data and more valid and reliable?

DR. BERGERSEN: When we audited the dataset that was in the paper for the measure development, for Level 3, 4, and 5 events at our institution, we captured all of them. In this group of institutions that have
been collecting data over three years, they do pretty good with the 3's, 4's, and 5's. As you would imagine, the 1's and 2's, there is some variation in what people would consider important enough to record.

But among the 4's and 5's, at least in a 10 percent audit, they reported all of them. And among the 3's, 4's, and 5's, it was as high as 92 percent. So, there were a few what we consider Level 3 events which were primarily respiratory events even before the procedure had started related to anesthesia. So, because they are clinically important, they tend to be captured.

CO-CHAIR HOMER: Okay. Thank you. David?

MEMBER CLARKE: Well, I have, I guess, a real problem with allowing so-called non-preventable events to be excluded because I think that, first of all, what might be preventable by one person is non-preventable by another. Second of all, you know, maybe it
is just because surgeons enjoy wearing hair
shirts, but, traditionally, when you are doing
surgery, if the patient dies within 30 days of
the operation, even being run over by a bus,
that is an operative mortality.

(Laughter.)

So, I really have a problem with
excluding these sort of big -- what is an
example of an unpreventable event?

DR. BERGERSEN: Yes, let me
explain to you why we did exclude it and why
it is important that we exclude preventable
events when looking at this outcome.

MEMBER CLARKE: You mean
unpreventable?

DR. BERGERSEN: Non-preventable.

And it actually goes towards the other
comment, which was the moderate events, as
defined -- let's see, how do you know that
they are not based on the patient's condition
rather than how the procedure was performed?

So, what we are trying to do here
is exclude those events that, because of the
patient's condition, that you could not have
avoided. So, for example, you get called in
the middle of the night to catheterize a
patient who is having ventricular tachycardia,
and there's a suspected anatomic problem. And
they are trying to manage them medically.

You bring them down to the
catheterization lab and they go into
ventricular tachycardia, and you have to do
CPR on that patient. There was no way as an
operator that you could have avoided that
event.

Whereas, I bring a patient to the
catheterization lab for an elective aortic
valvotomy. I cross the aortic valve and I jam
the catheter down in the LV, and the patient
goes into ventricular tachycardia. In that
case, maybe there was something that I could
possibly have done to have avoided that event.
And I want to make sure that I capture that.

MEMBER CLARKE: Right, and I am
not talking about events that occur and result in maybe a patient's death just on the way to the cath lab before you do a procedure.

DR. BERGERSEN: No, I am talking about, if I bring a patient to the cath lab and they have ventricular tachycardia, and I put catheters in them, and I am doing a diagnostic catheterization, and they go into their fatal arrhythmia, and I can't get them out of it, and it was a pre-existing condition where there was nothing I could have done in the cath lab to avoid it, then those are the events that we're --

MEMBER CLARKE: So, if a patient dies when you are opening the chest, it doesn't count, and that just isn't right.

DR. BERGERSEN: Well, it depends on what the outcome of interest is, I think.

MEMBER CLARKE: You know, I guess I look at it this way, from a broader perspective, and I said this earlier. If you exclude things that are not preventable from
the collection of the data, they are never
going to be preventable because you won't
identify them. Okay?

And what you can do, and I think
it is very reasonable in some cases to do
this, we do it at the STS database, congenital
database in several instances, is you collect
the data and you exclude them from the
analysis, which means you exclude them from
both the numerator and the denominator.

CO-CHAIR HOMER: So, you would
have the data, but the measure, then, would
still not reflect those data?

MEMBER CLARKE: Right. You could
decide to exclude those from the analysis of
the data. Then, at some point, you might want
to change your mind and put them back in for
some various reason, some events or --

DR. BERGERSEN: Yes. So, we are
not proposing that people not collect data on
not preventable events. And we feel,
actually, quite strongly that any event should
be recorded in your database and looked at.

But for this particular metric, we wanted to focus on events where there was a possibility for improvement of care.

So, the problem with putting preventable events in it, if you are going to look at different institutions, is, well, it wasn't the outcome that we were interested in.

CO-CHAIR HOMER: Can you remind me how you are determining preventability? Is it on the code or is it the judgment of the person who is entering the data?

DR. BERGERSEN: Yes. At our institution, we collect this data at a monthly meeting or more often. We review all of the events and, as a group, come to consensus on both the severity and the preventability.

There are fairly precise definitions. I think one thing that we would need to do, if this metric went forward, is look at within our dataset and the C3PO dataset what were those events that were not
preventable, so that we can have clear
definitions for the community of what we would
consider, like I did with the v tac example.

CO-CHAIR HOMER: I love the name, with your C3PO data.

(Laughter.)

That is just great.

Did the other sites do assessments
of preventability, and did you assess
comparability of the preventability
assessments across multiple institutions?

DR. BERGERSEN: You know, it is really interesting. When we started this project, many people said -- and it was part of one of the comments about I think feasibility -- "Oh, you're not going to get the community to tell you about their bad outcomes."

But, in fact, the project was met with a lot of enthusiasm. We actually had to limit the number of sites that could participate.
Like I mentioned, the audit, they report both their minor as well as their high-severity events. Two cardiologists review all of the events, and they have reported, similar to our earlier data, about a 30 percent rate of not preventables. So, they have not been liberal with the definition and rarely misapply it. So, we rarely change it when we review their preliminary classifications.

Did I confuse it? I'm sorry.

MEMBER PERSUAD: I just have one final comment about the issue of excluding non-preventable cases, and there's probably nothing to do about it now, but going forward I think, when we began this discussion about the overall importance, this is a ballooning area.

(Laughter.)

And your example, the example you described says to me really that you are doing more and more non-surgical corrective procedures on an increasing risk population,
is what it is. And that doesn't mean to me necessarily that, when they carry higher risk, it is non-preventable. It may mean that more sensitivity in the procedure has to be addressed for their risk to be lower because they are carrying higher risk. So, I just throw that out there.

DR. BERGERSEN: So, there is -- I'm sorry.

MEMBER McINERNY: No, no, finish.

DR. BERGERSEN: So, there is variation in rates of these events among even our practitioners at one institution. This variation has to do with different populations of patients being catheterized by different interventionalists.

When we sat down to look at procedures, we were able to identify 84 different types of procedures that we do with varying frequency, from 1 percent of our cases to maybe 20 percent of our cases. So, similar to what Dr. Lillehei had explained, we
couldn't adjust just based on one procedure type. So, what we did is we put all of those different procedure types into different procedure type risk groups, so then we could adjust for the case-mix complexity of a particular operator.

Then, this would apply an institutional outcome. You could apply this risk-adjustment model to an institution's outcome.

We haven't compared adverse event rates among institutions in cardiac catheterization because for many years you say, well, my case-mix complexity is more complicated than yours. But what we have shown by developing these models and looking at this measure is that you can do it fairly. You can do it fairly, and we should be looking at these outcomes.

CO-CHAIR HOMER: Tom, did you have a question?

MEMBER McINERNY: Yes. I, too, am
very concerned about putting some things in the non-preventable category. I think 10 years ago many people would have said that a central line infection was non-preventable. We know that, in fact, they are preventable.

I just worry that, when you do that, then it sort of becomes an accepted complication, and, oh, well, you know, yeah, we put a central line in and they get infected. Oh, well, we do this particular cardiac catheterization procedure and something happens, but it is non-preventable.

And that worries me because I think you stop thinking about, yes, well, maybe it is preventable if we did something else.

CO-CHAIR HOMER: So, maybe to delve a little further into this, could you walk through, because I am a little fuzzy on this, just kind of the algorithm for what actually happens in terms of the data collection and the categorization of adverse
events, and then adverse events as preventable or not?

So, I just don't understand the process right now by who it is going through and who is making which decision and judgment, and how it is being based. So, if you could just walk that process through, I think it would help the Committee. It certainly would help me.

MEMBER JENKINS: Can I say one background thing?

CO-CHAIR HOMER: Good.

MEMBER JENKINS: I just want to say that, as Lisa mentioned, we imposed these categories at the entire institution of the Children's Hospital, Boston, for all of our adverse event reporting. One of the things, the discussion is very important, and I don't want to minimize it.

I just want to make this point: one of the things that is the most difficult barrier to overcome for clinicians to feel
good about measurement is feel like they are being unfairly measured against something that happened on the day they were there, that there is absolutely nothing they could have done to prevent it, like the v tac example.

And it has been comforting and helped with our adoption of this concept to include a way out for that. How exactly to measure it or where the slippery slopes are is a real issue, and I don't want to minimize that.

One of the things that we have done at Children's is, if someone says something, we ask them to articulate what they could have done differently to prevent it. Okay? Because if there's absolutely nothing that anybody can think of that they could have possibly done differently to prevent it, it is different. Okay?

So, I don't know if that helps, but this is a field that is getting their hands around some of those issues, around
infection and the rest of it.

So, I just want to state that
point, that in order to have clinicians adopt
these measures, they do have to believe that
it is fair.

CO-CHAIR HOMER: I agree fully
actually. I think it is totally on target.
I think the challenge comes in when you are
sort of doing this high-stakes, potentially
high-stakes measurement and wanting to be sure
that different institutions are using similar
criteria.

So, that is why really just I
think, at least for me, I am not completely
clear on the process that actually takes
place. So, just walk that through, how you
are proposing that it take place with the
measure that you are proposing.

Could she respond or --

MEMBER CLARKE: Oh, sure. I'm
sorry.

DR. BERGERSEN: I think I have
articulated that this is evolving and that institutions, all institutions that perform these procedures are collecting adverse events. They are doing it in one way or another.

So, to make the measure work, they would need to collect their adverse events and record them using the definitions that are available through the International Pediatric Cardiac Code, which will be published this year.

They will have the severity as a qualifier for an event. They will have clear definitions attached to them, and institutions would have to adopt those definitions into their collection of their adverse events to be able to apply the measure to their institution.

There is a national registry that is starting for congenital cardiac cath through the ACC called IMPACT. That could potentially be a way to centralize data.
collection if they adopted the same strategy.

CO-CHAIR HOMER: So, again, the event happens. The institution records it. The institution categorizes it reliably, using the 1-to-5 scale.

And again, really, just a process question then: after that happens or before that happens, when does that preventability assessment take place and who is making that judgment? I believe it is really a process question.

DR. BERGERSEN: Well, I think there's two that I would like to answer. One, within C3PO, the registry of these eight institutions, when the event happens, they assign preliminary categories, and then those are independently reviewed by two physicians. Now that is a research project. So, that is how that registry works.

So, what the individual institutions would need to do is do what they are already doing, record their adverse event
and assign a severity category and a preventability category at the time of the event, and all institutions also review events as a group, I think in some format, whether that is weekly or monthly or bimonthly. I think it is pretty common in most institutions to review their adverse outcomes. So, then, they would have the opportunity in that venue to come to consensus and make sure that there is agreement and no operator bias in the classification.

MEMBER CLARKE: I just want to clarify that non-preventable adverse events are all collected. There is not an option to not report? Because, obviously, voluntary adverse event reporting is a problem, and if you give a provider a loophole, it is going to take it. That is the thing I am most concerned about.

But whether or not you determine after careful analysis that this event should not be included when we analyze the data and
present and report results, that is a totally
different matter and I fully understand that.
But you have to identify these adverse events
or nothing is ever going to be done about
them.

DR. BERGERSEN: I think in terms
of this metric and the usability by
institutions, I would like to just echo what
Kathy said in terms of physician buy-in and
feeling that they are fairly evaluated.

Also, I think we underestimate
sometimes physicians' willingness to be
transparent, especially when it comes to the
opportunity to improve their care. So, if you
collected your events and you had the
opportunity to apply this model to your
outcome and calculate your standard adverse
event ratio, and you looked in the literature
and you saw that Hospital X's rate was this,
and you were outside of the bar for
performance, you might try to do a little bit
better. That is the purpose of putting this
forward and sharing it with the community.

CO-CHAIR HOMER: I think we have had a great discussion. I would like to move -- one more question about scientific acceptability? Because then we want to be able to vote on it.

MEMBER McINERNY: I think what Charlie is trying to say is that, as long as there is a standard criteria for what is considered preventable that is applied across the board uniformly to all institutions, then that would go a long way to making us feel a bit more comfortable. But what we would be uncomfortable about is that Institution A says, "Well, that was unpreventable," and Institution B says, "Yes, it was preventable."

Then, you are not comparing equally.

CO-CHAIR HOMER: So, why don't we vote on the scientific acceptability criterion?

So, how many vote that it completely meets the criteria for scientific
acceptability?

Okay. How many would vote that it partially meets the criteria for scientific acceptability?

Okay.

DR. WINKLER: Eight.

CO-CHAIR HOMER: And feel that it minimally meets the criteria?

DR. WINKLER: I probably need to check. Did Marlene Miller join us at all?

(No response.)

Okay.

CO-CHAIR HOMER: Okay. Did we get everyone then?

DR. WINKLER: Yes.

CO-CHAIR HOMER: Okay. Good.

Then, on the usability, because I think we have discussed the usability a fair amount in the context of discussing the scientific. So, this relates to the issue of whether it is understandable, harmonization, and there is another element to it, which is
whether it provides added value.

        How many feel it completely meets
the criteria for usability?

        Feel that it partially meets the
criteria for usability?

  DR. WINKLER: Nine.

  CO-CHAIR HOMER: Okay. How many
feel it minimally meets the criteria for
usability?

  DR. WINKLER: Five.

  CO-CHAIR HOMER: Okay. All right.

We got everyone?

  Then, for feasibility, I guess
that is actually -- so, again, data being a
byproduct of care, available through
electronic mechanisms, exclusions
appropriately specified, not susceptible to
inaccuracies, and ease of implementation.

  So, how many feel it completely
meets the criteria for feasibility?

  Partially meets the criteria for
feasibility?
DR. WINKLER: Twelve.

CO-CHAIR HOMER: And minimally?

Two?

And then I think an overall vote.

And again, this one, again, I think would be, the vote would be for time-limited endorsement subject to conditions.

MEMBER LIEBERTHAL: With the conditions on the wording changes.

CO-CHAIR HOMER: Conditions for wording changes around setting and provider and even age, potentially age restriction to under -- age to be determined.

Okay, David, question?

MEMBER CLARKE: I would like to see some testing of non-preventable adverse events.

CO-CHAIR HOMER: I think, again, the question there is, is that something where we would want conditional approval for testing or more recommendation or suggestion to the developer that they sort of do further
evaluation and testing and come back at a
future date?

CO-CHAIR WEISS: And could we add
to that the objective being to standardize the
definition?

MEMBER CLARKE: Yes, I agree. I
think that the measure ought to be introduced
and used, but I think a lot of attention ought
to be paid to this non-preventable event
issue. That should be monitored and reported
back at a specific time in the future.

CO-CHAIR HOMER: I could repeat
what you said, but Helen?

DR. BURSTIN: Well, we were just
talking a little bit about there's enough
changes that you guys are recommending that
the question be, do you actually want to see
some analyses back before you make this
decision? I mean I don't know how much you
could look at these analyses.

I think one of my only concerns at
looking at this is, at least the way it is
written in the submission form, the
definitions of how anybody beyond your
institutions would use this measure would be
very difficult. This is intended to be a
measure of a national, that any hospital could
pick up and use, and at least what is in the
submission form is really fairly imprecise.

You may just have more of it that we haven't seen, but I think, given the number
of conditions, I just sort of wonder whether you actually want to take just a quick look-back.

DR. BERGERSEN: Also, what I have presented to you was based on a single
institution. As I mentioned, we now have three years of multi-center data that we could
insert your questions with.

CO-CHAIR HOMER: You know, one of my concerns is, again, my suspicion is that
the institutions you test this in are all very high-performing, highly-competitive, academic
institutions, and that if we were to apply
this broadly to the many places that are doing
cardiac catheterizations, they may not all
share the values.

I mean we can't assume some of the
transparent orientation of some of those
institutions. So, that is partly why at least
I have questions about standardization and
consistency that would not be as dependent on
the goodwill of the participating
institutions, to put it bluntly.

So, that is why I am not sure some
of those questions would be fully addressed
within the context of the collaborative
among -- I just want to say C3PO again, since
I like that term -- institutions.

(Laughter.)

So, I don't know. My sense is
there's a lot of interest and excitement about
this measure and the desirability for further
testing of it before we are kind of on record
endorsing it even in a time-limited manner.

But we don't want to give the perception of
kind of going back to ground zero and come
back at some indefinite future date.

MEMBER CLARKE: It sounds like the
testing may have already been done, and we
could just table it and then consider it at a
phone conference call, or something.

CO-CHAIR HOMER: So, rather than
vote, a suggestion that we table this and that
you provide additional information on, first,
clarifying the definitions, looking at how
this would be affected if you limited the age
criteria, for example, and coming back with
some of that information.

MEMBER PARTRIDGE: And, Charlie, I
think there were several suggestions about
actually --

CO-CHAIR HOMER: Clarifying the
wording?

MEMBER PARTRIDGE: Specifications
for the numerator and denominator.

CO-CHAIR HOMER: Yes.

MEMBER PARTRIDGE: So, that should
be in here, too. I would kind of like to see
the actual text of what we are voting on.

CO-CHAIR HOMER: Okay. Good.

Was there a comment on the phone?

Did someone on the phone say anything?

DR. MAIN: No, this is Elliott
Main. I am waiting for the next measure.

CO-CHAIR HOMER: Oh, okay. Thank
you.

(Laughter.)

Thank you for your patience. We
are a little behind schedule, I am afraid.

DR. BURSTIN: And just one more
analysis, since it was brought up, since you
have the data, if there is any ability to look
at the number of events that weren't
classified as potentially preventable and
preventable? Just to kind of give a sum of
how many are actually being excluded might be
useful, given the number of comments.

DR. BERGERSEN: I would be happy
to provide additional analyses. It would be
helpful for me if all of the comments and suggestions could be summarized in something.

CO-CHAIR HOMER: Okay.

DR. BERGERSEN: Great. Thank you very much.

CO-CHAIR HOMER: So, I think, yes, rather than calling a vote, why don't we recommend that and revisit this on a phone call?

Thank you very much. I think this was great.

So, what I would like to do is call for a 15-minute, well, I will say 10 minutes, but it will be 15 minutes, but a 10-minute break. We will try to reconvene at 11:10. Okay?

With Elliott. So, Elliott, you have a 10-minute break.

Thank you.

(Whereupon, the foregoing matter went off the record at 10:57 a.m. and resumed at 11:10 p.m.)
MS. McELVEEN: Let's go ahead and get started. We are going to reconvene, if we could have everyone come back to their seats. I promise you will get a lunch break, so you can chat then.

So, Elliott, are you still on the line with us?

DR. MAIN: Yes, I am.

MS. McELVEEN: Okay. Great.

Our next measure we are going to be reviewing is No. 31. This is the healthy term newborn, is the title of this measure. This is the percent of term singleton live births, excluding those with diagnoses originating in the fetal period, who do not have significant complications during birth or the nursery care.

I just want to also mention to the group one of the attachments I sent out last night was a visual diagram of this measure, which may help as we discuss it.

Elliott, just so you know, I also
have it here, projecting it, so the group can view it.

Elliott, did you want to take a few minutes just to make a few comments about the measure or introduce it any way? Or we can just open it up for discussion.

DR. MAIN: Well, I would like to say a few comments.

Thank you very much for allowing me to speak from San Francisco. It is a long trip back to D.C.

I am going to take you back to the beginning to pediatric care. Instead of looking at complications of sort of operations or procedures, this is really a reflection of both maternity, the summation, if you would, of maternity care and newborn care or regular nursery care.

A normal newborn is the most important outcome for us as obstetricians. I am perinatologist in the California Maternal Quality Care Collaborative, which is the
sister organization for the Perinatal Collaborative, led by Gould.

It actually serves as a balancing measure for most of the other measures that we have in the maternity realm. Ideally, you would like an institution that has an average or even a below-average maternity infection rate and a good, healthy newborn outcome rate, as opposed to a hospital that has a very high C-section rate and also a low rate of healthy term newborns. So, you really don't want to be in the position of pushing in one direction and having adverse outcomes in the other.

