The Expert Panel met at the
National Quality Forum, 9th Floor Conference
Room, 1030 15th Street, N.W., Washington,
D.C., at 8:30 a.m., Stephen Pitts and Suzanne
Stone-Griffith, Co-Chairs, presiding.
PRESENT:

STEPHEN PITTS, Grady University, Co-Chair
SUZANNE STONE-GRIFFITH, Co-Chair
TERRY ADIRIM, HRSA

BRENDAN CARR, University of Pennsylvania Health System
EMILY CARRIER, Center for Studying Health System Change
GABRIEL EDWARD, Office of the Assistant Secretary Preparedness and Response
WES FIELDS, CEP America

DAVID LEVINE, University Health System Consortium
ANTHONY MACINTYRE, George Washington University Medical Center
DAVID MARCOZZI, ASPR
GREGG MARGOLIS, ASPR
LINDA MCCAIQ, CDC

MELISSA MCCARTHY, George Washington University
RYAN MUTTER, AHRQ
ANNMARIE PAPA, University of Pennsylvania
SALLY PHILLIPS, Department of Homeland Security
MICHAEL RAPP, Centers for Medicare and Medicaid Services (via telephone)
KATHY ROBINSON, National Association of State EMS Officials
JAY SCHUUR, American College of Emergency Physicians
MANISH SHAH, University of Rochester Medical Center
MIKE STOTO, Georgetown University
SHELLEY TIMMONS, Geisinger Medical Center (via telephone)
ARJUN VENKATASH, Yale University
ELLEN WEBER, University of California San Francisco Medical Center
NQF STAFF:

HELEN BURSTIN

ANGELA FRANKLIN

ANN HAMMERSMITH

ADEELA KHAN

JESSE PINES
C-O-N-T-E-N-T-S

Call to Order and Welcome 8
   Angela Franklin

   Senior Project Director
   NQF

Welcome 8
   Steven Pitts 8
   Co-Chair

   Suzanne Stone-Griffith 9
   Co-Chair

Introductions and Disclosures of Interest 10, 42
   Ann Hammersmith
   General Counsel
   NQF

Project Introduction 24
   Angela Franklin 24
   Senior Project Director
   NQF
   Jesse Pines 27
   Consultant

Overview of Proposed Paper and Framing Questions 30
   Helen Burstin
   Senior Vice President for Performance Measures
   NQF
Emergency Preparedness: Discussion of Concepts and Measures

Jesse Pines
Consultant

Measurement Issues

David Marcozzi

Questions and Comments

Mike Stoto

Questions and Comments

Reconciling Daily Crowding and Disaster Surge

Opportunity for Public Comment

Crowding and Boarding: Discussion of Concepts and Measures

Jesse Pines
Consultant

Questions and Comments

"Quality measure developers should ensure the validity and reliability of the data used for ED crowding and boarding measurement."

Questions and Comments
Crowding and Boarding: Discussion of Concepts and Measures (Continued)

"Developers should explicitly define the timestamp."

Questions and Comments 220
Risk Adjustments 240

Questions and Comments 243

"Quality measure developers should consider setting time-specific recommendations for unadjusted or adjusted measures of ED crowding and boarding."

Questions and Comments 268
Recommendation 7, the same topic, but with the added element of standardizing by triage acuity

Questions and Comments 288

Measures of central tendency, medians versus geometric means

Questions and Comments 297
Recommendation 9, structural quality measures for crowding

Questions and Comments 301
Recommendation 10, measures of ED outflow beyond boarding

Questions and Comments 315
C-O-N-T-E-N-T-S (CONTINUED)

Committee Discussion on Additional Recommendations 322

Regionalization and Accountability 343

Pathway from REMCS Concepts to REMCS-Based, NQF-Endorsed Performance Measures 371

Arjun Venkatash

Questions and Comments 374

Measure Development Discussion 425

Administrative Matters 431

Opportunity for Public Comment 432
MS. FRANKLIN: Good morning, everyone, and welcome to the Regionalized Emergency Medical Care Measure Prioritization Task Force.

We will start this morning to discuss prioritization of measures for the field.

Before we get started, let's just go ahead and welcome our Co-Chairs, Dr. Steven Pitts, Suzanne Stone-Griffith.

My name is Angela Franklin. I am Senior Director for the Project.

I have with me Adeela Khan, our Project Analyst and, also, Jesse Pines, our consultant on this project.

With that, I will turn it over to our Co-Chairs, who will start us off with introductions.

CO-CHAIR PITTS: Hello. Am I on?

Okay. Great.
This is Steve Pitts.

Welcome to the meeting.

By way of introduction, I am an ER doctor. Most of my career has been spent working clinically in the emergency department at Emory University in Atlanta. In the last five or ten years, I have gotten involved with statistics-type stuff and spent a year at the National Center for Health Statistics, and have just dipped my toe into this business of crowding in emergency medicine. I apologize in advance for not knowing lots of facts, but I am going to do my job as a policeman.

Thank you.

CO-CHAIR STONE-GRIFFITH: Good morning.

I am Suzanne Stone-Griffith. I am delighted to be here, a bit of a departure from other meetings that I have been a part of. I am really looking forward to this process.

I think with that, and in the
spirit of time, Ann, I am going to turn it over to you for full introductions and disclosures.

MS. HAMMERSMITH: Okay. Good morning, everyone.

I am Ann Hammersmith and NQF's General Counsel. I am here to guide you through the disclosures of interest.

As Suzanne said, we combine introductions with disclosures of interest, for the sake of time. It is a little bit easier for everybody.

Several months ago, you received a form from us, a rather lengthy form, which you filled out. Thank you for doing that. This morning what we are going to do is ask you to go around the table and disclose anything that you think is relevant that will be done by the Committee today.

Please do not recount your CV. We will be here for a very long time if you do that. We know that you are experts. That is
why we selected you to serve on the Committee.

What we are interested in you disclosing is any grant funding, but only if it is relevant to the work that is being done today with the Committee; research funding if it is relevant to the topic today; speaking if it is relevant to the topic today. Some of you will have nothing to disclose, which is perfectly fine. Just because you disclose it does not mean it is a conflict of interest. It is simply a disclosure.

I am going to remind you that you sit as an individual on this Committee. Sometimes we have members who perfectly innocently will say, "I'm So-and-So, and I am here representing the American Society of" fill in the blank. Actually, you are not here representing anybody. You are not representing your employer. You are not representing people who may have nominated you to serve on the Committee. You sit as an individual expert.
And then, finally, I am just going to remind you that things that you disclose or conflicts that members may have are not necessarily financial. People often say, "I don't have a financial conflict of interest," which is great. But because of the unique nature of the work we do here, you could have something to disclose -- it may or may not be a conflict -- where no money has changed hands.

For example, you served several years on a Committee that looked at topics relevant to what you are going to discuss today, that would be something we would like you to disclose to the ground.

So, with that, I am going to have you go around the table, tell us who you are, what your day job is when you are not here laboring for NQF, and then tell us if you have anything you would like to disclose.

So, I always pick on the Chairs first to start us off.
CO-CHAIR PITTS: Steve Pitts. I am currently working clinically at Emory University Hospital in the emergency department.

I do have some contract funding through ASPR, the Emergency Care Coordinating Committee, which is a small portion of my salary. Otherwise, I spent one year as a Fellow at the National Center for Health Statistics. Other than that, I am pretty much just a regular doctor.

Thank you.

CO-CHAIR STONE-GRIFFITH: Again, Suzanne Stone-Griffith. I am the Vice President of Emergency Services, EMS, and Trauma for HCA in the Continental Division. I am actually based out of Denver now, although I think it says Nashville in some documents.

I have served in previous times on the consensus panels for the emergency measures and ambulatory care measures. I have
also been part of the ENA Crowding Task Force.

I have worked with the emergency department of
Benchmark Alliance on some of the measure
definitions.

MEMBER WEBER: I am Ellen Weber.

I am a Professor of Emergency Medicine at the
University of California, San Francisco, where
I also work clinically.

The only grant funding I currently
have is a very small part as an expert
stakeholder for an AHRQ-funded grant that is
coming out of Stanford to talk about quality
indicators.

I have served with Suzanne at the
ED Benchmark Alliance. I am a member of the
SAEM Interest Group on Crowding in Emergency
Medicine.

And the other thing, I guess -- I
don't know if it is a conflict -- but I did
have a grant from the SAEM to study what they
did in England to solve their emergency
department crowding, and got that grant both
from SAEM as well as the Bupa Foundation. But that is now finished.

MEMBER STOTO: Good morning, everyone.

I am Mike Stoto, on the faculty at Georgetown University. I am also an adjunct faculty member at Harvard School of Public Health and do most of my research at Harvard. I am the co-PI there one of the CDC-funded Preparedness and Emergency Response Research Centers. There are nine of them. Two of the Centers focus on measurement as a theme, and ours is one of them.

A lot of the work that we have been doing was summarized in a White Paper that was cited in the background for this. I will be saying a little more about that later today.

The other thing is that I am also the Chair of what is called the Model Design Working Group for a group that is preparing the National Health Security Preparedness
Index, which is obviously related to this as well. That group is meeting again on Friday to try to come up with some specifications for that Preparedness Index.

MEMBER FIELDS: My name is Wes Fields. My day jobs are as a Director of the largest partnership of emergency medicine in the country and part-time clinical faculty at UC-Irvine.

I have spent a lot of time the last few years doing policy development and advocacy pieces with federal regulatory reform and the AC rule-writing in mind. I am more of a bundler to create new research projects, which means I am very popular with a lot of people in and out of the room who do research in this sector. And I am very happy to be here today.

MEMBER MARGOLIS: Good morning.

My name is Greg Margolis. I am the Director of the Division of Health Systems and Healthcare Policy in the Office of the
Assistant Secretary for Preparedness and Response at HHS. And I don't have any other conflicts to disclose.

MEMBER SHAH: Good morning.

My name is Manish Shah. I am Associate Professor of Emergency Medicine at the University of Rochester, where I work as an emergency physician. And also, I work as one of the County EMS Medical Directors.

Most of my work has revolved around pre-hospital care, particularly of older adults, and how to improve the care we deliver. I have AHRQ, CDC, and other grants, but nothing directly related to this.

MEMBER ASPLIN: Good morning.

My name is Brent Asplin. I usually have a voice, but I don't. This will be the best contribution so far to an NQF committee. I will just sign it in.

(Laughter.)

I am President of Fairview Medical Group in Minneapolis, part of Fairview, which
is an integrated system there, and emergency physician and Immediate Past Chair of a Quality Performance Committee. And I don't have any relevant conflicts to disclose.

MEMBER ADIRIM: Good morning.

My name is Terry Adirim. I am a pediatric emergency physician. In the past, I have been in academic medicine work in pre-hospital and EMS; currently, the Director of the Office of Special Affairs of the Health Resources and Services Administration. I represent HRSA on the NQF Board. I don't think that is a conflict, right? And that's it.

MEMBER PAPA: Good morning.

AnnMarie Papa. My day job is the Clinical Director of Emergency Nursing at the University of Pennsylvania and Penn Presbyterian, a medical center.

A couple of volunteer things: I am the Immediate Past President of the Emergency Nurses Association and have had the
opportunity to sit on the SAEM Regionalization
Task Force and a number of committees that
worked on crowding and throughput with ACEP
and ENA.

MEMBER ROBINSON: Good morning.

Kathy Robinson. I appreciate the

opportunity to participate today.

I am a Program Manager for the
National Association of State EMS Officials
and, also, a Past President of the Emergency
Nurses Association. I have had parallel
careers in EMS and emergency nursing.

With NASEMSO, we receive some
grant funding from HRSA and the National
Highway Traffic Safety Administration, but I
really don't have any relative interest
business to disclose.

MEMBER CARR: Good morning,
everybody.

I am Brendan Carr. I am sorry I
was late.

I am an emergency physician and a
policy researcher at the University of Pennsylvania, and I work part-time a day a week in Greg Margolis' office at ASPR.

Other than the fact that I have had research funding for conference work around regionalized emergency care systems and a lot of my research funding from AHRQ and CDC are tied to these issues, I don't think I have pertinent disclosures.

MEMBER VENKATASH: Hi, everyone.

My name is Arjun Venkatash. I am currently a Robert Wood Johnson Foundation clinical scholar at Yale, trained in emergency medicine at Harvard.

And the only conflicts I believe I have are that I have serve on the ACEP Clotting Performance Committee and last year received support from ACEP to do a mini-fellowship here at NQF. That included work on the previous phase of this project.

MEMBER SCHUUR: Jay Schuur from Brigham Young Women's Hospital. I am a
practicing emergency physician. I also do health services research.

And conflicts: I chair the Quality Performance Committee for the American College of Emergency Physicians, and I think that is it.

MEMBER LEVINE: Good morning.

I am David Levine. I am an emergency medicine physician. My day job is I am Vice President of Informatics and Medical Director at UHC, which is the University Health System Consortium, the academic medical membership organization.

My only conflicts, with UHC, we do benchmarking and performance. And so, obviously, emergency medicine is one of those sets of measures that we present to our membership, although we do not create the metrics.

And then, in my previous life as an emergency medicine department medical director, I served on a number of SAM interest
group committees that dealt with crowding.

MEMBER McCAIG: I am Linda McCaig with the National Center for Health Statistics. I work on the National Hospital Ambulatory Care Survey, and I have nothing to disclose.

MEMBER CARRIER: Hi. I am Emily Carrier. My clinical training is in emergency medicine. My current day job is as a researcher at the Center for Studying Health System Change.

Relevant to this project, I have current research funding from the CDC for a study of regional healthcare preparedness collaboratives and have had previous funding, mostly a number of foundation grants, primarily Robert Wood Johnson Foundation, for a large, ongoing qualitative study that touches on many of these issues.

MEMBER McCARTHY: Hi. I am Melissa McCarthy. I am a faculty member at George Washington University in health policy
and emergency medicine.

And I guess the only relevant conflict is my career development work, but, through AHRQ, it definitely touches on crowding and its impact on quality emergency care.

MEMBER MacINTYRE: Hi. Good morning.

I am Anthony MacIntyre. I am also an emergency physician at George Washington University. In fact, Jesse and I commiserate frequently about our inadequate EHR that has recently been implemented, but this isn't about EHRs.

(Laughter.)

My academic career is focused on emergency preparedness and response at all levels, facility, state, federal, and international. I guess I am not sure I really understand the term "regionalized" in this context.

But I guess the two things I have
to disclose that might be construed as
conflicts are: one, I was the coauthor of the
MSCC document, which HPP currently uses in its
funding grant cycles. And then, the other is
I recently stepped down, serving four years as
the Chair of the Emergency Management
Committee for D.C.'s Emergency Healthcare
Coalition.

MEMBER PHILLIPS: Good morning.

I am Sally Phillips. I am with
the Office of Health Affairs at the Department
of Homeland Security, Deputy Assistant
Secretary. I guess I would say in my previous
life I led the Public Health Emergency
Preparedness Research Portfolio at AHRQ in
support of HPP and ASPR, and nothing else to
disclose.

Thank you.

MS. HAMMERSMITH: Okay. I may
have some people on the phone. I am going to
call on you.

Is Mike Rapp on the phone?
(No response.)

Is Rebecca Katz on the phone?

(No response.)

No? Okay.

And then, a Committee member just walked in. Two. Okay.

We are doing introductions and disclosures. So, if you could tell us who you are, who you are with, and if you have anything relevant that you want to disclose to the Committee in terms of research funding, grant funding, speaking, et cetera. But only if has to do with the topics before the Committee.

MEMBER MARCOZZI: Dave Marcozzi from the Department of Health and Human Services, Assistant Secretary for Preparedness and Response. No disclosures.

MEMBER MUTTER: Ryan Mutter, Agency for Healthcare, Research, and Quality, HHS. No disclosures.

MS. HAMMERSMITH: Thank you for
making those disclosures.

Do you have any questions of me or anything that you want to discuss with each other based on the disclosures this morning?

(No response.)

Okay. Thank you. Have a good meeting.

MS. FRANKLIN: Thanks, Ann.

So, we will go quickly to the project scope and activities, just quickly. So, today I just want to highlight to everyone that we are focused on methodological issues and not endorsement of any particular measures. We are looking to lay a groundwork for development, testing, and endorsement and implementation of measures in this topic area. We want to review measures and measure concepts that are available in ED crowding, boarding, surge, and emergency preparedness areas, and any other areas that the panel will identify throughout today.

I want to note that we are going
to look for gaps and barriers to fully-
testable, implementable measures that could
pass the NQF criteria. And we are looking for
this group to provide recommendations for how
measures could be aggregated at the higher
levels, such as regionalized and by
geographical unit.

The purpose of our final report is
going to be to tie together our concepts of
crowding, preparedness regionalization, and
specifically to how we can report those
measures of quality at the regional level.
And we also want to help inform and give
guidance to the field as to the path to
measure development that could be brought to
NQF.

This project is funded by HHS, as
you have heard earlier.

With that, I think I will turn it
over to our Co-Chairs. Or Jesse is next. I'm
sorry. Jesse is next.

MR. PINES: Great. Thank you.
I think I know most everyone here.

For those who I don't know, very nice to meet you.

Jesse Pines. In terms of my background, I am an emergency physician and health services researcher at GW, today representing NQF on this project.

And I have had several grants looking at the association between crowding and quality of care from various organizations, federal and university funding and also foundation.

I would like to thank everyone for coming today. Essentially, our goal today is really to work on this report and really focus on seeing what we can do to really provide a guidance to measure developers who are interested in developing crowding measures and preparedness measures; also, seeing if we can bring together those fields.

We have experts from the crowding world and the preparedness world. Aside from
the calls that we had leading up to this
meeting, those worlds don't really come
together very much. Really, one of our goals
today is to really help reconcile the folks
who measure crowding and the folks who measure
preparedness in the NQF framework and, also,
think practically about how we basically get
from A to Z with measure development and
really provide some very concrete
recommendations to measure developers.

So, really, our goal today is,
again, to help us work on this report, come up
with specific recommendations, and basically
set up a runway for measure developers that
would, hopefully, happen in the coming years.
We don't have a specific plan. There is no
specific contract in place for measure
development in this field, but, essentially,
our hope is that, by setting up this runway
and by giving some real guidance to measure
developers, that we can figure out who the
players are, who are going to develop these
measures, and essentially what specific
guidance they would need, taking the
perspective of crowding and preparedness into
account.

So, with that, I wanted to maybe
turn it over to Helen. Did you want to talk
a little bit about the NQF process and the NQF
standards.

DR. BURSTIN: Sure. Good morning, everybody.

I am Helen Burstin. I am the
Senior Vice President for Performance Measures
at NQF. I have actually always, as a health
services researcher, always loved emergency
departments and have always enjoyed doing
research in those settings.

I am happy Sally could be here,
since I was at AHRQ, and she oversaw our work
on emergency preparedness when I was --

MEMBER PHILLIPS: She was the
mother of that.

DR. BURSTIN: I was the mother, I
guess. I don't know. I always felt like the child.

(Laughter.)

But, essentially, we would love to, we are really pleased to engage again in this topic area. I think that what we are trying to do in this project, unlike I think a couple of the ones that preceded it around more of an approach around regionalized emergency care services, is really be very definitive about what is the pathway toward saying, how do we get to a set of measures that would really allow us as a nation to measure issues around crowded and preparedness.

So, we thought it might be helpful just to give a little bit of a backdrop about how NQF evaluates measures, as you are beginning to think through what measures might look like.

Next, please.

So, why NQF endorsement? For
those of you who don't know, obviously, from
where we sit, an important piece of this is
the fact that, if you have standardized
performance measures, there are tools that
allow us to assess quality in a way that
allows us to have comparable information to
really be able to compare providers and
others.

As many of you know who have sat
through our panels -- and many of you have
-- the NQF endorsement is intended to really
reflect rigorous scientific and evidence-based
review, input from patients, families, a whole
wide range of stakeholders, and people really
across the entire industry.

Next.

So, these are our evaluation
criteria. I won't do a deep dive, as we often
do before committees start working, because
you don't have any measures before you today,
but we thought, again, just as a backdrop,
just to give you a sense of it.
So, NQF has always used the top four criteria, but over the years they have gotten more and more precise and, in fact, a higher bar in certainly the last five years that I have been at NQF. And they are also hierarchical.

So, the first one around importance to measure and report is a must-pass criterion. If it doesn't pass that first one, we just stop our assessment.

And probably the cornerstone of that one is the level of evidence for the measure focus. We really focus there in on the quality, the quantity, and the consistency of the evidence, and consistency tends to be very important for both guidelines as well as measures.

We also want to see if there is an opportunity for improvement. We don't want to be measuring things that are topped-out or things where variation is not going to be seen across providers.
And finally, we do anchor ourselves to the National Quality Strategy and other high-impact areas. We want to make sure we are really looking at an area that makes a difference. I often describe this to groups as, you know, is the juice worth the squeeze? It is a lot of work to get these measures. Is it really worth it? Will it drive improvement at the end of the day?

Scientific acceptability is the second one, and I have got another slide following up on validity, because it is such a major concern. But, essentially, are the measure specifications precise enough that comparisons are possible? Is there reliability and validity testing of the measures at either the data element level or the score level?

Usability and use was recently updated. The idea here is to really ensure that audiences who want to use those measures, whoever they may be -- and in this particular
project, it is a very wide lens of who may look at these measures, from the community folks to regional folks, to people in EDs, and others -- can they use those results for both accountability as well as performance improvement?

Feasibility. Can the measure be implemented without a lot of burden? Can you capture it increasingly, in this day and age, with electronic data and moving towards electronic health records?

And finally, not so much in this space because there are so few measures, but if you look at areas like cardiovascular care or diabetes, for example, a lot of our efforts now are really focusing-in on trying to select the superior measure among competing measures or at least harmonizing measures across different sites of care. This may be relevant in this field; for example, if we want to be able to cascade up and down, to have measures that work at an emergency department, but also
could roll up to give you more regional or local assessments of services. So, something we will talk about as we get further.

Next.

We also over the years have been doing additional work on evidence and have moved towards really a significant hierarchical preference for outcomes. Outcomes, particularly those linked to evidence-based processes and structure, is really the place we would most like to go. We want to ensure there is at least a plausible relationship to process and/or structure. And if we are going to have process measures, in particular, they have got to be the ones as close as possible to the outcomes. The ones that are really distal and so far away that you could measure them and measure them and measure them and never move the outcome is not where we want to be anymore. We have really started eliminating many of those measures from our portfolio.
Next.

I mention this just briefly. I will leave it up here in a bit more detail.

Again, we do require testing of the measures for reliability and validity. We allow that either to be done at the score level, which is often used for claims data or things we have just a lot of data, just to look for signal-to-noise, for example, or at the data element level. If there is one particular element you really want to be able to capture, can you test it and show you can reliably collect it?

The rest of it here, we won't go through.

Next.

Threats to validity is a major part of what our Committee spent a lot of time on. This just goes through some of them.

Again, conceptually, is it related to an important area of care or strongly linked to an outcome? Again, a measure that is unreliable can't be valid. So, that is a
starting point for us.

    We want to make sure that patients aren't inappropriately excluded from measurement. Some of the most complex patients, we always complained, are left out of research studies and sometimes they are left out of guidelines. And so, sometimes they are left out of measures. So, we really prefer that those approaches be stratified rather than excluded.

    We also want to, whenever appropriate, measures, if they are outcome, should be risk-adjusted. We increasingly are in this world of measure scores being generated with multiple data sources and methods. We want to make sure there is some comparability if people are going to be using different data sources. And we want to avoid systematic missing or incorrect data.

Next.

Usability and use, I mentioned briefly. Just two quick points on this. The
last bullet there is we have updated this.

So, we look at it to see whether the measures have actually made a difference, not just if they are being used. Are they driving improvement?

And here, the idea is are we actually making progress towards improvement, but also is there any evidence of unintended consequences. Many of you in the ED space, for example, lived through the pneumonia measure, antibiotics within four hours, which I think we heard pretty clear indications from the field -- that was just when I came to NQF -- that this was actually causing harm in emergency departments. We quickly did what we call an ad hoc review, re-reviewed the measure. The measure was changed.

But we really want to get a handle as much as possible, as you think through these measures, of what is really important to measure going forward. It is also important to think about, if we systematically measure
that and it may be adopted for an accountability application, are there any likely responses that may result in unintended consequences, I think is something we want to just make the case of.

Next.

We oftentimes talk about measurement, measurement, measurement. We are NQF. But, again, the end goal here is improvement in either healthcare, provider-based healthcare, or population health, obviously, given the focus today.

This was some work initially Don Berwick had done, recently updated. Really just making the case we understand measurement has lots of uses for both improvement but also selection. It is important to remember that we want to try to get to a set of measures that can both drive improvement, but also be useful for accountability.

Next.

Feasibility we have talked about a
bit. Again, try to get the data elements in a way that is easily retrievable or collected as part of routine care. I have heard about Jesse's pain points on the EHR and EDs. So, I won't go there. But, again, is there a way to capture some of these data moving forward?

Next, and I think probably last.

This was some work, RWJ's project on Aligning Forces for Quality it put forward. I thought it was useful for today, in particular. Even if we begin thinking about eMeasures, the reality is there are so many different sources of data that we are going to want to pull into these that go way beyond what you are going to get just out of the EHR on your desk. Just a reminder for us.

Next.

And I always end with this slide because I think it is important for us to remember that we are in sort of a difficult place at the moment of lots of things we really want to measure that we can't quite
measure yet, but we also know we can't improve what we don't measure.

So, with that, I will stop and turn it back over to Jesse.

MR. PINES: Great. Thanks so much, Helen.

So, any questions for Helen?

MS. FRANKLIN: Dr. Gabriel, did you want to introduce yourself quickly?