This is applicable to all hospitals that do maternity care. We have been working on it for over 10 years, tweaking the codes, looking at ways of capturing data in settings where people don't choose to code diagnoses for medical legal reasons.

For example, a number of hospitals in California have given up coding for perinatal asphyxia because that is a marker
for plaintiff's attorney.

What we have learned over that
decade, though, is that they do code for
procedures because you get paid for
procedures. So, being on a ventilator, having
CT scans, et cetera, all get coded quite
accurately.

So, this measure is a mix of
diagnostics, diagnoses codes, procedure codes.
What we have had more recently as a failsafe
is a length-of-stay indicator.

To start it off, though, instead
of doing extensive risk adjustments, we did
exclusions from the denominator. The
denominator is chosen to reflect healthy baby
as the mother arrives to the hospital for
maternity care. So, we have excluded the
general anomalies, intrauterine growth
retardation, babies who have hemolytic disease
due to Rh, for example, or hydrox, or infants
of mothers who have drug addiction, for
example.
So, that is our starting point. That actually accounts for over 3 million babies in the United States. This is a very high-volume measure, which is important because bad outcomes in babies are still an uncommon event. So, the infants that we are looking at here are somewhere between 1 and 3 percent, is the range we see in the hospitals, which makes it still a reasonable number, given the maternity ward denominator.

The only other measurement in this domain that has been approved is the AHRQ measure, ES-17, for birth injury/birth trauma, and a version of that was previously picked by NQF to be a measure.

Unfortunately, that measure has significant limitations. It is very low incidence, about 2 to 3 per 1,000 births. It is highly dependent on coding.

An article came out this last month looking at the HCUP's experience with that measure nationwide and found that 75
percent of all the kids that meet this
criteria for birth injury/birth trauma are
identified with two ICD-9 codes that both
begin with "other", other specified birth
injuries and other non-specified birth
injuries, which are very variable diagnoses.
That is probably the reason that that measure
was not picked up by the Joint Commission or
Leapfrog for their measure set.

This is really trying to fill a
void of a neonatal measure that would go into
the basket of measures to support maternity
care, and maternity care that includes nursery
care.

I would be glad to take any
questions or I will be available. Thank you.

CO-CHAIR HOMER: So, first, if I
could ask members of the Work Group if they
have questions. So, any questions? David?

MEMBER CLARKE: I would just like
to comment that I felt that this was the best
worked-out, most complete, and probably
easiest-to-evaluate measure that I reviewed.

I really don't have any thing wrong with it.

(Laughter.)

MEMBER RAO: Just a question, Elliott. Could you comment on its use in other environments? I understand it is being used internationally in the UK and other countries.

DR. MAIN: There is a normal birth measure in the UK, but that actually is a maternity measure rather than a newborn measure. A normal birth there is one without any interventions at all.

Everyone has been looking for this kind of a measure for a long period of time. This is the Holy Grail of what we are trying to do. And it has taken a while to put together the different pieces of the different codes to do this.

One of the challenges is in past measures the charts included codes from the mother, codes from the baby, and that is very
hard to do on any kind of large scale because those two charts don't intersect, don't relate to each other, and no data assessment. So, we had to take some extra time to focus only on the codes that we could get from the newborn codes.

So, there are flavors or variations of this that have been tried elsewhere. There is not one in the United States that has gotten to this point.

CO-CHAIR HOMER: So, again, if you were explaining this in words, and I know you could do it either on the healthy side or on the non-healthy side, but, basically, you are saying this is a term infant who doesn't have -- so, I am just trying to think how you are explaining this to a consumer.

DR. MAIN: We wanted to frame it specifically so it would be understandable by the public. But it is the proportion of term live births without a diagnosis, without a complication prior to birth, who do not have
significant complications during the birth or nursery care. In other words, this is a good take-home baby.

CO-CHAIR HOMER: Well, I know that, but I may be the only one on the Committee who is having just a little trouble understanding, but I am still unclear. Because, again, if you are coming to the hospital, actually, you don't know whether you have a congenital anomaly or not.

DR. MAIN: In this day and age, you often do with the advent of ultrasound, but it is excluding diagnoses originating in the fetal period, is the other way of explaining it.

CO-CHAIR HOMER: Okay. So, you are coming to the hospital. Presumably, you have had an ultrasound or something like that. So, you know if there is going to be a major congenital anomaly. Then, you are saying you know you have made it all the way to full term. Then, you are saying, what's the
likelihood that everything is going to go okay in the hospital and you will come home with a healthy baby, excluding bilirubin issues and excluding a few other --

DR. MAIN: Yes, there's a few minor things like bilirubin, but things that are clearly -- we also have excluded if we go into details of social situations such as babies being put up for foster care that may have a long length of stay in the hospital, babies that have drug withdrawal. Conditions that originated before you enter the labor and birth process, these are the ones that would be excluded. Conditions that arise during or after the birth process are the ones that are included.

CO-CHAIR HOMER: Okay. Good.

All right. Any other questions specifically on the importance? Then, we can move to the others. Kathy?

MEMBER JENKINS: I was just curious about the variation that has been
observed in the measure.

DR. MAIN: We field tested in a large health system in northern California with 25 maternity hospitals. We have seen variations there. Of the full measure, almost 150 to 200 percent, a fair amount of variation.

We have more limited detail on subsets of the measure that we published in the past which show actually quite large variation looking at the State of California, and, again, subsets of the adults where there's probably three- to fourfold variation.

CO-CHAIR HOMER: Faye?

MEMBER GARY: I just wanted to ask a quick question. I am not clear how you would deal with low-birth-weight babies.

DR. MAIN: Those are not included in this measure. This is 37 weeks or beyond, because those, obviously, have a large number of complications. You know, the mother's expectations are quite different if you are
coming in in pre-term labor or premature.

We also do exclude term low-birth-weight babies, which I mentioned before, those with small birth weights or intrauterine growth retardation. That, again, is a condition that arises before the labor and delivery process. So, that is a specific exclusion.

MEMBER GARY: But you have here that have not been -- these are morbidities that may or may not be clearly related to medical care. I was just thinking about all of the conditions that might impact whether a woman has a healthy baby or not, such as nutrition, diet, where she lives, what kind of support she has. There are just tons of data that support that especially, let's say, with African-American women that even healthy, middle-class African-American women deliver more low-birth-weight babies and have higher mortality/morbidities than their Caucasian counterparts.
So, I am not clear how these measures will help us to get at disparities among different groups who have had poor outcomes for a very long time.

I was just commenting, well, Dr. Zimmer just commented that, if you are poor, then what happens if you need a sonogram, for an example? Or what happens if you can't afford your calcium and your milk, or whatever?

I like what you have written, but it seems like to me there's so many other issues that revolve around what you are trying to do here, and I don't see any discussion about it or any acknowledgment of it.

So, would you just help me with my confusion?

DR. MAIN: Okay. Of course. It is very well-known that African-American populations and other disadvantaged populations have higher rates of pre-term births and small birth weight babies. That
would be covered by other measures that
address our nationwide racially-associated
rate of low birth weight.

This is really focused on, once
the mother gets to term, what are the
complications that arise during the labor and
birth process, rather than the prenatal care,
which is a subject of a different type of
measure. This is, whether or not you have
ultrasound, if you end up with a birth defect,
you would be excluded from this measure.

So, this is really trying to set
up an apples-to-apples type of comparison. It
has been looked at in actually rural
hospitals, urban hospitals, and big and small,
that would compare really what happens in
labor and delivery as to the outcomes then in
the nursery.

So, this looks, for example, at --
the numerator, then, is full of the codes for
birth trauma/birth injury, including the ones,
actually, that were excluded from the AHRQ
measure, such as brachial plexus injuries and clavicle fracture, the diagnosis and procedure codes around hypoxia and asphyxia and respiratory complications.

We have seen a rise in newborn respiratory complications from the use of elective recent C-sections at 37 and 38 weeks. This is the measure that would identify those.

There is the partner in quality improvement arm. That is one, for example. The other partner in quality improvement arm is the IHI safety for oxytoxin, where this would be the neonatal measure that would go with that to identify babies that had perinatal hypoxia or asphyxia related to prolonged oxytoxin use.

In terms of disparities per se, though, it does not address the low-birth-weight issue or any really of the prenatal issues that occur in those types of populations, but it is focused on how you manage labor and delivery, which should be the
same for everyone.

CO-CHAIR WEISS: Elliott, this is Marina Weiss.

I may just be reading this wrong, but as I understood the measure, it was the absence of conditions or procedures reflecting morbidity, but you are going to the other side and identifying the morbidities or the procedural problems that may occur, is that right?

DR. MAIN: It is either you get the absence by identifying the presence and subtracting it. It is a nice way of terming, I think, for families, and that is why we chose to do it that way, which is to focus on a healthy baby outcome rather than an ill baby outcome. The two are mirrors of each other.

MEMBER GARY: And the use, the utility of this measure is stated in the positive from your perspective? Suppose you were able to say that at your institution 97.2 percent of the children are born healthy and
everything is fine, given the exclusions, and so on. What have we learned?

DR. MAIN: Well, in comparison to other measures, and alone this should be as close to 100 percent as you can get. So, we have worked with some focus groups on whether it should be positively or negatively. People are attracted to the positive nature of it.

When you get down to the exact numbers of how it is presented, is 98 different than 97.5 percent? It gets a little tricky.

As with a number of the measures we've included, they end up with stars, based on their quintile distribution and the statistics that have been applied to them. That is probably how it would be displayed in a public release mode.

MEMBER GARY: So, would it be fair --

DR. MAIN: It is better than expected or worse than expected or average.
MEMBER GARY: So, would it be fair to say, then, that what you are doing here is attempting to think in terms of presentation to the general public, but at the same time you are capturing information that will be relevant to clinicians who are providing care, in that you are, in fact, keeping tabs on the morbidities? Is that correct?

DR. MAIN: That is exactly correct. We wanted to have something that would be easy to use, and perhaps for clinicians we might flip it and say, what is the incidence of ill term infant outcomes, which should give you, then, around 30 per 1,000 on average if it goes through the AHRQ thing. As I said earlier, about 3 per 1,000, and that allows you a lot more play in the ability to statistically compare hospital to hospital. It allows you to look at more hospitals as well as bigger hospitals.

MEMBER GARY: And if you could indulge me just one more minute here, and then
I will be quiet and let others interact with you, you said in your opening description that the purpose of this measure was to be a neonatal measure to support maternity care. Could you explain to me a little bit more -- maybe I am just not getting it here -- that link?

I mean a healthy newborn is the ultimate positive outcome. We all agree on that. But how does that reflect on the care that is given to the mom?

DR. MAIN: One of the major concepts in maternity or elsewhere is that you want to have balancing measures so that you don't push too hard in one direction to the detriment of another direction. And here, in theory, you have two patients, the mother and the baby. One of the concerns, for example, with trying to reduce the various infection rates is that you may end up with worse babies. That is possible. Or any of the other interventions that we do in obstetrics,
really we have our eye on what happens on the fetus, and we haven't had an initiative to go with that.

So, it is a balancing measure where we do more or less things to the mother that may advantage or disadvantage the baby.

CO-CHAIR HOMER: Nancy?

CO-CHAIR WEISS: I have a question. Speaking about the caesarean rate, we have a high incidence of caesarean. It says in here that you see babies now at 38, 39 weeks that end up with respiratory problems because of caesarean section. Okay, I understand that. But according to this data, if you come in and your baby gets a respiratory problem, aren't they excluded?

DR. MAIN: No. That is one of the numerators where there is both TTN and RDS and all these procedures that go along with being on a ventilator. A test tube, for example, non-invasive ventilatory, those are also included.
MEMBER FISHER: So, on 2a.3, all of those are included?

DR. MAIN: Again, it is the framing of whether it is healthy or, you know -- so, those, if you go down to the measure calculation, those are in the numerator, that it excludes you from being healthy.

CO-CHAIR HOMER: So, basically, what he is doing is he is identifying the number of kids who have one of these complications like TTN or respiratory disease, and comes up with a number or a percent and then subtracts that from 100 percent.

So, what is important is he comes up with either half a percent or 2 percent or 3 percent of the population, but presents it as 99.5 or 97 percent. But you still have the challenges identifying that percentage, that small percentage, and then it is a question of marketing your presentation or what families want to know as to whether you present it as that 2 percent of kids have a problem or 98
percent of kids come out just fine.

        Kathy?

        MEMBER JENKINS: I just want to be
sure that I understand then. Everything that
you have basically included in the definition
of not a healthy newborn you believe is
preventable or avoidable by changes in
maternal care? Is that correct?

        DR. MAIN: That is one of the
topics that is debated. The neonatal births
that we work with, the procedures, that is
basically an offshoot of Vermont Oxford, in
our group we looked at these very, very
carefully. One example that I said before was
brachial plexus injuries. You know, AHRQ
excludes that, though that is a major
morbidity for babies. It can be prevented if
you do a C-section. It does not mean that
this is malpractice though. That is why it is
excluded, because people thought, well, you
can do perfectly normal or perfectly adequate
obstetric care and still get brachial plexus
injuries.

From the patient's perspective, though, that was an unexpected outcome, and it is a significant outcome, that you don't have a healthy term newborn if you have a baby with brachial plexus injury. And that was the philosophy that we ended up choosing to use in those borderline cases, balanced by trying to exclude as many diagnoses that were present in fetal life before we get into the measure itself by screening those from the denominator.

CO-CHAIR HOMER: Do you have a follow-up question?

MEMBER JENKINS: I asked that question when you mentioned TTN. So, I assume that there is a way that TTN can be avoided.

CO-CHAIR HOMER: Through a C-section.

DR. MAIN: TTN, the most frequent cause by far is C-section without labor. We don't have the squeeze on the lungs, and you
have often a little bit of early gestation involved at 37, 38 weeks as opposed to 40. That has a three- to fourfold increase rate just from that case alone.

Of course, our goal 100 percent. But, no, there's no center that will get 100 percent from this measure. There will always be something that gets through. But there is big variation and big opportunities for improvement here.

CO-CHAIR HOMER: If I could ask a question, I had a question. I was a little confused about your definition says that it is identified term signals in infants, and yet you said this would be sensitive to this issue of, quote, "late pre-term" births, which is, of course, the most important contribution to the increase in pre-term.

DR. MAIN: The term is, the normal is focused up the early problems at 37 to 39. Actually, there is a big project we are doing with the March of Dimes right now on
prevention of low-weight births, which we think will sort of spill over into the late pre-term population.

CO-CHAIR HOMER: So, I guess, again, I agree that is a critical or the critical thing to be addressing. It feels to me there are more direct ways to address that, like, you know, measuring the proportion of infants that are born less than 38 weeks or something like that. Do we have that measure already? Okay.

Because I was going to say that this seems like a rather broad brush to use to attack that specific thing that should be addressed. So, okay.

Allan?

MEMBER LIEBERTHAL: Yes, I have two questions. One is how you deal with intrapartum fever in the mother and whether those are excluded or not. And the second is, now that you have excluded so many of the things that cause neonatal morbidity and
mortality, even 150 percent difference among
institutions, what is the effect size of that
difference? In other words, does 150 percent
really mean anything?

DR. MAIN: Sure. Let me do the
last one and then I will go back to
intrapartum fever.

We are talking the differences
between basically 1 percent and 3 percent or
a little over 1 percent or a little less than
3 percent of the population. So, that is
still a significant effect size. Yet, when we
get into term babies, the biggest proportion
of morbidity -- this is a general anomaly --
but there is still a fair amount of morbidity
of babies admitted to NICUs, which in a sense
this is a surrogate for, babies that go into
the NICU and have other morbidities that don't
quite get you to the NICU, but it still
accounts for a real number of cases.

The trouble with anomalies is that
there isn't really much we can do at this
point to prevent them once they occur. We are all giving everybody a lot of folic acid and taking that route, but prenatal diagnosis doesn't actually cure your anomalies unless the family should terminate. So, that is a very different population, a very different issue than what we are dealing with in birth issues and counting managed labor and delivery and its consequences for the baby.

In terms of fever, that is one that the expert panel worked on a fair amount. There is very large variation in how infants are handled in all the nurseries around the country, and we have most of them in California, in terms of what kind of workup the baby gets after the mother has had a fever in labor. It goes from observation to IV antibiotics.

It is quite interesting that there is not a lot of difference in outcomes when we look at those. So, we are looking at encouraging mothers in labor with fever to get
aggressively treated in labor. That does appear to prevent a lot of the neonatal outcomes. So, there is the ability to affect that.

Now what is included in our numerator or in the, quote, "exclusion" set is babies that actually have sepsis, not babies who got antibiotics. So, that gives the obstetrician the opportunity to have that intervention. There will be some of the cases where IV antibiotics with the mother actually is significantly reducing sepsis rate in infants.

CO-CHAIR HOMER: Tom?

MEMBER McINERNY: I think if this really becomes widespread, that it may be one of the first things, if not the only thing, that would reverse the trend in increased caesarean section rates.

I don't know, do you anticipate that or have you actually seen any evidence of that since you have been using it?
DR. MAIN: Yes, that is one of my directions; I will have to put that out. Actually, what you would like to have is a good rate of good babies and a reasonable rate of C-sections. Right now, we have C-section rates that range from 15 percent to 50 percent in hospitals in California. There is not much variation in there, and everybody wants to have good babies. You don't get that much additional benefit, if any, on the baby's side for those kinds of variations in C-sections.

You may have been following in Sutter Health, which is, again, 25 hospitals in northern California, some variations of this. That includes Apgar scores, for example, 500 Apgar scores. That has elevated our C-section rate quite significantly. So, we are way below the State average and the national average. It still has increased. I can't say it is flat, and even though the quality effort is there, but it is much below the national and State rate.
I think you want to have data like this to really show what your outcomes for your babies are in your term babies. They have focused a lot in outcomes on prematures and survival rates for under 15000-gram kids, and so forth, but we haven't really had much attention looking at term babies, which this will fill the gap for.

CO-CHAIR HOMER: Can I just ask what the drivers are of these rates? I mean, do you have -- no, not the C-section rates, but this performance measure. You know, again, you have got lots of different codes that can get in there, but I am trying to see, basically, is it the TTN for the 37- to 38-weekers that is driving 80 percent of the variance here or is it everything together?

DR. MAIN: When you look at composite measures, you always have to look at which component has the biggest frequency within. I mean which drives the code. Respiratory is the main one. Birth hypoxia
and asphyxia is probably second or third in there. First would be respiratory. The second would be infections. The third would be hypoxia/asphyxia.

For more hospitals, it is transferred for care, you know, where you have to transfer the baby out to another facility. That is Part B on this schema. That is a major dissatisfier, a major negative for families to be put in that position where they are separated from their baby.

CO-CHAIR WEISS: Let me just observe that that is very interesting in that it correlates perfectly with the top expenditure codes in the Medicaid program. I mean there are four or five different categories in which expenditures for these kids fall that are pretty high, highest in the respiratory distress arena.

CO-CHAIR HOMER: Nancy, I think you had a question?

MEMBER FISHER: I had a comment.
I think you asked about -- I don't know about since this measure has been out there, but there are several studies across the United States with people in hospitals reducing the C-section rate. Especially I can think of one; it was in Akron, Ohio, and they talk about reducing the C-section rate by making sure that you have a protocol for induction and that the people buy into it and stuff like that.

In Washington, we are also working on that, but we are just taking that measure, not something this big. I was wondering the advantage over this because I believe Leapfrog is now going to start collecting information, too, on -- what do you call it? -- C-section rates and in what we call late-term babies, 38, 39 weeks. I mean late C-section is what they call it.

DR. MAIN: Yes, the risk adjuster in the low-risk term C-section rate is actually a measure from our institution. It
is an NQF measure. It is now a Joint Commission measure.

If we use that and implement it around both in systems and in states, the obstetric pushback is, what about the baby? You know, we may be high for C-sections, but we want to make sure we have good babies at the end. That is one of the drivers, to have this as a balancing issue.

I think the effective measure that you mentioned helped the elective delivery prior to 39 weeks measure, a little bit by C-section induction. It is a very important measure and it will change some of the practice. That is just measuring the frequency of births at that time period. That is going to be a very important measure, as I have said. This will allow us to say that this is actually includes outcomes for the babies at the same time.