MEMBER GABRIEL: I am, fortunately or unfortunately, not a physician, but Ed Gabriel, the Principal Deputy Assistant Secretary from ASPR. I am glad to be here and participating in the group.

MS. FRANKLIN: And I just have one more question.

Arnika, are you there?

THE OPERATOR: Yes, I am here.

MS. FRANKLIN: I wanted to check to see if we had a Michael Rapp or a Rebecca Katz on the line, and if their lines could be opened, if so.
THE OPERATOR: Okay. Not at this time.

MS. FRANKLIN: Okay. Great. Thanks.

MR. PINES: Great. Thanks so much.

So, essentially, what I wanted to do is get the discussion started this morning. We have people from a lot of different backgrounds, a lot of different areas of expertise.

One of the issues that I think came up on our earlier conference calls was sort of making sure that everyone was on the same page in terms of understanding what we were trying to do with preparedness measurement and crowding measurement.

So, essentially, one of the things we had talked about was having two really short presentations this morning from Dave Marcozzi and, also, from Mike Stoto to give us an overview of some of their work on
preparedness measurement. Really, our hope is in the next half hour or so to really get on the same page in terms of the goals of preparedness measurement, some of the major issues, so we can start getting into the major meat of this, which is going to be specific recommendations for measure developers.

So, at this time I am going to go ahead and turn it over to Dr. Marcozzi.

MEMBER MARCOZZI: Thanks, Jesse.

I appreciate it.

I first just want to recognize two staff members in the back who are part of ASPR, Peggy Sparr, who is actually in charge of evaluation for ASPR, and Dr. Rick Hunt, newly brought on from CDC to ASPR. Certainly, their expertise with regard to this discussion, we would solicit their advice if we break up to ask them any questions that further the discussion that I am going to present here today.

Let's just kind of couch where we
were with preparedness before 2012 and where we are heading for preparedness for the next five years.

As a result of a Presidential Directive -- and that was Presidential Directive 8 -- there was a shift from a planning-based scenario to a capability-based scenario or capability-based planning. That was an important change. You can't make plans for every type of event. There was recognition within the Administration that you have to establish some core capabilities, apply those capabilities to, hopefully, any event, and have an 80 to 90 percent answer. Certainly, there has to be some vectoring right or left for a chemical event versus a large-scale biological event versus a pandemic. Those are all different types of specific events. But the response and the preparedness response activities are some foundational core capabilities that we can project to any type of those events.
And that was a key shift from where we were before to where we are now. To that end, the release of the healthcare preparedness capabilities which are in line with the public health capabilities -- so, CDC in 2011 released 15 public health capabilities, and then we, subsequently -- ASPR, "we" -- but this was a consensus-driven capabilities document that was released in the beginning of this year that looked and spoke to eight specific capabilities.

I think we will have probably significant interest in two of the capabilities. And then, I am going to jump over to what we are going to discuss here today. It is in and around performance measures.

The eight capabilities are through coalition development, emergency operations center, the ability to mobilize volunteers, so a volunteer capability, a fatality-management capability, and what I think will have
importance here will be the medical surge capability.

So, there are eight capabilities, and I certainly didn't list them all, that we, then, had to think about how are we planning on measuring. One of the challenges we face in preparedness is, when we stand in front of the press or stand in front of the Hill with regard to testimony, is: are we better prepared than we were before?

In an effort to try to establish some sort of markers and marks on the wall with regard to preparedness, there were measures put forth for each capability. The first capability is community or coalition development. And previously, our measures had spoken to, do you have plans in place? Do you have people, and you count noses, who come to the table?

But, unfortunately, we found that that really does not establish a true performance measure because counting how many
people are at the table or counting how people
were involved in the exercise does not
necessarily translate into better
preparedness.

What we found was that, actually,
established some better preparedness were
formalized coalitions that allowed for IAAs,
MOUs, charters, business development plans, in
conjunction with different healthcare
entities. Let me just speak to healthcare
entities for one moment.

When I speak to coalitions and
healthcare entities, the Hospital Preparedness
Program is misnamed. The Hospital
Preparedness Program provides monies to
awardees, then gives monies to coalitions.
Coalitions are subsequently defined as
hospitals, long-term care, primary care, EMS,
emergency management, public health.

All of that nexus, that core
group, is what we define as a healthcare
coalition. And certainly, within this
document, there are others defined. So, now you understand where we are with regard to coalitions and the development of coalitions. And there are some synergies certainly within the Affordable Care Act and the Accountable Care Organizations that are standing up. We are looking at how to blend our efforts with regard to what the ACOs are going to be doing, and looking to leverage some of the efforts with regard to that work.

However, in its simplest form, we are trying to put money forth to have partners come together to work better to effect a response. And when we describe "partners," we are only not describing hospitals; we are describing other healthcare partners, truthfully, a health community, to come to bring to bear to effect a response.

So, that is the unit of measure. The next step, and I think where we have interest in this discussion, is really a
paradigm shift for what we describe as medical
surge today.

Classically, medical surge has the
-- pick a percentage -- 20 percent, it is
approximately 20 percent above whatever the
typical capacity is of what previously were
hospitals. And we have changed that paradigm.
We have done that for specific reasons.

The first is we had to put a
performance measure out there that allowed it
to be independent of an evolving healthcare
system. As our healthcare system changes and
adapts and becomes more modern and evolves,
the measure we put out there we hope lasts for
the next five years. That is the grant cycle,
and that is the high-water mark on the wall
that our awardees look to as to actually
establish success and define success.

The second was we had to allow
these capabilities, we had to approach it from
a sustainable model. And this is irrespective
of finances. This echos the door-to-balloon
time of 90 minutes, in essence. It is irre
respective of size. It is irrespective of s
cope. It is irrespective of capabilities t
that you have. You know your end goal. Y
know your deliverable, and that is what the e
expectations are for the Hospital Preparedness P
Program.

So, let me just speak to what that p
performance measure is and how the paradigm h
has shifted from the 20 percent above on a s
system of healthcare that is trying to get l
leaner and meaner every day with just-in-time s
supply chains and staffing that is trying to j
just right-size-fit the number of patients we h
have within facilities. So, there is no new s
staff that are waiting to receive patients. T
There is no new space that is awaiting p
patients to just be received. And a just-in-t
ime, as-lean-as-it-can-get healthcare d
delivery system, which is what the Hospital P
Preparedness Program stands on, this p
performance measure integrates within that.
So, what we are describing is the ability, and it is evidence-based. And I am forgetting that part. It is evidenced, and it is operationally-tenable, this new performance measure.

We call it IBA, Immediate Bed Availability. But, truthfully, it is immediate care availability. It is the ability to accept 20-percent higher acuity patients within your facility within four hours.

Now that is irrespective of the disaster. That is all-comers. So, the MI that just hit the door, vice, the explosive event that just occurred. Notice I did not caveat that. It is all-comers presenting to your facility -- pardon me -- to your coalition. That is the unit of measure. You have to have the ability to accept 20-percent higher acuity patients within four hours. We build it into the system. The tagline is: this is medical surge with no new staff, no
new stuff, and no new space.

The evidence base on which we stand, Gab Kelen out of Hopkins, a 2006 Lancet article talked about reverse triage and talked about the ability of our healthcare system to be able to accept higher-acuity patients with no adverse outcomes.

So, standing on that evidence base, we then shifted to an operational construct. What is the average discharge rate currently within our healthcare systems today? Our average length of stay is approximately 4.9 days, plus or minus. You can figure that out.

So, we average slightly less than approximately 20-percent discharge per day; again, slightly less than that because we know that there are more. But if we have a 4.9 length of stay, then we can kind of start to think that this is an operationally-tenable goal. And then, we had to put a mark on the wall within four hours.
Without four hours is a large stretch, and we think about three pillars to be able to establish -- our coalitions need to be able to establish this performance measure. The first is the ability to throughout the time assess acuity. Now that could be done from an evidence-based standpoint or the truth is, when the internist on the floor writes "Out of bed ad lib" or "Tolerate PO ad lib," that is a potential triage, surrogate triage marker, because those patients or at least that internist is assessing that they have the ability to walk around on their own and they have the ability to tolerate PO.

So, whether or not this has an evidence base -- and we can certainly cite different types of triage methodologies, and Dr. Hunt could speak to this much more than I -- or operationally and with a logic model behind it, but it is the ability to assess through time the acuity throughout their healthcare coalition.
Second, the second pillar of this is the ability to rapidly offload patients. So, just as we would do onsite in a disaster, we would ask, "Anybody who can stand up and hear my voice, please move over to Mr. Gabriel. He will be glad to help you with any of your concerns or issues." And all those walking-wounded would then get up and move over to providers.

The higher-acuity patients would, obviously, then, need to be subsequently triaged. We are kind of, in essence, doing the same thing and providing the same principles that we do within responding to an event, but within a healthcare coalition.

And that healthcare coalition then, if you think about it, this is not done ad hoc. This is done, when you sign and you come on and you are admitted to our facility, you are signing paperwork. Your first piece of paperwork is a HIPAA form. Your second is a "you are going to pay us" form.
The third is you are part of a National Healthcare Coalition. In the event of a disaster and that you are deemed a lower-acuity patient, we will make appropriate plan of care for you as an outpatient, so that your outcome is the same, or as close to the same as it can be.

And if you need to be retriaged, then this is a constant flow and back into the system. So, triage is not static, as we all know who have done operations. This is a continuous flow through the event.

Then, during the event that the trigger goes off and you hit the button and we have to execute, those patients already are understanding that they are deemed lower acuity. We have rapid discharge plans in place, and within four hours a Greyhound bus is pulling up, and that ankle fracture you were planning on pinning tomorrow doesn't get pinned, goes home with crutches and a splint, and gets pinned in a week. Or that soft-call
chest-pain ruleout that is awaiting the stress
gets discharged and gets his stress in a week.
And they get on their Greyhound bus or they
have their family members coming to the door,
and they get offloaded rapidly.

The third pillar of the execution
of this performance measure is the ability to
accept higher-acuity patients to lower-acuity beds. This is a difficult road to walk.
However, it is consistent with what we saw the
Institute of Medicine speak to with regard to
crisis standards of care. We move from
conventional delivery of care today to
contingency, to crisis.

IBA is the ability to execute
contingency care and to give the healthcare
coalition greater depth to provide appropriate
levels of care before they have to shift to
crisis standards of care. So, those three
pillars are what is needed to execute IBA.

That is the performance measure
that the Hospital Preparedness Program is
going to be focusing on over the next five years. It has an evidence foundation. It is operationally-tenable. And we think that this will be achievable, and we hope that we can stand 200 coalitions across the nation or 400 coalitions across the nation, with 1,000 beds per coalition, that can get the job done within four hours.

That is establishing local resilience, regional resilience, and national resilience. And that is the target we are trying to achieve.

I would be glad to take any questions.

CO-CHAIR PITTS: I lost you a little bit at one point. What are the three pillars again --

MEMBER MARCOZZI: Sure.

CO-CHAIR PITTS: -- for my concrete thinking.

MEMBER MARCOZZI: The ability to continuously monitor care across your
coalition, that is the first goal. Monitor acuity, not care really, acuity. The second is the ability to rapidly offload, and the third is the ability to rapidly onload. In short, those are the three pillars.

Jay?

MEMBER SCHUUR: How are you defining coalitions? And is it sort of self-defined? Can a group of hospitals that may be an ACO or private group do that? Is it going to be a governmental function?

MEMBER MARCOZZI: A great question, Jay. We don't have any defined -- we have no definition with regard to coalitions. We know what the measures are. We know what the partners must be. But, as the awardee, we tell the awardee we want a hard-boiled egg, but we don't tell them how to boil the egg.

So, we know that they have deliverables and expectations. We don't tell them -- you know, some places may have five
long-term care facilities involved. Some may have one. Some may have seven primary care. Some have five.

And we are seeing different coalitions are established per our awardees. Fifty states are awardees. For instance, we can’t project that. One of our awardees is Guam. Well, defining a coalition for Guam is much different than the coalition for New York City. So, you have to be very careful what the feds project out on what defining coalition is.

Will it blend with ACOs? We are hopeful it does, Jay. I think it needs to.

MEMBER ADIRIM: Thank you. I wasn’t sure how you were identifying people to speak.

This was very interesting. I think you brought up a good number of concepts that are helpful in moving forward and looking at how to measure preparedness or integrate preparedness into the work with regionalized
emergency care.

I think some of the concepts that I am hearing you are talking about that could be a challenge in measuring these things is that most of what you are describing really is process. You stated that there was evidence for your particular measure, the 20 percent of increased bed availability. I think, though, I would like to think more about outcomes. Like what are you trying to accomplish? That would be something that I think would be interesting to look at.

The other thing, too, is whether or not that measure that you are developing can be tested, which is a challenge, of course, in preparedness because, you know, disasters don't happen every day. So, those are a couple of things that I thought you may want to think about.

And the other thing, too, is I heard you talk about performance measures. So, I think a little bit of clarity on
performance measures versus quality measures would be useful as well. I mean, I have other comments, but those were some of the main things that kind of came to mind as you were speaking.

MEMBER MARCOZZI: Yes, I tell you, you hit the nail on the head with regard to the ability to test this. We are looking right now at exercising what IBA is and how to do it. Actually, this is the first year out of the gates. So, everyone needs to know this is a crawl, walk, run approach. If we deluge our awardees too much too fast, the cart is broken and the wheels come off the wagon.

So, this is incremental and staged. Everyone is trying to figure out their coalitions look like and how to execute those performance measures.

So, with regard to exercises, we have some mandatory -- certainly, the Joint Commission has their mandatory exercise and drills, requirements, and we have our own.
Our exercises are actually large-scale, but we have not been able yet to test, and we plan on testing, IBA as it evolves.

It is integrated within the daily delivery of healthcare today. It is process-oriented and not as outcome-oriented as -- we would love to drive to eventually an outcome-oriented approach where we look at the effect of mortality and morbidity on this process that we are trying to put in place. But, in essence, the ability to care for higher-acuity patients we hope, then, translates into better outcomes for those affected by disasters.

Brendan?

MEMBER CARR: I have two questions. The first is I am wondering about white space. I am wondering if there is anything as part of the HPP that suggests to the country that there should not be a lot of space that is left without membership in the coalition.
And the second is, if you can give it, your opinion about whether or not, as these develop, if they become appropriate denominators for boarding. You know, one of your pieces here, piece two or piece three is the ability to onboard patients. That is dependent upon whether or not you just effectively reverse triage in pillar two. But is piece three tied to boarding measures at a coalition level, at a regional level?

MEMBER MARCOZZI: Our hope is -- and is really a grant discussion -- but our hope is, unfortunately, we have a $350 million program and a $2.5 trillion industry. So, we have to figure out how to do this very smartly and most economically.

One of the things we are trying to -- if you spread that $350 million out too diffusely, then we actually don't have the ability to move the needle and affect an outcome or affect a process. In essence, if we spread it out, and some of our awardees are
considering doing this and now changing, had
previously spread it out to every hospital
within their state. That is about $60,000 per
hospital. If the average budget is about $200
million, plus or minus, $60,000 is really not
going to be able to move the needle too much.

This is a grant discussion, but
one of the things we are thinking about is our
awardees should think about consolidating for
effect and trying to address a white-space
question is I am all right if we have more
white space. What I am not all right with is
our awardees and our coalitions can't get the
job. So, I would much rather stand on 100 or
200 coalitions that can get the job with done
with slightly more white space and maybe some
hospitals that fall out because they are not
as engaged, and they are not prepared. This
is not something they would like to be
involved in.

That said, any hospital can be
involved with a coalition. They just may not
be getting the funding to be able to support. We are looking to try to get funds to each coalition, $1.5, $1.8, $2.1 million, so we can affect the ability for them to execute the capabilities in the performance measures. So, consolidating for effect is something we are looking at.

With regard to white space, one of the discussions that we have had is we are trying to look at covering about 80 percent of the nation with regard to our grants. That is our hope. That is what we are trying to achieve. And hopefully, we are trying to get those measures and get our coalition input in. What is your geographic region? Who are your partners? Who do you cover? What is your population size that you have the breadth to be able to affect? And with that data coming in, we will at least have some idea on how close we are to the 80 percent and whether or not we need to revector.

The second question was --
MEMBER CARR: The second question is about synergy with this initiative, synergy with boarding.

MEMBER MARCOZZI: Oh, yes, boarding. Sorry.

MEMBER CARR: You know, the synergy to add your $350 million to someday a metric that pushes coalitions to think strategically about their capacity.

MEMBER MARCOZZI: So, in truth, I am heartened at the fact that we are having this discussion today and trying to bridge and weave. Truthfully, what we need to do is we need to weave a thread of healthcare -- pardon me -- we need to weave a thread of preparedness within healthcare. That has to be done. Preparedness can't stand alone.

The opportunity to have discussions on crowding and preparedness need to happen. The truth is, Brendan, to that end, our measures did not specifically target and think about crowding as much because, if
there is crowding tomorrow and we address some
of the issues and it gets better, still our
mark on the wall is you can accept 20 percent
irrespective of places that have the ability
to accept and don't have a crowding issue and
places that do. So, we try to be independent
of operational constructs and crowding. So,
that was the prism we looked through when we
tried to establish the measure.

CO-CHAIR PITTS: Arjun?

MEMBER VENKATASH: I guess my
question kind of gets back to a little bit of
what Brendan was just asking about white
space, in a sense that I think there are two
ways to think about it.

One is in terms of what areas are
just not covered by coalitions. But what I am
thinking about is, when we think about
performance measure, validity. The question
I would have is, when a coalition comes
together locally, if it doesn't include all
the relevant players within that locality, you
could see a situation where an IBA-type measure looks really good for that coalition, but misses the mark because they just haven't included all the relevant hospitals, long-term care facilities, whatever else it is.

So, is there any capture within the system to ensure that a coalition actually has adequate coverage within however they define that locality, be it state, county, whatever, local?

MEMBER MARCOZZI: I don't know if I have my arms around your question. You are describing a coalition -- let me just see if I can break it down -- you are describing a coalition, then, for a large city that is only affecting 20 percent of the city, and the other 80 percent is left in this white space, in essence? That is what you are describing? So, you are saying that they could establish the IBA, but not actually have the ability to respond to their large-scale area?

MEMBER VENKATASH: Right. So,
they could form it. They could report performance. We could do an exercise, and it would look great, right? They would show that they are able to offload/onload both, meet all three pillars for their system or the coalition as it is defined, but it misses the target because the general population is missed.

MEMBER MARCOZZI: Yes, I follow you. I have to tell you, it is interesting. We have not had that posed at all, only because we have come from the construct that 100 percent of our population is covered by the Hospital Preparedness Program. We haven't shifted that pendulum way right.

Now it may, with coalitions, and that would be something we might have to revise this measure to say we revised the coalition measure, the first capability, that your coalition needs to cover 80 percent or each coalition within your awardee's region has to cover 80 percent of your population.
But, right now, I will be honest with you, there is blanket cover. The Hospital Preparedness Program is diffuse and touches the entire nation. We think that that may be depending on which way we have look, because we have paper tigers out there. A guy coming into a meeting and then leaving, that is not prepared; that is just a guy coming to a meeting, yes.

CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: Thank you.

At the risk of being shortsighted, and just again to I guess dovetail onto what you said, Arjun, I wonder as I look around the room do we have all the right players in the room. We can talk about this ability to 20 percent uptake, and to take all of this additional surge, but, again, we always are looking at the mirror at the emergency department, the one that really needs to manage this.

What we have to do is look at how
can we coordinate with our inpatient partners
and our outpatient partners and the ability
for us to offload those patients that we need
to offload. So, how do we really coordinate
that and then what is that 20 percent that you
are talking about, that inpatient offload,
that outpatient coming in?

Because, yes, the coalition would
say, "Yes. Great. We'll do it." What we end
up, having to have people in bunk beds in the
emergency department.

So, I just wonder if having some
inpatient partners and processes, even some
type of research or measure that you partner
with an inpatient unit and that is your babe,
so to speak.

MEMBER MARCOZZI: Yes, so in the
interest of transparency, I am an ER doc. So,
I get it. I hear you. And I would always be
challenged when I would have door patients
boarding, and I would say, "Well, this is not
a bolus of patients to the ER. This is a
bolus of patients to the hospital." For some reason, if you keep those doors closed, that conversation typically does not happen.

So, this effort and this measure is actually to affect the entire -- actually, the truth is I have actually had discussions with our critical-care colleagues around this. So, what is the MICU's and what is what is the SICU's perspective on this measure?

So, the guy who we are planning on weaning from the vent, and he has been on the vent for two weeks and is stable, and we have got to start to wean. And all of a sudden, we have an event. Well, in their MICU that patient is potentially is a lower-acuity patient. So, then, that patient potentially you keep on the vent; you just keep sedated, move to a floor, so that the nurses can manage that. So that, then, you can accept higher-acuity patients.

So, the 20 percent is not just for the emergency department. It is across the
spectrum of care. Now one key thing about the 20 percent, which this may be a wordy discussion and not for all those in the room, but if a coalition has five long-term care facilities at 200 beds per, and you have a couple of hospitals, the measure for the 20 percent is not for the entirety of the number of beds within their coalition. It is the number of acute beds within their coalition. So, it is not every long-term care, the 200 beds within every long-term facility. Their measure to get from an operationally-tenable goal is only for their acute care beds. And that is different. We had to let our awardees know that, that that was the expectation.

But, to address it, it is not only and the intent is not only to be for the emergency department. It is to be systemwide.

CO-CHAIR PITTS: Okay. Anthony, actually you were next. And then, Melissa.

MEMBER MacINTYRE: Thanks.

Dave, you bring up a lot of points
sort of all within 10 minutes. For me, it is
kind of confusing the picture in relation to
this project.

I think several of the points you
made bring up several questions for the
project managers. One is David is obviously
focused at the healthcare coalition level.
What is your unit of measure? Where are these
measures going to be applied? I think that
should be very carefully articulated because
it could be at the healthcare coalition level,
as I read it in your paper. It could be at
the individual facility level. And quite
clearly, there are going to be different
measures, I think, we are looking at as you
move forward.

I think another important thing to
articulate -- and I think your paper touches
on it -- is the extreme difference between
preparedness and response. In fact, it is so
important, I would recommend you change the
title. It isn't just about preparedness; it
is about preparedness and response. And you
do cite how it is much more difficult to
develop measures for a response, but we still
need to get there to have measures for that as
well.

The third thing is, when we look
at response, much of the conversation is
dominated on surge. Quite clearly, that is an
important thing for any healthcare system to
be able to do. But I would encourage you to
look at some other work out there, including
some that the Veterans' Administration has
done, where response and surge -- surge
actually takes sort of a tertiary priority to
two other things, the first being safety and
security. If you can't keep your facility
safe and secure, then you can't surge.

And the secondary sort of priority
is continuity of operations. If you can't
keep your operations going, then you can't
surge.

So, there is sort of a tiered
approach to this: safety/security, continuity
of the operations, and then surge. And I
think some of that might help shape this
framework.

MEMBER McCARTHY: I think I was
thinking on a very similar vein because I was
going to ask about these capabilities needing
to be prioritized. Because you can't have any
medical surge until you have a coalition
developed, and I am not sure we do have strong
coalitions developed. So, it seems to me that
you do have to kind of prioritize and start
there, and then increase the competency.

MEMBER MARCOZZI: Yes, that is
exactly what we are seeing. We are at step
one out of the gates. We are seeing forming
coalitions currently. Some places, I will be
honest with you, are already well-formed and
mature coalitions. In fact, Virginia already
has a very well-formed coalition. Seattle has
a very well-formed coalition. So, they are
actually moving beyond and now trying to
accomplish some of the other capabilities in
and around medical surge.

But we go from literally some of
our awardees have no coalitions to some of our
awardees are trying to actually to be the
exemplary, the A-plus students.

CO-CHAIR PITTS: Wes?

MEMBER FIELDS: I am really
intrigued in a couple of ways. One of the few
good things about a disaster response is that
it is one of the few times in a metropolitan
service area where market forces are
suspended.

And one of the interesting
corollaries on the public safety side is it is
also one of the few times when Medicare-
provider hospitals see EMTALA suspended.

Andthirdly, a lot of the surge
capacity you might need within an area of
impact, I am concerned, as Arjun has implied,
might be in the part of town that is not part
of a coalition.
So, I am kind of wondering if the best investment for the resources you have for this is to look at a different set of rules of engagement for Medicare-participating hospitals in these scenarios and essentially a different kind of EMTALA that wasn't focused on the needs of the patient, but of a population that was in harm's way.

Because it may be that that is what you need to begin to try to create and measure and promote, is the ability of all providers within a service area to respond and how that fits together.

I think there are some attractive alignment between EMS agencies and hospital systems in highly-consolidated markets like Seattle that might make that doable.

So, speaking in favor of using the resources to think about what all providers within an area, whether it was rural or metro, how they would collectively respond. I think that could be useful.
CO-CHAIR PITTS: Okay. I think we can take a break here and proceed to the next step, which is Mike Stoto, to give a presentation. Okay? Am I missing something? Ellen? I'm sorry. Ellen, I didn't mean to cut you out.

MEMBER WEBER: Maybe this is stating something people are thinking about or inadvertently saying, but it seems to me that this is a totally scalable idea. And I think Brendan was kind of getting at this earlier, which is, why not have a 3-percent or a 5-percent, and could that be at the hospital level? So that, in terms of getting together our two ideas about crowding, boarding, and preparedness -- because, first of all, you could potentially measure that for real because there is not going to be a disaster, but there are going to be 3-percent, 5-percent surges. So, does the hospital have a way to deal with that?

It is kind of one of the things I
think we need in this report, is how does this connect? How does your ability to handle a daily surge connect to your ability to handle a disaster? Although I think that there could be complete separations on that, I think the adaptability, the accountability, the flexibility of any organization to be able to do what you are talking about at a lower scale would at least start those conversations between the inpatient, the outpatient, and between perhaps a neighboring hospital when you have no more beds, the primary care clinics when you need to offload some of the lesser acute patients. So, it does seem to me that everything you are talking about would totally apply to the crowding issue.