MEMBER FISHER: I was saying, yes, we have the measure. What I am saying is that
we are doing something about the number of C-sections. So, it is the same thing. We have got a couple of hospitals that have a 50 percent rate for C-section. They are small hospitals. The average rate in Washington is 33 percent. We know that, and we know we need to reduce it.

So, we have five pilot projects going about looking at babies born at 38, 37 weeks, and we do things about induction. So, we have the numbers. We are implementing something.

I guess what I am saying is, why is this measure better than what is being measured out already?

MEMBER PARTRIDGE: I don't want to respond for Elliott, but I served on the Perinatal Steering Committee, and we debated the C-section rate measure endlessly.

(Laughter.)

I think that we need both. The C-section rate tells you you've got a rate
that seems way out of line. As I understand it, what Elliott is trying to say is people advance in support of a high C-section the danger of an unfortunate outcome for the baby. This measure is designed to give you some sense of, I think as Elliott said earlier on, if you lower the C-section rate, your rate of bad babies is going to go up.

Am I sort of right?

DR. MAIN: That is, well, there is a legal risk, there's all kinds of risks out there in terms of babies, but the reality is that the C-section rate has gone up, but the outcomes for babies has not changed. It has not improved with the higher C-section rate. But we don't have a measure to really show that.

So, it is a complementary measure that allows you to put it in the place of projects on C-sections and have it be the safety measure that shows that you are not being harmed. In fact, you may be actually
improving neonatal care by having a more moderate C-section rate.

CO-CHAIR HOMER: And this measure doesn't weight different complications differently, which is fine.

DR. MAIN: No, no, we decided not to do that. That is inherently objective one way or the other.

CO-CHAIR HOMER: Yes, I am not arguing with that. I am just thinking of the countervailing argument. When you reduce C-section rates and reduce with them, presumably, the respiratory complications, there may or may not be, but probably there won't be, there may or may not be some small increase in some other kinds of complications, which was the rationale for the C-section in the first place.

I think we have actually had a great conversation about this. I would suggest we could probably move on to voting, unless there are compelling questions. I
don't see any.

So, I would say the first threshold question is whether this is important enough for us to proceed.

So, why don't we have all those who believe this is sufficiently important to proceed, show of hands?

DR. WINKLER: Fourteen. That's all we've got now.


So, then, let's move on to the discussion of scientific acceptability. We have had a fair amount of conversation about this, but I don't know, Elliott, if you have any comments or there are questions from any of the members about validity, reliability of this measure and the various other elements of scientific acceptability. Or do we feel that it has adequately been addressed?

Some people do need lunch. Okay, but we are not quite there yet.
Okay. Any questions about scientific acceptability of the measure?

(No response.)

People feel good about it.

Okay. So, those who feel it completely meets the criteria for scientific acceptability show of hands.

And partially meets?

Good. Okay, that has got everyone.

Next is the area of usability.

And again, you have said you have done a fair amount of focus group work with this and efforts to communicate it.

DR. MAIN: And also, if it is straightforward administrative data, that would probably be nice. So, that is the gun I am under in California, is that it has to be, new quality measures need to be using administrative data as much as possible.

CO-CHAIR HOMER: And can you describe any use in your Collaborative or at
Sutter or Kaiser or any of the other places in terms of how providers have experienced this measure and how it has contributed or not contributed to quality improvement activities, et cetera?

DR. MAIN: We used earlier versions of this extensively in Sutter Health as the parallel to our C-section quality improvement effort and our oxytoxin quality improvement effort. We are starting the elective delivery for 39 weeks, and we will probably go with that, but it has been both a source of reassurance and, you know, it changes the focus of this to say, okay, what could we do to optimize the baby outcomes that is appropriate? So, it has been the patient measure that goes along with the other quality improvement measure.

CO-CHAIR HOMER: And is there any evidence of improvability? That is, I know there is variability across sites. Have you seen within single sites any trend data on
that?

DR. MAIN: We have seen trend data both for the components of respiratory and infection. We have been pretty good on hypoxia and asphyxia, the biggest categories. That is a third the big three categories. So, we haven't seen as much there.

But there are places around that have higher rates. We don't have quality improvement efforts that are just for show. We have improvements for the respiratory complications, for infection. So, there is opportunity.

CO-CHAIR HOMER: All right. So, in terms of usability criteria, those who feel it completely meets the usability criteria?

DR. WINKLER: Eight.

CO-CHAIR HOMER: And partially meets?

DR. WINKLER: Six.

CO-CHAIR HOMER: All right. So, that's got everyone. Good.
And then, feasibility, which is the one that does specifically get at the issue of availability of administrative data and ability to collect and generate reports, and all that sort of stuff.

So, how many feel it completely meets the criteria for feasibility?

DR. WINKLER: That's everybody.

CO-CHAIR HOMER: Okay. Good.

All right. So, I will call for a measure to recommend endorsement of the measure. This one would not be conditional or time-limited. This would be endorsement of the measure to go forward as a regular measure within the NQF.

So, all in favor of recommending endorsement?

DR. WINKLER: Fourteen.

CO-CHAIR HOMER: All right.

Congratulations. This is good.

DR. MAIN: Thank you.

CO-CHAIR HOMER: We've got two
more measures, guys.

(Laughter.)

DR. WINKLER: Thank you, Elliott, very much.

DR. MAIN: Thank you very much.

CO-CHAIR HOMER: Thank you.

MS. McELVEEN: Okay. We are going to move on to our next measure. It is Measure 48. I hope we still have folks from AMA PCPI on the phone, who have been waiting for this measure.

MS. FEI: Hi. This is Kerri Fei, staff from the AMA PCPI, and we also have Dr. Barbara Fivush, who is our Co-Chair.

MS. McELVEEN: Okay. So, again, this is Measure 48. The title is plan of care for hemodialysis. This is the percentage of calendar months during the 12-month reporting period in which patients age 17 years and younger with a diagnosis of ESRD receiving hemodialysis have a single-pool Kt/V greater than -- yes, okay -- or have a single-pool
with a documented plan of care for inadequate hemodialysis.

So, I will allow you guys to kind of explain that. Sorry, I butchered the description a little bit.

(Laughter.)

MS. FEI: Oh, no, you did fine.

Did you want me to give a little, brief description or --

MS. McELVEEN: Sure, that would be fine.

MS. FEI: Okay. So, we developed this measure I think about two years ago, after we had developed the same measure for the adult population, which was actually developed prior to this one, wanting to have the same measure for the pediatric population as well.

So, really, there's really not much difference between this one, and the RB panel did endorse the adult measure, which is actually, we just gave testing results for.
I think we will be going to the CPAC sometime next month for potential full endorsement.

We did provide the testing results from the adult measure. We have not had the uptake for the pediatric measure as of yet.

So, really, Dr. Fivush, was there anything else you would like to add?

DR. FIVUSH: Yes, just because this is a highly-specialized field within a field, so we are really talking about a small population of patients. Probably in our country maybe 800 pediatric patients maintain on chronic hemodialysis, but it is a very vulnerable population in that it has a fairly high mortality rate, which we are trying to address in other ways.

But there is a gap in care here, in that we think about 12 percent of patients in previous datasets have not met what we think is adequate dialysis. That is measured by a Kt/V which looks at the way urea moves, to simplify it.
So, we have good evidence that a Kt/V of 1.2 is a dialysis prescription that is adequate, and it is a really minimal prescription. We have linked a low Kt/V to poor outcomes. We have a high mortality rate, and we think this is an easy-to-capture measure.

It is reported on a monthly basis, physicians have coverage. Doctors can easily get to this number, and we will be able to closely monitor how patients are getting dialysis in the country that are pediatric. Hopefully, we will be able to use this data long-term to really link it to more long-term outcomes. This is an intermediate outcome.

The measure is both a process and an outcomes measure in that we are looking at a standard of 1.2, but we are, additionally, looking at a thought process that, if you do not dialyze this patient well enough, what would you do to change that? So, we think it is a good combination measure that is going to
give us important information in a vulnerable population that has a high mortality rate that we think is easy to capture.

CO-CHAIR HOMER: This is Charlie Homer.

Could you explain again why it is an outcome measure rather than a process measure?

DR. FIVUSH: Kerri may want to help me.

It is an outcomes measure. It is single-pool for a Kt/V of greater than 1.2. So, we aren't looking at an outcome specifically, but we are looking at this as long-term.

Do you want to clarify that?

MS. FEI: Sure. The measure actually is a combined process and outcome measure. So, when the measure results get reported out, you are going to know your patients meet the outcome, and for the patients who don't meet the outcome, that they
I have a documented plan of care.

So, the users of the measure would get all pieces of the measure reported back to them. They would have patients with a single-pool Kt/V greater than or equal to 1.2, patients who have Kt/V less than 1.2 with a documented plan of care, and have patients with a single-pool Kt/V less than 1.2 who don't have a documented plan of care, which would be your measure failure.

CO-CHAIR HOMER: Could you explain Kt? I mean it has been a long time since I did dialysis or nephrology. Just in laymen's terms, what Kt/V is?

MS. FEI: It is urea kinetic modeling. As I briefly alluded to before, it is really the movement of urea and how long you are clearing it from the body over the course of the dialysis procedure. We use that as a measure of adequacy, with the idea that if we are moving urea, we are moving any pools, you know, through the process of
dialysis, and then if we are dilating someone well, the movement of urea going through results in a higher urea kinetic modeling. It is going to result in a higher Kt/V than if we do not.

So, higher would mean more dialysis, either longer dialysis, a different cartridge, higher blood flow, but it would indicate with the Kt/V, the higher the number, the more dialysis a patient is receiving by measuring the way urea moves.

CO-CHAIR HOMER: Kathy?

MEMBER JENKINS: So, can you help us understand why a patient would not have an adequate Kt/V and why the measure wouldn't just be having an adequate Kt/V as opposed to if you didn't have the plan?

DR. FIVUSH: I think I don't feel that it would be simple to just dial up the dialysis or to make everybody have a Kt/V greater than 1.2. It is hard for me to speak to specifically why people wouldn't try to do
that, and my assumption is they would.

But there are patient
characteristics and catheter characteristics
that lead to the inability to dilate the
patient adequately. For example, in a
pediatrics population, one of the things we do
think is a problem is that most of our
patients we chronically dilate have external
catheters as opposed to internal fistulas or
grafts, and so they have a higher risk of
infection, will try to move in that direction.

But if you have a catheter, they
may not get the best blood flow. There may be
recirculation of blood within that catheter,
and you may not be able to adequately dialyze
this patient. So, there are some factors, and
then there are some patient factors about
their ability to tolerate how we dialyze them.
If we are dialyzing them three times a week
and trying low fluid, we may be unsuccessful;
they may get hypotensive during the procedure.

We may not be able to do what we would like to
prescribe.

So, maybe that patient would have a Kt/V on a single session of less than 1.2, but the nephrologist would be bringing them in for a fourth treatment a week. And another plan of care might be to change the access in the patient. Another care plan might be to try to change blood flow by changing the way you actually expose the patient to sodium.

So, although it sounds that it would be easy, in this many patients we can't always get the blood flow rates we want. We have recirculation. We have patients' vulnerability. They can't tolerate how long we want to dialyze them for.

So, sometimes, to get to that 1.2, we have to be creative. We have to put thought into, and we may have to change an access. We may have to work with the family and our surgeons to move towards a better access. We may have to do dialysis more often or differently.
And it actually is allowed in our care plan to say, well, you know, we are going to change this. We are going to change the rate of flow. We are going to consider more frequent dialysis. We are going to change to a different dialysis. We are going to change our modality.

It just gives us the ability to address the fact that, although it sounds very easy -- I would just use the example of when we talk about target hemoglobins, and we say they should be 10 in our patients, that we can give them a lot of erythropoietin-simulating agents. And many times, we can't reach that 10 anyway.

So, there are just patient variables that prevent that from always getting to be adequate, to what we think is needed. The care plan will let us look to make sure that physicians are addressing the adequate Kt/V.

MEMBER RAO: I just wanted to echo
what Kathy's concern is. I think with only 800 children going through hemodialysis, I am concerned that the numerator, the number of patients who don't have a documented plan, is going to be very, very small.

What constitutes a documented plan? It sounds like it would be complete lack of recognition that the Kt/V was less than 1.2. I mean, if somebody wrote down, well, increase frequency of dialysis, would that be adequate?

If the standard of 1.2 is so well accepted, it is hard to imagine too many physicians not documenting something to that effect.

DR. FIVUSH: I think until we look at this -- I mean we have looked at Kt/V through the KTM dataset. We have been fortunate that the government -- because overall the pediatric part of the this is not in Medicare; the adult part is. So, we have had scrutiny for a long time in data
collection, for a long time, and we know there is a gap in care, you know, in terms of Kt/V. I am not sure we know why yet. This measure will allow us to better understand practice around it.

MEMBER RAO: Right, and I understand there is a gap in Kt/V. It is the documentation of plans that I am not sure there would be such a big gap for.

MEMBER DOCHERTY: I was wondering what the evidence was of the relationship between documented plan of care and better outcomes for these patients.

DR. FIVUSH: Well, you know, as I said, in the United States we have a very, very high mortality rate in the first six months of patients placed on hemodialysis. It is about 22 percent. It is very high, and it is high in pediatrics as well, and going up, but probably not that high.

We have never really been able to capture the data looking at individual
physician practice patterns. We have looked at it in the CPM. It has been looked at more as -- it is not being broken down regionally because the cells are too small. But I think it is important, you know, to really improve care, to start looking at this as a physician measure to see if there are practice patterns that can change, because, clearly, there seems to be in the literature the suggestion -- we know we are looking at intermediate outcomes in our patients. I mean a payment of a dialysis prescription is an intermediate outcome; it is not a true outcome.

But there is in the data evidence to suggest that needing the intermediate outcomes results in fatality and hospitalization. So, it is a complex, it will be a complex analysis because there are other intermediate outcomes that we have to do as well.

You know, I mentioned hemoglobin before, but there are nutritional outcomes.
There are a lot of intermediate outcomes we have to meet, but this is one that we thought we could target, start educating physicians. That will be important for this measure and its linkage to mortality, and additionally, to start having them submit their care plans because I think that is critical to start thinking about how people are addressing it.

MEMBER DOCHERTY: I think that is just the piece that I am having a hard time understanding, not the physiologic outcome, but that a documented plan of care will lead to that physiologic outcome.

DR. FIVUSH: Sometimes I guess I think if a physician, because we get licensed in the State of Maryland, certified for a dialysis unit on a yearly basis here, we have to write care plans. Most states do not have yearly licensing of dialysis facilities as we do. So, they may not be licensed or certified for seven to ten years.

But we write down a care plan for
every patient that doesn't meet standard
targets here. And actually, I have found it
has -- and we have looked at our numbers over
time -- it has driven quality improvement.
Because if you continue to report that you
have 2 percent of your Kt/V's less than 1.2,
or 5 percent, you have to justify each six
months what you are doing. It has really
started to -- not just dialing it up in
dialysis; it is really pushing our unit toward
start using in-dwelling lines, to move away
from external catheters, which is critically
important.

So, I think, looking at your
numbers and reporting them, and looking at
your inadequacy in dialysis, and documenting
what you are doing about it, is going to be
very important for driving improvement.

CO-CHAIR HOMER: Kathy?
MEMBER JENKINS: I am sure that
that is correct. I guess the question I still
have is, first of all, there's general issues
about composite outcomes, but this is even more complicated because it is a composite outcome and a process outcome at the same time.

So, just to state it in the extreme, if there was one site that met the outcome by having all the patients meet the physiological outcome, and another center who met the outcome by having none of the patients meet the physiological outcome, but have all of them have a documented plan, I do not consider those two to be equivalent.

So, it almost feels to me like you are trying to have all the sites like look fine or be able to achieve 100 percent, and I think it is the variation, and then, to your point, you know, the steps they take to achieve the 100 percent on the physiological outcome which is actually the relevant outcome.

And if there are intractable patient factors that make it much harder to do
that, then that becomes a need for risk adjustment for the outcome variable, as opposed to adding in the process, at least the way I am hearing this.

DR. FIVUSH: I think one of the things that we can do, and we are moving towards, when we have those types of patients, it is to move to more frequent dialysis, which is a move across the country. And again, I think we will, you're right, the way they state it clearly suggests that those outcomes would be equal, but I think those outcomes are not equal, and I agree they are not equal. But the way they will be reporting back to the physicians will include which of their patients had what adequacies and how many were over 1.2, but how many weren't over 1.2 and had a care plan.

I think, clearly, having a care plan and not having an adequate dialysis means that that is something that needs to change over time. You have to figure out a way to
have adequate dialysis. You can't just report that you are trying. So, I think it is going to be very valuable because it is reported back to people, because they are going to see those numbers and that detail.

If we just left it at greater than 1.2, I think as just an outcomes measure, I think we wouldn't be giving an opportunity for the kind of improvement we are hoping to see. Because in many patients it is going to be difficult to get to 1.2 because of the factors we have discussed.

I think, again, when we started the conversation, if it were easy to achieve a 1.2 in everybody, I don't think we would have a gap of 12 percent. So, I would agree with your point that it is very important, but it clearly is not the same to have 10 patients who have met your Kt/V of 1.2 and another unit has 10 patients who have not met any adequacy measures but have a plan. Those would be very different outcomes.
CO-CHAIR HOMER: Helen?

DR. BURSTIN: This is Helen Burstin. I just want to weigh-in.

Having lived through the first round of ESRD measures in 2007, this is essentially -- just correct me if I am wrong -- the same measure with a different level. It was 1.7 for adults; it is 1.2 here. Yes?

MS. FEI: It is the past hemodialysis measure that is 1.7. The adult hemodialysis measure is also 1.2.

DR. FIVUSH: Right. So, we are really aligned with that adult measure.

MS. FEI: And with this measure, you can have a rate report out of patients between 1.2 and 1.7. That is done through the administrative coding for the adult measure as well.

DR. BURSTIN: All right. I guess my question was trying to understand, is there any reason you couldn't potentially take the initial measure that is already endorsed and
just extend the age down to children?

MS. FEI: Actually, we did talk
about that. However, the plan-of-care
definition for the pediatric measure is just
a little bit different than --

DR. BURSTIN: You could stratify
the measure and have that information in
there. It just doesn't necessarily seem like,
you know, if it is really very, very similar,
do we really need another measure in this
case?

My second point was just that,
when we went through this the first time, we
had a lot of discussion about this exact issue
that you are grappling with today of adequacy
of dialysis and plan of care. One of the
requirements that came out of that process was
that the expectation was the measure would
reported with two rates, so that you would be
able to see the adequacy of dialysis and,
then, you would be able to see, if not
adequacy of dialysis, is there a plan of care?
I just want to be sure that that
-- I mean, certainly, we would hope to be
internally consistent as best as we can at
NQF. So, that would certainly be the
expectation for this one as well. I just want
to make sure that that's your understanding as
well.

MS. FEI: Yes, and that is how we
have it set up.

CO-CHAIR HOMER: Okay. So, it
really is, in essence, two measures under one,
or at least reported as two linked, paired
measures in some sense.

And again, is there an assessment
of the adequacy of the plan or it is simply
they have a plan? Now how does that work?

DR. FIVUSH: I think we put down
in our description there are various plans
that we would consider acceptable, and we
listed examples of plans that we would say
were acceptable plans.

I think the level at this point
would reveal that there was a plan. This is
a little different than the adult language in
what is an acceptable plan.

CO-CHAIR HOMER: I am trying to be
a little consistent with some of our earlier
conversations when we gave another group a
very hard time about the categorization of
preventability or not, and things like that.

So, is the idea here that all the
plans would come to a single place? It would
make a judgment based on criteria as to
whether a plan is adequate or not? And again,
I may have missed it in the specifications.
So, how will you be judging the adequacy of
the plan?

MS. FEI: I don't think through
the use of the measure we would be able to
just have the adequacy of the plan.

CO-CHAIR HOMER: Okay.

MS. FEI: It would be that there
is a plan of care in place.

CO-CHAIR HOMER: Okay. So, any --
MS. FEI: In the definition, we have a definition of what the documented plan of care may include.

CO-CHAIR HOMER: It may include any of those things, and if it doesn't include any of those things, but says, you know, I don't know, "I will see them back more frequently" or "I will call the mother to make sure he is doing okay," or something like that, or less frequently? I mean, you know, in asthma we talk about the importance of a written care plan. So, maybe that is kind of similar to what we are talking about here. But, to be honest, there is also at least a little bit of evidence in that case that --

MEMBER DOCHERTY: So, is it just a dichotomous variable? Either it is there or not?

Then, along with that, I was just wondering about your Kappa statistic. Then, it looked like it ranged from 42 percent all the way to 93 percent. I guess it appears,
then, that there might be some differences in
definition of whether there is an adequate
plan of care.

MS. FEI: There is a lot of static
on the line. So, I am not sure --

CO-CHAIR HOMER: The question
really was what the reliability of the
assessment -- there was a Kappa statistic that
was presented that has a pretty low bottom
number of .4, you said, a pretty wide range,
and didn't know --

MS. FEI: Right, and the Kappa is
from the testing of the adult measure.

CO-CHAIR HOMER: Okay.

MS. FEI: And really, through that
experience, what they found was that at
different sites the manner in which the plan
of care was documented was different, found at
different places or not present at all. Or it
was either in the physician's office or at the
dialysis facility, depending upon where the
physician was seeing the patient.
CO-CHAIR HOMER: So, Ellen, you had a question?

MEMBER SCHWALENSTOCKER: Yes. It just relates to the plan-of-care specification. I am drawing a parallel, as you did, Charlie, to the children's asthma care measures, which actually has components of what should be in the plan of care.

That has problems of its own, but I am wondering if you have looked at that, and if it would be possible to get a little bit more specific around what must be in the plan of care in order for it to be adequate.

DR. FIVUSH: There's a difference between what a documented plan of care may include and what a documented plan of care should include.

I think, looking at our plan of care, I mean our measure, I know that the reason we didn't say "should" is because, for example, one of the things that could be in a documented plan of care would be increasing
the blood flow or increasing the dial at the site. That is not possible for some patients. They may not tolerate that. So, if we say "should" -- we can't say "should" because the same "should", if the patient is big enough, it should, but we can't say that in a patient who is hyposensitive because that would make the patient sick.

That is one of the problems we had in creating the measure. You know, certainly we should say it should include documenting revisional renal function because that is easy. But many of the things we can't say that is the way to fix it. We can't say, for example, changing the access because it is possible that that patient isn't a surgical candidate for better access.

So, I think there are things that should be in a plan of care, but I don't think we could standardly say they must have this in a plan of care because it wouldn't allow for any patient variability. Do you know what I
mean? The patient per se couldn't have a better access because they had sort of used all their blood vessels. The patient couldn't tolerate a higher blood flow. Those are really very real scenarios.

MEMBER RAO: Once again, in the interest of simplicity, and I know you have addressed this, if you just switched to a simple Kt/V measure up and down, is there any reason to think that some of those other factors, children with poor access, are distributed any differently across the country among those 800 patients? I mean you are going to get those people everywhere. So, as a quality measure, wouldn't it be simpler just to switch to the 1.2?

MS. FEI: The other thing that we don't know precisely is this is a pediatric measure we talked about, but we are not sure how many pediatric patients are dialyzed as adult. We know that we can tell something about provider types, but in a study that we
did several years ago trying to figure out how pediatric patients were dialyzed and where, we think that at least one-third of children under the age of 17 were dialyzed chronically or dialyzed by an internal medicine nephrologist.

So, we are not sure, as we go forward, if practices are different regionally, if they are different, say, in provider type, if they are different based on care as in a pediatric unit versus an adult unit. I think we will find out some of that information when we start looking at a physician-level measure that we don't have right now.

CO-CHAIR HOMER: Kathy?

MEMBER JENKINS: Can I just ask if most of the issues of essentially patient factors that make it impossible to achieve the goal, is that only in the little babies? I mean, is there a way that you could perhaps not go down all the way to zero here and get
rid of some of the challenges? Or else, I
guess alternatively, create an age
stratification or a risk adjustment by age or
size?

DR. FIVUSSH: Well, you know, one
of the things about this, we just haven't got
the simplicity. It is not a large population.
When you try to take out or look at the small
children, you end up going into more and more
subgroups and losing your ability to look at
children, although, clearly, the babies, the
infants, they are different than the
adolescents.

But I think that even knowing that
it is harder in an infant, it is probably more
important for the younger children to have the
dialysis, if we were to say, where is it more
important, because of issues in growth and
development.

So, I really don't want to take
out the infants, even though there aren't
many, and say, okay, we're not going to look
at how you dialyze babies. Because if people are doing dialysis in young children, they need to be very aware of their adequacy.

I agree it is hard. As in all pediatrics, we are dealing with different patient issues as children grow. And certainly, there is an impact on growth in terms of if we can use blood flows. But I still think we need to look at the young children because they probably are the most vulnerable patients.

CO-CHAIR HOMER: The last comment and then I think we could probably move towards voting. Faye?

MEMBER GARY: I just wanted, before I vote, to clarify that there will be some determination about where the care takes place, and thinking about university centers and where they are all, let's say, intensive research-oriented university center versus, let's say, private facilities that might be in rural areas, for an example.
DR. FIVUSH: I think, certainly, because this is a physician measure, I will be able to find out who is providing care for these patients. I don't know that we will be able to tease it out at this level yet.

Kerri, you can help be with that.

This will probably also go in, we are hoping, as the facility-level measure, as part of the clinical performance measures, but they don't have physician-level measures throughout. So, hopefully, if we can get these measures in place, we will be able to address that very important question: who is the primary provider? Is it an internal medicine, a pediatrician? That may really have no difference; we don't know.

And where is that care being provided? In a hospital unit? In a freestanding pediatric unit? In an adult unit that takes care of children? In a private practice facility? I think those are very important questions.
CO-CHAIR HOMER: All right. So, I would suggest -- this has been very helpful -- that we move towards voting on the measure.

MEMBER LIEBERTHAL: Have we decided whether this is, indeed, an outcome measure or a process measure?

CO-CHAIR HOMER: I think my sense is it is a combination, that the Kt/V is an outcome measure, but it is a paired measure, both outcome and process.

MEMBER LIEBERTHAL: So, it meets our scope?

CO-CHAIR HOMER: I think it would fit within our outcomes scope.

DR. BURSTIN: We have basically been saying any composite measure that included outcomes was in. So, I assume a paired measure that included an outcome would be within scope, too.

CO-CHAIR HOMER: Okay. So, voting on the importance of the measure. Remind me the criteria for importance? Okay. So,
clearly, in terms of relation to outcome, it seems strong. For the Kt/V, it is challenging because we've got one where I think we have a lot of confidence in the relationship between the intermediate and long-term outcomes.

But, okay, without more editorializing, let's vote.

All those who believe it meets the importance criteria?

DR. WINKLER: Eleven, 12.

CO-CHAIR HOMER: Okay. Those who believe it does not meet the importance criteria?

Two? Okay, good.

The next one is the scientific acceptability of the measure. How many would believe that it completely meets the criteria for scientific acceptability?

How many feel it partially meets the criteria for scientific acceptability?

DR. WINKLER: One, two, three, four, five.
CO-CHAIR HOMER: How many believe it minimally meets the criteria?

DR. WINKLER: One, two, three, four, five, six, seven, eight.

CO-CHAIR HOMER: Has that got everybody?

MEMBER PERSUAD: I'm a none.

DR. WINKLER: No, I am missing one.

CO-CHAIR HOMER: Okay. Not at all? All right.

The next one is the usability of the measure. Does everyone remember the criteria, the elements of usability?

So, again, understandable harmonization and added value. From a harmonization, just simply the point is there is an adult measures that is almost precisely the same. And understandable, I think we should view this again as a paired measure. That is, it is really reported as two different components of the measure rather
than a single item.

So, how many believe that it completely meets the criteria for usability?

None.

Believe it partially meets the criteria for usability?

DR. WINKLER: Six.

CO-CHAIR HOMER: And minimally meets the criteria for usability?

DR. WINKLER: One, two, three, four, five.

CO-CHAIR HOMER: And then not at all?

DR. WINKLER: One, two, three.

CO-CHAIR HOMER: Okay. All right. And then, feasibility, which is, again, data is a byproduct of care, electronic exclusions, inaccuracies, and implementation.

How many believe it is completely feasible?

One.

How many believe it is partially
feasible?

DR. WINKLER: One, two, three.

CO-CHAIR HOMER: How many would say minimally feasible?

DR. WINKLER: Nine.

CO-CHAIR HOMER: Okay. And not at all?

DR. WINKLER: One.

CO-CHAIR HOMER: Okay. Good.

All right. Then, why don't we move to an overall recommendation? I think this would a time-limited, given that the adult measure is time-limited, and with conditions that would relate to -- what conditions would we want to put on it? Do we need to?

DR. WINKLER: I don't remember any conditions.

CO-CHAIR HOMER: Well, do we want conditions related to --

MEMBER JENKINS: The two conditions I heard, one had to do with age
stratification and one had to do with
specification of the elements of the plan in
more detail.

CO-CHAIR HOMER: Again, this is
where, just as a comment, it is not that we
would be dictating what the plan is, but that
it needed to address those elements.

MEMBER RAO: And I thought age
stratification wasn't possible because of the
small number. That is what she said.

CO-CHAIR HOMER: We want to see
the data reported, I would suggest we would
like to at least potentially look at that. It
may be impossible.

MEMBER JENKINS: What I heard her
say -- maybe she could say what she said
instead of what I heard -- is she did not want
to exclude the babies, but that is different
than reporting the results by age
stratification or risk adjustment by age of
baby.

DR. FIVUSH: Yes, and I'm the
"she".

(Laughter.)

I'm sorry, it is Barbara Fivush.

I think that is a very good summary. Yes, I didn't want to exclude them because I didn't want to lose the importance of them, but was concerned about the numbers. We could report it out that way and see how it looks.

CO-CHAIR HOMER: Okay.

MEMBER FISHER: Can I ask --

CO-CHAIR HOMER: Yes, please.

MEMBER FISHER: There is no way for us to do what was suggested, is extend the age group under the adult endorsement?

DR. BURSTIN: It sounded like they said the plan of care was different.

MEMBER FISHER: Oh.

DR. FIVUS: The plan of care was different, and the other thing is we really have specified in our measure, our numerator -- and please tell me -- I know I
have already had opportunity to speak, and I
know you all have been working hard.

I just will quickly say the other
difference is we want this to be a single-pool
Kt/V, which means it is precisely measured at
a certain time after the dialysis session, as
opposed to the adults who are less concerned
about when they measure that Kt/V. That has
to do with body size in pediatric patients and
the way things may rebound.

So, those were the two things that
came up about harmonization. I think the
measures are very close, though. It is
possible that over time, if we get time-
limited data on this, we could really think
about harmonization. So, I don't want to say
that is not close with the issue of
harmonization when it came up earlier, but
harmonization can be very valuable, if we can
do that.

CO-CHAIR HOMER: And my
understanding, so I am just thinking of
advantages or disadvantages to having this an
extension in age group. Dialysis is covered
through Medicare on the CMS side. So, one
reason we sometimes would like to be under the
common element would be because we want CMS to
use this. But in this case, we know CMS is
paying increasing attention to the Medicaid,
and this would be consistent with their
longstanding emphasis on Medicare quality.
So, by having it a separate measure does not
decrease the likelihood that CMS would use
this.

DR. FIVUSH: And I would just
point out that these patients are Medicare-
eligible, but one of our problems is that
often their parents may have other insurers.
So, they are not necessarily covered by
Medicare, even though they could be covered by
Medicare. That really ends up making it
difficult for us to just enter a Medicare
database and see claims and reporting. That
is why this is a great opportunity for an
additional reporting system that we can perhaps see this information, with Medicare supporting the concept.

CO-CHAIR HOMER: Okay. So, again, I think the vote is for a time-limited endorsement with the conditions that Kathy so well articulated.

So, all those in favor of a time-limited endorsement with the conditions that were mentioned?

DR. WINKLER: Six.

CO-CHAIR HOMER: Okay. All those opposed to a conditional endorsement?

DR. WINKLER: Eight.

CO-CHAIR HOMER: Okay. I think the measure did not pass muster.

Anyone want to reconsider their votes?

(Laughter.)

No, that's fine. No. So, okay, the measure didn't go through as is.

I do want to thank the stewards
for presenting the measure, and I do look forward to -- well, I would encourage you, nonetheless, to continue to collect these kinds of data and bring it back.

DR. BURSTIN: Great. I just want to point out as well that we are planning a ESRD/CKD project starting in the late summer or early fall. So, if any of this input makes you want to think about a new submission, that would be a good time.

CO-CHAIR HOMER: Right.

DR. BURSTIN: With a committee filled with nephrologists who understand all this Kt/V stuff.

CO-CHAIR HOMER: I would love, also, to see an ongoing learning collaborative amongst these institutions that share these patients. Then, we could actually see whether you could refine further the issue of this plan. But that would be outside the scope of the current --

DR. FIVUSH: I want to thank you
for giving us the opportunity to present.

   It is a moving target. I think we are all trying to improve care, and we will just keep these measures. Thank you.

   CO-CHAIR HOMER: Thank you.

   MS. McELVEEN: Okay. We are going to go ahead and take a very brief break for lunch. If you could take maybe 10 to 15 minutes and get your food and come back, and we will have to reconvene.

   We are adjourning around three o'clock, and we have about six more measures to go through.

   (Whereupon, the foregoing matter went off the record at 12:37 p.m. for lunch and resumed at 1:02 p.m.)
1:02 p.m.

CO-CHAIR HOMER: While Marina is getting a little bit of food, I think we could probably get started.

I would like to ask, the measure that we are going to be addressing next is the validated family-centered survey questionnaire for parents' and patients' experiences during inpatient hospital stay, if I got that correct.

Nancy, are you okay?

She is still sitting upright, seems to be breathing. I just wanted to make sure you are okay.

MEMBER KIBORT: That is what I was asking about.

MEMBER FISHER: I got this horrible virus. I have had all my flu shots. Okay? Then, after it -- I hadn't had this happen to me since I was in medical school -- I got bronchitis with an asthmatic component.
Okay? Or some people say you have reactive airway disease.

(Laughter.)

And then, I am getting better, and something went down the wrong way, and then I kept coughing.

CO-CHAIR HOMER: It triggered the reactivity.

MEMBER FISHER: Yes.

CO-CHAIR HOMER: So, I wonder if I could ask the stewards from Children's to tell us about this measure, Boston Children's. That would be wonderful, the developer. That would be great. Not the steward, the developer, yes.

DR. ZINIEL: Okay. Does that work? I think so.

So, I am just going to give you a brief overview over the measure. We have high goals with this measure. We really hope that this survey becomes, so to speak, the pediatric H-CAHPS.
In the work I have done at Children's Hospital, also in collaboration with CHCA, I have seen the great heterogeneity in patient experience or patient satisfaction surveys, however you would like to call it. And I have also seen the quality of these surveys with regard to survey methodology principle. I was quite appalled as the survey methodology, what I have seen.

So, we basically did this project to really get a set of survey items that could be used like H-CAHPS as benchmarking across institutions, across departments, within the institution, for several dimensions of the care of patients.

Due to the third-party involved in pediatric settings, it is not really possible to just rephrase the H-CAHPS questionnaire. There are certain aspects that have to be taken into account. So, we have several dimensions that this instrument that we propose addresses.
There are experiences that parents report with regard to nurses, doctors, admissions, discharge, care coordination, medications, and there are, of course, a set of demographic items in order to be able to look at differences between ethnicity, et cetera.

For all items, reliability and validity data are available. So, we have test/retest reliability. We have predictive validity. We have validity for items within a certain domain. We have calculated Cronbach alpha to make sure that there are no redundant items in there to minimize the respondent burden.

We have validated, and I should say that these are items are a subset of a 120-item questionnaire that we selected due to their good performance with regard to missing data, validity, ceiling effects, and reliability.

The survey is validated for mail
and phone. We have also mode effects calculated. The reason why we were able to do that was because we had a very strict protocol when we started with this project. So, patients that were recruited were randomly assigned to either mail or phone mode. So, on average, we would really expect that the differences we observe are due to the mode and not to any other aspects of their care.

We also have really rich frame information. We have kept data, processed data, in order to be able to look at non-response bias. We have medical record data, so that we can stratify for different categories in complexity of care.

So, we can relate it to clinical outcomes. And what we are doing right now is that we are proposing within the framework of CHCA to field the survey at other institutions in order to use their data to get the survey down to about 30 questions.

We wanted to do this with other
institutions to make sure that the questions
that we select are really the ones that allow
good validity and reliability across national
institutions, and not just one hospital. So,
we basically used our hospital to get to the
62 items that really perform good in terms of
psychometric properties, and now going to go
and use other hospitals as well to sort of get
the survey shorter.

We also plan to have the survey in
other languages as well as an adolescent
version.

So, the sampling approach that we
proposed was a random sample of all patients
that were discharged within a certain time
period. It is, obviously, possible to
stratify for race and ethnicity.

We found, looking at the non-
response across the different modes, that it
is actually important to use a mixed-mode
approach for patient experience because
Hispanics and other minorities were
significantly more likely to answer the phone survey than the mail survey.

So, I think we have enough data to look at outcomes across race/ethnicity, if this was the first hospital stay for that child, if it was not the first hospital stay, if it was medical/surgical, how complex the procedure was.

So, based on the data in the survey as well as frame data, we can evaluate how the experiences of parents and patients differ across these dimensions.

**CO-CHAIR HOMER:** Can you describe the domains it mentions and the measures that derive from the survey?

**DR. ZINIEL:** So, we have not derived composite measures per se for the domains. Also, it is possible. So, the domains are experiences with nurses, experiences with doctors, experiences with regard to how they work together, if the parent felt that there was communication.
We asked about the admission process, about the discharge process, the care coordination after the discharge, medications during the hospital stay, as well as medications that were provided when the child or prescribed when the child was going home.

Then, we have about 12 items that are demographic of nature.

CO-CHAIR HOMER: Again, so there are composites that are calculated? It is done as an item-by-item reporting?

DR. ZINIEL: It is an item-by-item, but it is completely possible to calculate composite scores.

CO-CHAIR HOMER: Okay.

DR. ZINIEL: Because the scales are fairly similar. And there is, of course, I forgot to say, an overall rating. There is a section with a few overall ratings.

So, composite scores would be added, summative scores. The scales are usually from 1 to 5.
CO-CHAIR HOMER: And could you compare and contrast with the H-CAHPS, I mean realizing that H-CAHPS you would have to either alter the questions, so it would be your child rather than you, and things like that, but as you look at the structure of this compared to the H-CAHPS survey?

DR. ZINIEL: There are domains that are the same where questions are very similar. There are also domains that we realized are more significant for the care. So, for example, parents, with regard to how they rate their experiences at the hospital, are really, really -- or they feel it is very important with regard to the communication. So, the items that we have are more in number or higher in number than with regard to H-CAHPS just by the fact how predictive they were with regard to how the parent rates their experience in the hospital.

CO-CHAIR HOMER: I'm sorry. So, there are more items because there was a more
diverse number of issues?

DR. ZINIEL: Right. There are aspects, I think, in a pediatric setting that are important to consider with regard to the overall satisfaction. They were highly predictive of overall satisfaction, but the correlation among them was fairly low. So that we can assume that they measure different dimensions.

CO-CHAIR HOMER: Lee?

MEMBER PARTRIDGE: Could you just tell us a little bit more about the domain? You talked about care coordination after discharge. Is that care coordination between whom?

DR. ZINIEL: So, we have items in there that ask if they have seen their primary care physician right after they went home. I mean that's, I think, one of the -- we also ask about if they felt comfortable to go home with regard to the information they had, things like that.
MEMBER PARTRIDGE: In some of the work that we have done, focus groups with families across the country, the care coordination element turns out to be very, very important to them and a lot of the areas in which they feel it doesn't work very well.

DR. ZINIEL: That's correct.

MEMBER PARTRIDGE: So, you are going a little bit beyond the hospital here.