MEMBER MARCOZZI: Can I jump off of that? Sorry.

CO-CHAIR PITTS: Sure. Yes, go ahead.

MEMBER MARCOZZI: So, yes, I think that, again, I am hopeful that we can actually
jump right off of that and try to figure out how we could blend our efforts here with the performance measures that we are trying to shoot for. So, that is great feedback.

The second thing I would talk about from EMTALA's standpoint for one second. So, I mean, here are the triggers for EMTALA, right? So, the hospital has to declare a disaster. Then, the Secretary of HHS needs to declare a public health emergency. Then, the President of the United States need to declare a Stafford Act to execute an 1135 waiver, which is what you are speaking about, about patient dispersal and EMTALA waivers. So, you are talking about high bars.

Now, and again, transparency, that is an "and," right, public health emergency and Stafford Act. There is actually floating out there a law that actually makes it an "or". So, even if we make it an "or," public health emergency or Stafford Act, we still have a high measure. The Secretary still has
to come in front of everyone and say, "We are
declaring a public health emergency." And she
will look to the Assistant Secretary for
Preparedness Response for that advice on an
event.

But this speaks to what I tried to
hint at with regard to local, regional, and
national resiliency. But I think that this
measure speaks and tries to accomplish, that
if we have a 50-car pileup and there are 100
patients presenting, the region has the
ability to respond, and it creates regional
resilience. But the Secretary of HHS does not
have to stand up and say, "We are declaring a
public health emergency to be able to execute
IBA."

Now IBA allows -- pardon me --
PHEs and 1135 waivers allow the dispersal of
those patients appropriately to affect that
care, and EMTALA and allowing those 1135
waivers, if they give bolus to the closest
facility, which we saw certainly in Madrid,
events like that, to be able to push those
patients back out to other facilities if we
allowed those 1135 waivers.

But the truth is IBA works if an
1135 waiver is accomplished or even if it is
not accomplished. So, it is integral within
the system, and it can be used in either way.

Thanks for the 1135 comment. I am
all about 1135 waivers.

CO-CHAIR PITTS: We can do a
couple of short things.

MEMBER ADIRIM: I just didn't want
Dr. Weber's point to be lost because I think
it is really probably one of the best points
that was made, that a way to integrate these
concepts into the work that is being done
here, I would imagine would be to develop
measures that could help you measure whether
or not you are prepared, but also are related
to other everyday measures. So, I just
thought that her point was right on target.

CO-CHAIR PITTS: Sally, do you
want to get your fair share?

MEMBER PHILLIPS: Yes, I mean, it will come up again. I think one of the things, as we are trying to develop this sort of a measure complex is, having worked in this area for a long time, we are sort of leaving the individual clinician out of this. As we discussed in many ways, these are performance measures for a program and for a system. But one of the things is we are potentially asking them to go in a way that is contrary to a lot of the quality measures we have developed as far as quality of care and ED delivery and timing.

When we talk about, well, we will delay that treatment or that surgery, then that reflects back on their quality measures of how they sort of set up their practices and how hospitals are being measured. So, somewhere in the middle there is a culture change of getting people to understand under extraordinary times those quality measures.
So, it sort of puts this juxtaposed. It is a little bit of where you were going with the systemwide. But if you bring it down to the clinician level, full participation in what we are talking about is going to require a little bit of tweaking because we finally have gotten people to start instituting quality measures into their care and measuring performance, and in many ways asking them in this first step of, when you are surging, sort of making these alternative decisions, it puts them in conflict with a culture of measures that they have had in place. We get kind of ratcheted up around the system and forgetting that the system is made up of a lot of clinicians who have just kind of come onboard really well with this.

CO-CHAIR PITTS: Okay. We shall proceed with Mike Stoto then. Thank you.

MEMBER STOTO: Okay. Thank you.

Do someone have the slides?

The things I say really come from
some work that I have been doing with my colleagues at Harvard through our CDC Preparedness Research Center, and it really reflects the work of a lot of people and conversations we have had with a number of people, including people like Anthony, who was on our advisory panel. I will try to sum up some of the things, the thinking that we have been doing -- if you could just go to the next one? -- that I think has some importance for what we are here today.

So, I think it is important to begin, and I am also happy, if we have time as we go along to do that, that might be a more efficient way to do it.

I think it is important to recognize some of the challenges -- public Health Emergency Preparedness is what PHEP stands in our lingo -- that are somewhat different from a lot of the work that NQF does. One of them is that public health emergencies are rare. That has two important
factors. One is that you can't measure outcomes directly. If you don't have a stream of heart attack patients coming into the emergency department, then you can't measure what fraction of them get asked.

Secondly, because they don't happen very often, also, it is hard to study what works. So, the evidence base is somewhat thin.

Second is that an effective response we know is complex and multi-factorial, and it is hard to know what is the right way to respond to any given response. And we usually don't have the counter factuals. We don't know what would have happened if we had responded some other way.

What we do know is that we need to have system-level measures. So, this is a point of contact with Dave's point about the unit of measurement has to be bigger than the patient or the emergency department, and so on.
That really gets us into the third point, this idea that the public health system is fragmented. We use that term "public health system," it really draws on the Institute of Medicine report that set up the PHEP research programs that think about not only the official governmental public health agencies, but also the healthcare delivery system and Homeland Security, employers and businesses, education, and so on. That is the whole system that we needed to get to work together.

Well, that varies quite a bit. It works at the city, the county, and the state, and the national level. Sometimes there are regional structures embedded in that across state lines like we have in the Washington area. And basically, it is different almost every place you look in the country.

And then, the partners to public health vary quite a bit. Clearly, healthcare is in there, EMS, and so on.
I think that this concept of regionalized emergency care systems obviously comes into play there. But, again, building on what Dave said, that is the unit, but you also have to think about who they relate to outside of the healthcare delivery system as part here.

Another complication that comes up that I have seen in my work is that, increasingly, hospitals and healthcare delivery in the U.S. are parts of chains. So, our Georgetown Hospital is part of MedStar that has about a dozen hospitals and a lot of other healthcare facilities between here and Baltimore. A lot of those hospitals will think about coordinating first with the people in their chain, rather than the other hospitals in D.C. So, that is another complication.

Ultimately, you have, who is responsible for what? You need to think about that before you can come up with different
measures.

So, our goal in this paper, it was cited in the Draft Report. There is a slightly more updated version of it online at that URL. Our goal was really to kind of think about how do we apply the science of assessment, the kind of stuff that NQF does so well, so this area of public health emergency preparedness.

We think in terms of a measurement development cycle that involves, first of all, clarifying the purpose of the measurement, the accountability, QI, or is it for research? Identifying the concepts to be measured, and then developing specific indicators, and then assessing validity, reliability, practicality, and utility.

I think if you go back to the points that Helen made about the criteria, that is all embedded, but represented in a slightly different way.

And there is often a tension, in
particular, between accountability and quality improvement, the kind of measure you want for both. For one, maybe it is not the best for the other, and I will come back to that point in a moment.

So, let's think about the second step, identify the concepts to be measured. You have to begin by, first of all, thinking about what do we mean by preparedness. Anthony, this gets at your point about preparedness versus response. And I am going to say a bit about that as well.

We start with a consensus statement developed by Chris Nelson and others five years ago. The capability of public health and healthcare systems, communities, individuals to prevent, protect against, respond. You can read all of that.

I think that implicit in that is the goal that in a public health emergency we can do things to mitigate the mortality, morbidity, psychological, and social
consequences, but 100 percent prevention is impossible.

But particularly when you are talking about a contagious, infectious disease, there are things that can be done to really reduce the consequences across the board with an effective response.

So, then, the question is, what does it take to do that response? So, on the next slide we talk about outcomes, capacities, and capabilities.

Because emergency is rare, we typically can't measure outcomes, which NQF likes to do, but we can't do that in general, particularly at the system level.

So, it is important to think through the capacities and capabilities, and this reflects the thinking that Dave mentioned about moving from capacities to capabilities. And I think, Anthony, this is what I think gets to your point here, is that the capacities are what I think of as what
preparedness people do now to get ready, so
that when the time comes, they have the
capability to respond as needed.

Most of the effort we do is in
preparedness, in getting those capabilities --
excuse me -- getting those capacities in
place, so that we will have the capabilities
to do what we need to do to respond when the
time comes.

That is a point that I think is
easy to say, and it is a lot harder to
operationalize because, when you actually see
what people do, what is the capacity, what is
the capability, there is sometimes not a lot
of consensus there.

You know, we see it in the
assessment world when we talk about what is
the structure and what is the process and what
is an outcome. And everyone knows exactly
where things fall, but that problem is a lot
more complex in this realm here.

So, what it takes, we think, is
building a logic model that really connects up what we do now in terms of building capacity to what we can do during an emergency, to have the capability to respond, and how does that meet the goals that we are trying to reach?

So, the common ground preparedness framework that is in the background document is one example of this. On the next slide, we have a model that we have been using that thinks about this.

There are different ways of thinking about this, different ways of categorizing this. But this idea that you have to think about how these capacities lead to capabilities and help you meet your objectives I think is fundamental because we are not going to be able to measure the objectives or the outcomes. What we want to do is make sure of the capabilities, and oftentimes we are reduced to measuring the capacities. So, I don't want to go into the detail of this, but I think that that is the
fundamental point.

So, on the next slide, we review in our paper some of the things that have been done. In fact, if you look back over the last decade or so, where people have been worried about preparedness and response, a lot of what has been done really falls into the capacity world, inventories, capacity assessments, and so on. I have got a list of all these things that we have seen before that fall in there.

And in fact, if we just hit -- there we go. Most of the Joint Commission standards that are referred to here are for hospitals, but they actually fit into this capacity assessment as well.

So, on the next slide, we talk about some of the strengths and the weaknesses. One of the strengths -- and this was really an important one early on when no one really knew what to do -- is that these capacity assessments communicate standards and expert guidance. They tell hospitals, they
tell health departments what the experts think need to be done. If you have done that, well, that is good.

But the problem is that the evidence to support this guidance is often lacking, and there may be other ways to achieve the goals than the ones that the funding agencies actually put forward.

In addition to that, many of these capacity assessment in practice aren't clearly operationalized or consistent in the kind of way that NQF likes to see things done. They are hard to summarize across place and time. And a lot of the question about who is responsible for what in these things is unclear.

So, on the next slide, what I have done here, this really draws on the work that CDC has done, what Dave mentioned about the CDC's public health emergency preparedness capabilities. These 15 are the capabilities that the public health world is coalescing
against as the critical things, the critical capabilities that we need to have in place here.

I have marked in red and with an asterisk the ones that also are in the Hospital Preparedness Program. So, there are some conversions there as well, which is a good thing. They really reflect the latest collaborative thinking about what should be done with preparedness on funding.

And another thing I think about capabilities that is important is, if you say, "Here is what we want the system to do," that allows different localities, different hospital systems, different health departments, and so on, to figure out what is the best way to do it in their location to achieve that capability.

Part of the problem with the capacity approach is it tells you should do it this way. And I think that it makes sense. I think that this is reflected in the kind of
things that Dave was saying as well. We need to allow these different systems and localities, and so on, to figure out what is the best way to do things, given the resources and the structures, and so on, that they have, to achieve comparable capabilities across jurisdictions. Again, that is an easy thing to say and a lot harder to actually operationalize, but we have been thinking about that.

So, on the next slide, I have got this is the way that these things are structured, these capabilities. They have a definition. They say what are the critical elements and the functions, the performance measures. As it turns out, so far, many of them don't yet have performance measures.

And then, these response elements -- if you just click one more time -- what they really do, I think these are the things that build the links between the capacities in the left column of the logical model to the
capabilities. They say, what do we need to do
now that is going to get us to those
capabilities?

So, this document, which is a very
useful document, and the same thing for the
HPP document, is useful in communicating
consensus about what we think we need to do
now, so we have those capabilities.

The challenge in many of these
things -- and I think this is true for the HPP
-- is how do you actually measure those
performance measures. So, on the next slide
I have got some definitions here.

Two of these CDC capabilities deal
with the general area of biosurveillance, and
they have to do with -- Capability 12 is
public health lab testing and 13 is epi and
surveillance.

And what they really have to do
with, if you look at those things, is how long
does it take before people can get together?

Have you passed proficiency tests in your
labs? How many infectious disease outbreaks have been reported? Have you done after-action reports and learned from them?

Those are all useful things, but in the last couple of years we have actually been looking at the public health system response locally up to globally, to H1N1, the 2009 H1N1 outbreak. It turns out that these things really don't predict what was necessary. What was really necessary during that point, during that outbreak, was, could you figure out that you have all these things going on in different parts of Mexico and in California and New York, all part of the same phenomena? That is a critical capability, the ability to really integrate this information and think about it and understand what it means.

I think that is fundamentally what we mean when we talk about biosurveillance capabilities. But a lot of these measures that we have really represent capacities, and
how do we go from those capacities and capabilities is a challenge that we are struggling with in the public health work. And I think it is also true in the health system world as well.

So, I have just put up here the preparedness capabilities for mass care and medical surge, which I think are the ones in public health that most directly relate to the stuff we are talking about here today, just so you see what is there.

I guess I have too many words here. But one thing you see in both of those, and underlined and in italics, is that there are no performance measures available for these at this time. That really is the state of the world that we have to wrestle with.

So, one think that I would offer as an alternative to think about going forward is using exercises and actual events as a way of measuring capabilities. We have had some experience with this.
The first part of this deals with the work that Paul Biddinger, in particular, my colleague at Harvard -- I am sure you know him, Jay -- has done to develop this exercise program. They have now done about, I think, 30 or 40 different exercises where, rather than using the exercises as a training opportunity, we use it as an assessment opportunity to measure how well different kind of systems -- some of them were hospital systems; some of them were public health systems; some were combinations, and so on -- can respond to certain kinds of events.

We do that by asking not only the participants to evaluate how well they did, but having some trained external evaluators. We have studied this carefully and developed checklists and scores. So, we think that we actually have some valid and reliable measures that can be used in this setting, to get a sense of how would the system respond during a particular kind of emergency.
Coming back to what Dave was saying about the measurement, about your ability to expand acuity beds by 20 percent, well, how would you actually measure that? This is a possible way of measuring that, where we can say, well, here is a particular scenario; what actually would you do? I mean, who actually would you move, bump to next week, and so on? How would you share resources across these different hospitals and in other parts of the healthcare system within a region? What actually would you do to see whether or not you can meet that goal?

The other thing that we have discovered along the way is that sometimes for some purposes it is more useful to have qualitative rather than quantitative measures, and particularly because I think so little is known about what works in this area. Some of the qualitative analysis about what went wrong, what could have been done differently, what would have gotten a better result
actually is a more useful thing that comes out for some of these quality improvement purposes. Now, for accountability, you need something like these quantitative measures, I think, but some of these qualitative things can work best for that.

And then, the other thing I want to mention is something that we are just working on now as part of our CDC Preparedness Center grant, is the idea about learning from actual events. We are trying to develop a registry where health departments and public health systems can do something like the kind of root-cause analyses that are now relatively common in the healthcare system and learn in a deep way about what happened, what could be done better, and to share that with others, so that others who have similar circumstances can learn from that, and so that researchers can look across similar incidents and see what patterns there are.

As part of that, we are trying to
develop a kind of peer assessment model where we take some of the approaches that we do from this exercise program and use them to look back at critical events and have valid and reliable measures that come out of that.

But I think that, again, this is a potential model to look at for assessing how well the system might respond to future events, by seeing how well it responded to the current events.

So, to wrap up, I would like to say there is no assessment approach without problems, but I think that everything has something to contribute here in the PHEP work, and I think in the emergency response world as well.

What we think we need is a portfolio of measures that are useful for both accountability and for quality assurance, that address both capacities and capabilities, and that balance the detail-specific quantitative measures with more holistic qualitative
measures, and balancing objectivity and professional judgment, and perhaps varying as the needs, what we are looking, changes.

We need measure systems. It is not just going out there and sort of doing this ad hoc. We need to sort of think about a systematic approach to measurement and developing measures. So, the fact that we are having this meeting is a big part of exactly what we are talking about.

Thank you.

CO-CHAIR PITTS: Thanks a lot.

We are making up the agenda a little bit as we go. I know it is almost time for a bathroom break.

So, let’s have a couple of rounds of comments. It is a complex subject.

Being a relative newcomer, it is intriguing to me to see how people come from their separate spheres and have different approaches and looking at the same thing from a different direction. We will eventually try
to reconcile all this stuff and put it in the NQF framework.

For now, does anybody want to say anything about the presentation, any comments specifically?

Yes, go ahead.

MEMBER CARRIER: Well, this may cross over a little bit into the NQF framework. But just thinking about both of the presentations, and wondering how they would translate, if we were providing guidance to measure developers, how they might translate into quality measures in terms of things like feasibility and what the data sources that could be used to measure these things.

So, thinking about the capability.

And I think we heard in the beginning about the importance of outcomes measures, and you were mentioning the challenges of measuring outcomes and how we may have to go back to capabilities.
But, even in terms of capabilities, to take an example, let's say you want to move 20 percent of your acuity beds. Ultimately, you would like to know that not only can I, as CEO of Hospital A move this 20 percent of patients, here's the ones I would move, but that the receiving entities, whether they be long-term care facilities, other hospitals, even a patient's continuity provider, who would need to follow, someone who you might otherwise have kept in the hospital, might be willing to accept them.

So, that is how in an ideal world we might want to measure these. In the actual world, in most cases our data is limited to the hospital level. We don't have the kind of unified data systems that would allow us to really follow the hypothetical path a patient would take and really know that path is clear.

So now, we would be stuck with at the station, which to me is not very satisfying. I wonder what the group thinks,
if that would be sufficient, or if people think that it is worth trying to develop a novel data collection system.

MEMBER STOTO: We would be stuck with what?

MEMBER CARRIER: Attestation. You know, I check a box "yes".

MEMBER STOTO: Thanks. "Yes," I can do that, right. yes.

MEMBER CARRIER: I can move this 20 percent of patients.

MEMBER STOTO: Yes.

MEMBER CARRIER: Which is sometimes easier to do than actually doing it.

(Laughter.)

MEMBER STOTO: Right.

MEMBER CARRIER: And then, we talked about measuring at the coalition level, which would require a whole novel data collection effort, the development of instruments, the validation of instruments. I mean, which do people think is the path
forward, going with measuring at the
individual hospital level, hoping that some
integrated systems will have the capacity to
give us a clearer picture, or pushing for
novel data collection that can truly capture
regional efforts?

MEMBER STOTO: No, I think that is
a real important problem. I would say two
things. One is I think that these exercises
that actually are done quite a lot, because
they are required for Joint Commission
accreditation, and so on, are an opportunity
to not just use them for the basic purpose,
but actually use them as a measuring purpose.
But there needs to be an infrastructure built
around them.

And then, secondly, I think that
you actually can do more than attestation.
So, I will give two examples.

One is we used this approach early
on. I worked on RAND, and this was, I think,
2003. We did tabletop exercises in California
to see how they would respond to a smallpox outbreak.

But what we did beforehand was we went out there and we interviewed them, understood what their resources were. So, if they during a tabletop said, "We do" so-and-so, we knew enough to say, "No, no, you don't have that." And so, we can actually have some judgment about this.

The other thing is that I was involved in evaluating, Anthony knows, the DC Healthcare Coalition, and some of the work in Boston. A lot of that really involves doing these exercises and asking the participants to evaluate how well it worked.

But I think we can add on to that. So, for instance, when you are doing a drill that involves communications, they often say, "Here's how many beds we have" and report that in, for instance.

One thing that you can do is have somebody receive it and see whether or not the
people who received that information understood it the same way that the people sending the information did. And I don't think that that is always the case. I think that just because you can say we have got this many beds, it may be understood in a different way. Sometimes they say this is how many beds we need or it is confused with how many we have available.

So, I think you can look at fidelity in that way. But, again, that takes an effort. It is more than just doing exercises. It is really developing a measurement infrastructure to go with them.

CO-CHAIR PITTS: All right. Wes?

MEMBER FIELDS: Yes, I want to sort of extend Ellen's analogy here. One way to get to 20 percent is showing how 10 hospitals can pick up 2 percent. I think that is truly valuable. In this context, in terms of measurement, if we could connect everyday problems, or not everyday problems with
crowding, but somewhat more dramatic problems with crowding on a particular day, especially if it is related to, say, things that are predictable like flu season and that, I would far prefer to measure how real hospitals respond to real surges in demand as a preview of their ability to respond in a logarithmically-larger event. I would much rather use the present-day response of hospitals within a community or a metro area than to measure what happens on a benchtop exercise.

I acknowledge there are probably a lot of things that are unique to your work that probably have to be done on a tabletop. But I think if we are really trying to connect the dots between crowding and the real-world capacity of hospitals and emergency departments and EMS agencies, it would be useful for our measures to look at what happens when you get a 10-to-20-percent pop in demand at a particular department or within a
particular healthcare system.

MEMBER STOTO: No, I agree, and I probably was misleading when I talked about those real events by referring to H1N1. In fact, there are things that happen all the time that stress emergency response systems, and so on. And we should learn from them.

The problem is that they don't stress quite the same way, the big ones we are really worried about. So, I think it has got to be a combination of evaluating these kind of day-to-day big events and doing some of these exercises.

CO-CHAIR STONE-GRIFFITH: I know that I would like to sort of speak from a recent event that has been in the news and been talked about a lot, the Aurora incident. One of those was our hospital; the other one was the University. We have heard the statistics.

During ACEP, right before the Dark Knight presentation, ENA was there, and we
happened to do a debrief at Aurora and included the folks at the Children's Hospital and the University, which many of you heard in a different light. We also had their disaster preparedness folks and security folks and plan ops folks there.

It was fascinating what came out of that discussion. Our hospital, half of the hospital was under construction and closed, the ED, and so, all of our trauma rooms and all of our code rooms. So, we had to deal with that. We received 18 folks.

The University, what we did not know, which was very interesting, or wasn't really illuminated at the time, was that they were only divert. They had 25 inpatients holding at the time. They received 23 patients.

Now, when we said, "Well, gee, what happened to those inpatients that were boarding," within 45 minutes all of those patients were out of the ED and they had been
effectively moved and absorbed and taken care of.

My question was, "How did your leadership respond after this incident to the day-to-day diversion and holding that you all are doing?" And their answer to me was stunning. It was, well, they have really put on fast track the new building tower and the expansion of the ED.

(Laughter.)

And I thought to myself, we have just completely missed the boat. Because that incident was over in four hours. People ask me all the time, "Gee, what did you do?" I said, "I didn't even know about it. It was the middle of the night. I was asleep." By the time 4:30 came, I mean, both hospitals will tell you it was over. People who came on for day shift didn't even feel that. It was business as usual.

So, I am struggling with this issue of it is not just about a big disaster.
It is about the little day-to-day disasters on top of this surge. I think we have to look at it across the system. We are competitors. We are an exit away from each other. We did not have good communication between each other.

And I have heard a lot about EMS, but you all know that EMS did not bring those patients. They came in the back of swat cars and police cars.

So, I think we need to think about it in the context of those kinds of situations because I think they are so much more common. We could learn a lot about that. We have had brief after brief after debrief and community responses, but are we really getting to what happens when that happens next?

It will be a different incident. We won't know what it will be. But are we going to respond any different? Are we going to be in any different circumstance?

MEMBER STOTO: Well, you know that is an issue in the public health world as
well. One of the things that we have suggested is that one measure could be, did you learn from this incident? Did you actually do a deep analysis that got beyond -- what? Did you make changes based on that learning?

That is a capacity measure.

Excuse me. That is a capability measure.

That really is a structure measure. But I think, given the state of the field, that is important, that we learn from these things. That is something we should consider as a performance measure.

CO-CHAIR PITTS: Just to pile on, when we had the Olympic Park bombing in Atlanta way back when, it was very similar. By the time morning shifts got there, the 20-odd people that had rolled in were gone. It was pretty amazing.

Jesse wanted to say a couple of things before we take our next break.

MR. PINES: Great. Thanks.
I just wanted to thank everyone.

I know, sort of looking around the room, and what has happened in this first hour and a half here, and I can see a lot of people squirming. This is really the goal of this, really to bring these two worlds who speak completely different languages together and really get onto the same page.

You know, I think we have made a lot of progress, sort of thinking about linking daily crowding to disaster surge. I do want to continue that discussion and, also, hone in on something that Emily said just a few minutes ago, that really our goal here is to guide measure developers and to have practical recommendations for people who want to do measure development in this area. What do they actually need to get some of these preparedness measures through the NQF process?

And finally, I wanted to make a clarification on Mike Stoto's presentation and, also, to clarify a question that Anthony
had about the level of measurement specifically for this report. Our goal is really to look at the health system and potentially healthcare coalition level, as essentially there is a whole world of public health emergency preparedness that really sort of is outside of our scope, you know, managing casualties, managing bodies and things that really fall under the bailiwick of a local public health or under true public health.

Really, our focus today, what sort of data do we need in order to get some of these healthcare system measures potentially which would go up to the healthcare coalition level or hospital-level measures through the NQF process?

I do want to make sure that we are able to take a break. But, essentially, one of the things that I wanted you to think about is, after the break, I want to see if we can take some of these preparedness measures for a little test drive, and maybe we can think a
little bit about the IBA measure and essentially what it would take to get the IBA measure through the NQF criteria.

Thank you.

CO-CHAIR PITTS: Fifteen minutes.

(Whereupon, the foregoing matter went off the record at 10:30 a.m. and went back on the record at 10:54 a.m.)

CO-CHAIR STONE-GRIFFITH: Hi, all. We would like to get started again and move on to reconciling daily surge and disaster surge. So, if folks could take their seats? Thanks.

CO-CHAIR PITTS: All right. We shall continue here. We have a couple of nominations for topics to discuss.

The first topic, it has already been discussed a bit. A couple of people thought about this, and I think it is an important issue because daily surge is meant to be different from disaster surge. The question is, does good performance with daily surge translate into good performance with
disaster surge? I think, obviously, most of it is speculation because we don't have enough disasters to have a sample size.

But Melissa McCarthy sort of has some opinions about that, I think, and has written a paper about it.

(Laughter.)

So, I will let her start the discussion.

MEMBER McCarthy: I think they are not completely different animals, but they are really quite different. I think it a little bit goes back to Sally's comment about just very different mindsets. You might be able to handle daily surge quite well, but not handle a SARS patient that comes in. I mean, we have talked about trauma. That is really kind of doing what we already do, you know, just adding a few more trauma patients. But the kind of threats that we are likely to face, you know, we could face, and these different kinds of disasters, I don't know that
hospitals that do a good job with daily would
do a good job at all with those kinds of
disasters.

I think Sally's point that you
want in a disaster, the personnel to respond
in a completely different manner. It kind of
goes back, I think, to Wes' point about now we
are thinking about the population as a whole,
where in daily surge we are talking about
fighting for each individual patient, right?
So, I don't think they do actually have a lot
in common. So, I am a little uncomfortable
this morning about this. That is my own bias.

CO-CHAIR PITTS: Feel free to
think some of questions here.

As a professed ignoramus on the
topic, does anybody know of any studies in
which that has been looked at in peered data?
I mean, it would be hard to imagine the
scenario where you might have done that, but
no?

MEMBER McCARTHY: Actually, one
other comment about it. I was actually thinking about, from Harvard, there is a business person, a famous MBA, who talks about kind of different corporations and sustaining, how they sustain their operations versus they face kind of emerging and novel situations.

The way he has described this is the best way to do this is to have like a separate group that handles these emerging and novel situations because what you need to sustain, and the way you need to think about sustaining and handling daily operations, is very, very different than the kind of creativity and mindset that you need in these. And it is totally different area, but I think it applies here.

Frank, do you have any thoughts on this, too? I know you are sick, but --

MEMBER ASPLIN: I pretty much need to go home. I apologize.

(Laughter.)

And will be as soon as I think I
I think there is overlap in the Venn diagram. What I am struggling with a little bit is, even if operationally they would be very different, would some of the measurement framework still apply in kind of a cascading fashion, like Ellen alluded to earlier?

I think that I agree that the operational approach is going to be different. There might be components of a measurement framework, though, that could scale up and still apply, depending on how it is constructed.

CO-CHAIR PITTS: Everybody here knows about black swans? I mean, it is kind of the way these things happen, nobody ever thought of this happening before. It is tough to prepare for something you have never even thought of happening.

Mike?

MEMBER STOTO: Well, I think that
the possible answer is to think, what are the common response capabilities in both the day-to-day and the emergency setting?

So, in the public health world, we need to, whether it is a routine food-borne disease outbreak or a smallpox outbreak, we need to be able to have people identify cases, report them to the health department, track the numbers, and so on. Those are kind of core capabilities that are tested on these routine events and in these bigger events.

I don't know enough about emergency medicine to know what they are, but I suspect there are some core things that are tested in these routine events that also would be critical in the more extreme ones. That may be a way to structure making the bridge between the two.

CO-CHAIR PITTS: Melissa, I know you have done, you and your collaborators have done a lot of work looking at individual institutions and their day-to-day variation.
I wonder if it is possible to look at performance during those peak days? Has somebody done that kind of stuff? You know, there must be days when the school bus arrived or when there was some event of that sort that led to a huge surplus of patients. I wonder if that has been studied as to how do we perform in those days compared to other days.

MEMBER McCarthy: Yes, we have looked at when acuity increases, but they are so rarely really, you know. You figure you have 60,000 visits in ED, and the one day that you might have a couple of extra school buses arrive, you know, we just don't -- I tried very hard when we were looking at crowding on length of stay and stuff to see whether like an uptake in acuity would matter and just could not show it at all.

I know, Brendan, you have done a little work, right, in this area around like trauma and its effect, if you do have a couple of additional trauma patients, how it is
affecting emergency care, right? But I haven't seen it.

MEMBER CARR: You know, it strikes me that we are, it feels to me like we are talking about the distinction between capacity and operations, though. I don't know if that is right or not. But, in my mind, the day-to-day is different because it is operationally different, but maybe what Brent was suggesting is that we can find a common ground that says 2 percent is still 2 percent; 20 percent is still 20 percent. Whether you do it right, once you can make room is an operational issue, not a capacity issue.

I mean, do the boarding folks think about it in those two separate silos?

CO-CHAIR PITTS: Anybody here boarding folks?

MEMBER CARR: Yes, I am looking at them, and they are all looking at me.

MEMBER WEBER: I will jump in. We have, and I think that has been one of the
problems I was just seeing. You know, when I think about disaster preparedness, it is like somebody else; it is not me. I don't participate in the disaster drills because I see myself as the person that is going to be in the emergency department taking care of the people that eventually somebody sorts out and says need to be in the emergency department.

But, yet, when I think about it, there is a lot of parallels. I mean, what you were describing earlier about some people are going to have to get sent home and not get their stress test today. We do that every day. We triage every -- we don't send people home, but we make people wait a long time because their problems are not as, using our clinical judgment and our ESI scale, we have decided that this person has lower likelihood of a bad outcome if they wait than another person waits.

So, the concept of triage and prioritizing patients, and figuring out who
needs to move, and making capacity, you know, is exactly the same, it seems to me. I don't know if it would actually improve your preparedness, but I would think that sort of couching that, couching the two together would increase the interest in preparedness and possibly have some bearing on improving the boarding issue. It just seems to me that they are both kind of -- one is kind of sexy, but a lot of us aren't involved in it, and the other group is like, oh, that is day-to-day operations. But we need to get more people involved in that.

CO-CHAIR PITTS: So, I had a great experience for six months working as an ER doc in New Zealand. Just to illustrate this plasticity of one's thinking, when I got off the plane in New Zealand, I discharged all my chest pain patients, regardless. When I got off the plane in the U.S., I admitted all of them to the obs unit.

(Laughter.)
So, what you do in a disaster, you just change your moral compass, basically, and you do things differently because it is a disaster. I think every individual working doc would do that, regardless of what you told them to do, once people started to pile up.

MR. PINES: Just to extend a little bit on what Melissa said, I think one way to think about this is in terms of flexibility. So, essentially, I think that the way to bring this together from the crowding side and the preparedness side, essentially, what both are asking for is a health systems flexibility to maintain the same quality of care, not actually changing those standards of care on days when you have, like Brendan said, 20 percent more versus 2 percent.

So, essentially, when you have this really abnormal day, can you maintain the same level of service in the hospital? I mean, I think that sort of brings it together.
What we know, I think, empirically, and I think most of you in a lot of your work you have shown this, there are days when there are more people in the emergency department. As there are more people there, the length of stay for everyone goes up. Essentially, that is an example of a system that is not flexible; that, essentially, on those busy days we consistently can't flex-up our services to basically maintain the same level of service, which would be length of stay in the emergency department or quality of care.

I mean, I think that is a lot of what we have shown, is that during those really busy days, that quality of care suffers. Essentially, extending that to disasters, what we are trying to do is on one of those ultra-busy days, when a disaster happens where we actually need external resources to really help us, can we maintain the same level of service.
I think, conceptually, that is what we are trying to do, is really link these concepts together. How do we maintain the same quality during a surge of patients? I mean, basically, it works on a micro-level where you may have the bus crash, and also at a macro-level when the Aurora happens.

MEMBER FIELDS: To your point, I think you are probably right to be uncomfortable about the way parts of the hospital system would break down that are used to being very efficient. I am thinking about scheduled surgeries, acute rehab. Anything that is predictable and probably happens during business hours in the hospital, I think that part of healthcare breaks down very rapidly in these extraordinary black swan situations. And probably a lot of those resources and those providers become irrelevant.

But I think, at the risk of being political about this, emergency nurses and
emergency physicians and the people that back them up, and acute hospitals participating in Medicare, have basically been working with a profound capacity problem, a profound mismatch of resources and services since at least 1986, if not a lot longer.

So, one of the parts of the Venn diagram where you will have overlap is I have absolute confidence, just thinking about who is around the table, that when you have an extraordinary event, the people who will most predictably be there to help with the response at the grassroots level will be emergency nurses, emergency physicians, pre-hospital personnel, the rest of public safety, and a coalition of the willing from medical staffs in the community, using whatever they can bring to bear in terms of resources.

I think the part that is probably political about this is I think that, since it is pretty well-established that crowding is sort of a weird metric that demonstrates the
lack of equity and smoothness about the way services get provided in the acute care sector, I think if you acknowledge that and you can see how delays in treatment result in bad outcomes, and that is also a metaphor for lack of access to coverage, I think what we are asking for is to recognize that there is a kind of calculus here, and that if you are willing to look at some of these kind of ordinary, everyday disasters in both rural and metropolitan areas, that the summation of those represents what the society has to do to respond to a black swan event.

One of the things that is universal about that is that the people around the table and the people in the emergency department will be the interface between the event and the response, along with pre-hospital personnel. And I feel really good about that. But I would feel better if we could use the high priority society has given to disaster response for extraordinary events.
to help solve everyday problems in hospitals
operating in marketplaces with mandates to do
better for their own populations of the
communities.

CO-CHAIR PITTS: Yes, Mike?
MEMBER STOTO: I don't like the
black swan analogy.

(Laughter.)

I think that the author mixes up
black swans and black squirrels.

(Laughter.)

Black swans are totally different
from everything that has happened before.
Black squirrels are just a little bit grayer
or maybe a lot grayer than normal squirrels.

I think that a lot of the kind of
emergencies that we are concerned about are
more like black squirrels. They are just a
big version of some of the things that we see.

If we think about it that way,
then we really can build on the day-to-day
lessons. If we think about it and there is a
black swan that is just totally different from anything we have ever seen, there really is no way to learn from that. I think that the only hope we have is thinking about these things as black squirrels.

CO-CHAIR PITTS: So, squirrels are smaller than swans. Is that the difference?

MEMBER STOTO: No, no, squirrels are gray, but black is an extreme form of gray.

(Laughter.)

CO-CHAIR PITTS: Would you consider Aurora a squirrel or a swan?

MEMBER STOTO: A squirrel, yes. I mean, unfortunately, people are shot all the time. It is just that a lot of people were shot.

CO-CHAIR PITTS: What are some of the scenarios -- I know there are a number of scenarios that you are interested in -- that are not squirrels, that are big things? You are talking about influenza?
MEMBER McCARTHY: SARS.

CO-CHAIR PITTS: SARS?

MEMBER McCARTHY: SARS is a perfect example of we had a lot of resources, but we couldn't contain. They could not train personnel to contain an infection. They actually had to set up one hospital to treat all SARS patients because they couldn't have all hospitals treating SARS patients because they kept getting infected.

So, we learned from that, that even though we had resources, we didn't have the processes in place to handle that kind of thing that we weren't used to doing. So, that is what I worry about a little bit here.

Trauma is our easiest-case scenario, but what happens if it was infectious or it was some kind of chemical hazard or radiation?

CO-CHAIR PITTS: Something that can hurt the providers themselves.

MEMBER McCARTHY: Hurt the
providers themselves. We don't deal with
those situations.

MEMBER STOTO: That is not unheard
of or unusual. To deal with SARS, they did
surveillance. They did infection control
procedures. They did it at an extreme level,
by setting up extra hospitals. But the
fundamentals they did were the same kinds of
fundamentals we deal with in any kind of
infectious disease.

MEMBER MacINTYRE: But these
fundamentals changed the way in which the
healthcare organizations operated completely.

MEMBER STOTO: Yes.

MEMBER MacINTYRE: They changed
the ways in which these healthcare
organizations operated completely versus a bus
accident or the Aurora shooting; you heard it
was over in four hours. So, it is a much
different animal.

MEMBER McCARTHY: But it took
months for us to get SARS. We were lucky that
it wasn't highly infectious. Had it been --

MEMBER CARRIER: Maybe it would be
helpful, because disasters can present in such
diverse ways, maybe it would be helpful to
think sort of generically. Like if you
imagined a plot where you are X-axis was the
level of surge and your Y-axis was, let's call
it not just acuity, but maybe level of threat
to take into account. You know people are
shot. It is obviously tragic, but it is over,
versus something that could spiral and spread.

There are definitely disasters
that are high on both axes. In that case,
maybe, yes, all bets are off. You really want
to think -- you know, you may need to abandon
processes you thought would be useful. Maybe
the only benefit that the kind of coalition-
building we are talking about would have in
these situations is to build strong
relationships, creating a framework where
people could think creatively.

But there is a lot of stuff done
on this end of the axis, high surge, low acuity. You know, I trained in New York City. Anthrax in New York City was three incredibly-sick people and a million people wanting Cipro. I mean, to me, the overlap there between disaster and surge is very strong, and I could definitely see a validity to combining measures or thinking about measures in the same framework.

But maybe in other areas of this plot maybe trying to measure preparedness through the lens of surge is not such an effective approach. Or maybe we need to think about different ways, measuring the strengths of the relationships generally, rather than the creation of particular structures or processes that may be less relevant, given how weird things can actually get.

CO-CHAIR PITTS: I think that is really a cool observation. I wonder, it is so cool, that I wonder if it hasn't already been thought of. Has somebody --
MEMBER CARRIER: It is very possible.

(Laughter.)


Brendan?

MEMBER CARR: So, I don't know if this will help or hurt. But I continue to be intrigued about whether or not it is crowding/boarding that we think doesn't fit into some sort of larger, summed-up to the level of we still sort of don't know geography coalition or if it is everything about measurement of emergency care outcomes that we think doesn't fit.

So, I guess I would say, to move this into a totally different space -- and I get that we are not talking about this today -- but to maybe take us outside of where we are to say, do we instantly say that doesn't apply, either?
What if I looked at -- I am from Philadelphia -- what if I look at the City of Philadelphia. Right now, I care about my outcomes at the level of the hospital. If I were to now say, well, you know, I travel all over Philadelphia. Sometimes I am up north; sometimes I am in the center of the City; sometimes I am in the west, and sometimes I am just on the outskirts, you know, in the sort of southeast.

I would like to believe that there is some coordination across that region, so that if I have unplanned critical illness, my outcomes are similar. I would like to think that the hospitals work together to make sure that, if I show up at a hospital that can't take care of me, that they have thought of a plan to transfer me or that they have thought of a plan to bring resources to me in some capacity, so that I do okay.

So, I guess what I am saying is, are we totally uncomfortable with the idea

Neal R. Gross & Co., Inc.
202-234-4433
that we should be measuring outcomes at the population level? Or are we saying, okay, that might be okay; it is okay to measure things at a population level, to make hospitals in a region mutually-accountable for my outcome, but boarding is different?

Maybe that complicated things, in which case we could just go on and pretend I didn't speak.

(Laughter.)

But, to me, it feels like we are sort of saying we understand that it is a competitive market. We also think that there are times where you need to cooperate. And hanging up a billboard that says, "Best cancer care anywhere at our place. Come to our place" is different than unplanned disease.

CO-CHAIR PITTS: I heard you and I understand you, I think, but I am not sure what the answer is.

David?

MEMBER MARCOZZI: Yes, I mean, I
just want to echo, in the interest of transparency, Brendan and I have had this discussion before. I mean, the concept that he just promoted in and around trying to figure out a regionalized system and trying to get to the measure for unplanned acute illness, so that the measure affects that community, that health community, and that they can respond appropriately, deliver the right care at the right time to the right patient, that I think is one of the pivotal pieces.

Because, I will be honest with you, we stand on, what preparedness stands on is the ability to deliver -- you know, STEMI care is now, we know where we are going to transport, the closest cath place. So, we stand on, the preparedness folks stand on that ability, that system of care of delivery today. If we get that piece right that Brendan just spoke to, our job gets a lot easier.
Because you don't all of a sudden pull the plan off the shelf, turn to page 4 and go, okay, now that we are in disaster mode, these are the things that I need to do. You would stand on what you do today right now when we are responding to a disaster, likely. I mean, granted, there are shades of gray or squirrels of gray or black. But, to get that right, that is the foundation of what preparedness stands on.

CO-CHAIR PITTS: Am I missing somebody? Yes, Anthony?

CO-CHAIR STONE-GRIFFITH: I think there is a temptation to look at this as sort of two ends of a spectrum, daily crowding surge versus disaster surge. I am falling into the camp of they are very much different beasts.

The assumptions, the motivators, the willingness to bend the rules, if you will, are completely different during a disaster, however you are going to define
that, than day-to-day surge. The motivators
during day-to-day surge are, quite frankly,
very often economic. I am sure Jesse can
speak to that more elegantly than any of us
can, and we will hear about this this
afternoon.

But when you had your Aurora
incident, you bent the rules. And suddenly,
it was magically okay to have patients in the
hallway on the floor as opposed to in the
hallway in the ED.

I think when we look at those
assumptions and motivators and willingness to
comply with regulations, it is going to be
very hard to have measures that equally
reflect preparedness versus daily surge.

The other piece that is
interesting to me -- and it gets to your gray
squirrel concept -- is if we are responding to
these smaller disasters adequately, we should
have that different management team that
somebody was talking about. It would be awful
for me in my hospital if I had a mild power outage, which we did have two years ago, and I tried to manage that through our usual committee method, where we are all going to get together and sort of causally discuss things, and we will have an answer in a week's time. That doesn't work. You have to have different methods of managing those types of incidents.

And so, for me, the real interesting, almost academic, somewhat operational question is, at what point does daily surge become a disaster? And there are some interesting stories out there of people, not a big bus rollover, not a plane crash, but just a high-volume; they activate their EOP. They have a different management team in place, and they start to bend the rules and they start to put people in hallways upstairs. I think that, for me, is an interesting point worthy of further investigation.

CO-CHAIR PITTS: Terry?
MEMBER ADIRIM: Yes, great. Thank you.

I actually agree with my colleagues, Anthony and Emily, about the whole gray squirrel. I don't even know what those mean, but that these may be separate issues and it may be useful to think of them differently.

But, to get back to a little bit of what Brendan was saying, because, actually you brought up a couple of interesting points when thinking about how we measure whatever it is that we want to measure, I am just wondering the two things that I got out of what you were saying perhaps is maybe we need to look at some of the previous work done in coordination of care, seeing if there are any measures that may be applicable or can be adapted for this particular use.

And then, the other concept, too, is determining quality across a region. You know, I would imagine that you would want the
same quality of care across the region, which
is what I thought I heard you saying.

I am wondering if we are going
about this the wrong way. It may be better to
think of it as, what are we trying to
accomplish? What is the outcome? What is the
goal that you are trying to accomplish, and
kind of work backwards. I don't know if that
makes any sense, but that is sort of how I
would approach it.

CO-CHAIR PITTS: Are you talking
about looking at disaster responses that
failed?

MEMBER ADIRIM: Well, no. I mean,
well, there are two separate issues here. I
am talking more about the crowding and the
type of care that you are expecting within
region, either during a disaster or with high
numbers of patients or just every day. What
is it that you are trying to accomplish? What
are the outcomes? These would go to
population kind of measures. So, kind of
applying the same standard, for lack of a
better word, across healthcare institutions
and kind of looking at it from what are you
trying to accomplish, you know, coming to
outcomes, as opposed to looking just
distinctly at the processes to some outcome
that you haven't defined yet.

Does that make sense?

MEMBER MacINTYRE: At least it
does to me.

MEMBER ADIRIM: Yes.

MEMBER MacINTYRE: And I would
piggyback on that and say, what are you
anticipating to respond to? Again, I think
much of the conversation trends towards surge.

Our facility where Jesse and I
work is much more likely to experience a power
outage, for those of you from Montgomery
County and the surrounding area, than it is to
have a surge. I am sorry, Jesse, but you are
not going to be able to order that CAT scan
when we don't have power because our CAT scan
is not hooked up to backup power.

So, I think we need to broaden this conversation because there are things that are much more likely than the bus rollover and they are just as, if not more, important to address.

CO-CHAIR PITTS: Brendan?

MEMBER CARR: I am tempted to sort of respond to that because I do think that we have to work within the confines of what we been asked and funded to do, which is to think about whether or not there is intersection in this space. There are lots of important things to address. I am not sure that we can change the mandate of what NQF has been asked to do here.

But I guess the other piece, to talk to Terry, that I wanted to say, was I think that we all agree that we want quality for the region, quality for the U.S. to improve. The way that we have gone about that is by looking at the thing that was easiest to
look at, to some things at the level of the hospital. That is where the data comes from. It is where the accountability is. It is where you can set some sort of financial incentive or financial penalty.

But we believe that the rising tide will rise all ships. So, we are trying to incentivize regional health, national health, community health, by incentivizing hospital-based performance.

It is unclear to me why it feels so Herculean, not from an operations standpoint. I agree, as Dave said, telling you how to boil the egg is not the game plan here. The game plan is to sort of say, at the end of the day we would like you to please produce a boiled egg.

So, operationally, day-to-day disaster, day-to-day crowding is very different than I am now going to bend the rules. Crisis standards of care exist for a reason.
I am still stuck on why it feels so different to say, "I want all the hospitals in my city to have good performance metrics on X" versus "I want my whole city to have good performance on X."

CO-CHAIR PITTS: Jay?

MEMBER SCHUUR: I guess I am going to comment on Terry's comment, which I really agree with. I think we are going to talk about a number of these issues.

I think it is very interesting to think about outcomes for regionalization because I think we will probably find things that are very different if we look at outcomes than if we look at processes.

And so, just thinking about my own health system in the region where I practice, I think there are a lot of things we are doing because of market forces, because of hospital integration, which are quite costly and add very little value, and if looked at at a regional level, at an outcome level, would
sort of change the perspective. And I am thinking about transferring patients, even the way we transfer patients for STEMI, transferring certain patients and things like that.

So, I think if it is within the scope of our charge, thinking about measures for regionalization, about population health, but particularly thinking about value and making sure that we are measuring outcomes as opposed to processes, because if there are process measures, that is what a lot of these systems are doing right now, is responding to the process measures that hospitals have and doing things that may shorten the door-to-balloon time, for example, by using helicopters, send a patient a few miles to get to a cath lab, where from a regional perspective for the outcomes for that patient, there are different processes that would be set up and probably have the same outcomes at lower cost.
CO-CHAIR PITTS: So, you are saying that from the standpoint of performance, I mean preparedness, the performance measures ought to look at integration issues and not individual hospital issues?

MEMBER SCHUUR: This is very much about Brendan's frame of sort of regionalized health. Those sorts of measures, I would emphasize measuring outcomes and thinking about value.

I fall into the camp that I think the preparedness issues and the measures for preparedness are going to be different than those for regionalized care. I think for lots of us who practice in the emergency department that comes from personal experience.

I was working the night of the Station nightclub fire in Rhode Island. If you had gone to the hospital the day before, even earlier that day, it would be very low on the level of operational efficiency, surge
capacity, all of those things, and the outcomes were remarkable. Of the 50 or 60 patients who showed up critically ill, one patient died, essentially, who wasn't dead when they arrived.

And that had to do with the people who happened to be working that night, in particular, the emergency department attending and the trauma surgeon who sort of organized everything, and there was a lot of luck involved. But I think the measures to capture that are going to be different than the operational performance measures.

CO-CHAIR PITTS: Arjun?

MEMBER VENKATASH: I think this distinction in the measures between preparedness and some of the operational performance measures that Jay just alluded to and what Brendan just said, which was our goal is to think about this intersection here, I think one of the things that is making that challenging is the directionality of the
issue.

So, if we took the STEMI example,
we could measure some sort of regionalized
ability to have transfer agreements. And
then, we can measure at a hospital level their
door-to-balloon time. And then, we have this
outcome of mortality. All of these things
kind of go in the same direction.

In the same way, we have been
talking about how, when crowding becomes surge
as a spectrum, and that goes in one direction.
When Jesse says, well, we should think about
how much flex we have, I think what becomes
challenging is that the operational measures
may not go in the same way of direction of
performance as the preparedness measures.

By that, I mean if we think of
whatever the unit is, be it hospital, be it
collaborative, be it region, that is high-
performing on some of the process measures we
have for crowding or boarding, that doesn't
necessarily mean that they are high-performing
for preparedness. In fact, high performance for something like boarding or crowding could just mean that you have got all your just-in-time processes down to a point where you have maximized where you want length of stay; you have minimized your amount of boarding time, but you actually have no flex in the system. And that is a system that doesn't have a high degree of performance on the preparedness side.

And I think getting at some of that tension of whether or not these measures are actually going in the same direction or against each other is important in thinking about how measures get developed kind of down the line from this.

CO-CHAIR PITTS: Brent?

MEMBER ASPLIN: I wonder if one of the reasons we are struggling at the low end of emergency or disasters, kind of low-level disasters and daily surge is just the whole fact that we have accepted boarding as a fact
of life in the country. We are talking about bending the rules for low-level disasters. There are no rules that prevent boarding. So, I think that is why it is just such a different mindset for us, because we have accepted this, I think, deviant system response, that it is okay to just stack patients in the ER. I think if there was a different mindset about that not being acceptable, we would not be struggling with the overlap of daily surge and lower-level-volume disasters.

CO-CHAIR PITTS: Suzanne?

CO-CHAIR STONE-GRIFFITH: Yes, I can't help but agree with you on that, Brent. I mean, we have been measuring this for a long time. Over the last two years, our number of holding hours or boarding hours has really risen, this last year, potentially on track for a 25-percent increase, where right before 2009 we had a 16-percent decrease.