DR. ZINIEL: Right. So, the other thing that one of the comments mention sort of as a point was that we do not collect the data sort of during the hospital stay. The reason why we do not collect the data during the hospital stay is that we also want their experiences with regard to discharge and sort of right after discharge. That is the reason why we can't. I mean either we would then have two surveys, but then it is really hard to link them together and to get responses from the parent in both. So, that is why we are doing it after the child has left the
hospital.

CO-CHAIR HOMER: So, are there questions from the Work Group. We had started already, but other questions from the Work Group that reviewed this?

DR. WINKLER: I just have one question. Do we have a copy of the survey tool?

DR. ZINIEL: Yes. I submitted it.

CO-CHAIR HOMER: It was filed in the wrong -- no, maybe it was. Where was it?

DR. ZINIEL: Yes, we submitted the current survey tool when we submitted the measure.

MEMBER PARTRIDGE: Can I ask one more question?

DR. ZINIEL: Yes.

MEMBER PARTRIDGE: You are talking about developing an adolescent tool.

DR. ZINIEL: Yes.

MEMBER PARTRIDGE: And this is an issue that came up frequently for those of us
who were on the stakeholder group way back when H-CAHPS was being developed because we were concerned particularly about the teenager who was hospitalized, most often for maternity care, but also for other reasons, you know, like they skied downhill into a tree.

And we really wanted the adolescent patient assessment of care rather than the parents' assessment of care.

DR. ZINIEL: Yes.

MEMBER PARTRIDGE: And you don't have that subset yet. So, you are putting adolescents in here?

DR. ZINIEL: No. So, this survey will be for parents 18 years and older of their child. The reason why we did this is because we really wanted to develop an extra tool just for adolescents.

MEMBER PARTRIDGE: Right, but for the interim, if my teenaged child is discharged, you are going to ask me my opinion of the experience and not that teenager?
DR. ZINIEL: Oh, sorry, I misunderstood you. Yes. I mean we definitely could go down to maybe 15, 16 years. I wouldn't go down to like, not that I know how this happens, like 13 years, because from a scientific point of view we don't know enough about the response formation process in adolescence, and there is a lot of research to be done.

CO-CHAIR HOMER: So, just for the members of the Committee who maybe hadn't seen the survey, it was misfiled. It is under Work Group 1, Measure 27, and it is a PDF document. So, if you happen to have your flash drive, that is where the item is.

I am still, I guess, a little maybe -- your writeup, I guess more the scientific characteristics, the writeup says you describe things like Cronbach alpha and dimensions and things like that, but I am still asking the question of dimensions because, typically, with the CAHPS survey that
is typically what people report out. At least it used to be in the old days when I used to work with surveys.

DR. ZINIEL: I mean the dimensions are basically the headings in the survey. So, we have a report about 300 pages long that describes all of the results.

I was a little unclear how to sort of attach that, or I mean not attach that, but to describe that in the application. So, I am definitely happy to submit that one.

So, we have done factor analysis, et cetera, of the items that we had where we selected these 62 from. It is nursing, doctors, medications, admission, discharge.

CO-CHAIR HOMER: So, I think we would want to see that, yes.

MS. RAUSCHER: Just from a perspective of this tool, the possibility of reporting out by composite score was able to just --

DR. ZINIEL: Yes.
CO-CHAIR HOMER: Okay. Good.

So, Ellen?

MEMBER SCHWALENSTOCKER: I'm sorry, I'm going to share the microphone with you, even though I am sitting right next to you.

I guess two questions. One, you mention the importance of having both modes. Have you found a difference in response in mode influencing the response, whether it is phone or mail?

DR. ZINIEL: So, there are some differences with regard to distribution. So, at Children's Hospital, Boston, we have the problem that I think no national survey has, that like everyone is always super-satisfied. So you have like this ceiling effect, and it is really hard to track something over time if everyone is always satisfied.

So, we try to extend the scale in a way, based on focus groups, and during the survey what people actually reported, in order
to get sort of the differences.

In the telephone survey, which is known from a scientific point of view, people are more likely on average to rate it higher. However, the items that had significant differences, and I think there is only one item left in the set of 62. So, the reason why we started out with 120 was really to figure out what are the items that have high percentages of missing data, that have a great ceiling effect, where tracking change is hard, that have low test/retest reliability. And that is exactly why we excluded them.

So, another factor was, if the mode effect was very strong, we also considered the item to be excluded in order to minimize that exact problem.

MEMBER SCHWALENSTOCKER: Then, the only other question I had is it sounded like, I think you mentioned earlier, that you are hoping to reduce the number of items in the survey?
DR. ZINIEL: Yes. Yes.

MEMBER SCHWALENSTOCKER: So, kind of what is the plan going forward, the timeline for doing that?

DR. ZINIEL: So, we are right now talking with CHCA about how to set all of this up. We had talked with CHCA about a year ago, and there are a number of hospitals that are interested in fielding this survey to compare it to the current survey that they have. So, there is interest there.

I think the steps forward that have to be figured out is from a methodological point of view what I would really like is I would also get data that is at the same time collected using the current tool from the hospital as well as the scores. It would have to be randomly selected, what patient gets what tool, or what parent gets what tool. So, that we really can assess if there are differences across hospital with regard to the validity of items, how these
items fall within a dimension.

   So, just really we didn't do that
for the current version. Because of the
importance to really look at several
institutions and see if we want to use this
nationally, then we really should use items
that are applicable to all institutions and
not just to the Children's Hospital, Boston.

   MEMBER SCHWALENSTOCKER: Right.

   DR. ZINIEL: So, that is why we
felt, okay, we start out with 120. We get the
items out that perform badly from a
psychometric point of view and from a survey
methods point of view. Then, we basically go
national and say, okay, let's collect data;
let's collect data to compare it at the same
time. So that we can really make sure that
the ultimate tool with about 30 items, that
the measures that are in there are really the
ones that are applicable and good for every
institution, if I can say it like that.

   MEMBER DOCHERTY: That was sort of
my question, but your factor analysis, you
have done more limited factor analysis? And
you are going to do more later?

        DR. ZINIEL: Correct.

        MEMBER DOCHERTY: Okay. That
makes sense.

        DR. ZINIEL: Yes. I want to get
the data and figure out, you know, is there an
item that is really important? Or, based on
the current analysis, seems to be really
important for our situation, but that might
not be that important if I take other data
into account.

        CO-CHAIR HOMER: Allan?

        MEMBER LIEBERTHAL: You mentioned
the H-CAHPS before. Who is the owner of
H-CAHPS, and have you talked with them about
an H-CAHPS version that would be for children
and one for adolescents, so that non-
children's hospitals would be dealing with one
organization or one set of questionnaires?

        DR. ZINIEL: So, the measurement
owner or developer is AHRQ. We have not been in contact with them yet. The last thing that I have heard, based on their statement on their website, is that they are not currently working at a pediatric version.

I am not quite sure if behind the curtain, so to speak, there is something going on.

(Laughter.)

This survey will, nevertheless, be able to be used in hospitals that just have sort of a pediatric department and are not freestanding.

So, the way we phrased the question was that we really wanted to make sure that it would be applicable for all situations.

MEMBER LIEBERTHAL: Maybe the "not created here" wouldn't apply and AHRQ might welcome working with you.

CO-CHAIR HOMER: Bonnie?

MEMBER ZIMA: I probably have a
less interesting question than Marina.

(Laughter.)

But I was wondering in sort of your preliminary analyses whether you explored the impact of variable length of stay.

DR. ZINIEL: Yes. So, length of stay, we explored length of stay, medical/surgical, if this is the first time they are at the hospital or not.

So, generally, I mean it depends on the item, but overall I can say that people where this not the first hospital stay are overall less satisfied. The people who have like a longer length of stay are less satisfied. Minorities overall seem to be less satisfied, and surgical, no, medical are less satisfied as well.

MEMBER ZIMA: How did you think about the impact of the severity of the illness and the child's prognosis?

DR. ZINIEL: This is a really good question. The problem with surveys in general
is that they are measuring something that usually cannot be measured otherwise. So, from a provider perspective, we would hope that -- I mean they should be satisfied, no matter how they go through the hospital, no matter how long they stay, no matter how often they have to come back. The service that we provide should be satisfactory.

The other thing is it is always based on expectations. So, parents that, for example, have been in the hospital previously have other expectations than parents that have been there the first time.

So, there will always be a subjective, based on just the experience that you had, there will always be sort of an influence of expectations. That is what surveys basically measure.

It is really hard to sort of get people to set to an expectation. They come in with an expectation, and these expectations vary, but I think from a hospital point of
view, no matter what these expectations are, our goal is that parents have a good experience.

MEMBER ZIMA: I just have one more question. That was, with a response rate of 25-35 percent -- I know this is kind of generic question.

DR. ZINIEL: That is actually a comment.

MEMBER ZIMA: Oh, okay.

DR. ZINIEL: Yes, go ahead.

Sorry.

MEMBER ZIMA: How are you thinking about the selection bias? How do you avoid overrepresenting happy campers?

DR. ZINIEL: So, this is a common phenomenon in satisfaction surveys. The concerns are that happy campers and really, really unhappy campers do not answer.

So, what we found is that, on average, in this survey it equals out. So, it doesn't really affect the score. We can say
that because we have frame data. So, if you have frame data, you can actually adjust for it. So, you can use non-response weighting, which that is another part of this project, to calculate non-response rate to see how that affects, actually, the differences of the scores.

You need a really good protocol to make them participate. The unfortunate thing is that the survey climate nowadays, I mean everyone is completely oversurveyed. You really have to write a letter that convinces the participant or the parent to participate.

CO-CHAIR HOMER: Marina?

MEMBER ZIMA: Just one more issue, and that is you only have your English speakers, as you had said. So, I was wondering if you could speak a little bit more. Particularly something like this could not be used in California.

DR. ZINIEL: And that is where, basically, the plan is so we started out with
the English version. We plan to develop this into other languages. I mean that is one goal, to be able to use it as a sort, if I can frame it like that, pediatric H-CAHPS tools.

CO-CHAIR WEISS: Okay. We may not have caught some of the questions that are intended to get to this, but going back to Lee's point about care coordination, and particularly the handoff from the inpatient to the outpatient setting, this is a really sensitive area, and an area that gets a lot of attention from the consumer community. But I don't see, as Reva and I have been scrolling through your questions here, the questions appear to be more oriented toward parent satisfaction that they understood --

DR. ZINIEL: Right.

CO-CHAIR WEISS: -- something about medication, and so on, but not specifically toward the issue of did they get adequate instruction about what to do with the child once they left the hospital. How do
they hand off from the main, as you call them,
the main physician in the hospital to the
office-based practice? Is there a different
instrument that does that or do you just
presume that every child who leaves your
institution has a care plan, so that is not
even a question that should be asked?

DR. ZINIEL: No. So, if you are
sort of really interested in that domain, I
refer you to Jay Berry, who is actually
working on that right now.

(Laughter.)

So, I have the honor to work with
him on that as well with regard to the survey.

So, this is really an experience
survey. While I completely agree with you
that that might not be the case for every
child, what we are really trying to measure is
the satisfaction. If parents see that the
care, that there is something that they are
missing or it is difficult, and the left hand
doesn't know what the right is doing, they
will express that in dissatisfaction.

So, it is really to measure the subjective view of the process. So, patients can be really, really satisfied or parents can be really satisfied, even though the medical care itself might not have been optimal. But it is hard for a parent to judge that because the parent doesn't know the standards. So, this is really to get at the subjective opinions of the parents.

CO-CHAIR WEISS: Okay. I would just say that, particularly with parents who have children with chronic conditions, and who are in and out of the hospital on a regular basis, a part of satisfaction is going to be feeling confident that they know what to do once they leave the institution, who to call, where to go.

MS. RAUSCHER: That is a very important piece of the. This is more general. As Dr. Ziniel said, Dr. Berry is working on one for complex care.
I just wanted to add another bit of a detail about how we got into this process, which was that we always intended to do this, but about two years ago one of our payers came to the table and said, "You will do H-CAHPS for a p-for-p contract," a huge piece of it. We said to them, "There is no pediatric H-CAHPS."

So, that has been the impetus for this, of trying to develop something that could be used across the country, and would carefully reflect the domains specific, not just changing from you to your child, a lot of rigor into that measurement process. That is what we have been doing.

CO-CHAIR HOMER: Tom, did you first have a question? And then, Ellen and Faye.

MEMBER McINERNY: Yes. You know, our hospital has been doing the Press-Ganey surveys for years. Obviously, it crosses over to the children that get care in our hospital,
and I suspect other hospitals do that. I don't know whether you have taken a look at your survey versus the Press-Ganey survey. I suspect there may be some overlaps, and there may be a way of sort of trying to help form which of your questions are the ones that are most important, based on Press-Ganey as well.

DR. ZINIEL: Yes. I mean that is one reason why we planned the multi-center. So, we are not using Press-Ganey. It is really hard to get data from hospitals, you know, to basically say we would love to have your data to be able to analyze it with regard to patient satisfaction. That is one of sort of the conditions I would like to put on this sort of more national project, to say I really would like to see the data that you currently collect during the same timespan with your instrument, to be able to see how they actually correlate.

CO-CHAIR HOMER: Ellen, and then Faye.
MEMBER SCHWALENSTOCKER: Oops, sorry, I hope I didn't just turn someone's computer off.

So, I just wanted to make the point, first, I really want to applaud you for this work because I think it is a huge gap that we don't have a pediatric H-CAHPS. I think the survey, Marina, to your question, in my view, it goes beyond satisfaction. It includes parent reports on how well-prepared they were. So, it may be perceptions of care, but, in my view, it is more than satisfaction.

What I am struggling with a bit is it sounds like it is still being developed. I guess I need to understand a little bit from the NQF staff perspective, you know, what the implications of endorsing this are, given that you are looking to maybe reduce the number of items.

Then, kind of knowing a little bit about the history of H-CAHPS and the vendor involvement in that, I am struggling a bit
with, well, you know, nobody has stepped to do that, although there are instruments out there. I guess I am struggling with kind of what the path forward in terms of process needs to be, but I also feel like this is the first opportunity we have had to really look at a great step in the direction of developing a survey.

CO-CHAIR HOMER: Helen, do you want to respond to that? Then, we have Faye.

DR. BURSTIN: Sure. I will just respond briefly.

I mean, certainly, the group would have to decide if they feel like it is ready for primetime. That is sort of the issue. We do routinely get measures that get updated. We have a three-year maintenance policy. So, that if you made a significant change to the survey, you would have to bring it back to us for our re-review.

DR. ZINIEL: Yes, we know that.

DR. BURSTIN: So, that is fine.
DR. BURSTIN: I mean I don't see that as a problem. My major question was actually more about harmonization, and I know you can't harmonize completely with a CAHPS tool.

DR. ZINIEL: Right.

DR. BURSTIN: And I give my bias here as an adult-only doc, but a whole lot of these items look really similar to H-CAHPS.

DR. ZINIEL: Yes.

DR. BURSTIN: I am imagining myself in my old days that I used to run quality measure for a hospital. If I had to look at the H-CAHPS responses on some of these, and then look at these, the response categories aren't aligned. You have five; they have gone to three.

I am just trying to think about what a hospital who is not a freestanding children's hospital would have to sort of think through to make it work, if you had an adult survey. I mean we used to try to parse
it by adult surgery and OB. We didn't have kids at the Brigham.

But how do you imagine this kind of working in the real world, I guess?

DR. ZINIEL: So, the problem with the response category, where from a scientific point of view sort of what you would really like to measure is with a three-category scale, based on the ceilings effect that we just observed in our hospital, there is no way you would be able to really measure a change. I mean, if 85 percent are in the top category, how would you measure change? So, I mean, this tool is basically really to be able to measure change. Not that I was in the AHRQ group and want to criticize their work, but when I looked at H-CAHPS, I didn't understand why they have three. I mean three is really limited.

So, the problem that we have seen in the focus groups is it is really hard to get someone who is almost always satisfied to
completely satisfied.

DR. BURSTIN: I think some of this actually truly is the nature of adult care versus kids care. I mean I have seen H-CAHPS scores, and it is remarkable how much of a splay there is between those categories. It may just be that maybe kids truly -- Lisa Simpson always told me, "They're not just little adults." Maybe they are really different.

(Laughter.)

And maybe those parents have very different perspectives on their care. Your kid is sick; everything is great.

DR. ZINIEL: Right. I mean that is why we, for example, selected the five-point scale because with a three-point scale, I mean there would be no chance --

DR. BURSTIN: That is very helpful. Right. Maybe just some of those responses back formally, if we put this out --

DR. ZINIEL: I mean the other
thing is what I just don't know is on a national level, when we give this instrument to other hospitals, what the range is that is there.

With regard to what you report, I mean boards usually like to see the percentage where everyone is super-satisfied. From an improvement point of view, I want to see the percentage that has really problems because that is where you actually can do something about it.

DR. BURSTIN: I'm with you. What I will tell you, though, is H-CAHPS actually shows that, only because you would be amazed at how poor this is when patients report on their care.

DR. ZINIEL: Right.

DR. BURSTIN: It is not satisfaction. There is only one satisfaction item on CAHPS. It is really the very similar patient reports of care. "Did somebody explain your medications to you in a way you
"Can you understand?"  "Did somebody explain your discharge instructions?"

DR. ZINIEL: Yes.

DR. BURSTIN: That always/sometimes, those categories remarkably show lots of dissatisfaction.

CO-CHAIR HOMER: Because it is the percent always that -- and it is hard to get --

DR. ZINIEL: I mean, you know, based on the data that we have, I can tell you, I mean we have items that have like 85 percent always, very satisfied, extremely. I mean that's where we started developing the survey. So, how do you measure something if you have 85 percent?

DR. BURSTIN: I think you just justified it, but I think those are probably some of the explanations we would need when this would go forward.

DR. ZINIEL: Okay.

DR. BURSTIN: Otherwise, people
will look at this, particularly people who know CAHPS well, and --

MS. RAUSCHER: We also wanted to just share that we did do an assessment of the freestanding hospitals. One-third used Picker, one-third used Press-Ganey, and actually one-third have a hybrid, which makes it very interesting.

CO-CHAIR HOMER: So, there were some other --

MS. RAUSCHER: Or some other --

CO-CHAIR HOMER: I think, Faye, you were up next.

MEMBER GARY: I just have several quick questions. No. 1, how do you explain to the parent who the physician is or who the nurse is, and how are they going to use that as the base to make the decision about their satisfaction, especially in a teaching hospital?

DR. ZINIEL: So, we actually had items, well -- sorry. Go ahead.
MS. RAUSCHER: So, originally, what we did when we planned this out was hold focus groups. When we asked the question about satisfaction with your physician, the very first thing they asked was, "Which one?" because we are in an academic medical center.

So, the team put together a whole battery of test questions specific to three --

DR. ZINIEL: If it is a resident or if it is the attending.

MS. RAUSCHER: And now you can tell what the results were.

DR. ZINIEL: So, the results were interesting because, if it is the first hospital stay, about 80 percent, or I think it was 80 or 85 people could not tell the difference. So, it was like I don't know what the difference is. So, they said, like I didn't have a resident or a fellow. I mean this is a teaching hospital, like there is no child that goes through there that doesn't see someone who is in teaching. So, we knew that
they didn't -- it's all doctors, all white coats. However, in the people that are frequent flyers and are there more often, they can make the difference.

So, now, if I can sort of take a step back, that is why we didn't include it here, but we at Children's would like to go to a modular system to have this as a core and add on modular questions that like rotate throughout the year that will allow us to get to certain areas and have, for example, 10 questions. So, there will be surgery, ER, ICU. And one of these modules will be the question with regard to the difference of attendings and fellows.

So, kind of the criteria would be that it would not be a person who was staying there the first time because they can't -- it is really the people that have been there before know the difference; the other people don't.

MEMBER GARY: Well, yes, I think
that to determine the difference sometimes can
be quite a struggle for even seasoned
people --

DR. ZINIEL: Correct. Yes.

MEMBER GARY: -- in hospital

settings.

The same question could be also
related to nurses. What nurses are you
talking about? Because they have three shifts
or two shifts --

DR. ZINIEL: Yes.