So, we are actually back where we
started. And yet, operationally, we have sustained some pretty incredible improvements in things like length of stay and door-to-doc and volumes and things like that. So, there is some day-to-day acceptance that we board. And, oh, by the way, it is not just the emergency department; it is throughout the hospital.

MEMBER ASPLIN: Yes, imagine the overlap --

CO-CHAIR STONE-GRIFFITH: Right.

MEMBER ASPLIN: -- if, as part of its daily operations, hospitals always had to do some form of surge to prevent people from staying in the ED.

Now Ellen can point out the unintended consequences of that policy in the UK, but it certainly did affect operations. And from that standpoint, you would be using a lot of the surge capabilities on a daily basis to respond to folks in the emergency department, which just doesn't occur today.
CO-CHAIR PITTS: David Levine?
There are two Davids, and I was looking at
Levine. Sorry. You're next.

MEMBER LEVINE: Thanks.

I just wanted to echo what Jay and
Arjun and the piece of what we are looking at,
these right metrics that we are looking at.
The difference between like the day-to-day and
the disaster, the patient population also
changes. So, we are looking to pick those
metrics and some of those door-to-doc times,
or whatever.

We have to be very cautious
because some of those less-urgent patients
that are crowding our waiting rooms that we
have deemed are okay to wait, actually, when
things are going on, like possibly an Aurora
-- definitely it has happened in Chicago when
we have had bad shoots, and whatnot -- is, all
of a sudden, the waiting room is relatively
empty of the non-acute patients. They have
already self-selected not coming to the ED.
We have to factor that into the care provision metrics that we are going to be using or be very cognizant of especially those lower-acuity patients that may self-select and not come when a disaster is happening.

CO-CHAIR PITTS: Jesse?

MR. PINES: I just wanted to make quick comment. Just to sort of refocus us, again, the task today is really to come up with tangible recommendations for measure developers. Essentially, what I am hearing is, you know, I think that we are not totally there, thinking that crowding and preparedness are not really the same. But I think there is sort of broad agreement that they are related.

So, as you start formulating a lot of these comments, also think about, after this discussion for the last hour, what sort of recommendations can we make to measure developers who are thinking about making these measures, specifically in the context of really linking these concepts quantitatively?
CO-CHAIR PITTS:  David Marcozzi?

MEMBER MARCOZZI:  Yes, I mean, I

guess that jumps right off of what Jesse just

spoke to around tangible recommendations for

measure developers.  I think I am hearing

three different areas of interest within the

group.

The first is really what Brent had

spoken to around regionalized accountable

care.  Let me just echo what Brent said.  I

think that there is a Venn diagram here and at

the least they are associated; at the most

they are interrelated with regard to

overcrowding from regionalized care to

crowding issues to emergency preparedness

issues.

But let's speak to one very

tangible or try to bring a very practical

example.  Let's do a man-down drill right now.

So, who owns that patient right now on the

street of D.C.?

The reason why I think we are
challenged with this discussion around regionalized care, regionalized acute care, is because there are so many parties involved with that. The public owns a piece of that. How fast? Law enforcement owns a piece of that. Is there an AED close by and are they in the rigs. EMS owns it, and EMS is not engaged as much as it should be with regard to this. And EMS really owns a majority of the man-down drill issues.

Then, lastly, the piece of the man-down drill is our healthcare facilities. So, that speaks to the audience or the potential components in and around those regionalized care systems.

And then, where is the data? Because if you want to develop what Jesse just spoke to, tangible recommendations for measure developers, well, then, where is the data? We are close. We certainly have a lot of -- and Ryan could speak to this better than I -- a lot of hospital-associated data that could
potentially look at outcomes. We are there. We are close. We could probably improve, but we have got some good data there.

The EMS piece we actually have, and no one has really had a discussion. It speaks to what Jesse rolled out in his -- NEMSIS is already there. We can, and we have the ability to, start to collect data from the pre-hospital environment. So, we actually have to make probably NEMSIS more robust and right-size-fit-it and ask the right questions to it, but if you link the NEMSIS data with the outcomes data that Ryan has got visibility on or that we all have visibility on, then we can actually speak to that regionalized accountable care -- regionalized accountable care? -- regionalized acute care, providing acute care in a regionalized format. So, that is first.

The second is crowding and the linkage between crowding and preparedness. As I said, I think there is a Venn diagram here.
But I don't necessarily, not entirely, although I am hearing, I am not entirely in the camp that they are two separate and distinct things. I think that they are interrelated, but what I think Jay just spoke to on a busy day and the ability to transition from a busy day in the ER to a disaster that he just spoke to and responded to, I think speaks to the change that we are having trouble grappling with. And that is a change from individual-based healthcare systems and conventional delivery of care today to population-based systems.

And you have a bunch of clinicians in the room. It is very difficult historically for that conversation to figure out really, in essence, that is what we are talking about. The tipping point of when we go from individual-based healthcare and overcrowding, and the pneumonia who is on the floor who could have gone home, but he is going to keep the bed, and the ER still is
overcrowded, and we potentially even have a
sicker patient in the ER who still can't get
the bed because the pneumonia could go home,
but he is up on the floor, and the doc hasn't
written the discharge instructions yet, versus
breakpoint, trigger point, population-based
care where that patient immediately gets
discharged. Change in care, the sicker
patient then goes upstairs and deserves the
bed.

And that is why we are having such
a challenge in this discussion, I think. We
are waxing and waning from really the
fundamental issue here of a paradigm shift.

CO-CHAIR PITTS: So, we are going
to sort of migrate back to Jesse. He had a
couple of comments before we -- I guess we are
breaking at noon, right? That is my
understanding. So, let's a little bit more
about this topic, and then we will let Jesse
take over for a few minutes. I was going to
let everybody put in their two bits and then
we will sort of fade out toward Jesse.

MEMBER ADIRIM: I just have a

quick question, going to what --

CO-CHAIR PITTS: Yes.

MEMBER ADIRIM: -- not a question,
a comment with regard to what Jesse has asked
for, ideas for how to frame the
recommendations.

You may want to consider more than
one set of measures for regionalized emergency
care because what is coming out of this to me,
what it sounds like to me is that there is
multiple different sets that you could create
for this issue. So, that would be No. 1.

No. 2, I would encourage a look
at, with regard to those sets, specific
emergency department sets with regard to
coordination of care, whether or not you
include preparedness response in that
coordination-of-care set. You know, you can
kind of think through that. But it would be
coordination of care.
And then, also, really, again, repeat looking at this from kind of an intermediate outcome or outcome-based-type measure. So, that would be across a system, as opposed to just institutions. So, those would be a couple.

And then, I don't want Brent's point to be lost on the boarding issue, not that I think that should be a separate set, but that point shouldn't be lost and that there should be a development of measures that go to the patient's experience. Because I know NQF is also concerned about patient experience from their point of view, and boarding may not be in really good emergency departments an issue with regard to the quality of care because we do excellent care for those that board, but it really is a bad experience for the patient.

So, I think it is something that is definitely worth measuring, and if it is measured, may affect changes in how hospitals
operate if there are carrots and sticks associated with it. So, I didn't want that point to be lost.

CO-CHAIR PITTS: Ryan?

MEMBER MUTTER: So, I have been thinking about Jesse's question about recommendations sort of in the context of the Einstein and Deming quotes at the very beginning.

Something I am not clear on is, should we recommend that developers have a path to incentivize improvement in the things that they measure? So, for example, I am thinking about unplanned critical illness and sort of care at an area level, not just how good this hospital does, but how good the area does, including hospitals in that area who are competitors and may question why should they take steps to improve care at the area level, which is going to make their competitor look better and make them look better, but it will make their competitor look better. It may not
be worth it.

And so, I don't know important in
terms of recommending -- are we okay
recommending measures that are important, but
where an incentive to improve may be hard to
get to, hard to identify, or do we feel that
it is important that there is an identifiable
incentive that can be used to effect
improvement in that thing we measure?

CO-CHAIR PITTS: Helen?

DR. BURSTIN: Okay. Thank you.

So, I may be the token primary
care/general internist in the room, and just
a thought off of Terry's comment about
integration. I think there is a real
opportunity to think very broadly about some
of the measures we already have and bring them
up a level in terms of level analysis. So,
just a few thoughts on that.

There are a set of transfer
measures that were submitted years ago looking
at transfers for patients with STEMI. I mean,
those might be very good measures to think about going up a level in terms of analysis to community or region. They already exist. Again, thinking about that as a starting point.

You know, all the measures -- and Ryan may be able to speak to some of the newer work on the avoidable ED measures, if that is still happening. But there are avoidable hospitalization and avoidable ED measures I hope in development at AHRQ that are community-level indicators. Again, those are not necessarily something -- I think we have heard lots from providers who feel that that is not directly something for which they are accountable solely. It has got to be a community-level indicator of access.

So, those are the kinds of measures I think would also be useful both for just day-to-day operations, but also surge. Because if you are in a community that is doing a better job of not having patients who
don't need to be in the ED in the ED in the
first place, then you potentially have more
room on a regular basis to bring in other
patients.

The other thing is those ED
throughput measures, which I am sure many of
you don't terribly love, that were endorsed by
NQF as well, about time in the ED to admit
decision, but there is a piece of that puzzle
that is missing, which was we don't have the
second part of it from admit decision to being
on the floor. I mean, those are very logical
ways to get at some of the boarding issues.

Just one analogy, that the folks
at the Office of the National Coordinator have
been working on a measure looking at closing
the referral loop. So, you refer a patient.

Did they get a note? Did they send a note
back, et cetera? It just seems like there is
a lot of the sort of missing pieces of the
loop here that we haven't yet factored in that
we could take existing measures and build off
of them, and perhaps get that admit time to boarding time.

And then, lastly, you know, going back to Terry's point about patient experience, admission through the ED is a variable on HCAHPS. So, there is a real opportunity and a nice research study there even as well just looking at, if you can actually figure out from a hospital level patients who came from the ED, are there variable impressions for them overall in terms of their impression of the hospital, based on admission through the ED, another way to potentially get at some of those issues?

So, I think the issue of integration and thinking broadly about measures we already have that could be adapted and modified to help satisfy these issues, and perhaps not always assume they have to be at the provider level of analysis, which gives people hives sometimes, but think about bringing them up a level, where I think you
won't have perhaps as much of the pushback
that we get about the issues of shared
accountability.

CO-CHAIR PITTS: HCAHPS, what is
HCAHPS?

DR. BURSTIN: Oh, I'm sorry. It
is the hospital experience-of-care survey that
is mandated.

CO-CHAIR PITTS: Okay.

DR. BURSTIN: It is actually now
incentivized for every hospital.

CO-CHAIR PITTS: Right. I have
heard about it.

Let's finish these things, and
then I will let Jesse take over and say a few
things.

Ellen? Mike? Sorry.

MEMBER STOTO: Okay. Thanks.

Two points. One is that most of
the NQF-endorsed measures really have to do
with, are defined in terms of some proportion
of patients having a good thing happen to
them. That works fine in most cases.

But I think that Dave is right.

When we talk about regional preparedness, the unit really has to be at a higher level than that. I think it is a different paradigm that we really need to come to grips with.

The other point goes back to the question about developing measures. We have to think both in terms of what to measure and opportunities to measure. When we have lots of patients, we don't worry too much about opportunities to measure. We focus on what is the right thing to measure. But when we are talking about emergencies, we don't have a lot of opportunities to measure things.

So, that is why I think, to the extent possible, that we can measure things in more routine settings that have a bearing on how the systems respond in emergencies, that would be helpful, because there are more opportunities there.

CO-CHAIR PITTS: You didn't want
to talk, Terry? Okay.

Ellen?

MEMBER WEBER: This is not specifically to recommend particular measures, but I did want to say something about the idea of the individual versus population-based care, because I actually think in many ways many emergency physicians and probably others would like to be able to think that way. Having some kind of measures that allow you to say, "This person really doesn't need to be in the hospital, and the government stands behind me and the cardiologists stand behind me" -- (laughter) -- and everybody is saying that this is not an indication for a hospitalization, would take a lot of onus off that individual physician.

I mean, when I was in England, there was a lot of, in the emergency department, is this a good use of NHS dollars? Okay? The people actually thought about the fact that, when you have like this four-hour
target, that actually one way to deal with getting patients out of the emergency department within four hours was just to admit everybody, and then the hospital has to deal with it.

But they realized, one, that was a really bad use of resources and, secondly, in the end it would wind up blocking their beds. So, they didn't go that cynical route.

I think that idea is, you know, one of the things that I think the American healthcare system -- if I may wax poetic for a minute -- we always think we have unlimited resources. We are always dealing with limited resources, and we are always letting somebody wait, so somebody else who is sicker can go forward.

I think we actually have to agree -- I don't think in a disaster we are going to be able give the same level of care to absolutely everybody the way we do when there is no disaster. But, nevertheless, I think
the principles are the same, that we should be having more mindset all the time about what does the admission have to do with the surge capacity that I might need any day. So, I think there is a relationship there that we should be sort of maybe thinking about in our measures.

And getting to the issue of process versus outcome, I completely agree that the outcome measures might be different for preparedness and for boarding. I believe a lot of the process measures, what do you have in place to anticipate a problem, goes for all of this. And so, that may be where the distinction is.

CO-CHAIR PITTS: Okay. Last comment, Ryan?

MEMBER MUTTER: I think Helen invited me to give a very brief update on some of AHRQ's work in this area under our Quality Indicators mechanism.

So, we completed a project where
we took AHRQ's inpatient prevention quality indicators, which is basically a measure set that uses what is going on in the hospital to get a sense of what is happening in the ambulatory care setting. We took sort of those inpatient-oriented Quality Indicators and tried to see if we could expand them to be applied to ED data. When I say "ED data" in talking about AHRQ, what I am talking about is sort of administrative data based on bills. That work has been completed. We are going to have a working paper that we are going to post on our Quality Indicators website very soon.

AHRQ has just begun a second project -- the contract has been awarded to Stanford -- to do some more measure development work looking at community. It is ED Quality Indicators, but, again, it is not looking at individual hospitals and looking to assess care in the ED. It is basically using the ED as a window into the healthcare system to basically look at what is going on inside
of the ED as an indicator of what is happening outside of the ED.

As I am sure many of us do, I really like Brent's model here, which is on figure 1, page 7, of the Draft Report. As I think about the work that we are about to be doing -- is it about to magically appear on the screen? That would be amazing.

(Laughter.)

Page 7. Scroll up just a little bit, a little bit more. Oh, too far. Okay.

This contract has just been awarded. We have only just had our first preliminary meeting. So, what I am saying now is preliminary and should be taken as such.

But my take is that our focus is mostly going to be, if you look under the input column, is mostly going to be in that safety-net care and unscheduled urgent care space, is probably what we will be doing. And again, it is going to be community measures.

So, that is the update.
CO-CHAIR PITTS: Okay. And,

Jesse, did you want to be last before lunch?

MR. PINES: Sure. Just some brief comments, and I never want to stand between a big group and lunch.

But, again, I just wanted to thank everyone. I mean, just such a great discussion, and I think we have really a lot of very tangible recommendations that are going to come out of this.

What I think I have heard so far is, thinking about preparedness from a measurement perspective, there are some potential ways to measure whether or not we are prepared. Potentially, tabletop exercises, thinking about some structural measures of what kind of stuff that we have in the event of a disaster, and, also, thinking about, what Anthony said before, this is sort of measuring response, which I think really, from a crowding perspective and a preparedness perspective, is really very different and
actually may use very different methodologies.

From a preparedness perspective,
the measurement of a response may be actually
very similar to a lot of the ways that we
would actually measure preparedness; for
example, using validated survey instruments,
you know, sort of after an event happened.
Because, like Mike Stoto was saying, we really
have no counterfactual, we never really have
a control group for that. So, really, you do
need some sort of a qualitative or rigorous
qualitative assessment that can be calculated
in a quantitative way, and that would be
potentially through some sort of a survey
methodology to make an assessment of that.

And then, on the crowding side,
really thinking about our crowding measures in
the context of preparedness. So, the
afternoon is going to be about thinking about
crowding measures, but, also, I don't want to
stop the discussion of preparedness, you know,
thinking about how we could potentially link
the existing measures of crowding to the extent of what Helen said, I think, which is fantastic, thinking about current measures that NQF has that could potentially be taken to a different level that would start to think about like transfers at a higher level and really linking the concepts together from a measurement point of view.

And then, I think those were my basic comments. Essentially, I just wanted to say that I do want to continue this discussion in the afternoon, really talking about specifically boarding and crowding with a preparedness lens, which sort of the morning was preparedness in a boarding and crowding lens.

And without any final questions, final comments --

MS. FRANKLIN: No, I just wanted to pick up on Dr. Adirim's comments about the framing questions. As we continue to think about this through the day, what are the
outcomes? What does good look like that we expect to see from the measures that we are going to be making recommendations about? So, just keeping that in mind as we continue our discussion.

CO-CHAIR PITTS: I think you are okay to go to lunch now.

MS. FRANKLIN: Lunch has not quite appeared.

CO-CHAIR PITTS: I'm sorry. I noticed there is a public comment.

MS. FRANKLIN: Yes. I'm sorry. If we have members on the call or public on the call, now we would like to hear some comments from them.

THE OPERATOR: At this time, in order to ask a question, press *, then the number 1 on your telephone keypad.

We will pause for just a moment to compile the Q&A roster.

(Pause.)

At this time, there are no
questions.

MS. FRANKLIN: With that, I guess we can go ahead and break for -- oh, sorry, Terry.

CO-CHAIR PITTS: Yes, go ahead, Terry.

MEMBER ADIRIM: Very quickly, with regard to framing, I would encourage somewhere in this document that, whatever measures are developed or whatever sets are measured, that they do keep children in mind because we are so used to having to be retrofitting with regard to these kinds of things. I just want to make sure that, especially when it comes to capabilities and measuring stuff, it is different in kids. So, that is all for me.

CO-CHAIR PITTS: Okay. Lunch will be here soon.

(Whereupon, the foregoing matter went off the record for lunch at 11:57 a.m. and went back on the record at 12:48 p.m.)
12:48 p.m.

MS. FRANKLIN: Before we get started, I would like to check the line.

Arnika, I wanted to check to see if we have a Dr. Timmons on the line.

MEMBER TIMMONS: Yes, I am here.

MS. FRANKLIN: Oh, great.

Dr. Rapp on the line?

MEMBER RAPP: Yes, I am, Angela.

MS. FRANKLIN: Okay.

MEMBER RAPP: How are you?

MS. FRANKLIN: Great. Thank you.

I just wanted to make sure.

If either of you want to weigh-in, please feel free to do so.

MEMBER TIMMONS: Thank you.

MEMBER RAPP: Sure.

MS. FRANKLIN: Okay. Thanks.

CO-CHAIR PITTS: I think Jesse will start off talking about a subject he has already got prepared. Thank you.
MR. PINES: Great. Thanks.

So, essentially, what I would like to do for the next few hours here is talk a little bit about some of the measurement issues in crowding and boarding, but also really not lose the frame of preparedness. Essentially, we are going to basically go through some of the current measures that are out there, talk a little bit about what happened back in 2008 and the measures that were not endorsed.

We did have a fairly robust discussion around Brent's input/throughput/output model. Actually, I think it is fantastic that AHRQ is actually planning on looking at the input side, specifically the unscheduled urgent care and safety net, and developing some measures around there.

And then, finally, what we are going to do is, and I think really the bulk of this, is going to go through each of the
recommendations in the current draft and essentially think, are these the recommendations we really want to have for measure developers? In the context of the discussion this morning, is there any way we want to modify that, modify those recommendations?

And also, at the end I do want to spend some time going around the room and thinking about other recommendations. And other recommendations can be in the context of this morning, and are other recommendations that we didn't think of for the preparedness section? Or are there other recommendations specifically on crowding and boarding and things we would want to develop for measure developers?

So, with that, I would like to basically start with thinking about crowding measurement in general. This is something that I have spent many years thinking about and studying.
There are basically two different ways to measure crowding. One is from the perspective of the emergency department looking at measures such as occupancy. We all know, coming into an ED shift, that if you come to the waiting room and there are 25 people in the waiting room, that it is going to be crowded in the ED. That is A, and there are going to probably be potentially issues getting patients out of the emergency department.

But the issue with measures like that is it is difficult to generalize across hospitals. And really, those are point-in-time measures. A lot of the work that Melissa has done has really looked at the association between various levels of census and length of stay. Essentially, what the literature shows is that at different censuses between different hospitals that is associated with highly-variable differences in length of stay and sort of gets at the ultimate question:
what is crowding?

My opinion on that is that, really, the best way to measure that is really looking at it from the patient perspective and looking at issues of, basically, the timestamps, potentially time to provider, but also looking at broader timestamps such as overall length of stay and the boarding time, and the other measures that actually were previously endorsed in 2008. Really take it from the patient's perspective as opposed to from the hospital's perspective in a point in time.

Really, I think that is where the future of measure development should go. It would be to look at potentially re-endorsing those measures; like Helen said earlier, potentially thinking about where the gaps exist. And do we want to start thinking about other intermediate timestamps such as when a patient was seen by a provider, when the decision was made to admit a patient, and
think about time intervals.

Some of the work that we have done has actually looked at specifically to admit to departure time across hospitals. And actually, as it turns out, a lot of hospitals will have a very long length of stay but a very short decision to admit to departure time, and also vice versa.

So, I think we have also got to think about making recommendations to measure developers that really prevent gaming and really thinking carefully about what boarding actually means. I think we are going to be having a discussion about the definition of boarding, which is really if you ask the Joint Commission, which is very different if you ask the stakeholders in the emergency care community. So, I think coming up with a uniform definition for that and a specific recommendation for that will be important.

With that, any questions or any other issues?
(No response.)

So, essentially, why don't we go ahead, then, and I would like to again sort of review Brent's input/throughput/output model.

Maybe, Adeela, you could go ahead and put it up.

So, this is really our framework for this section, which is basically using this conceptual framework to think about, from a measurement perspective, what measures we have and in an ideal world what measures we would want to have.

On the input side, Ryan had mentioned a number of measures that AHRQ is developing on the input side. Actually, that was one of the recommendations that came from the conference call. I don't think Ryan was actually on that.

I don't think, Ryan, you were on that conference call, but that was actually one of the recommendations, to think more broadly about input measures and output
measures, some of the measures that Helen had
mentioned regarding transferring patients.
And I think that one of the major things we
could think about would be taking some of the
existing measures and trying to fit it into
this framework, thinking about taking some of
the transfer measures and potentially
aggregating that at the level of the region.
Later on, we are going to be talking a little
bit about accountability and regionalization.

So, next, what I wanted to do is
talk a little bit about the recommendations
and start really opening up for discussion
around looking at the recommendations sort of
one-by-one, going through the document. We
are on page 9 here.

So, essentially, the first
recommendation we have is that: "Quality
measure developers should ensure the validity
and reliability of the data used for ED
crowding and boarding measurement." I guess
the broader question is, should we add
preparedness into that recommendation? And how should we frame that in a way that is understandable to measure developers? I mean, I guess that is sort of a general recommendation.

Go ahead.

MEMBER ADIRIM: I have a question about that. When you develop a measure, I mean, is it required that, before putting it through any kind of process, don't you have to ensure validity and reliability a priori? So, I am just wondering about this as a recommendation.

MR. PINES: I think that is true, but I guess for this recommendation, I mean, this is sort of a general recommendation that, in order to go through the NQF criteria, all measures would have to be reliable and valid. Are there any sort of crowding-specific ways we would want to modify that recommendation or just sort of leave it?

Yes?
MEMBER STOTO: I mean, I don't know the field, but what I get out of that is the sense that there are measures there that no one has looked at this yet, validity and reliability yet. So, is that right? I mean, is your sense that there are potential measures there that just can't get through the process because they haven't been studied with respect to validity and reliability? That is the implication of this to me.

MR. PINES: So, actually, in the 2008 process -- and I wasn't a part of that -- several measures actually did get through the process. I am not exactly sure, Helen, if the standards have changed so dramatically since then.

DR. BURSTIN: They are probably a bit more specific, a bit more precise. I don't think they have changed much. Actually, Suzanne was the Co-Chair of that project. So, if you have questions there --

MEMBER ADIRIM: Speaking more to
the measure developers should identify
already-existing measures that could be
improved and validated, to be ready for NQF,
I mean, something like that.

MEMBER MUTTER: Yes, see, I was
thinking maybe sort of along the same lines.
I was thinking what I think is along the same
line, which is sort of quality measures should
use data that is valid and reliable. So, it
doesn't sound like that you are sort of --
part of the process is do this big data
validation, but use data that is valid and
reliable.

CO-CHAIR PITTS: I will continue
my policeman role here.

(Laughter.)

Arjun?

MEMBER VENKATASH: Brent I think
is actually the coauthor on a paper that
looked at timestamped data, comparing actual
charts within the ED and a tracking system.
I think that actually follows this section
well, which is probably the guidance we do need to give measure developers, which is that some reliability testing needs to be done between the data source used for measure development and the intended data sources for application. Meaning that if it is developed out of chart review or if it is developed from an electronic tracking system that they use, we have to know that there is some fidelity of that in comparison to claims in which it could be derived from or another data system or manual abstraction. I think that is probably a more specific recommendation for a developer, and there is evidence base to suggest that it is not always reliable.

CO-CHAIR PITTS: Emily?

MEMBER CARRIER: I had a question maybe for Brent, or maybe for others, about the literature on timestamps. Have the timestamped chart data been studied in the setting of a quality measurement or is it more a study in general, just looking at process in
the absence of an ongoing quality measurement and incentive program?

MEMBER ASPLIN: The paper Arjun is referring to is the latter.

MEMBER CARRIER: Okay.

MEMBER ASPLIN: It is really around process improvement and not a quality measure per se. We just looked at active-versus-passive timestamps, and there is a lot of error in our active stamps that we have to do by signing up or doing something on the electronic record, and the gap between when we do that and when it actually happens, if you are using various active steps as proxies as seeing a patient, for example.

I don't know how we comment on this, and Mike is on the phone, but one of the issues on the admin decision time to departure, that a number of different parties, and most recently the Measures Application Partnership, had questions about was the ability to game that admin decision time. I
I don't want to relive -- this is an approximately 55-minute discussion. But we could get into this. I don't know that we really want to go down this whole pathway.

But there was a specific CMS directive that asked them not to use the order for a bed request as the admin decision time. Not universally, but by the time we grind out every last ounce of energy in the room every time I have had this conversation -- (laughter) -- people kind of begrudgingly come around to that probably is the closest we are going to get, is the actual order, because there is no way to quantify when a decision is made in our heads. And so, that might be one other piece.

If we look at reliability as part of a directive, and a recommendation from this project is to look at, compare alternative methodologies for determining when the admin decision time is for purposes of that particular measure.
I do believe that, because it is a nested measure within the overall ED length of stay for admitted patients, it is a subset of that larger measure, to me, the gaming is a little less of a serious problem.

Nevertheless, if we are going to have people have some confidence in these measures, we are going to have to have better data and understanding of how they perform.

CO-CHAIR STONE-GRIFFITH: So, Brent, if I could add to that, I recall that our group talked a great deal about the order to admit being a very specific time; whereas, decision is much broader. Bed management can be very specific and measured, but the bed management process and the process in the ED is not always in parallel.

I guess I would agree with you; you talk about it being nested. And yet, we are measuring these as individual things as opposed to looking at them in the context of that continuum. So, the length of stay, to
Helen’s point, from the time I make that decision until they actually get to their bed or their place where they are going next, wherever "next" is. I mean, we don't have the whole continuum. We have really focused on the ED and not the entire hospital process in that as well.

So, those would be my thoughts.

CO-CHAIR PITTS: Wes?

MEMBER FIELDS: Yes, I want to follow on what Suzanne just said. I think all of us understand that crowding at best is a trailing measure of hospital capacity. If we really want to connect the dots between the discussion this morning and your ability to actually more effectively use the resources in the hospital, you probably need this framework of measurements to be inside of something that is hospital-wide, whether it is in eDocs or something else.

We had a really entertaining experience. We passed a bill through the
legislature in California three years in a row
asking hospitals to use the eDocs because it
has been established as one methodology you
can use to dynamically manage system status
and bed status. And two different Governors
have vetoed that bill.

So, I am just speaking to the fact
that, as long as you are looking at ED-based
measurements, in a way you are misrepresenting
the management challenge as something which
exists inside that department.

CO-CHAIR PITTS: Brendan?

MEMBER CARR: I was wondering --
again, I will call them the "crowding folks"
because I am not really one of them -- to help
me understand what we missed with the three
endorsed ones. Because I feel like we could
get very far into the weeds on what
recommendations they need to make crowding
better, crowding measurement better, but these	hree, to me, you know, they feel like we are
80 percent of the way there.
Help me to understand why we need --

CO-CHAIR PITTS: Of the three different measures that CMS is proposing or is using?

MEMBER CARR: Using, right?

CO-CHAIR PITTS: Yes. Yes, using.

MEMBER CARR: 495, 496, 497. Do we need 10 more recommendations about crowding measures or should we move towards crowding to disaster, crowding to population?

CO-CHAIR PITTS: Yes, I will bow to the people who are expert.

Yes, Helen?

DR. BURSTIN: I was going to just reiterate what I said this morning, that I think what is missing there is admit decision to being in a bed upstairs, because that is the other piece of the bottleneck that we are never looking at. We only looking at the ED, when, in fact, a lot of those patients are staying in the ED because we can't make room...
for you upstairs. So, I think without that piece of the puzzle, it is hard to get the full picture.

I also think it would be interesting to see what we can learn about taking these metrics and actually trying to look at them at a regional level. Can those data do well at a level that is beyond the individual provider level, which I think is really important as well?

MEMBER CARR: The latter I agree is critical, but I am still sort of not understanding why my decision matters that much if I know ED arrival time and ED departure time.

MEMBER McCARTHY: Well, it is the hospital, what Wes is speaking to, I think, about it is really like Ryan's idea of these primary care access. You are measuring in the ED, but you are really measuring primary care access. The boarding time is measuring hospital access, right? We are capturing it
in the ED, but it is all about the hospital side. So, I think we do need that information.

MEMBER ASPLIN: Helen, what about 497 --

DR. BURSTIN: Because the admit time oftentimes doesn't necessarily translate. I think you mean the decision to -- I am not sure I know what "admitted" -- is admitted time when you are in a room upstairs in a bed or not? I think that is not clear to me.

MEMBER ASPLIN: I know the intent was the departure time to --

DR. BURSTIN: The proxy.

MEMBER ASPLIN: The proxy of being in a bed. The delta would just be whatever transit time that you had. We wouldn't call them depart from the ED until they are actually on their way to the bed.

MEMBER CARR: Can I have one more?

CO-CHAIR PITTS: Sure.

MEMBER CARR: Is the intent, then,
to tease out what is me, right, as a practicing emergency physician, ordering too many tests, ordering too many labs, being too slow to make a decision versus tease out the hospital, and the hospital not making capacity for me to admit my patients?

And I guess, if that is the intent, is that the right message? Are we all in this together or should we be splitting to decide who is at fault?

CO-CHAIR STONE-GRIFFITH: I would say our original intent was really to try to understand at the point where the ED has done everything that they can now what part is owned by really a variety of leadership forces within the hospital and that hospital capacity management.

I think what has changed sort of between the time we started this work to now is where holding and boarding was so significant in the emergency department, now it is in a variety of places. It is more than
just about the ED. It is about the PACU. It is about the cath lab recovery unit. It is about the interventional radiology. It is about wherever we can find a nook and cranny and put patients.

If this really needs to transcend to being patient-centered and about the availability of resources on the hospital side, should this even be broadened? Because now the ED is to some extent competing with boarding patients in the PACU or in some other place. So, then, there is another pecking order in play here a bit.


MEMBER PAPA: It is no problem. I think it is also, too, about the coordination of care, Brendan. We do have to own the piece in the emergency department where, if the physician or the provider or the nurses aren't coordinating their care, so that you can get it done in a timely manner -- and that piece
takes a long time, as you said earlier; some
hospitals have a long time from the time they
see the provider until the time the decision
is made, and others, the other way around.

How do you coordinate that? And
who does own it, so that you know exactly what
you are measuring and what you are going to
improve? Because there are ways to game the
system. You make a decision to admit, and
your decision is made at two o'clock. By the
time you get somebody to accept the patient,
it could be two hours later.

So, who owns that? Does the ED
own that because that two hours was a struggle
for you to find a service that wasn't capped
or didn't have this, or whatever, until they
have to leave?

So, I think it is really a
combination of the two and how we can best
coordinate it, so that for the patient the
experience that they have is the best that we
can give. And then, of course, the outcome.
CO-CHAIR PITTS: It sounds like we are having that 55-minute --

(Laughter.)

MEMBER ADIRIM: Well, what I was going to say, I was going to get back to what Brent was saying because I was struggling like Brendan was about why it has to be when the physician makes the decision. So, it sounds like that what you have moved to is that you want to look at it from an institution standpoint.

Then, it gets back to what Brent was saying; it is when the bed was ordered. I mean, right?

CO-CHAIR PITTS: I can tell you, from pouring through the NHAMCS database, which we have talked about HCUP -- NHAMCS, for those of you who might not know, is a sample -- it is not a sort of census -- of all ER visits. It is a sample that is nicely representative. As a consequence, it is a much narrower database, not as many cases, but
a much deeper database in terms of the amount of information per case.

It includes time intervals. Among other things, it includes time of arrival, time of departure, and, in theory, time seen. Time seen is very difficult to get at, as you can imagine, if you are looking at a true national sample, not people who report to you, but an actual sample of all ER visits. Because in a fair number of ERs, you have a hard time getting that data.

I think we decided on the time of bed ordering. Yes, bed ordering was the instruction to the surveyors. It is present; you can find in about 80 percent of the cases. So, that means 20 percent of the time you can't, and that means it is probably hard to find. That is what the national level of compliance might be with that item, if it is defined as the time the bed was ordered.

MEMBER ADIRIM: You know, not to keep hammering this in, but, again, if we are
looking at it from the patient perspective,
then the other two measures are more
important. It is when they get there, when
they see the doctor and when they get
dispositioned. So, I don't think they are
cognizant of when the decisions are made, when
beds were ordered. So, I think the other two
measures kind of -- right? I mean, wrong? I
don't know. Does it kind of capture what we
are looking --

CO-CHAIR PITTS: I will leave that
for the record, I am sure.

Okay. Jesse?

MR. PINES: So, I think one of the
reasons why CMS thought that the boarding time
was potentially game-able is based on some of
the data that we actually showed them where we
did a field test of these measures and
actually found that some hospitals will have
an average length of stay for admitted
patients, essentially, which is what the
patient sees, of six hours, but will have a
boarding time of two hours; whereas, some places will have an average length of stay of 10 hours and will have a boarding time of one hour.

So, really, it is sort of an artificial number in the middle. I think that it is actually very useful for hospital QA to see where the delay is. Really, the purpose of NQF-endorsed measures is for public reporting. I guess the broader question is, how meaningful is that internal QA measure, really comparing between hospitals? Because if you compare the hospital with two hours versus one hour, it looks like the hospital with the 10-hour length of stay is actually performing better.

CO-CHAIR PITTS: Sorry, I am blind. Wes?

MEMBER FIELDS: I just want to ask a dumb question that we can sort of answer later, but it relates to observation status, which I also think is something which has
changed a lot since 2008. Heavy pushback on appropriateness for admission from CMS for Medicare patients, in particular. So, it is not really clear to me whether or not the 495 and 497 are really only be used for patients being admitted to inpatient status or if there is an intention around observation services.

Because, potentially, I think there is a lot of good things that could be done both in longer-term observation units as well as short-term observation units in the ED. But I am really not sure how they fit with the measure and with the reporting.

CO-CHAIR PITTS: Ellen?

MEMBER WEBER: Yes, I just sort of wanted to speak to the value of this 497 because, although it may be sort of not what the patient sees, it is very important to changing the process. I think if you have to report it, and it looks like in our hospital three hours to evaluate a patient and another three to four hours to get them to a bed, that
looks pretty crazy.

Actually, Australia has a three, two, one rule, which is three hours in the ED and two hours for a consultant, and one hour from the time of admission to a bed. So, they just came up with it. I have no idea if it is evidence-based.

But it does suggest that, if you are doing 50 percent as the workup and 50 percent is getting to a bed, your internal processes are really messed up. I think there is a value to not just reporting them, but just identifying. Because, otherwise, what will happen is people will try to shorten the whole visit, and that is not what we want to do. We don't necessarily want to make the time in the ED where you getting the workup less valuable. So, I think it is important to say this is a bureaucratic part, and that is the part that also should be measured.

CO-CHAIR PITTS: Kathy? Kathy, yes.
MEMBER ROBINSON: Thank you.

I guess in thinking about this particular recommendation, I am struggling to think about crowding and boarding when there really has not been any discussion about another piece of that, which to me is ambulance diversion, the amount of time that it takes to offload patients from an EMS stretcher to an ED stretcher sometimes. And if we are talking about patient experience, a greater picture really encompasses those other elements.

CO-CHAIR STONE-GRIFFITH: Yes, so very true.

Brent?

MEMBER ASPLIN: This is a little response to Jesse's question. I mean, to me, if we could figure out a way that we were more confident about the reliability of the admin decision time, that would help us sort out trends over time in output-related factors, in hospital access, boarding, et cetera, and this
growing phenomenon that you and Steve and
others just nicely pointed out in the new
paper this year around how the throughput
section of the model is starting to drive a
lot of our congestion and problems. And that
is probably growing at least as fast, in many
settings faster, over the last six years,
three to six years, than the boarding piece of
it, with all the imaging, intensity of
workups.

I think that bucket is actually
going to have tremendous pressure on it with
the readmissions reduction program and with
the move towards global payments. Because I
can tell you that, as a Pioneer ACO, we are
going to try to work as hard as we can to
create alternative pathways to
hospitalization, kind of like we talked about
at our conference a couple of years ago.

I think this whole diagnostic
phase is going to go faster than the boarding
problem over the next 10 years. And knowing
admin decision time and having both 95 and 97 helps you sort out, because that is the pendulum, as Ellen points out, or the marker at which point there is a transition between the diagnostic pieces and the waiting-for-the-bed pieces. There might be some value in seeing that over time. That is my best guess as an answer to your question.

CO-CHAIR PITTS: Okay. So, that was about -- I am trying to get my own mind organized here -- that was about validity and reliability of these measures, or the data, I suppose.

And Recommendation 2 was that: "Developers should explicitly define the timestamp." So, which exact interval should we be talking about? And we actually talked on it already, I think.

Suzanne, I know I have heard you talk about that, that within HCA you have a fairly-uniform way of defining that, that is not the time that bed was ordered. Is that
right?

CO-CHAIR STONE-GRIFFITH: We actually separate it from bed-ordering time because we would like to encourage bed notification in the system. So, if I say, "Hey, don't tell anybody upstairs that we are going to need four ICU beds until the patient is wrapped up with a bow and I am allowed to put that order in," how do we commission resources upstairs? Back to that just-in-time, I might need to call a nurse in or not let a nurse go home early or think about other factors.

So, we think of them as very interdependent but running on a parallel path. So, as opposed to ordering a bed, we use an order to admission or an order to admit, which is what we think of as a complete statusted order. I have to have a date and a time. I have to have an accepting provider. I have got to have a bed type and the status of that patient. When I have that, then I put in the
bed order.

But that is not what is everywhere or what is being used. And I would even say in our own company it is a nine-step process to make this happen. So, it is by no means easy.

CO-CHAIR PITTS: Right.

Helen, how exactly is that defined in the NQF -- or you are looking for it?

Okay.

DR. BURSTIN: I have got the specs open right now.

CO-CHAIR PITTS: Okay.

DR. BURSTIN: It is very unclear, but we will share them with folks. There is an entire algorithm that is associated with the measure, but we will see if we can figure it out.

CO-CHAIR PITTS: So, I mean, it is really hard, if you are doing a national survey like NHAMCS, to instruct the surveyors on what to look for. It could lead to pretty
radically-different results.

I think there is a time that you decide to admit. There is a time the bed was requested. There is a time that a bed was received. And there is also the time when you have got hold them and intending to admit the patient.

So, the one that we have actually tried to use is the time when the bed was requested, right? Right.

And keep in mind, you know, this is not the UK or it is not Australia. We have a million different ways of documenting stuff in the U.S. And so, which thing is most likely to be found in a chart? I think in a paper chart, the old paper chart, we used to write "admit to," you know, time it.

What is everybody's feeling about that? There are quite a few people who still use paper charts. I mean, typically, nurses, the documentation is better than the doctors. You know, doctor in the room, doctor out of
the room with a timestamp there.

What is your impression about the
best way to get universal responses, given our
current system?

MR. PINES: So, also, just to
clarify, in the current version we did mention
two consensus groups that did actually come up
with consensus measures for a lot of these
timestamps, one of which was convened through
the ENA and actually a separate one through
EDBA. That will be included as part of this.

But, essentially, I think it is
important to have explicit definitions. I
think there is no disagreement there. I think
in the 2008 measures it is not totally clear
exactly what the decision to admit is, whether
that is the administrative bed or you see
someone, an 80-year-old with chest pain pops
up on the tracking screen, and you sort of
know that you are going to admit him. What
does that actually mean?

And also, just to expand the
discussion a little bit, I want to talk a little bit about the boarding time because this has been something that has been very controversial. Specifically, when does boarding start and when does this group really want to start boarding?

There have been a number of definitions out there. The Joint Commission Patient Flow Standard, the definition is four hours after the decision to admit. And that is also varied between these other documents from two hours to some of the documentation says that it is right after the decision to admit.

CO-CHAIR STONE-GRIFFITH: Jay?

MEMBER SCHUUR: So, just as a point of information, I think what 497, the current CMS specs are essentially that the timestamp for decision to admit represents the physician's decisions and actions thereof, but it is not the admit order. If you read the definitions, it is inherently difficult. It
is inherently subjective.

    So, my specific comment would be, could the recommendation be that the NQF or some organization comes up with consensus standards around this? Because I think if we leave it up to measure developers, the issue is there are these two papers. I was on one of them. A number of people here were on one or the other.

    We came up with different standards for a couple of timestamps. And so, if it left up to measure developers, are we going to have a whole bunch of well-designed measures, but still not agreement on the actual timestamps?

    DR. BURSTIN: And that could be one output of this group. I mean, if you think that is appropriate to try to put in this report for ASPR, we would be delighted to try to use this group to try to hone-in on those definitions.

    CO-CHAIR PITTS: Okay. Great.
MEMBER McCarthy: I think it is worth it, even though it may not be as precise as we want it. We should just try to come up with it. Boarding is really important.

Dr. Burstin: And keep in mind, these things could be subject to change. As something changes in the environment, we can do it. We did it probably about three or four years ago around definitions and calculations of medication adherence because there are so many different definitions. We just had our mid-management committee spend hours, and they just put it out there. At least for now, we seem to be getting all the measures in the same format, which does help at least reduce the noise in the measurement system, as opposed to trying to get the real quality signal.

Mr. Pines: And also, I think one of the things that we are planning as part of this document is to do a side-by-side of the two systems that are out there, the ENA and
the EDBA documents. We can potentially come up with some recommendations about reconciling those documents.

CO-CHAIR PITTS: Well, just a quick question. What were the pros and cons of the four hours versus the two hours? I think NQF does have some input on that. Was there a recommendation by NQF at all to do that? No? It is just JCAHO? Okay, yes. Then, I won't get into that. Thank you.

Brent, do you have something?

MEMBER ASPLIN: Well, along that line, I mean, to me, if we could get the admin decision time figured out, that is the beauty of 497, which is you are just measuring it.

Because, again, taking it from a patient's perspective, you know, once you are told you are being admitted, as far as you are concerned you are waiting to be admitted. There is nothing magical that is going to happen at two hours, that you are suddenly going to go, "Wow, I was waiting for a bed,"
but now I am boarding."

(Laughter.)

So, that is where I have always been in the -- you know, there is going to be a certain amount of boarding with each admission. "Boarding" doesn't have to be a bad word. I mean, we just want to track it, record it, and really long times aren't good. If you are thinking about it from a patient's standpoint, as soon as you start waiting, that is when it starts.

But I don't know; that has fluctuated depending on group, and various groups have chewed on it. Some people really want to have the boarding term itself connotate something bad. And the folks that are in that camp -- and several people in the room may be, so that is fine -- that is where the drive has been to -- you can't have it start at zero because, then, everybody will have a boarding for transition time. I just have not gotten caught up in that personally.
CO-CHAIR STONE-GRIFFITH: I think the Joint Commission has stated on several occasions that they are worried about the negative consequences of the transition in care from, let's say, the emergency department to the floor. But we really haven't wrapped any quality around that handoff or that transition. We haven't said, if you are -- I don't know -- an abdominal pain and you boarded four hours in the ED to go to a telemetry floor, what is the downstream consequence of that versus two hours? We really haven't wrapped anything from a quality around that.

I mean, we have in some patient types, like getting someone quickly to a cath lab or some of the impacts of long boarding. But, in terms of putting a timestamp on it, I am with you, Brent; I would say let's just look at it and see what that looks like over time. I don't know how magical four hours or two hours or one hour is.
theoretically, for any given condition -- and
the shape of the curve will be different --
but you could see an optimal outcome,
transition time, to allow for transition of
information and exchange and all that, where
transitions that occur prior to that time
could have adverse outcomes because there
wasn't enough time to prepare or exchange
information.

And then, obviously, we have seen
data around really extended time periods
before they move up, where, again, quality
starts to fall. So, it is sort of an arc of
outcomes, which is going to depend on
diagnosis as to what the optimal shape is.

CO-CHAIR PITTS: I think that the
pharma criteria also had subsets for mental or
behavioral categories. I will just note this
as an aside, and based on analysis I have
done, that it looks to me like almost that the
entire anomaly with psych disorders has to do
with transfer rather than admission. The way
it is phrased right now, it looks like
transfer is not a consideration. In fact,
there is no difference between admitted
patients with psych disorders and regular
patients without psych disorders. But if you
look at transfer, there is no admission, so
you don't think of boarding. That is where
all the difference is.

DR. BURSTIN: Just one more point
of information. It turns out that this was a
measure that was endorsed as time-limited.
They actually had additional time to test the
measure. We are actually expecting the
results any day now. So, it might be really
useful to share back with this group those
testing results and see if, in fact, there is
an opportunity to see whether they are getting
it right.

I did, actually, pull up the
detailed specs. You are absolutely right, it
does have strata, one of which is the global
score, one of which is the psychiatric population, one of which is patients formally admitted to observation, and all those. So, there was those strata. Again, we will have to see how that plays out in testing.

CO-CHAIR PITTS: Okay. So, the next topic here would be about risk adjustment.

MEMBER STOTO: Can I --

CO-CHAIR PITTS: Oh, sure, yes. Sorry.

MEMBER STOTO: Sorry. As you know, I am an outsider to this field, but I am sitting here thinking, what does the discussion have to do with regionalization, which I understand to be -- and maybe it doesn't. But, then, the last comments here you made about transfer, it sounds like maybe it is. So, I think it would be helpful to be more explicit about these issues.

CO-CHAIR PITTS: Well, yes, it is certainly related to our ability to transfer
patients within communities. I mean, the community capacity for handling psych problems determines to a great extent the amount of boarding in my hospital.

Wes?

MEMBER FIELDS: Yes, I just want to follow without perseveration here. But I really think we are, in 2012 and looking forward, probably in jeopardy of measuring things we shouldn't be measuring or putting out metrics which people will respond to in the hospital industry that they probably should ignore.

For example, I think there are probably three components of crowding that have been pretty well-established. One is the low-acuity patient who potentially could be seen in a community setting. The other is the patient who is typically a Medicare beneficiary waiting for an inpatient bed.

But the third, which I think we really need to be encouraging researchers to
look at, is other dispositions within the community. I think in my practice, whether or not a senior is capable of going to assisted living who previously was living independently but failing, that is a pretty big deal. Working through that transition of care takes time.

I think that it is instructive that England has pulled back from the four-hour rule. I think trying to tell hospitals to hurry up to make their decision is not necessarily the right incentive.

There is a substantial entrepreneurial hospital operator in California whose name will go unmentioned, who OIG is currently investigating, because he does such a great job of admitting a whole lot of patients who meet InterQual criteria within two hours. It just turns out that the gaming is on the inpatient documentation about the medical necessity for the admission.

So, I just feel like there are
other dispositions which are valuable and
which would actually reduce cost that we are
not measuring. I think that any population,
and the behavioral health population is
another great example, really what Brent and
others need to work on is the reiteration of
the throughput model that looks at population
subsets and looks at payer classes. Because
what the appropriate transition of care is
really depends on both where they fit in a
subpopulation by diagnostic category or
disease or degree of comorbidity, but it also
depends on their payer class.

I just feel like these are not
adequate measures for a complex problem. I
also feel like potentially, even if it in the
short-term aggravates the crowding problem, we
may be able to add value substantially by
doing a lot of services in the emergency
department or in an observation area that were
previously done in inpatient status. I think
that is worth a really hard look.
CO-CHAIR PITTS: Ellen?

MEMBER WEBER: I was just going to say, regarding the question of boarding, I wanted to make sure I understood, we are talking about when does boarding start, not what the harms of boarding are. And I think that we need to keep that distinction clear.

I think, actually, the greater harm is the lack of clarity about who is in charge of the patient during that transition. That is why the four hours is just awful, because it is like, well, if you are not boarding until four hours, who is in charge of you?

We actually, I think, several years ago had a Joint Commission inspection, and they said, "Well, who is in charge of the patient during that first couple hours?" We really had to kind of sit down and decide at what point does the admitting doctor actually take over the care. So that, if there is a problem, someone is in charge.
So, I think it is an argument for making the time to begin the definition of boarding short. It doesn't mean that boarding for three hours is bad. It probably is, but it is probably like, well, okay. But we shouldn't allow a four-hour -- because we don't want to penalize somebody on the other end. Because the lack of clarity at the beginning of the transition is the highest quality risk.

And I would totally agree, I mean, one other way to think about it is that, if you are going to be put in the hospital, it is true that there might be an alternative and we get you home. Great. But that time could be short because you are going to be in the hospital; we don't have to do a whole lot of stuff right now. Whereas, if you are going home, you know, that might be a longer period of time, and that is where our value-added is. So, let's figure out where we can do something versus where the hospital can do something.
CO-CHAIR PITTS: Emily?

MEMBER CARRIER: So, I think this is just taking what Wes and I have been saying one step further. I mean, what we have been talking about in these matters today are processes. So, are there outcomes that we could identify that would capture what is bad about boarding, what is bad about being stuck in the ED for a long time, when the length of stay is not being driven by specific clinical issues that are being addressed as efficiently as possible? Are there outcomes that could capture this, so we are not stuck with this blunt instrument of processes?

CO-CHAIR PITTS: Patient satisfaction?

Jay?

MEMBER SCHUUR: So, a quick response to Wes, I think the answer is not to not measure these processes because, still, at many hospitals the performance on these, it is not where we think it should be, and it is
just not visible within the hospital administration.

I think the answer is to develop measures for the examples you give. So, if there are measures for how we care for transitioning older adults to home, and have measures around that, then we can measure those important pieces of care.

MEMBER FIELDS: I actually think that is also the outcome question answer that Emily raised. Ultimately, this stuff makes a lot more sense in terms of actually improving the quality of population health and reducing cost if it is diagnosis-specific or condition-driven. I don't think it is one-size-fits-all anymore.