MEMBER GARY: -- and many people
who provide many different services. How do
you differentiate them from the people who
come up to do the x-rays, to take the blood,
the respiratory therapists? Because in a care
mode, that is a lot to ask people to separate
and to understand conceptually what the
difference is.

DR. ZINIEL: That is correct. And

we had items like that in there, too, and they
have really high missing value rates because
people just cannot -- it is one of these things that I think surveys have to deal with because you can only ask questions where people know something about it. If they don't know the difference, there is no point in asking a question. And people really have difficulties. I think it is just sort of how compressed everything is, too. You know, they go from one department to the other. It is like they can't remember who was what and who had what title and what procedure they got.

MEMBER GARY: Absolutely.

MS. RAUSCHER: But we are not saying that that is not important. It was an "aha" moment for us.

DR. ZINIEL: It was an "aha" because we had this in there, and we asked them, how did these technicians do and those technicians. And I mean people sometimes, they had procedures and the parent would indicate they didn't.

MEMBER GARY: The other follow-up
question to that is this is about satisfaction. The way I am looking at it, it is about the child, big children, because they are over 13, so they are big children.

CO-CHAIR HOMER: No.

DR. ZINIEL: No. It is all children.

MEMBER GARY: You said you are not asking anyone who is younger than --

CO-CHAIR HOMER: No.

DR. ZINIEL: No. So, the question that came up here was if we should give these surveys to parents that are teenaged parents.

MEMBER GARY: Okay.

DR. ZINIEL: So, this survey is for parents of all ages of children that are at the hospital.

MEMBER GARY: Okay.

DR. ZINIEL: We would like to develop a version for adolescents for the patient itself.

MEMBER GARY: Yes.
DR. ZINIEL: But I think that this instrument is completely feasible for, for example, 15 years, a parent of 15 years and older.

MEMBER GARY: Yes. Okay. That clarifies one part of the question. But the other part of the question is that it seems to me in many ways for the older children, at least it is a proxy measure, and have you thought through what happens when the parent wants to participate and the child does not? Or is that a problem, that you ask them, the parent, about the child, and the child would prefer not to have parents respond on his or her behalf about the care?

DR. ZINIEL: So, I don't think that we have like clearly thought through, and I think this is a great opportunity for a scientific study with regard to proxy measures.

I have looked at other data from -- let me phrase it like that. There are
areas where the parent is much better as a reporter than the child. There are also areas where the child is a better reporter than the parent.

So, if I construct this sort of adolescent survey, I am pretty sure that one of the items for the adolescents would be, "Was I able to sleep in?", has a clear impact on satisfaction for an adolescent in the hospital, which we would consider as fairly unimportant in the grand scheme of things.

Like I have never encountered that an adolescent was not happy because the parent rated on their part. I mean the survey was never introduced that way. It was really, what were your experiences in the hospital?

MEMBER GARY: What was the parent's experiences in the hospital?

DR. ZINIEL: Correct.

MEMBER GARY: Not the child's experiences?

DR. ZINIEL: Correct.
MEMBER GARY: Okay.

CO-CHAIR HOMER: So, we could talk at great length about the survey.

(Laughter.)

Go ahead.

CO-CHAIR WEISS: Just a very quick question. Are you planning to make available to the public the results of the surveys on a regular basis?

DR. ZINIEL: Yes.

CO-CHAIR WEISS: And how do you do that? Do you do that in each of the question categories or is it just selected questions? Or how do you handle that?

DR. ZINIEL: I mean, theoretically, it is possible to display every question.

CO-CHAIR WEISS: But what have you done with CAHPS, for example?

DR. ZINIEL: Well, we don't have CAHPS.

CO-CHAIR WEISS: Well, you don't,
but have you discussed how you intend to make
the information available to the public?

DR. ZINIEL: So, I definitely
think that it would be on the web page. The
other thing --

MS. RAUSCHER: Excuse me. I think
it is just a little bit of a different
question. You are talking about, if I am
understanding you correctly, the question is,
how would this be available to everybody?

CO-CHAIR WEISS: Right. If I am a
parent considering your institution, and I
went on your website, would I be able to find
the answers to these questions?

CO-CHAIR HOMER: Or could you go
to any website and find out comparative data
on, should I go to --

CO-CHAIR WEISS: Across
institutions?

CO-CHAIR HOMER: -- Boston

Children's compared to --

MS. RAUSCHER: Our goal is
definitely to try to make this the pediatric H-CAHPS. At that point, it would be available in the public domain to whomever.

DR. ZINIEL: Yes.

MS. RAUSCHER: At that time, and your question about contacting AHRQ, it would also be about contacting the individual vendors who are going to basically be able to pick this up and move ahead with it.

DR. ZINIEL: Right.

MS. RAUSCHER: But from a perspective of maybe you could just share the experience that we have done with the children's hospitals based on whole system measures, which is our first step of taking a high-level measure and agreeing that we are going to look at it together.

DR. ZINIEL: Do you mean with regard to the differences and --

MS. RAUSCHER: Well, just the process of trying to get people to accept the measure.
DR. ZINIEL: Oh. So, CHCA has this initiative about whole system measures. I don't know if you know about it or not. So, there was a group formed about service excellence, and we had, I think, 15 representatives of hospital in there. We were trying to figure out what question to use to be able to compare across these 15 hospitals.

It was a rather difficult discussion because people do not want to change their measure because they always measured it that way, and like how could you compare it if you changed it? And just administering the two-service profile at the same time to be able -- how to sort of recalculate one score or the other didn't seem as a valid option for them, either.

The questions are sometimes very different. There are sometimes, if I might say from a scientific point of view, some are horrible. I mean, how likely or unlikely would you be to recommend this hospital to
families and friends? And the answer
categories are poor. Poor? Yes, I mean
hello.

(Laughter.)

Anyhow, so it was a real battle to
get 15 hospitals to agree to choose the
question, how satisfied or unsatisfied are you
with the quality of care at this hospital? We
discussed this for over a year.

MS. RAUSCHER: But it is being
trained.

DR. ZINIEL: It is being trained,
exactly.

MS. RAUSCHER: The thing rolls out
and it is finally accepted. So, that we
anticipate is going to be part of moving us
forward.

DR. ZINIEL: But we really hope, I
mean based on this experience, what we really
hope is that there will be a national measure
that everyone will use, and that really allows
us to compare across hospitals and states.
I mean right now it is really hard
because the questions are different, the modes
are different. There are not adjustments
recommended whatsoever.

CO-CHAIR HOMER: So, I am going to
take the Chair's prerogative here and first
tell a brief story, and then move this along.

So, the brief story is I was
involved, as you may know, and I guess it is
disclosure, in developing the previous
iteration of the Boston Children's Picker
hospital survey. My first presentation at
Children's Hospital in 1991 as a presenter
was, "Do you know who your child's doctor is?"
And the answer was, of course, no --
(laughter) -- very consistent with what you
were reporting. So, it is interesting how
some things change and some things don't,
because it is hard in a complex institution.

So, having said that, we do need
to sort of wrap this conversation and come to
a decision about where we are going to go with
This survey. We have the process we need to go through. So, the first would be we need a series of votes on these.

So, the first one is, is this concept or construct or measure sufficiently important for us to proceed? And I would like to call the vote.

All those who believe this is sufficiently important show your hands.

Good. Everybody.

So, everyone, they were all yeses?

Okay.

DR. WINKLER: They were all yeses except Tom.

CO-CHAIR HOMER: Who is crawling under the table.

(Laughter.)

MEMBER McINERNY: Sorry.

CO-CHAIR HOMER: So, the next question is scientific acceptability of the measure. To be honest, I would contend that we haven't, because the developers didn't know
how to send us the information, were concerned we would be overwhelmed by a 300-page document. I am a little concerned we don't have sufficient information to actually make that judgment.

So, I guess I need to call for a vote as to whether it is -- I guess, really, where I am going on this is, rather than proceed with the next series of votes on this, do we want to, again, request the developer to provide us some of that additional information?

MEMBER DOCHERTY: Charlie, I just have a question. You know, in methods measurement, with a new scale, we tend to accept indices that are slightly less than our older, well-established scale. Could that also be true for our assessment here, that we recognize that it is a very new scale, and it is under development, and that the author or the developer is willing to continue to provide us with reliability and validity in
development --

CO-CHAIR HOMER: The caveat is, as Helen pointed out earlier, if we endorse or recommend endorsement, that would be basically anybody else can pick this item up and use it, and that we would have some assumption of comparability across institutions.

DR. BURSTIN: Charlie, it might just be that we would ask you to actually submit that document, and perhaps just give a brief summary of the reliability and validity based on the statistics that are in there. And you could vote on it today, conditional upon approval of that plan, just so you don't have to get into yet another spinning game.

CO-CHAIR HOMER: Okay. So, should we proceed, then, with the different votes on the different elements, and then come back? Okay.

So, in terms of scientific acceptability, then, how many would feel this is completely meets the criteria for
scientific acceptability?

I see none.

How many feel this partially meets the criteria for scientific acceptability?

All right. So, then I will move to minimally meets the criteria. I guess we would say minimally. Okay.

Okay. Good.

So, then, the next one is usability, which is how interpretable are the results, as well as -- why can't I ever remember the other elements? How understandable they are, whether they are harmonized.

And again, we have got this issue of comparability with CAHPS and where it is and isn't, and the added value again. And there is no H-CAHPS for pediatric, but there are different scales and things like that.

So, how many would vote that it completely meets the criteria for usability?

Okay. How many would say it
partially meets the criteria for usability?

And how many believe it minimally meets the criteria for usability?

Has that got everybody?

Or not at all? Because we don't have any comparative data and things like that, and English only.

Okay. Good.

And then, for feasibility, again, data clearly are not a byproduct of care. This needs to be just a survey. But it is feasible, electronic, exclusions, potential for inaccuracies, and experience with or capability for widespread implementation.

MEMBER GARY: I wanted to ask one point before --

CO-CHAIR HOMER: Related to feasibility? Sure.

MEMBER GARY: I think it was Marina who asked about how this data might be used by consumers. But I wanted to also know, have you all thought through in your focus
groups, or whatever, how this data would be used at the hospitals among the providers to improve care? That is No. 1.

And No. 2, do you have a standard definition that you share with people who participate about what quality of care means? Because that, even providers, don't have any clear idea about the qualities. How do you grapple with that? Do you give us a scenario? Or how are you going to do that?

DR. ZINIEL: So, from a standardized interview point, my answer to your question would be whatever means to you. It is really hard to give definitions for a concept because, once you give a definition -- I mean, how complicated would that definition be? Would people understand it?

And with regard to scenarios, there is enough scientific evidence that scenarios actually bias the way you answer because people will only think about the scenarios you provide.
So, it is really what the parent encompasses in that quality of care for themselves, as subjective as satisfaction.

Your first question, can you repeat your first question? Or go ahead.

MEMBER GARY: I am just concerned that people in general without literacy issues, many, many people will not have an understanding for quality-of-care use. So, I am wondering if it is quality of care you are measuring, that one's own experience in terms of interactions with staff --

CO-CHAIR HOMER: I do think that test, that question has been subject to very extensive -- I mean there have been a lot of focus groups, there have been a lot of cognitive interviews across a variety of socioeconomic -- even though the term is very abstract, people are able to make judgments with this poor-to-exceptional or 1-to-10 scale around rating quality of care.

DR. ZINIEL: I mean, you know,
sort of my "rebuttal", quote/unquote, would be, if the providers can't decide what's quality of care, like how should we explain it to a parent? I mean, if you, you know --

CO-CHAIR HOMER: I think it is the first line, actually, of "Zen and the Art of Motorcycle Maintenance".

(Laughter.)

It's exactly about that term "quality".

MEMBER GARY: The other question is, how you are getting the agreements among the professionals --

DR. ZINIEL: Oh, right.

MEMBER GARY: -- to improve the care?

DR. ZINIEL: So, I mean, definitely, there is a long-term monitoring of how rates change. The other thing that I would personally like to see with this instrument is that there is a clear linkage to data with regard to department. So, that if
the percentage of people that say they really
had a problem, you know, sort of below the
standard, like poor to excellent, where people
say poor to average, that the department sort
of really has to address if that percentage,
for example, goes up.

I think that that tool is really
to monitor how it goes overall, and that if
this percentage increases or the percent of
satisfied/very satisfied drops, that that is
really the point where the department, or
whatever area it is that shows these changes,
has to start investigating what is going on.

CO-CHAIR HOMER: Thank you for the
question and the response.

I am going to go back to voting on
the feasibility question.

How many believe that this
completely meets the criteria for feasibility?

Again, the components of
feasibility are, they don't -- again, it is a
little challenging because the data is not big.
because it is a survey. But, basically, how feasible is this to implement? What is the burden, the hassle? How well-specified is it? How easily could this be picked up and done in a consistent manner?

So, how many believe this completely meets the criteria for feasibility?

We said that?

How many believe it partially meets the criteria for feasibility?

How many believe this minimally meets the criteria for feasibility?

Anyone in the not at all?

Okay. All right. Now I think again comes the question whether we move to endorse it or not. So, I think there are several options that we have on the table.

One is, as we did I think in one of the early ones, is not move that question, but, rather, recommend or request that we have additional data provided to the Committee.

Isn't that what you were basically suggesting?
DR. BURSTIN: Oh, no, no, no. You could just move it with conditions, if you would like.

CO-CHAIR HOMER: So, we could --

DR. BURSTIN: Conditions on the satisfactory analysis of the tome.

CO-CHAIR HOMER: Okay. So, we could do one of three things, but really one of two things, in my view.

One is not vote and request further information. Second is vote conditionally, and I would still say vote for time-limited because, again, this is only in English. This hasn't been applied across in one institution. We haven't been presented with domain scores or mechanisms really for reporting out.

So, we could either make it conditional -- we could either request more information or we could vote a conditional, time-limited endorsement. I think those are really the options.
Lee?

MEMBER PARTRIDGE: I think probably several of us are struggling with the problem that the work has been largely done in cooperation with children's hospitals.

I think I heard Allan say that --

CO-CHAIR HOMER: With just one children's hospital.

MS. RAUSCHER: Just one children's hospital.

MEMBER PARTRIDGE: But you also had conversations with other children's hospitals.

DR. ZINIEL: Correct.

CO-CHAIR HOMER: But that is, yes, only one question really, that satisfaction dimension.

MEMBER PARTRIDGE: Right. I think my basic dilemma is I want an H-CAHPS for pediatrics.

CO-CHAIR HOMER: Right.

MEMBER PARTRIDGE: What I don't
feel comfortable with, as we have it in front of us today, how well it would also work in Kaiser's hospitals or in community hospitals in more rural areas. I don't know whether we can get that in a reasonable period of time. I think it is difficult to get it through the NQF process without having a little better sense of how it would work outside the children's hospital.

MEMBER LIEBERTHAL: I didn't speak directly to Kaiser, but knowing how it works, I think that they would respond better to a pediatric questionnaire that was under the H-CAHPS title, which they already use, than a totally new questionnaire.

And also, the issue of similar rollups, so they could have some comparison of their pediatric services to their adult services, recognizing the differences.

And I am usually not one who advocates that children are small adults.

(Laughter.)
DR. BURSTIN: I did just email the CAHPS team. It helps that I spent seven years there. I can do that stuff.

So, they wrote me back. "The CAHPS team acknowledges the importance of this population, but they have limited resources to do it at this time."

So, I think the reality is it is not there now. There you go.

MEMBER PARTRIDGE: And we all are very aware of that. Therefore, you don't want to stifle progress.

DR. BURSTIN: And at some point, if a pediatric H-CAHPS came in, those measures could be harmonized or one would be determined to be best in class. But, at this point, there's not a competing measure on the table. There is a theoretical one on the table, but it doesn't exist.

CO-CHAIR HOMER: But I also know what was required of the CAHPS team to get through both NQCA approval and then NQF
approval, which was vastly more data --

DR. BURSTIN: The first time. It hasn't been that way since, and the first time -- I was at AHRQ at the time.

CO-CHAIR HOMER: Yes.

DR. BURSTIN: I mean H-CAHPS was more of a political battle than anything else. Getting it through NQF actually wasn't the problem.

CO-CHAIR HOMER: But I did sit, I sat, as you know, on the Hospital Review Committee, and there was a lot more data presented. So, I am just a little worried that we haven't yet seen comparative data. We don't have domain scores. But that is my own stick.

DR. BURSTIN: So, why don't you just defer the vote until you have set a time? That is fine.

CO-CHAIR HOMER: I am just speaking --

MEMBER JENKINS: I would just like
to say, if anyone here has any suggestions or advice for us, we would more than entertain them. Because we started by having the adult measure tried to be stuffed down on us, which is how we got this far, and you guys are seeing exactly how far we have gotten.

MEMBER LIEBERTHAL: Recognizing that H-CAHPS doesn't have the resources to start from scratch and write a pediatric questionnaire, they might welcome working with you and merge the two. It would require much reduced resources on their part and acknowledge the extensive work that you have done. So, I think that might be a compromise that would be more satisfactory to many people.

CO-CHAIR HOMER: But I don't want to put on the team -- they are not in control of whether the H-CAHPS people will work with them or not. So, that seems, it is a wonderful suggestion, but, you know --

DR. BURSTIN: But, Charlie, to
your point, I think based on what you just said that the discomfort is, I think we should actually wait and get the methods piece back to assess reliability and validity, and just vote on another conference call. I think doing it now would feel premature, it sounds like, for too many folks here.

CO-CHAIR HOMER: Well, it would be for me. I am just speaking more as an individual than as the Chair.

So, my motion would be that we request additional information specifically on the domain score issue and really I think even crisper specifications, then, tied to what the reporting would look like, and then bring that back for a vote for time-limited endorsement, based on that.

Then, I think the conditions probably at the time of the time-limited endorsement would be application across multiple institutions to look at the feasibility of use.
DR. BURSTIN: Right. Which is in the works. Yes.

CO-CHAIR HOMER: Then, that would be the criteria for coming back within 12 months, more or less, after we do the time-limited endorsement.

So, that is my motion, is that we defer a vote, pending additional information. Do I have general agreement? I am seeing a lot of heads shaking that we do that. Okay. Good.

All right. So, why don't we go forward with that?

DR. BURSTIN: And I am also happy to play the matchmaking role, if you would like, with AHRQ.

CO-CHAIR HOMER: Well, I guess I would be specifically interested in AHRQ, in where there are disparities in scales, in particular, like, again, you've got this different five-point scale than the standard CAHPS 10-point overall rating scale, which I
know was a very intensely-researched topic.

I would at least like to hear a little bit more -- this is probably my old survey researcher coming out -- on some of those items.

DR. ZINIEL: Yes, I am happy to do that.

CO-CHAIR HOMER: Good. Great.

Thank you very much. Wonderful discussion. Wonderful work.

Also, I guess, on behalf of the child health community, I express my appreciation to Boston Children's for investing in developing and moving this measure forward.

DR. ZINIEL: Thank you.

CO-CHAIR HOMER: That is great, as well as the other one.

MS. RAUSCHER: Thank you.

MS. McELVEEN: Okay. We have five measures left and about 50 minutes.

CO-CHAIR HOMER: So, 10 minutes
each.

    MS. McELVEEN: Yes, I think we all know that that's probably not going to happen. This is my suggestion: the five measures left are the individual metrics that are, again, part of this larger survey measure submitted by CAHMI. My suggestion is, either out of the five, if we could quickly look through them just based on maybe title description and the reviewers who looked at it and do a scope call. Because I know a lot of the other individual metrics we viewed them as out of scope for various reasons.

    And also, taking up the first measure to look through more in-depth, if it does get that far, is the one on measure of a medical home for children and adolescents, only because I think that one will probably have a little more discussion than the others.

    Is that okay with the group to do that first and go from there? Any objections? None. Okay.
So, the first out of the five is 41. And again, this is a brief, you know -- I'm sorry. This is Work Group 4.

Sorry. I apologize. I am just trying to get through these. I am probably talking, working faster than I should.

So, this is Group 4. Unfortunately, Tom had to leave early, but he did provide his feedback. So, that is probably what you see up on the screen, is mainly his comments and ratings.

But, first, Measure 41 is children who attend schools perceived as safe. This measure ascertains the perceived safety of the child's school. So, again, just looking at that description, and based on the reviewers who did look at the measure more in-depth, if we could kind of give a call as far as importance and scope, whether it fits within scope of the project.