CO-CHAIR PITTS: Okay. Jesse says he knows something about risk adjustments.

MR. PINES: Great discussion on boarding.

Next, I want to talk a little bit about risk adjustment and really how the data
should be reported. We did a paper a couple of months ago in Annals that basically looked at the previous NQF-endorsed measures in the NHAMCS data and actually found that at the hospital level there were actually a number of exogenous factors that really went beyond ED volume. For example, the case mix in the emergency department and many other factors were directly associated with the length of the stay and the waiting time and a number of other measures.

So, really, the question here is, how should these data be reported? In the current version of Hospital Compare and the 2008 measures, they recommended unadjusted median as a way to report the data.

So, what I want to talk about next is what sort of risk-adjustment methodology should potentially be developed. Can we use existing data? For example, one of the things that they do in Canada, Canada reports data using stratified by the CTAS score, which is
basically the triage criteria. In Canada, it is basically required that everyone uses CTAS, which is different in this country. There is a lot of heterogeneity in the triage systems. So, from the patient's perspective, they come in, they spend five hours in the emergency department. Regardless of whether or not they go to a big, tertiary care academic center or a small, rural hospital, the prospect of reducing length of stay in a big, tertiary care center may be more difficult. So, essentially, that is why the recommendations are basically as written. The current version says that both unadjusted and adjusted data should be reported. In order to report adjusted data, however, there would need to be the development of some validated risk adjustment methodology, which does not yet exist, or alternatively, potentially make a recommendation about maybe some sort of standardization of triage classification.
CO-CHAIR PITTS: Suzanne, you are fixing to hit the button? Do you have something to say? No? Oh, I'm sorry.

Yes, Jay?

MEMBER SCHUUR: So, I would suggest, my comment here is I would suggest, rather than risk-adjusting, just reporting stratified data. It is maybe just sort of semantics, but rather than adjusting the actual numbers, just make people report it based on whatever metric you are going to stratify by, because I think it is going to be very difficult to truly risk-adjust.

And the second comment is I would not recommend using a triage criteria to do that because I think there are a lot of operational improvements that have essentially gotten rid of triage. Either it means people have to use traditional triage or the data you are going to get is actually not particularly important.

MR. PINES: Yes, and also, just to
clarify the stratification, we actually tried to create a simple stratification system. What EDBA uses is just is it volume-stratified in like 20,000-visits-per-year categories. And essentially, what we found was that that is predictive of length of stay in other measures, but actually doesn't capture even a fraction of the variation. Actually, the case mix was more important.

CO-CHAIR PITTS: Yes. So, I was really excited by that. What might you use instead of triage category? Have you considered actual potential things?

MEMBER SCHUUR: I would suggest visit volume and case mix index or some measure of disease acuity. I think it is going to be tough to get, from the datasets we have now, to get severity.

But I wouldn't want to overadjust because I think this whole issue with hospital readmission or all these things, how much do is the hospital on the hook for these
processes? My personal bias is that hospitals should be more on the hook than we should risk-adjust for patient factors.

CO-CHAIR PITTS: I'm sorry. Case mix index, is that a formal term? I don't know what it means. Or what is that?

MEMBER SCHUUR: I mean, there is a formal CMI classification that is used in calculating Medicare rates and other things. So, that is one method that can be used.

CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: And I was going to ask you about that as well because the CMI for the hospital is the hospital CMI. Depending on what your admission rate is from the ED, it could really fall through.

How about your facility code, which is really how your facility is billing for the acuity of care that you are providing in the emergency department as opposed to your triage rate? That probably is a better measure of exactly what resources that the
1 patient used.

   I mean, I think CMI is fine, but
2 in a hospital like ours at Penn we have a high
3 CMI, but we have a lot of cardiac surgery. I
4 rarely see those cardiac surgery patients in
5 the ED. So, just a thought.

   CO-CHAIR PITTS: And it is only
6 CMS patients, right, or everybody?
7
   MEMBER PAPA: Only CMS patients
8 for -- CMI is your Medicare reimbursement.
9 That is how Medicare --
10
   CO-CHAIR PITTS: So, it is only
11 calculated from the Medicare patients?
12
   MEMBER PAPA: For that. But your
13 facility codes, every patient in the emergency
14 department has a facility code.
15
   CO-CHAIR STONE-GRIFFITH: I think
16 in some places now the facility code is
17 probably only -- while it may be consistently
18 applied, it is probably only 40 percent of the
19 story. I think some of the other folks have
20 used the physician E&M code for that very
reason, because the hospital is only 60 percent of the HCPC. But you can't get to that.

CO-CHAIR PITTS: Ryan?

MEMBER MUTTER: One thing you have got to watch out with CMI, too, is that critical access hospitals, of which there are 1200, don't have it. So, you may end up having to be a bit more blunt and use hospital characteristics, sort of teaching status and things like that.

CO-CHAIR PITTS: Jay? Oh, I'm sorry. Brent?

MEMBER ASPLIN: So, is this recent thread in an effort to do a stratified cohort, use cohorts to report the data? Or is this thread to actually stratify it at an individual patient level? Because I agree with Jay.

MR. PINES: So, again, just to clarify what we did, we actually tried to create a simple stratification system with
NHAMCS data that actually had and actually used the reason for visit, common reason for visit classifications as proportions to basically see what really drove performance on these measures, and actually found that there were so many factors that were independently predictive that, unless we made the strata tiny, there was really no simple stratification system, which is sort of the reason why I think really the next step is to come up with some sort of a valid risk adjustment methodology that really takes into account factors that the hospitals can't control.

One of the things that we have in the report that is predictive of performance is things like percent Medicaid and percent Black and other minorities. Essentially, in our recommendation, those would not be actually in there. It would be more issues like a case mix, the MSA, ED volume, and things like that.
But that, in the final reporting,
just to make it understandable for patients
and consumers, that you would both see the
adjusted and the unadjusted data.

MEMBER RAPP: This is Mike Rapp.
Could I just make a couple of points about the
risk adjustment of these?

Hello? Can you hear me?

CO-CHAIR PITTS: Yes, go ahead.

MEMBER RAPP: Oh, I'm sorry.

Well, I guess there are a couple
of points. We are talking here more about the
regional aspect or the system aspect. So,
when you talk about risk-adjusting, I think
one of the points we made on a preliminary
conference call for this was, if there are
factors that are, quote, "predictive," it
would seem it is the hospital's job to try to
deal with those factors. In other words, you
put on more resources to deal with patients or
particular types of patients and that sort of
thing.
To risk-adjust it, so to speak, means that you basically will disguise the results of what is the amount of time that it takes to accomplish one thing or another. So, studies, and so forth, that connect up worse outcomes with crowded situations, it is not a risk factor for the patient. You don't think in those terms. You think of do whatever is necessary to be able to expeditiously take care of things. So, I am just generally opposed to the idea of this.

And secondly, when you do that, if you are talking about at a regional level, I think where you risk-adjust like this, you think in terms of, well, the hospital, and so forth. Although these are hospital measures, we are trying to think about how do you roll them up, but at a system level.

So, I just wanted to make those points. And then, I heard the discussion about how do you define what the decision to admit is, and I haven't necessarily followed
this. But, generally speaking, when CMS implements measures, the specifications of the measure make clear what the factor should be. Now, apparently, it may not be clear enough, and that could be worked on. To me, it would be the admission order, but somebody said that, apparently, some CMS documents may suggest that it shouldn't be that. But I think it is worthwhile to try to pin that down.

Certainly, in the hospital measures in general, CMS has meetings with the Joint Commission virtually every week. When people raise questions about this, they make an effort to answer them and, ultimately, come up with more specifics as to how you should approach those sort of definitional problems.

CO-CHAIR PITTS: Ellen?

MEMBER WEBER: What I thought, I see risk adjustment as sort of counterproductive to do what we are trying to do here, which is to make sure that everybody
gets the same level of care. Although I believe there are a lot of things out of the hospital's control, this is actually a way potentially for them to get resources. So, I don't know what we would really be accomplishing by risk-adjusting.

Of course, I am at a teaching hospital, and part of me would like to say, well, it is going to take us longer, but, okay, maybe is there value to that? I don't know. If there is no value to that, to the patient, they should know that. Maybe they want to be at a teaching hospital, and it is going to take longer. But if they don't want to be at a teaching hospital, maybe they should go somewhere else. But don't tell my CEO I said that.

CO-CHAIR PITTS: Yes, Brendan?

MEMBER CARR: And isn't there also a distinction -- I mean, it seems to me that risk, I echo what Mike and Ellen are saying about the diagnostic side, right? But the
administrative delay to boarding that you
described before seems to me like absolutely
it shouldn't be risk-adjusted. We are
proposing risk-adjusting the entire length of
stay, right? But we all think that it takes
longer to work up a sick person. Are we okay
with the fact that it takes longer to find a
bed for a sick person? I mean, that feels to
me like a different animal. I get risk-
adjusted workup. I don't at all get risk-
adjusting placement.

MR. PINES: All right. So, I
think the broader question is, should a 10-bed
rural emergency department be compared to a
100-bed innercity public emergency department
as apples to apples? That is, I think, really
the question that we are talking about here.
And it is good that we are hearing a lot of
different opinions on that.

So, what I am hearing, is the
group thinking that we should not make a
recommendation for risk adjustment. Can we
take maybe a straw poll on how many people
don't want to recommend risk adjustment?

MEMBER CARR: Can I just ask, when
you say risk adjustment, it feels like a very
loaded word. So, that means comorbidities to
those of us who live in that world, how sick
the patient is.

You just gave an example of the
size of a facility, which is a different
animal, I think.

MR. PINES: Right. So,

essentially, in that risk-adjustment model it
could be something like ED volume; it could be
something like the case mix. If you have a
higher proportion, for example, of trauma
cases, maybe your overall length of stay would
be longer or shorter.

Essentially, what risk adjustment
does, it basically allows you to really
compare apples to apples. So, let's say you
a Penn would be compared to a Jefferson as
opposed to Penn being compared to a tiny,
little, rural hospital.

MEMBER CARR: Yes, but you don't have to have them all in the model. You can put in hospital factors and leave out patient factors.

MR. PINES: So, essentially, we are not making any recommendation about what actually would go into the model, except for in the current version of the report we think that socioeconomic factors should not be in that model. But, essentially, we would recommend that some validated risk-adjustment methodology in the current version would be developed. But, essentially, the group could make a recommendation about what actually should go in there.

CO-CHAIR PITTS: Okay. This is important. I think we should let everybody speak.

So, Emily?

MEMBER CARRIER: I mean, this might go into the process of developing a
validated risk-adjustment model, but the only
data point I have seen was the paper that I am
sure everyone is familiar with that came out
in JAMA a year or so ago, looking at length of
stay and comparing safety-net to non-safety-
et hospitals. And that didn't find a
difference.

So, although I agree, Jesse, that,
intuitively, comparing an urban hospital with
crazy sick patients and volume and lots of
issues versus a much smaller, suburban or
rural hospital intuitively seems wrong, the
only data point I have seen doesn't show that
there is actually a difference to be risk-
adjusted for.

MR. PINES: Yes, we published a
paper a couple of months ago in Annals that
used the NHAMCS nationally, tried to create a
simple stratification system. We were
actually not able to do it because there were
so many exogenous factors that did impact
length of stay significantly. It was a
different paper, yes.

MEMBER CARRIER: But, I mean, was
the outcome that certain classes of hospitals
had systematically longer lengths of stay that
you felt like --

MR. PINES: Yes. Yes, basically,
ED volume, MSA, teaching status, and,
actually, even more importantly, the case mix
based on the reason-for-visit classification
was a big factor.

MEMBER CARRIER: Okay. So, I
understand from your conversation before that
that those were lots of little factors. I
mean, for me to risk adjust for something, I
want to see something that applies very
broadly. I was understanding lots of little
factors contributing in different ways at
different levels to different hospitals.

MR. PINES: So, essentially, what
we initially wanted to do is basically take
the stratification system that is used by
EDBA. Our hypothesis was that it was all
going to come down something like visit volume. So, essentially, you could do a stratification system. But, essentially, what we found was in the adjusted model, having all those things in the model together were all independent predictors. So, even after adjusting for ED visit volume, case mix and all these other factors were still very significant.

MEMBER CARRIER: I mean, is there anything unique about that data or --

MR. PINES: Yes. Yes, they were very, very significant. So, essentially, even adjusting for other factors -- we can send out the paper -- but that, actually, case mix, some of the case-mix variables were actually a lot more important than ED volume.

CO-CHAIR PITTS: All right. So, on the one hand, there is should we adjust at all in principle, and the second, if we think that you should adjust, is it even possible to do so? So, you are saying, basically, it is
possible to do so. But another question is, should we do it at all? Did I get that right?
Yes?

MEMBER WEBER: To Brendan's point, which I think is a good one, do we do it for all the measures or do we do it for some of them? Because the issue of complexity, that is a value-added. If I can do that and send the patient home, that may be better than having a short length of stay and admitting them.

So, I think we have to think about that is a real unintended consequence. So, I back off a little bit on what I said. But I think there are some measures, like the admit time to bed, should be not risk-adjusted. But the length of stay before that decision may be risk-adjusted. But the time to providers should not be risk-adjusted, because if you just have that kind of caseload, you have just got to get more providers or figure out your system better.
MR. PINES: And just as a quick comment, also, next we are going to talk about time targets. The way that Canada does time targets is actually by CTAS. They basically have specifically time targets for specific classes of patients. We don't have the benefit of CTAS. So, I think if we did want to do time targets, if we didn't one time target like they had in the UK, we would need some sort of risk-adjustment system.

CO-CHAIR PITTS: Arjun? And then, Jay.

MEMBER VENKATASH: I know Michael alluded to this earlier, but I am going to go back to it for one second, which is the task at hand to some degree was to think about the implications of these measures for understanding preparedness in regionalized emergency care and that intersection. From the perspective of that intersection, what we are thinking about and what we are discussing at hospital-level operational measures may not
necessarily inform that to the same degree.

What I am almost thinking is it is very different thing to say what is unadjusted boarding measured at a regional level. I could see how that is a window into capacity that we have on the inpatient side within some form of community. And that could be very helpful for understanding preparedness.

Understanding hospital-level adjusted or unadjusted boarding may not really inform that regional side at all and the preparedness part of this equation at all. So, I would almost say that, to some degree, perhaps the recommendation is that you don't measure at the hospital level, that it needs to be measured at a regional level unadjusted, because, then, you actually can say, okay, within this community, we know that we have long length of stay or we have high amounts of boarding. And then, that community can use that in terms of actually have something to track.
Because I think if you adjust it away, if you adjust away the characteristics, you have lost the whole concept of what you are trying to capture within a region, which is the variability in terms of what kind of capacity exists.

MEMBER SCHUUR: I agree with what Arjun just said, and then I have one other comment, which is I think it is important to think about what the measures are used for, for either public reporting or for accountability and payment.

While I tend to think that we shouldn't adjust away, definitely shouldn't adjust away hospital factors, I also wouldn't want to punish safety-net hospitals and under-resourced hospitals, which may have the worst outcomes, by implementing a measure that is going to punish them for having long length of stay. It would be the exact opposite consequence than you would want from this, I think.
CO-CHAIR PITTS: And being relatively new to this, the NQF ultimately attaches to CMS and becomes a -- there is a sanction that applies, 1 percent of Medicare reimbursement or something of that sort? No?

DR. BURSTIN: It is not nearly that direct.

CO-CHAIR PITTS: Okay.

DR. BURSTIN: Not really. Basically, all measures that are endorsed by NQF, you know, the committees have deemed them as being appropriate for a full range of accountability applications, whether that is public reporting or pay for performance. So, there is a wide range of groups of applications.

There is another group called the Measures Application Partnership that also works that helps say which measures are appropriate for which program. But, again, we do want to make sure that it is appropriate for any of those accountability applications.
And again, this is where level of analysis might be important. It is really also endorsed measures for specific levels of analysis. So, there may be some of these that are at the provider level, and there are some of these that may be at the system level. Those are considerations that I think it would be helpful for this group to think about as well, if the bigger systems issues are really what you are trying to drive to.

MEMBER TIMMONS: This is Shelly Timmons on the phone. May I make a comment?

CO-CHAIR PITTS: Yes, go ahead.

MEMBER TIMMONS: I just wanted to say, about the hospital-specific length-of-stay issues with socioeconomics, and so forth, it really does have a major downstream effect on throughput from the ED and beyond, because patient resources used in a given region are going to necessarily affect a length of stay.

If you have a large population of patients who don't have resources for home
care, rehab, or even family support in the large trauma system, for example, the length of stay is adversely affected. And that, then, backs up the whole entire hospital and system from a preparedness standpoint.

So, I don't necessarily think we should completely discount hospital-based, safety-net-hospital-type data because those things do impact the care and preparedness of the region.

CO-CHAIR PITTS: Thanks a lot.

Wes?

MEMBER FIELDS: Yes, if we are working on the straw poll, I would not be in favor of risk-adjusting for consumers. I think, in that case, raw data about length of stays, it has got to be something that they deserve see unfiltered or unadjusted.

I think I agree very much with Arjun that the same thing is true if we are really serious about trying to move towards status of a region, and how well all hospitals
perform collectively within the region.

And then, the final sort of thumbs-down would be that, if this is a matter of reimbursing hospitals, you don't want to disadvantage safety-net facilities, although I am not sure what that means anymore in our Brave New World, but I am quite sure you don't need to protect rural hospitals because they will almost always look great on HCUPS and great on length of stay. And that is the only real benefit of having low volume through their emergency departments.

MR. PINES: So, just to clarify, the current way that the recommendations are written is that both unadjusted and adjusted data would be reported. So, at the level of the region, I agree that I think in the later discussion on regionalization, I think taking these measures to the regional level and creating some incentive for hospitals to cooperate, to reduce systemwide boarding and crowding, I think are important.
But I think that we do lose a lot of information just by reporting unadjusted data. Particularly, certainly from the perspective of knowing how a hospital is doing, you are sort of are able to better compare apples to apples, and it doesn't uniformly make the small hospitals, rural hospitals look better. Because, just by the nature of their size and the way they are set up, it is a lot easier to have a shorter boarding time when you have got five or ten beds in the ED and 20 beds in the hospital.

And also, when we move on to talk about time targets, I think that without that risk-adjustment methodology, we would not really be able to any sort of stratification.

CO-CHAIR PITTS: Great. It looks like we have exhausted the risk stratification.

So, let's go on to time targets.

Ellen is the received expert here.

The recommendation states,
"Quality measure developers should consider setting time-specific recommendations for unadjusted or adjusted measures of ED crowding and boarding." Pros and cons of that?

I think it has been abandoned in the UK. No?

MEMBER WEBER: I will speak to both the pros and the cons, I guess, because I am fairly familiar with it.

So, the four-hour target went in 2004, kind of graduated to the point that 98 percent of patients needed to be out of the emergency department in four hours. The 98-percent figure came from the government. Actually, the emergency physicians wanted it to be 95 percent. So, the idea is that 5 percent of patients would be exempted from this. So, the sicker people, or whatever, you would have 5 percent of those patients did not need to meet those targets.

Just sort of in the background, what this really did for the emergency
departments was they created clinical decision
units. So, the patients they wanted to keep
longer who had sort of a clear-cut diagnostic
pathway, like chest pain or cellulitis, or
whatever, they kept them in the clinical
decision unit. They were off the clock; they
went home. So, what it really did was force
the hospitals to find beds for the patients
that needed to be admitted.

In 2010, when the Labor government
was voted out of office, the new government
came in. And part of this was to undue what
Labor had done. That is the UK way.

But, also, they were concerned
that the focus on time was taking away focus
from quality. So, what they did was they said
we are going to have a dashboard of measures.
One of them will be the time in the
department, and the four-hour target was
reduced to 95 percent. But it is not the kind
of cutoff it was where everybody comes
charging into your department and says, "We
are not going to meet it this quarter," and so forth. So, it is one of their quality measures now.

And there really wasn't a lot of evidence one way or the other for whether it was bad. In fact, we are about to release a paper that shows, at least from the administrative data, that there was not an increase in admissions, not an increase in resources.

But one of the things we weren't able to do is follow like the patients who got admitted and have been just sent to some ward because it was the only available place.

The upside was that the admitted patients got beds. The EDs were far less crowded. They also got a lot of resources, as the hospital did, to either redo their processes or build, or whatever. So, there was some money involved, at least at the beginning, although not later on.

The patients for the most part
I have really liked it and have said, "This is great." In fact, there was some concern that patients would like it so much that they wouldn't go to see their GPs.

The downside was, the potential tension from the emergency care was having a sick patient who really shouldn't go upstairs and kind of thinking, well, is this going to be my 99th-percent patient that I really can't allow to stay down here? So, there was a lot of pressure, and the physicians were able to kind of say, "You know, this is more important than the target, and I am going to keep the patient here."

But that did create a lot of stress in the departments, a little bit, not a lot, but there was some degree of the rest of the wards were resentful of the emergency departments because they thought they were getting all this money, and they had their own targets, and so forth.

But, for the most part, I would
say it worked to get rid of crowding, and it
doesn't seem to have been a negative effect.

It is a very blunt instrument, but, on the
other hand, it is an instrument that is pretty
easy to put in, and people manage to make it
work. And I think they made it work to the
benefit of the patient.

CO-CHAIR PITTS: So, I think I may
have created a scenario in my head that
explains how it happened, and it is probably
incorrect. And that is that they massively
expanded the obs capacity. Did that happen at
all?

MEMBER WEBER: Yes. Actually,
that is a good point. Many departments had
these, but those who didn't started to expand
what they called their clinical decision
units. So, they basically kind of did what we
are talking about doing here anyway, which is,
for other reasons, to expand your observation
capacity, so that people don't have to be in
the hospital. Their observation units were
mostly ED-run, but some of them were run by
their internists to move people in and out
fairly quickly.

CO-CHAIR PITTS: Don't they have a
specialty called acute care medicine which
is --

MEMBER WEBER: Yes, they do, yes.
Well, it is sort of a fledgling -- it is a
little bit different. The clinical decision
units were largely run by the emergency
physicians, and they would take the chest --
those were the people who they said had a low
risk of a high-risk condition. Okay? So,
that is where they would observe their head
traumas, and so forth.

The acute physicians work in
something called the admission assessment, no,
the admission unit -- or, no, I forget what it
is called -- the admission assessment unit.
It was a medical assessment unit and a
surgical assessment unit, but they are
actually admission units where there are acute
physicians whose job is to sort them out, get
them through all the testing, figure out
whether they can go home in 23 to 48 hours, or
whether they need an "ology". You know, do
they need a specialist now and need to move
upstairs to one of those wards?

CO-CHAIR PITTS: Brent?

MEMBER ASPLIN: So, I would
suggest one intermediate step to having a
static target would be just going back to the
measures from 2008, and I think you talked
about it on the call that I was not on, and
reporting not only median, but 90th
percentile. And the data would really start
to shine a light on how skewed the data are
and where the performance is when the wheels
are coming off versus just the median.

If I had it to do over again --
that suggestion was made about a week after
our meeting -- I wish we had done that. I
think median is helpful, but just median and
90th rather than a static target. Plus, I
mean, who are we kidding? Politically, there is not going to be a static target in this country anytime soon.

MR. PINES: Just to clarify, in the next section, I think Recommendation 8 does talk about the median and the 90th percentile. But let's make sure we also talk about time targets and beyond the politics. Is this something that we want to recommend to measure developers?

CO-CHAIR PITTS: Emily?

MEMBER CARRIER: I just had a question for Ellen, thinking about how the UK experience might not bond to the U.S. When patients were admitted to the clinical decision unit -- like let's say I was a patient who, prior to four-hour time target, would have spent a really long time in the ED having various things done, and now I am sent to the clinical decision unit at 3 hours 30 minutes. Is there an additional charge in the same way an obs admission would carry an
MEMBER WEBER: Because it is the NHS, I don't think there is any specific charge. What I understood, the way it works is the acute care hospitals contract with the primary care trusts, and they basically classify their patients as either simple or complex. If they are complex, the hospital gets a certain payment for them. If they are simple, they get a different payment.

So, anybody, basically, who went into the CDU was likely to be a complex, but so would be an admission. Or it would be somebody who got everything done in the first three hours, who had a fairly complicated history of abdominal pain, had a consultant, went home.

MEMBER CARRIER: Okay. So, sending someone to obs in the U.S. would have different implications in terms of resource use than, it sounds like, in the UK in terms of --
MEMBER WEBER: Well, it is an interesting question. I guess, yes, I mean, sending somebody to obs rather than just keeping them in the ED might, if you have an observation unit set up with separate billing.

The question is, there are a lot of issues there because you have to be there a certain amount of time. You have to be there overnight for anybody to make any extra money. So, it would really be more likely that a hospital might want to do that to avoid the readmission issue, to avoid unnecessary admissions they are not going to get paid for, where they would sort of provide strategic support, shall we say, for that observation unit, even though they couldn't bill separately for it.

But a lot of places do bill separately. They have figured out how to do it. It is still much less expensive overall for the healthcare system, yes.

CO-CHAIR PITTS: So, in the U.S.,
if you have obs patient, they don't generate
a second H&P to the ER obs? They just
generate a discharge fee, basically, I think.

Whereas, when I was in New Zealand they had an
obs unit, and all the medicine obs patients
got a second H&P the next morning when they
made rounds. It was nice to be on that
service because it was always morning rounds.

Whereas, in an ER obs situation, you might
discharge somebody at night.

So, there was a little bit of
qualitative difference between an ED admission
to obs versus the UK model obs admission, at
least in my brief experience.

MEMBER WEBER: Well, you are
talking about the UK obs admission where the
ED kept the patient was different than the obs
where you were on a separate service that
rounded on them the next day, right?