MEMBER PARTRIDGE: I was on Work Group 4, and I was a negative on this one,
primarily because I had difficulty with evidence of relationship to child health.

We had a long discussion about this issue in the context of Measure 2 above, which is children who live in communities perceived as safe. So, I guess from my point of view right now, I would still consider this one out of the scope, but I am willing to be convinced.

MEMBER PERSUAD: I was on the subgroup that reviewed this. I would concur with that. This is a single item on a larger questionnaire. It is a single item. It is very general.

The only thing I could think of that would be of interest in this would be bullying, and I don't think that the statement on linked to outcomes was strong enough as is written here. I am not sure that it is not, actually, because it is a child's school.

But I would be fine if it is out of scope. It is a single item, and it is very
CO-CHAIR HOMER: Ellen?

MEMBER SCHWALENSTOCKER: I was not part of that Work Group. So, I hope I am not out of turn speaking.

CO-CHAIR HOMER: No, that is fine.

MEMBER SCHWALENSTOCKER: I agree with what has been said, but, then, it seems inconsistent to me that we would endorse the safe community and not endorse the safe school.

CO-CHAIR HOMER: The amenities was out, but we did -- the safety was in.

MEMBER PERSUAD: Which is one thing I was thinking about. I did want to ask us to review maybe a little bit of the discussion about the safety in neighborhoods. Right? That was a safety in neighborhoods. Communities, safety in communities.

CO-CHAIR HOMER: We felt that the experience of being safe in your community was an important stressor, health-related.
stressor. So, the question is whether we feel that the perception of safety in the school is --

MEMBER JENKINS: There is the link to physical functioning and obesity --

CO-CHAIR HOMER: Yes.

MEMBER JENKINS: -- because of safety --

MEMBER PERSUAD: In this document, the summary of evidence for linkage was that children who attend schools that are usually or always felt as safe are much more likely to be in better overall health than those who attend schools which are never safe, 85 percent to 59 percent. That is the only piece of evidence that we have in this document that they listed.

MEMBER KIBORT: Charlie, when you made the comment that for the communities it seemed to correlate, their sense of safety correlated with their health, but since the child spends so much time in school, wouldn't
it sort be the same logic?

CO-CHAIR HOMER: You could make that argument.

MEMBER FISHER: Why isn't the school part of your community?

MEMBER KIBORT: It is.

CO-CHAIR HOMER: So, Faye?

MEMBER GARY: One of the --

MR. STUMBO: This is Scott Stumbo on the call.

CO-CHAIR HOMER: Oh, good.

Faye, I'm sorry. Please.

MEMBER GARY: That is okay.

One of the struggles I had is that some schools are in very blighted neighborhoods, and children feel very unsafe when they are walking from home to the schools. When they get to the schools, they may feel relatively safe in the schools, but when they walk out on the sidewalk and head home, they don't feel safe. Lots of things happen between school and home.
MEMBER KIBORT: Faye, it may be the opposite, too, though, right?

MEMBER GARY: Yes, it might be the opposite.

MEMBER KIBORT: That the community is safe, but the school is not.

MEMBER GARY: Well, then, in some of the schools they have police in the schools and there are surveillance secret men in the school to make the children feel safe.

And those neighborhoods, I would suggest, are relatively unsafe. So, in my mind, it is hard for me to separate out a safe school and not a safe neighborhood, and no information about how a child feels safe in the neighborhood. Then, conflicts in the school just spill over in the neighborhood.

So, I am having difficulty with this.

CO-CHAIR HOMER: Scott, have they looked, have you looked at the correlation between these two items?
MR. STUMBO: They are highly
correlated, yes.

CO-CHAIR HOMER: Okay. Is there
additional information in one compared to the
other or they so highly correlated that they
are really no added value?

MR. STUMBO: That I don't know. I
am not sure.

CO-CHAIR HOMER: So, I guess for
consistency -- I mean I am not going to
revisit yesterday's vote -- for consistency's
sake, it is hard, I guess in light of this
conversation, for us to view this out of
scope. Is that an accurate -- so, it is
within scope. So, that means, how do you want
to deal with that? Do you want to go to the
other ones that we think maybe are out of
scope?

MS. McELVEEN: Yes.

CO-CHAIR HOMER: Then, we will
either come back to this one now or come back
to it in a conference call.
MS. McELVEEN: Yes. Okay.

Okay. The next one on the list is 42. It is children who receive the mental health care they need, and this is the percentage of children age 2 to 17 who have an ongoing condition which requires mental healthcare who actually have seen a mental healthcare professional in the past 12 months.

CO-CHAIR HOMER: So, I guess the question would be, is this a process measure or an outcome measure?

MEMBER PERSUAD: I was two minds about this. I think it is at face value is a process measure, but, as a general pediatrician, I know that that is such a critical early make-or-break, and that is really the thing that we are dealing with that I was trying to figure out if it moves over to being somewhat of a proxy measure.

CO-CHAIR HOMER: Bonnie?

MEMBER ZIMA: I had some concerns. But my initial impression was that there was
too much diversity on what a mental health professional was, and I could not link sort of whether the condition that child had connected with the right provider.

CO-CHAIR HOMER: Could we confine the responses to this issue of whether we consider it a process or -- because that is really going to be a question of whether we consider this in scope or out of scope, as opposed to the validity of the measure.

MEMBER LIEBERTHAL: Seeing a mental health professional may lead to better outcome or may not. So, I see it as a process measure. It is just one step on the path. It may lead to better outcome.

CO-CHAIR HOMER: Lee?

MEMBER FISHER: I felt it was partially important. I am perfectly happy ruling it out as a process measure.

CO-CHAIR HOMER: Okay. So, all that means is it just goes to our meeting in July, you know.
(Laughter.)

So, I think this one is considered a process.

MR. STUMBO: Can I ask a question?

CO-CHAIR HOMER: Yes.

MR. STUMBO: So, it is not purely based on an item saying did you or did you not see a mental health professional. There is an identified need. So, this I would think it would fall under the same category as any other unmet need for access to healthcare, and so in my mindset, makes it much more in the realm of outcome. If it was just did you or did you not see a mental health professional, but it is clearly based on a two-item measure, and based on the first item, the child has been identified as having an ongoing need, not just a crisis, but an ongoing need for mental health care.

And then, completely unrelated to a different part of the survey, it says, by the way, did you happen to see a mental health
professional? It does define what could be included as a mental health professional.

So, those who said no have somehow indicated earlier that their child did, indeed, have a need. That is the risk --

MEMBER ZIMA: Certainly in the title we see that there is a need. In the title it implies appropriateness.

CO-CHAIR HOMER: Certainly, that was a good point, but I think I am still in the category, I think, that it is a process.

So, the first question is, is there a need? To some extent, that is a health status indicator. That is an unmet need, it is a combination of a process and an outcome.

MEMBER PARTRIDGE: I assume if we deal with this later on as a process measure, we are not totally precluded from identifying some element of it as also an outcome measure. Just as we have talked to the payers here, I think we all feel there needs to be something
that addresses the question of assessing the unmet need.

    MEMBER FISHER: Of mental health problems.

    MEMBER PARTRIDGE: Of mental health, yes.

    MEMBER ZIMA: It mirrors a little bit of the discussion we had yesterday with Dr. Murphy around the pediatric symptom checklist.

    CO-CHAIR HOMER: But at least the pediatric symptom checklist was a direct indicator, at least intended to be a direct indicator of the health status, you know, whether you had symptomology that you had indicated you had --

    MEMBER FISHER: But this says that you have a diagnosis, you have a need, so you have a diagnosis.

    CO-CHAIR WEISS: It also says they received the care.

    MEMBER FISHER: And so, to me, if
you didn't receive the care, that is pretty bad. I just think that we are so used to taking mental health and taking it away from physical health, that we forget you get bad outcomes if you don't get to see -- if you don't see the cardiologist about your arrhythmia, you know, you get into problems.

So, I am thinking if you don't see a mental health person about your, let's say, bipolar disease, you can get into problems. Of course, it might be fun to go out and spend a lot of money, but you know what I am trying to say. If there is a problem there --

CO-CHAIR HOMER: Yes, Donna?

MEMBER PERSUAD: I guess one thing is, I notice in the measure specification it is children who have a mental health condition and saw a mental health professional in the last 12 months. I guess I didn't see that as meaning necessarily that they got the total cure for their condition or the level of care they needed. I saw it more as process on that
point. They just got there.

         MEMBER RAO: And there is the
other issue that I think a significant
proportion of mental illness is treated by
primary care physicians, too, and it is
ongoing as well.

         CO-CHAIR HOMER: So, are we,
again, are we deferring this, then, to our
process conversations in a couple of months?
Okay.

         CO-CHAIR WEISS: On the basis that
one would expect that most mental health care
is going to be given over time, I think
process makes -- I mean I think there are
elements of both, but I would put it in the
60/40 process basket.

         CO-CHAIR HOMER: So, let's defer
this until our next long meeting and
conversation. Good.

         MS. McELVEEN: Okay. The next one
up is No. 44, and this is children who have an
adequate insurance coverage for optimal
health. The measure is designed to ascertain whether or not current insurance program coverage is adequate for the child's health needs, whether the out-of-pocket expenses are reasonable, whether the child is limited or not in choice of doctors, and whether the benefits meet the children's healthcare needs. So, it is a lot of components.

CO-CHAIR WEISS: Let me just say that I think this one needs to be thought about in context of health reform, and health reform is going to phase in over time. So, I think we should be thinking about a broader timeframe than just the year 2010.

But I do believe that this particular set of questions needs to be measured. It just really important, particularly for people who will be getting their health insurance coverage through the exchanges and for adolescents and others who will be getting the stripped-down, Spartan healthcare plans. It is just very important
to monitor whether those plans have an adequate scope of coverage. So, whether outcome or process, this is super-important, in my mind.

CO-CHAIR HOMER: Lee?

MEMBER PARTRIDGE: I think you can tell by the scores up here that we had some unanimity on the importance of this measure. I would argue that it is not a process measure. I would argue it is an outcome measure because the flip side is you cannot have access to your healthcare in many instances if you don't have the capacity to pay for it. So, I put it in the outcome bucket comfortably. Probably on a slippery slope, but --

MEMBER FISHER: Also, because you have insurance doesn't mean you have access. I am just adding that --

MEMBER PARTRIDGE: I am well aware of that.

MEMBER FISHER: Okay. I am just
adding that to what you said. I am not arguing this.

MEMBER PARTRIDGE: You are talking about mental health?

MEMBER FISHER: Talking about any kind of insurance, health insurance.

CO-CHAIR HOMER: So, we think this is in scope. That seems to be -- I think this is definitely in scope.

So, all of our efforts to expedite this conversation are not really being very productive, but that's all right.

(Laughter.)

So, that is in scope.

And then, the last one?

MS. McELVEEN: That one is in scope.

CO-CHAIR HOMER: The measure --

MS. McELVEEN: Medical homes. So, the next one -- and we probably could take some time to discuss this one a little more in-depth -- is around the medical home. This
is the measure of the medical home for
children and adolescents.

It is basically a composite
measure that assesses whether children and
adolescents receive healthcare within their
medical care. This is according to the survey
respondent. Then, it looks like there the
measure is based on six of seven components
that are proposed by the American Academy of
Pediatrics, I think, it looks like defining
what the medical home is.

CO-CHAIR HOMER: So, again, I
think the key question is, is this in scope,
first. Is that what you wanted?

MS. McELVEEN: Yes.

CO-CHAIR HOMER: And then we are
going to have to prioritize which of these
various ones we are going to cover during our
discussion.

So, the first question is, do
people feel this is an outcome measure within
the scope of our deliberations?
DR. BURSTIN: I don't see it as out of scope, but I am not sure I see it as in scope. I mean, when we put it together at the start of the Outcomes Project, we explicitly put patient self-report on the list of outcomes. So, from that perspective, I think it is potentially in.

CO-CHAIR HOMER: Yes.

CO-CHAIR WEISS: Could I weigh-in with maybe a different way of looking at it? That is, it seems to me that currently it is an outcome because not every child has a medical home, and we are driving in that direction.

There will come a point in time where every child does have a medical home, at which point it becomes a process.

MEMBER FISHER: I like that, yes. And I think, also, it is that we have got to think differently about medical care, and that is the problem; this is different.

MEMBER JENKINS: Or maybe another
way of saying it is it is an intermediate outcome.

CO-CHAIR HOMER: Yes, I am just wondering about our consistency with our unmet mental health need and whether that would also fit the same criteria, but let's not go there. So, this is also, clearly, within scope.

MEMBER FISHER: Remember it when it comes up again.

CO-CHAIR HOMER: And then the last one for us to decide if it is in or out of scope would be Measure 50. So, then, who receives standardized developmental -- yes.

MEMBER LIEBERTHAL: Yes, that is clearly a process.

CO-CHAIR HOMER: Okay. We took care of two, one. So, I think the question is, which one do we want to pick up and which ones do we think that we have a reasonable likelihood of being able to complete within a half-hour
conversation? Do you think we can medical
home and insurance?

DR. BURSTIN: Go for it.

CO-CHAIR HOMER: Okay.

MS. McELVEEN: I won't object to

that, of course.

(Laughter.)

CO-CHAIR HOMER: So, let's do the

insurance one first. Since there was a lot of

enthusiasm that I saw around the room, maybe

that will be an expeditious measure.

CO-CHAIR WEISS: Shall we time it?

(Laughter.)

CO-CHAIR HOMER: So, the insurance

item is item 44.

Again, either does the steward,

the way we were doing it today, want to make

a brief, any brief introductory comments about

this item?

MR. STUMBO: Sure. Well, this is

one that we are particularly fond of.

This is a national survey, first
of all. This is being used by the Maternal
and Child Health Bureau for quite a while. It
is a relatively-new measure. It was
introduced into the 2007 survey, but we are
still getting publications out about it.

Like a commenter on your panel, we
believe very strongly that saying whether a
child has coverage or not is actually not the
whole picture. It is when you actually start
to dive a little deeper, we do find that, even
among children who are reporting or their
parents are reporting that they are in current
coverage, 15 or more percent, and it is
actually much worse among the private, are
stating that they do not have adequate
coverage, based on whether they have
unreasonable out-of-pocket expenses, not able
to see all the providers they need, and/or the
benefit does not talk to the child's needs.

So, we think that's a really
important story to tell.

CO-CHAIR HOMER: Great.
So, any other questions about importance issues on this particular question?

(No response.)

Okay. So, let's vote.

How many believe this is an important item sufficient to proceed?

DR. WINKLER: Eleven.

CO-CHAIR HOMER: Good. There were none opposed. Good.

All right. So, then, the scientific validity of the items, again, I think we reviewed the characteristics of the survey overall quite a bit.

Any comments on the testing, the questions themselves, the cognitive interview, the testings, and also, any assessment of these items and how they fit together, how well the algorithm works?

MEMBER LIEBERTHAL: The questions are very subjective. This is parents' perception of their insurance plan. As being subjective, it can be all over the place as to
what is adequate coverage. Again, somebody may perceive adequate coverage as no out-of-pocket expense; whereas, somebody else may be happy with some out-of-pocket expense. I don't know how to draw conclusions on that as to whether a child has adequate insurance.

It also depends so much on the family's inherent finances and socioeconomic status.

CO-CHAIR HOMER: Okay.

CO-CHAIR WEISS: Let me just say that that is a point that is debated and has been for many, many years. Five percent of adjusted gross income is one measure that has been used. Of course, you know, the Internal Revenue Service has used different measures. So, I don't know that we are going to be able to even come close to settling that issue. It is a subjective judgment.

MEMBER PERSUAD: I guess the measure steward can comment on this. They do recommend stratification based on
vulnerability. So, that may be a way to get at the issue of whether there's comparative relationship between what they would think is unreasonable and what they really have.

I would actually argue that what the parents' perception is of unreasonable is unreasonable. That is face validity to me.

MR. STUMBO: Right. In fact, this measure has so much face validity. Basically, it is an incredibly low bar. In fact, I think you guys were discussing the previous measure prior to our measure. There is an immense positive bias on all these questions. So, all they have to do is there are these three components: whether or not the out-of-pocket costs are unreasonable, whether the plan provides for everything the child needs, you know, that the benefits provide for the child's need, and it never, sometimes, usually, always -- this is a usually-and-always measure on all of them. So, all you have to do is fall into the sometimes or never
on one and you become adequate, which is
actually a relative low bar.

Of the three domains, it is the
out-of-pocket expenses which drive a bit of
it. To give a little flavor on the
stratification of the face validity, it is the
private insurance which is actually doing
much, much worse on the overall measure and on
that component. You know, public-insured kids
are actually doing better.

And when you stratify by income,
it is not related to income the way you might
talk it is. In fact, the people doing the
worse are the folks in the 200 to 400 percent
poverty rate, which often fall outside of that
SCHIP. And the lowest under poverty and the
highest 400 percent above are equal on whether
or not the insurance is adequate. So, it is
not being driven entirely by the out-of-pocket
expenses, but can be for the privately-insured
kids.

CO-CHAIR HOMER: So, Kathy,
please.

MEMBER JENKINS: I thought what you just said was that any of those needed to be usually or always, but the way it is written, it is a series of "and" statements. You actually have to meet all the criteria. Isn't that true?

MR. STUMBO: Yes, in order to have adequate coverage, you have to be usually or always in all three of the components. And nationally, without any stratification, 15 percent of kids are not usually or always meeting those criteria.

When I have talked to both the National Caucus of State Legislators, and we have brought these numbers for a lot of other folks, the preschool and regional, they can't believe that it is not significantly higher than that. Most people's personal experience is that they can't believe everyone says they don't have adequate coverage.

MEMBER ZIMA: This is probably a
fine point, but I am looking at the 2a.21
calculation algorithm, and it looks to me like
it is needed to see the healthcare provider.

How are you handling mental health?

MR. STUMBO: I'm sorry, I don't
actually have the form, submission form, in
front of me. Can you explain a little bit
further?

MEMBER ZIMA: Yes. It says,
"Current insurance offers benefits or covers
services that meet the child's needs. Current
insurance allows the child to see needed
healthcare providers."

Does healthcare provider include
mental health or not?

MR. STUMBO: Objective to parent
interpretation.

CO-CHAIR HOMER: I'm sorry, what
was the --

MR. STUMBO: If the coverage does
not cover mental health coverage and the
parent thinks it should, then maybe, like we
said, they were not happy with that.

CO-CHAIR HOMER: Okay. All right.

Do we have sufficient information to move on the scientific validity, scientific acceptability of the measure? I think so.

So, how many --

MEMBER JENKINS: Can I ask one more question?

CO-CHAIR HOMER: Yes, of course, Kathy.

MEMBER JENKINS: Could I ask you, then, what I heard you say is that there is this positive response bias, and that people are shocked that the measure does as well as it does. Is there a potential unintended consequence that the problem with the positive response bias could be misleading in the opposite direction?

MR. STUMBO: I'm not sure I could comment on that. We do find that, in general, parents tend to be positive on everything. How well are your kids doing? If anything --
I have three.

(Laughter.)

Yes, I was surprised on the positive bias myself.

MEMBER JENKINS: So, you are saying you are not worried about that, that families are inaccurate and that could be an unintended consequence? Because what you are really saying is you are not sure they are accurate.

MR. STUMBO: I cannot reach inside a parent's brain and understand if they are confused by the question or the world around them.

CO-CHAIR HOMER: But you said 85 percent of parents report that their child has adequate insurance, meaning they meet all five of those criteria, that it covers their needs and that they don't have too high out-of-pocket expenses, et cetera?

MR. STUMBO: Right.

CO-CHAIR HOMER: Okay. And that
includes people who have no insurance at all?

MR. STUMBO: No, it does not.

This is all children who have current
insurance, regardless of the type of
insurance.

MEMBER JENKINS: That is actually
a question. It is one of the five criteria.

CO-CHAIR HOMER: Yes.

MEMBER JENKINS: So, I assume that
if they said they don't have insurance, then
they are not excluded, I don't think.

MR. STUMBO: I'm not sure they
have that, but --

MEMBER JENKINS: You are just not
in the numerator.

MR. STUMBO: -- but it is just
children with insurance.

CO-CHAIR HOMER: Now I'm sorry,
are they in the -- I am still confused. I'm
sorry. Are they included or not included?

MR. STUMBO: They are not included
in the denominator. It is of children with
insurance --

CO-CHAIR HOMER: Okay.

MR. STUMBO: -- how many have adequate or inadequate --

MEMBER JENKINS: Well, for a child to be included in the numerator of having adequate insurance, criteria from the following five questions must be met. Child has current health insurance coverage.

CO-CHAIR HOMER: But, then, the denominator excludes -- so, if you have 5 percent uninsured in your community and you then have 15 percent who say they have inadequate insurance, so the total would be 20 percent if you were speaking to the legislature, how many children are, quote, "underinsured" in your community, you would say that would include the 15 plus 5. Right? Okay, that would be the way to interpret it. Okay.

Okay, scientific validity then or scientific acceptability, how many feel this
completely meets the criteria for scientific acceptability?

Pretty good, actually.

How many feel it partially meets the criteria?

Good. Did that catch everybody?

DR. WINKLER: No.

CO-CHAIR HOMER: No?

How many believe this is minimal?

Okay. All right. I think from the usability, which is how understandable this is as well as issues of harmonization with other measures, and whatever the third one is which -- added value.

So, any questions about this?

(No response.)

If not, we will move on to voting.

To what extent does it completely meet the criteria for usability? Yes?

How many believe it partially meets?

Okay. That is a lot of other
people.

DR. WINKLER: Okay, that is everybody.

CO-CHAIR HOMER: That's everybody.

Okay.

And then, feasibility as part of the national survey.

So, how many feel it completely meets?

DR. WINKLER: Nine.

CO-CHAIR HOMER: Okay. And how many feel it partially meets?

Good.

All right. Then, to move the question, recommend endorsement of this measure?

DR. WINKLER: That's everybody.

CO-CHAIR HOMER: Good. So, we got one.

Now we have 14 minutes to do medical home, which I think is going to be pretty hard because that is a very complicated
measure.

(Laughter.)

I don't think --

MEMBER PARTRIDGE: I wonder if I could raise an issue.

CO-CHAIR HOMER: I would rather not, actually.

MEMBER PARTRIDGE: Yes, I would like to raise an issue here that might make our discussions a little shorter. That is, and these were my comments on the measure.

CO-CHAIR HOMER: The medical home measure?

MEMBER PARTRIDGE: The medical home measure. Since this survey was developed and used in the field, the definition of medical home has become multiple definitions. I am not sure, therefore, that this question in quite this form with these characteristics is as timely today as it ought to be. I don't quite know how to deal with that on a procedural basis because CAHMI can't go back
and rewrite their survey.

But if you put this out as the standalone survey and a practice was graded very highly against this definition, it would not be consistent probably with the Minnesota definition of a medical home or health home. It might not be consistent with the definition that comes out of NCQA, PPC-PCMH, which the revisions will go public next week.

CO-CHAIR HOMER: But isn't that a harmonization question?

MEMBER PARTRIDGE: It is a harmonization question, and I don't know quite, from a procedural point of view, how to deal with it. I guess maybe we discuss it and deal with it in the harmonization context.

CO-CHAIR HOMER: I think we would have to discuss it in the harmonization context.

MEMBER PARTRIDGE: And maybe, since we are not going to get to it today --

CO-CHAIR HOMER: I just don't
I think, in fairness --

MEMBER PARTRIDGE: No.

CO-CHAIR HOMER: -- to the complexity of this measure --

MEMBER PARTRIDGE: Right. I wonder if our measure developer might want to look at that issue a little bit and give us any further guidance about how completely this really would be consistent. I don't know. I just worry about conflicting standards out there.

CO-CHAIR HOMER: There are really different definitions of what a medical home is.

MEMBER PARTRIDGE: There are quite different definitions, yes.

CO-CHAIR HOMER: So, the question might be how well this concept maps to the joint principles --

MEMBER PARTRIDGE: Yes.

CO-CHAIR HOMER: -- that have been adopted by the primary care associations.
MEMBER PARTRIDGE: Yes.

CO-CHAIR HOMER: As part of that presentation.

MR. STUMBO: We are submitting the measure because we would like to create a national standard based on the American Academy of Pediatrics. So, especially in regard to the question of, does Minnesota or Oregon or California, different communities' definitions -- we would actually say that that is the whole reason why the national survey was revised, to measure it across states in a systematic way.

CO-CHAIR HOMER: Well, I am quite sympathetic to that, but I do think we need to have a longer conversation. So, I don't think it is fair to do that in 10 minutes. I think I might actually suggest we adjourn 10 minutes early rather than rush through another one in the last --

DR. BURSTIN: Right, and it just might be helpful, if you are going to do this
measure on a subsequent phone call, perhaps
for the measure developer to specifically look
towards the updated medical home survey and
come back with some responses around
harmonization, so we are that much closer.

CO-CHAIR WEISS: Right, and as
long as we have the measure developer
listening, are there any other things, aside
from the issue that Lee raised, that we want
to put on the table right now for the measure
developer to think about?

MR. STUMBO: I'm sorry, was that a
question?

CO-CHAIR WEISS: To this group.

CO-CHAIR HOMER: I, for example,
would be interested in knowing from the
members who sat on the SNAC CHIPRA Committee
why this measure basically, which was
recommended by me to the Committee to be
adopted, why it was turned down, and whether
there is anything that the steward could do
that would help address any of the concerns
that Committee had. I realize that is a different process and it was using different criteria, but I don't know if there were any specific issues raised that would inform our further conversations.

CO-CHAIR WEISS: Well, I think one of the issues that colored all of the conversations had to do with how widely the concept is used in the Medicaid and the CHIP programs currently, and the ease with which states could move to universal application in those programs.

CO-CHAIR HOMER: So, the context of the recommendation here, by the way, for the medical home measure is not that these items would be used in a different context, but, really, again, this is more in the context of using the national survey in the way that we are using it on all the other. So, as a measure of population health, okay.

MEMBER PARTRIDGE: Yes, it is a population health measure.
CO-CHAIR HOMER: It is a population health measure. Okay.

MS. McELVEEN: Okay. Well, thank you all for plowing through as much as we could to get the day completed.

I can say we certainly accomplished a lot in the past few days. We definitely got through a lot of these measures. Many of them were different measures than what NQF is traditionally considered to be looking at. So, applause and hats off to you all for getting through that information.

And also, thank you to Charlie and Marina for leading the discussion over the past few days.

So, quickly, next steps: I did a quick count on the tabled measures. There's about seven of them, which in my mind and experience I don't think we can do that on one conference call, even if it is for two hours. So, just thinking out loud right now, I
I suspect we could possibly have two conference calls coming up.

Again, we are going to work really closely with the measure developers to try to really narrow down and get down the exact items and information that you all would need to inform your decision and to expedite the process, of course. But I just want to put that on your radar.

Again, we will be following up with a summary from this meeting and get your feedback on that to make sure we have captured your thoughts accurately.

Also, following up on your involvement and participation with the CHIPER project.

MEMBER JENKINS: What about scope, the unmet needs part?

MS. McELVEEN: I'm sorry.

MEMBER JENKINS: That part about the unmet, the gaps.

MS. McELVEEN: Oh, the gaps, yes,
that is a specific deliverable as part of this project. So, we will set aside time to identify gaps as well.

DR. WINKLER: We are having the same issue in other parts of the project, getting the measures done. So, a lot of what you talked about, we are capturing and we will probably start drafting some things for you to review and add to, and all of that, as we go along. But our first priority is getting through the measures.

DR. BURSTIN: But, as long as it is fresh in your mind, on your plane rides home, or whatever, feel free to write them down and send it to us.

DR. WINKLER: Yes, if you've got anything, yes, send them in.

DR. BURSTIN: We will start compiling them.

DR. WINKLER: Yes, compiling them.

MEMBER PARTRIDGE: Nicole, have you got any sense of the timeframe for the
conference calls and when you want to get this part completed before we start phase 2?

MS. McELVEEN: That is a good question. I was looking at my calendar. I think that, well, probably we will give the measure developers at least two weeks to get this information together and work with them.

So, looking at probably the last week of May, first week of June -- this is just, again, off the top of my head -- for a call, factoring in vacation time and then holidays and that sort of thing.

So, we will have to really talk about it internally because, on our timeline, we are trying to go out for comment in June. So, we will have to figure out the best way to adjust our timeline and, obviously, meet the needs of the project.

CO-CHAIR HOMER: I just wanted to express my appreciation to staff, to Nicole and Reva and your teams, for all the hard work that you did. The materials were excellently
presented. You did a superb job. On behalf of the Committee, I want to thank you. We couldn't have gotten as far as we have without all of your hard work and the excellent preparation. So, thank you.

(Applause.)

All right.

MEMBER PERSUAD: I have a housekeeping question. Where do our receipts go again?

(Laughter.)

MS. McELVEEN: The receipts are sent to Leslie Reader-Thompson. I can forward you her information, yes. The receipts, and I believe there is a form that has to be filled out for reimbursibles.

CO-CHAIR HOMER: Thank you very much.

(Whereupon, at 2:54 p.m., the proceedings in the above-entitled matter were adjourned.)
Neal R. Gross & Co., Inc.
202-234-4433
implies 318:8
importance 3:10
3:12 4:2,3,18,19
5:9,10,24 6:2,13
7:12 19:15 26:15
26:20 27:13 28:20
28:22 34:2,3,8
64:5,15 67:18
85:4 86:2 111:20
112:12,18 113:9
120:22 121:4
132:16 162:19
216:11 225:21,22
226:9,12 231:6
252:8 255:4 301:5
308:19 323:8
330:2
important 17:13
19:21 26:12 27:3
30:12 31:15 37:11
40:2 45:11,22
67:5 76:4,6 85:18
85:22 99:20
104:14 107:2
111:12,18 112:7
113:6 125:5,13
126:12 136:18
154:20 157:4
173:14 176:17
185:13,17 189:4,6
198:1 206:5 207:4
208:13,18 211:17
222:16,18 224:13
224:22 242:20
245:15 246:4
247:5 256:9,11,12
264:20 266:7
278:14 288:6,9
310:22 316:18
322:17,22 329:21
330:6
imposed 136:14
impossible 221:19
230:14
imprecise 148:7
impression 27:18
100:18 315:22
improvability 30:20 80:11
191:20
improve 78:20
80:18 142:14
206:5 236:3 293:3
295:15
improved 187:15
improvement
61:14 84:9 85:4
101:12 112:4
130:4 167:10,11
176:10 191:4,9,10
191:18 192:10
208:4,18 211:9
272:8
improvements 192:11
improvement-rel... 63:4
improving 188:1
inability 201:4
inaccuracies
107:18 145:18
228:18 292:13
inaccurate 337:7
inadequacy 208:16
inadequate 5:21
7:2,9 195:1 339:4
339:14
incidence 157:18
170:13 172:10
incidental 92:9
include 88:9 103:7
115:14 137:8
210:15 216:3,4,5
218:16,17 219:11
276:6 335:14
339:18
included 36:22
41:20 66:17 74:3
75:19,21 91:19
93:2,18 102:9
141:22 159:21
162:16 163:18
169:14 172:22
173:2 174:5 180:5
225:17,18 318:2
338:20,20,21
339:6
includes 84:20
123:21 158:13
181:15 185:19
267:10 338:1
including 55:7
105:3 113:13
117:17 166:21
income 331:14
333:11,12
inconsequential 113:5
inconsistent 310:9
incorporate 44:8
increase 21:10
50:17 124:8 176:3
176:18 188:16
204:10
increased 180:18
181:19
increases 296:9
increasing 132:22
218:22 219:1
233:7
incredible 68:2
96:3
incredibly 332:10
indefinite 150:2
independently 140:17
Indian 70:14
indicate 200:9
278:21
indicated 318:4
319:16
indicator 19:14
83:21 156:11
318:14 319:13,14
indicators 20:20
indices 289:16
individual 81:10,11
94:6,8 102:14,17
105:6,7 140:20
205:22 284:7
304:10 307:5,12
individuals 52:2
60:14
induction 184:8
185:13 186:10
indulge 170:22
infant 160:15
170:13 222:15
infants 64:2 75:1,9
156:20 157:6
176:14 177:9
179:12 180:13
222:12,21
infected 22:17
135:10
infection 21:9,20
21:22 22:7,7,10
22:12 23:3 50:18
71:20 79:4 135:4
138:1 155:7
171:19 192:4,12
201:11
infections 24:12
183:3
influence 83:5
101:4 259:17
influenced 101:2
influencing 252:10
inform 347:4 349:7
informant 40:5
information 16:2
46:18 62:2 79:10
86:22 150:9,13
170:5 184:15
198:1 213:7
221:13 234:2
241:11 246:21
283:2 289:1,4,12
298:11,20 304:12
305:8 313:15
314:4 336:3
348:13 349:6
351:7 352:14
infrequent 66:1
inherent 331:8
inherently 188:7
initial 16:15 35:22
212:22 315:22
initially 18:17
19:10 33:11 35:17
36:22 50:7 112:11
initiative 8:20
172:2 285:2
injections 92:6
injuries 158:5,6
167:1 174:15
175:1
injury 166:21
175:6
injury/birth 157:13
158:2
inpatient 6:9 52:22
62:22 237:10
262:9
inpatients 57:6
input 235:8
insert 148:17
insertion 21:1
inside 22:14 337:11
insignificant 111:10
instances 129:7
323:13
institute 43:1
institution 43:3
49:10 60:2 62:10
65:11 67:10,11
69:12,12,20,20
81:14,18,18 95:15
100:2 105:12,13
118:7,8 119:6,13
120:3 122:5
124:21 130:14
133:13 136:15
139:18 140:3,4
143:14,16 148:15
155:6 168:21
184:22 239:14
255:21 263:6
264:17 283:12
287:19 298:15
institutional 42:18
134:8
sequentially 64:15
series 288:3 289:9
334:5
serious 75:17 84:20
90:1
seriously 23:4 56:6
serve 83:2,4
served 186:17
serves 155:3
service 259:7 285:4
331:16
services 54:3 106:13
75:17 84:20
several 277:9 10,10
Sharron 214:15
share 69:21 75:16
sharing 114:2
setting 93:11
104:17 146:11
246:3 262:10
settings 155:18
239:17 277:6
settling 331:18
seven 207:21 227:4
301:2 325:8
348:19
severity 122:6
123:5,9,15 130:17
132:3 139:12
141:1 258:19
shaking 305:10
share 70:16 111:19
149:3 235:17
252:4 274:4
284:12 293:5
sharing 143:1
Sharron 2:4 10:16
shifting 106:19
shifts 277:9,10

shirts 126:2
shocked 336:14
shorter 242:9
342:10
shots 237:19
show 121:4 163:10
182:2 187:16
189:7 190:7
192:10 273:6
288:9
showed 68:21
75:16
showing 38:20
114:12
shown 134:16
shows 62:9 187:21
272:14 296:12
shunt 3:8 18:2
19:20,22 20:20
21:1,2 22:14,15
22:17,19,22 23:1
23:3 24:1 25:20
27:15,15 28:11,15
35:8,18 22:36:1
38:15 39:5,8,9
40:9 43:19 45:3
50:1 54:19 57:14
60:3 80:22
shunts 18:21 20:8
28:22 32:18 36:9
36:17 37:17,18
80:1,2
sick 56:18 219:8
271:14
side 83:20 160:13
160:14 168:7
181:10 233:3
323:11
sidewalk 312:20
signal 26:1 30:1
37:14 44:12,14
51:2
signals 176:14
significance 69:15
significant 31:18
32:4,5 40:21
71:16 93:3 95:2
114:22 117:15
153:16 157:17
161:1 175:4
178:12 245:11
253:5 268:18
321:3
significantly 51:1
68:7 180:12
181:17 243:1
334:18
similar 33:14 54:4
66:5 74:21 85:11
123:14 132:4
133:21 138:11
213:9 216:13
244:17 245:10
269:9 272:20
300:16
Simon 54:15
simple 200:19
220:9
simpler 220:15
simplicity 220:7
222:7
simplify 196:22
simply 94:17
214:15 227:17
Simpson 271:8
single 36:16 43:8
95:14 148:14
191:22 199:4
202:3 215:10
228:1 309:12,13
309:22
singleton 153:13
single-pool 194:21
194:22 198:12
199:8 232:4
sister 155:1
sit 302:10
site 47:17 209:6
219:2
sites 47:17 48:9
69:7 131:8,21
191:21,22 209:14
217:17
sitting 237:13
252:5
situation 35:4
256:11
situations 162:8
257:17
six 114:17 123:11
192:20 205:16
208:7 227:4 228:7
234:11 236:12
325:8
size 57:14 178:2,12
222:4 232:9
sizes 95:12
skied 249:6
sleep 281:8
slightly 289:16
slippery 137:9
323:15
slope 323:16
slopes 137:9
sloppy 56:7
small 30:11 31:16
32:3,6 41:19 50:9
50:20 69:17 78:11
78:14 114:19
164:4 165:22
166:15 173:19
186:4 188:15
196:10 204:5
206:4 222:8
230:10 300:21
smaller 111:3
SNAC 346:17
social 162:8
Society 54:6 97:6
socioeconomic 81:19 294:18
331:8
sodium 202:9
somebody 79:6
115:5,22 116:2
204:9 272:21
273:1 331:13
someone's 267:2
somewhat 33:14
315:19
Sonja 2:16 6:11
12:12
sonogram 165:7
sorry 11:10 14:6
48:1 59:2 84:11
91:2 93:16 132:10
133:9 138:21
195:4 231:3
245:21 250:1
252:4 260:12
267:2 274:22
288:18 308:3,4
312:12 335:5,18
338:18,20 346:12
349:19
sort 17:10 25:7
26:5,8 30:3 31:14
39:10 44:11 47:13
54:2 57:4 60:9
61:10 62:2 64:13
64:14 65:6 78:21
87:21 92:2 106:22
108:9 120:7 126:8
135:7 138:9
146:22 148:10
154:14 177:2
187:9 193:5 220:2
242:8 247:11,13
247:16 251:8
253:1 255:22
257:12 258:3
259:16,19 262:3
263:9 266:5,14,16
268:15 269:20
270:7 276:5 278:7
281:5 285:15
287:21 295:1
296:2,4 312:1
316:2 351:12
sorts 71:2,9 83:11
105:15
sounded 231:16
253:19
sounds 42:4 63:11
91:7 150:3 202:10
203:9 204:7
267:14 304:6
source 191:13
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30</td>
<td>1:19</td>
<td>8:3</td>
</tr>
<tr>
<td>8:31</td>
<td>8:2</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>79:14</td>
<td>106:16</td>
</tr>
<tr>
<td>80</td>
<td>182:16</td>
<td>275:15,16</td>
</tr>
<tr>
<td>800</td>
<td>196:12</td>
<td>204:2</td>
</tr>
<tr>
<td>800</td>
<td>220:13</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>133:18</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>4:3</td>
<td>66:19</td>
</tr>
<tr>
<td>85</td>
<td>103:12</td>
<td>270:12</td>
</tr>
<tr>
<td>85</td>
<td>273:12,16</td>
<td>275:16</td>
</tr>
<tr>
<td>85</td>
<td>311:14</td>
<td>337:15</td>
</tr>
<tr>
<td>86</td>
<td>4:4</td>
<td>89:14</td>
</tr>
<tr>
<td>87</td>
<td>89:4</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3:4</td>
<td>79:16</td>
</tr>
<tr>
<td>90</td>
<td>25:10,22</td>
<td>68:19</td>
</tr>
<tr>
<td>90</td>
<td>106:15</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>125:9</td>
<td></td>
</tr>
<tr>
<td>92.92</td>
<td>68:19</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>216:22</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>4:5,6</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>173:17</td>
<td></td>
</tr>
<tr>
<td>97.2</td>
<td>168:21</td>
<td></td>
</tr>
<tr>
<td>97.5</td>
<td>169:11</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>169:10</td>
<td>173:22</td>
</tr>
<tr>
<td>99.5</td>
<td>173:17</td>
<td></td>
</tr>
</tbody>
</table>