CO-CHAIR PITTS: Although where I
was, it was mixed together geographically,
yes.
MEMBER WEBER: Yes. Well, who was in charge is the question.

CO-CHAIR PITTS: Five beds for ER, 15 for the inpatients.

MEMBER WEBER: Yes. Okay. So, when you have a single payer, it is very confusing.

(Laughter.)

CO-CHAIR PITTS: Yes. And we were always trying to get one of their beds for our patients.

Wes?

MEMBER FIELDS: Yes, I think this is more like measure development, but I think it is worth trying to think through what the crosswalk would be. I think it would be fairly straightforward, but a lot of it does require measurement in terms of what is most cost-effective and what is going to do the best in terms of outcomes for patients by diagnostic category.

But I think the ED-oriented
observation would figure to be shorter stay, more intensive, probably more imaging or more ancillary. But what is the right length of stay for that category? What are the right set of diagnoses?

The thing that I think is maybe the most nuanced about this is what they call acute care medicine in the UK, we would call a hospitalist service. What is interesting to me about that, thinking about the whole first contact care/primary care debate in terms of acute care continuum problems, is that it would be an internist by training, but it is certainly not the primary care physician and it is certainly not a provider who is based in the community. They likely practice full time in the hospital. I think the argument would go in the American system that makes them more efficient and probably a little bit more rapid.

But I think it is worth beginning to figure out the measurements, both in terms
of cost and outcomes, about when aggressive short-stay observation that is ED-oriented is appropriate and when that 24-hour-plus thing comes into play, that it is more likely to be done in an inpatient setting where the hospitalist is the primary provider.

And I see both of those things as an important part of that gray scale that falls short of the statutory three-day stay for a Medicare inpatient. I think these are really worth understanding and measuring, even though they are probably a little bit beyond the scope of what we are doing today.

CO-CHAIR STONE-GRIFFITH: Ellen, I remember you speaking about this before. In my mind's eye, I almost see a third door that people are now coming through that is not the emergency department. It is a place where they get those fact-track services. They get treated and assessed rapidly, but it is not always the ED.

So, it speaks to who gets observed
out of the ED as someone who comes through that ED, and we need to spend that time, versus now a different portal altogether.

MEMBER WEBER: Just to speak to that, that is somewhat what happened there. This acute assessment unit that was run by the internists, hospitalists, whatever, was another entry into the hospital. Some places a GP could call up and send their patient to actually a unit that was nurse-run. They would say, "I want you to do this test and this test and this test, and then I want you to call this consultant." And that was even separate from this other.

But it was a way of bypassing in a good way the emergency department when a doctor wanted to keep control of what was happening with the patient, wanted to hear directly back from the consultant, didn't just turn them loose over to the emergency department. And then, if they needed to be admitted, they didn't go to the emergency department.
department to then get admitted. They got
admitted through the acute assessment unit.

And that is a big issue, I think, that we haven't talked too much about here in
terms of the input, is the difficulty now that
a lot of people are having figuring out a way
to directly admit patients to the hospital
without using the ED, because our systems are
so dysfunctional.

MR. PINES: Also, I just want to
clarify and I want to make sure I am getting
the right read on this particular
recommendation, because it sounds like there
could be some unintended consequences of using
time-based standards.

I think there are a lot of
different ways to do time-based standards.
And we wouldn't be really saying that we would
have to follow the UK model, you know, four
hours and 98 percent, but that, broadly, we
would make a recommendation that time-based
standards could be potentially meaningful.
I think that we could write that in the reports in such a way that we could recommend time-based standards, but say that that would be basically up to the measure developer.

MEMBER WEBER: I would agree. I think even though the UK took away the four hours, they would have been happy, the ED physicians would have been just keeping it at four with the 95 percent. And we clearly know now New Zealand and Canada are picking this up. So, I think the trend right now is towards people having these targets, not shying away from them. And those are more nuanced.

CO-CHAIR PITTS: Brent?

MEMBER ASPLIN: Yes, if we were going to move in time-based standards, I would be more open to it on just the boarding time, the 497 measure, given the changing role of what we are doing from a diagnostic standpoint, the time we are going to be
spending looking at alternatives to admission, et cetera. That might be a place to start, is
in that aspect of the overall measure, provided the gaming issue could be addressed,
to the extent there is one.

CO-CHAIR PITTS: Was there a lot of use of the obs unit for getting a CT scan?

MEMBER WEBER: Actually, no, because they do very few CT scans.

(Laughter.)

They do zero abdominal CT scans.

MEMBER ASPLIN: I think they observe instead of scan.

MEMBER WEBER: Actually, the thing about resource utilization, including obs unit patients did not go up. In other words, they didn't throw them into the obs unit and then order a CT scan. They threw into the obs unit truly to observe.

They actually are very anti-labs.

You know, you don't go into the abs unit just to get your labs back. You do it to get your
treatment for your cellulitis, for your
Tylenol overdose, or whatever. You follow an
algorithm, and so forth.

So, yes, there was not quite so
much what you were going forward with.

CO-CHAIR PITTS: So, boarding
sounds like a good interval.

MEMBER WEBER: And I agree, I
think boarding is a great way to start this
because it is exactly what we were getting at
earlier, which is we don't want to penalize
the complex workup. That is what these CDUs
in England did. They just gave the EDs an
opportunity to do the complex workups they
were doing anyway. But what we do want to
penalize is the long evaluation period after
the ED has now done a three-hour workup.

CO-CHAIR PITTS: Recommendation 7
touches really on the same topic, but putting
in the added element of standardizing by
triage acuity. I assume that what that means
is that it is four hours, depending on your
circumstances. I am not sure I understand the intend of that recommendation.

MR. PINES: Sure. So, this was actually something that came up on one of the first Work Group calls, where the Canadian system basically stratifies by triage acuity, where the more serious patients should stay longer. I think it is eight hours versus the more minor patients can stay four hours.

In order to do something similar here, we would need either our own risk-adjustment system or, alternatively, to recommend a standardized triage system. We know that ESI triage is the triage scale that is most commonly used. That is not used everywhere.

In order to truly standardize by triage scale in this country, we would have to make a recommendation that hospitals that are not on ESI, or whatever triage scale we recommend, would move to that. So, that was the source of this recommendation.
CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: You know, I agree on a standardized triage score, absolutely. But I can say, from a nursing perspective -- please don't shoot me -- but some nurses, when we are triaging, we consider who the provider is in the back.

(Laughter.)

And we know Dr. Smith is going to order two tests and move the patient, and Dr. Jones is going to keep them there for eight hours and do Q one-hour testing and get serum porcelain levels on every patient. So, we take that into consideration. There is a lot of subjectivity with that.

Plus, within the nursing -- and we have to own this as well -- some of the nurses are much better at making that prediction because they have got a little bit more experience. Certain hospitals you can triage after being there for six months. Our place, you have got to be there two years before you
triage. So, there is a lot of that interrelated piece with the nurses as well.

So, I would really tend to look more at another score. I don't know. Prefer the other nurses in the room to speak up. But sorry.

CO-CHAIR PITTS: It is a real issue. In Australia, I think everybody gets a score, right, and they actually tally that and see how compliant you are with those time intervals.

But I agree with you. And I guess it was you, Jay, who was saying that, who needs to triage? Wasn't that you? I think it was.

(Laughter.)

On the other hand, we would love to have some sort of national standard of severity classification. It would be really important to compare.

But, anyway, Ryan, you have something?
MEMBER MUTTER: I am not a nurse, but the economist's perspective, ESI tends to take on an institution-specific meaning. A lot of times what you see happen is the average severity in an institution is a 3. But a 3 in one institution is not comparable to a 3 in another. So, just something to watch out for.

CO-CHAIR PITTS: Yes, I have looked at heart rates. You know, that is objective. It varies a bit.

Yes?

MEMBER PAPA: And I just wanted to say, I don't know, I don't have as much experience with the CTAS method, so I am not sure that that is quite as subjective, for those of you who may have had the opportunity to work with it, I don't think that is quite as subjective as the ESI, just from my limited experience with it. I don't know.

MR. PINES: There actually have been some studies looking, comparing ESI to
other time-based triage systems. And actually, ESI is the most reliable system, but that doesn't get at Ryan's issue, which is the between-hospital differences. But, basically, within a hospital, when they tested, ESI is more reliable than other triage scales in terms of just inter-rater reliability, one person saying what is a 3 or a 2.

CO-CHAIR STONE-GRIFFITH: Wes?

MEMBER FIELDS: Just something quick and obvious. Acuity and length of stay don't always correlate. A patient walks through the front door with chest pain, and EKG data is obtained within five minutes. A STEMI is present. The patient is in the cath lab in less than 30 minutes.

So, I don't think you want to feel good about that patient being in the department for three hours because somebody forgot to do the EKG or they thought it was a stomachache.

MR. PINES: Well, you know, I
think the other that would dependent in this, if we do recommend time-based measures and do want to have some sort of a stratification system, the benefit of having that by triage level, and essentially why I think that is by triage level in Canada, is that the physicians and providers actively taking care of the patient sort of know what the target is for that individual upfront rather than after the fact.

So, if we did some sort of a risk-adjustment model and we said you should have been in the ED six hours, the providers may not know that until later. So, I think the benefit of having some sort of triage classification, actually, there is sort of active knowledge, when that patient is in the ED, how long they should be there.

CO-CHAIR PITTS: Jay?

MEMBER SCHUUR: The first concern I have about the triage as a standardization is just the fact that a lot of operational
improvements in many EDs are sort of getting rid of it. So, I think it is sort of already outdated, the idea that everyone is using it. But second is I am all for finding a risk stratification mechanism to look at an accurate thing. I think that is in the research realm right now. There is not a good one for the ED with datasets we have available.

But I would be concerned, if you think about value and cost, which we all should be thinking about, that it will bake, if you use the triage example, it will bake in the overuse we are all doing. And so, if we are doing way too many CT scans already, and you build in the adjustment for abdominal pain, and say abdominal pain patients should be there for "X" number of hours, it sort of adjusts that into the system. And so, I would not encourage that.

CO-CHAIR PITTS: Okay. There is another? I'm sorry. Ellen?
MEMBER WEBER: I would like to just reiterate again about the boarding thing is really what the time target dealt with. And so, we would not need to use any kind of risk adjustment for a boarding time target. That is the thing I agree with Brent, that that would be a very good recommendation because it is really what all of these targets have been about. It is not about getting the ED physicians to work faster, and they will become that if we don't make it very clear as to what part of the system we are really doing. Yes, there is a lot of overuse, and we need to work on that, but I am not sure that is the issue we can address this way with a time target.

CO-CHAIR PITTS: Okay. Oh, more? Yes. Arjun?

MEMBER VENKATASH: I think one of the challenges I have with this concept of adjustment is, if it is meant to either describe capacity on the preparedness side or
if it meant to drive a lot of improvements,
the problem with adjustment is you can improve
adjusted times, but those aren't real minutes.

So, it only helps, I think, with
an institution-to-institution comparison,
which I think is right now kind of, as I
alluded to, much more of interest from a
research perspective, understanding patterns
of utilization and care like that, but from a
perspective of, what do I do with those
numbers?

A patient certainly can't use an
adjusted time. I don't think even a
department itself knows what to do with our
adjusted length of stay is or our adjusted
boarding time is two hours and four minutes
versus our unadjusted time is two hours and 40
minutes. What does that mean?

What do you know is, if your
unadjusted time is two hours and 40 minutes,
you could put in place certain improvements or
try to understand how that changes with
certain improvements, but understanding how
adjusted time improves in the setting of
improvement I think is pretty useless at an
operational level right now.

So, I think from that perspective,
for all of these, leaving them unadjusted to
me just seems to make more sense for actually
being able to track this and make it
meaningful over time.

And I think the other thing I was
going to say is that this, to some degree,
kind of wraps into the discussion we were
having above. So, I don't know if maybe
combining the recommendations or putting them
somehow together just smooths it out. That
was Brendan's idea.

CO-CHAIR PITTS: Okay. Actually,
the adjustment bit, I mean, the way I have
encountered that problem is when my hospital
says, "Oh, we are at the bottom of this
ranking. We don't want to be here. It is
because we have sicker patients." I mean, I
am sure that UHC has that problem. That is where we always are compared to other hospitals within UHC. The answer to our failures is always it wasn't properly adjusted.

MEMBER VENKATASH: But even when it is adjusted, you will still say it is not properly adjusted.

(Laughter.)

CO-CHAIR PITTS: That's true.

Okay. Was there another? Okay. Let's go on to the next one. That is measures of central tendency, median versus et cetera. We have already heard that the 90th percentile was maybe a better choice.

Anybody have any comments about that?

I will just have a pointy-headed comment and apologize for making it. If you are in a small ER like when we used to work at Emory University Hospital, it was a tiny place. It used to be called "treatment room".
And you know that you didn't sleep all night
or you would have a disastrous night.

So, if you have a very small
volume, your chances of getting that to the
90th percentile are higher than they are at
Grady, which is tons of patients and the 90th
percentile won't bother you at all.

So, there is a difference, that
P9-to-the-P50 ratio will vary depending on
your sample size. I don't know if that is
really important clinically, and I am not sure
it has been looked at. I don't think it has.
It is just one thought that came to mind as I
was looking at the data and thought maybe HCUP
could address that at some point. But that is
truly a pointy-headed comment. Sorry.

Any other comments about medians
versus geometric means versus any other
measure of central tendency?

MEMBER McCarthy: Is that what you
mean, Steve --

CO-CHAIR Pitts: I am not sure
what I mean. I think that it is pretty obvious why you shouldn't use the mean, because of extreme values on the right side of the distribution. So, the median would be more reasonable.

But Brent was making the comment that the place where you struggle might be when you are really crowded. So, the P90 would be, the 90th percentile would be a more useful value for you to report.

Did I get that right, Brent?

MR. PINES: Right. Sort of the thinking is that just the median value. So, meaning the hospital on their average day is very different than this whole flexibility issue; measuring the hospital on their worst day, how do they perform?

MEMBER McCARTHY: And I think it is true, Steve, that smaller hospitals have less ability to absorb surge, because they are smaller, than large hospitals do. That is just kind of naturally -- that has been shown
1 statistically.

CO-CHAIR PITTS: Any other thoughts about central tendencies?

(No response.)

Okay. I am sorry, I don't know what the next point, Recommendation 9, exactly means. Maybe, Jesse, you can talk about that.

MR. PINES: Sure. So, the thinking behind Recommendation 9, there have been some specific ways to structure the emergency department that have been associated with differences in length of stay.

For example, having a fast track, physician in triage, there is a fair amount of literature around that. And there are some other things that some hospitals would call best practices.

And our recommendation here would be for those structural elements that have been associated with differences in length of stay or performance or quality, that those could potentially serve as structural quality
measures for crowding.

MEMBER FIELDS: So, the concept would be voluntary reporting if they are participating in one of these alternatives?

MR. PINES: Right. So that, if there could be some specific structural things like having a full-capacity protocol in place -- I think Anthony mentioned having some protocol in place where, when you get to a certain point, that your hospital actually does something different. That could potentially serve as a quality measure if in multiple studies that has been associated with differences in performance.

MEMBER ADIRIM: What are you measuring?

MR. PINES: So, I guess the easiest example would be the presence of, let's say, a fast track. There is emerging literature looking at different structural elements in the emergency department and the association with length of stay and other
outcomes.

The recommendation here is just to say that those could be considered as potential quality measures, as structural measures.

CO-CHAIR PITTS: Emily? I'm sorry.

MEMBER CARRIER: Maybe others here are much more familiar with the literature in this area than I am. Jesse, you said that there is a lot of it.

The few studies that I have read that have looked at things like that, I wouldn't describe them as of sufficient quality that I would say quality measure. It is more like our center was motivated to do this, and we did it, and our pre/post data shows that things have improved.

MR. PINES: So, the literature on this currently is pretty sparse, but, essentially, this recommendation would be, if in the future there is some sort of best
practice that is evidence-based, that that
could potentially be a structural measure for
crowding.

CO-CHAIR PITTS: Jay?

MEMBER SCHUUR: Sort of to follow
up Emily's comment, I would disagree with the
recommendation as written. I wouldn't want
structural measures that were associated with
improved flow. I would want structural
measures that were associated with improved
patient outcomes.

And I think about structural
measures as helpful if you don't have good
outcome measures or good process measures.
But we think we have pretty good process
measures for flow. So, why not just measure
the flow and people can implement whatever
strategies they want, unless the specific
strategies have been tied to outcomes?

CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: Yes, I would kind of
agree with that as well because it is not just
having the strategy. It is effectively implementing it. Because many people have a capacity management protocol, but how many people, how many hospitals really utilize it, and utilize it effectively every time?

And some people have a fast track, but they can only get it staffed between these hours and these hours. That may not go with your flow. So, I don't know how we would tie that in.

MEMBER FIELDS: I really want to try to have some continuity with the morning discussion because I think it was potentially really powerful. I think the idea of looking at populations and looking at regional services, and those kinds of outcomes, is a way we can lead.

And so, in that context, I think if Recommendation 9 was strategies that are deployed within a service area or across hospital systems, I think that is really what you would want to encourage. To the extent it
might help you with surge capacity or gray squirrels, I think you would like to know if they are doing it.

MR. PINES: And also, thinking more broadly about structural measures, I know that the literature on specific structural measures in the ED and length of stay is not particularly robust right now. But when you think about structural measures, that could be specific protocols in place. It could be transfer agreements between hospitals. Essentially, the purpose of having this Draft Recommendation in there is just to say that these should be considered, that structural measures should also be considered.

CO-CHAIR PITTS: Mike?

MEMBER STOTO: Actually, my comment follows up on that. I think the reason that I like this one is because I think that some of these measures will also help with the preparedness aspect of things. Well, particularly if they can be associated with
outcomes, but even if only with flow, I think that would be the kind of thing that would be useful for preparedness, too.

CO-CHAIR PITTS: Gregg?

MEMBER MARGOLIS: Actually, I was waiting for my comment until after we got to Recommendation 10, but looking at time and following up on my colleagues here, what I would really like to emphasize here is to make the connection between our morning discussion and this one.

That is really, while I think that boarding and crowding measures at a facility level are very important, also are boarding and crowding measures at a regional level. It gives us a sense of a communities emergency department capacity overall.

And I would suggest that perhaps a Recommendation No. 10, or I'm sorry, No. 11 might be something to the effect of figuring out ways that facility boarding and crowding variables could be aggregated in a way that is
meaningful to provide information as to the
capacity of emergency care in a given
community.

I am not sure whether that is a
regionalization question or a boarding and
crowding issue, but I wanted to make sure that
it was brought up in this context, especially
in the light of the Chairman’s comments.

MR. PINES: Sure, and just to
comment on that, I think we are going to have
that discussion after the next section.

CO-CHAIR PITTS: Okay. All right.
Well, Recommendation 10 then. I'm sorry, one
more comment. Ellen, yes, go ahead.

MEMBER WEBER: This is kind of
related to both 9 and 10. I wrote in here
some notes to myself.

Should we have some kind of
recommendation about what the hospital does in
this planning thing about high ED census?
Because we are talking about very specific
implementations of certain things that people
are talking about, but it kind of was a little
bit like the JCAHO flow thing.

   But since we are talking about
this this morning in terms of the surge,
should we specifically say a plan for daily
response to surges and capacity responses to
overcrowding per se as opposed to the
throughput measures? But what do you have in
place? What is kind of your early-warning
system and that sort of thing? And actually,
really make it very clear what the connection
is here between this and the continuum to a
disaster.

   CO-CHAIR PITTS: Okay. Gregg, are
you still up for a question? Or no? Brent?

   MEMBER ASPLIN: I would support
what Ellen just proposed if it was tied to a
boarding standard or some outcome measure
where we are actually going to do something
about it. If it is just a plan, I am not too
excited. But if we are going to couple that
together with you have to meet this standard,
so how are you going to plan to meet it, then
it would be I think more helpful.

CO-CHAIR PITTS: Jay? Wes?

MEMBER FIELDS: I am just going to
keep pounding this nail because it seems like
somebody gave me a hammer.

(Laughter.)

But I really do think, if there is
a way to serve communities and be able to
quantitize surge capacity, it is because we
can find a way to agree that on a regular
basis hospitals, individually and
collectively, have to actually have measurable
impact on ED crowding as their daily fire
drill that prepares them collectively for the
black swan.

You know, I really think there has
got to be a way for you to create a
measurement that demonstrates not that they
have strategy, but that they can implement it
and that they can do it in real time. I think
that is exactly why the JCAHO measure that
begins at four hours for boarding is inadequate for this purpose. Because, as we have heard I think three times I can remember during the day, there have been a number of very significant events for communities that come and go like the tides or a tsunami in four hours.

So, I think the metric needs to put some level of accountability around the ability of facilities within a region to respond and for this to be viewed finally, and hopefully forever, as something other than an emergency department problem.

CO-CHAIR PITTS: Arjun?

MEMBER VENKATASH: I guess the only thing I am thinking about when we think about structural measures is I really agree with what Jay said, which is when you have process or certainly outcome measures, the utility of the incremental value for that process measure is very little.

So, when I think about structural
measures, I think about what are either areas
that we can't measure via the other
mechanisms, and so structural measures have
added value, or is it necessary to help
balance the measure, that you need to have the
process measure with the structural measure?

In this case, I think Emily
mentioned earlier about thinking about what we
have done in care coordination. I was
thinking about this during the break. There
was only one structural measure I have seen
recently go through an NQF process that was
sort of interesting. And that was the NCQA
measure to tier medical home levels as NCQA
Level 1, Level 2, Level 3 medical homes. That
was a massive structural measure, right? It
included a ton of structural elements with
some research that was done that looked at how
well that tied to qualitative assessments of
patients feeling that they were part of
medical homes.

And I think thinking of that model
as a way for some of these boarding/crowding structures, as well as some of the preparedness structures, to fit into a list of multiple structures that, when done in concert, could be associated with this perception within a health system of either preparedness or being able to manage flow, and would allude to some of these ideas of, does that system have for flex, may be a way of doing it.

Because, in isolation, any one structural measure is going to look really weak, and it is going to be tremendous -- I can never see it getting through a consensus-development process. But, coupled together with some qualitative assessment that says that those structural measures in concert make sense, I think it is good guidance for developers, and that is probably the way we would want them to develop it.

And then, it gets at a process, again, like I was saying before, we are not
going to be able to measure outcomes, right, for preparedness? So, maybe this fits later in the discussion later the afternoon, but my guess is that this structure map includes boarding, crowding, traditional preparedness processes all in one.

CO-CHAIR PITTS: Great. That is actually interesting.

I'm sorry. Brent?

MEMBER ASPLIN: Well, what Arjun just said triggered in my mind this might be our opportunity to sunset the whole term "ED crowding," to begin to sunset it, because that was probably our biggest strategic mistake way back whenever.

CO-CHAIR PITTS: Call it something else?

(Laughter.)

MEMBER ASPLIN: Well, what we are really talking about -- and this does combine the two topics -- is system capacity and response, right? From a regionalized
emergency care standpoint, we are looking at
system capacity and response, and forms of
that involve daily operations and how we
manage daily surge. That is where flow,
delays, targets around boarding become
important. And then, that morphs into the
larger capacity and response issues of
emergency preparedness. That might be a way
to kind of tie this together and stop talking
about crowding. Just a thought.

MEMBER PAPA: And just to dovetail
off that, the biggest mistake we ever made was
to put that first patient in the hallway many
years back when we were trying to fix things,
because no other unit does that.
Unfortunately, we are the victim of our own
success. I agree with you, it is not ED. It
is not an ED issue, and this will help move us
forward.

CO-CHAIR PITTS: Yes, we did a lot
of things for a lot of people in the old days.

(Laughter.)
All right. This is the last recommendation before our break. So, then, Recommendation 10, measures of ED -- we have discussed this a bit -- measures of ED outflow beyond boarding. For example, hospital length of stay for specific conditions may be considered by quality developers to impact ED flow. I guess one question is, how easy would it be to get this kind of information?

CO-CHAIR STONE-GRIFFITH: Well, I will start off by saying we didn't think it would be that easy to get the information out of the emergency department once upon a time. But I think if we don't contextually look at it across the entire hospital continuum, then we are really going to miss the other patients that are being boarded elsewhere or not being managed as effectively and as efficiently as they can be. So, I would certainly support at least some guidance around that entire hospital experience.

CO-CHAIR PITTS: Brent, are you
up?

MEMBER ASPLIN: Yes. I like this one. I think most hospital administrators will know more about this than they will about their ED length of stay. So, these data are out there, and it kind of ties back into the same concept of system capacity and response.

CO-CHAIR PITTS: Ellen?

MEMBER WEBER: Yes, actually, I am going to steal this idea from Peter Marcello. But one of the things that he mentioned at our crowding interest group was something that you could track was how long it takes to get an order to put in for a needed study when a patient is admitted or to get the neuro consult for the patient who isn't going to get the TPA.

Also, bed days, how efficient are you with your beds? Of course, then you would need some risk adjustment.

So, beyond just the length of stay for specific conditions, there are probably
some other process measures of hospital efficiency that ought to be being tracked, not just for economic reasons.

CO-CHAIR PITTS: There is also just a simple, how do you schedule your surgery at this particular hospital?

AnnMarie?

MEMBER PAPA: Thank you.

The other thing you can look at is your discharge time and your discharge order. Your discharge order happens maybe at 11 o'clock. The patient doesn't leave until six o'clock at night, and we all know that because we feel it. You can't get a bed all day long, and then, all of a sudden, at six or seven o'clock at night you take a bolus of patients upstairs. The upstairs nurses are complaining. They can't take them all at the same time.

So, what is that process from beginning that discharge order being written to the patient leaving? Because that is a
huge issue as well.

CO-CHAIR PITTS: Brendan?

MEMBER CARR: Yes, I guess I just wanted to echo that I think this one is also really important. It starts to get at the systemness of this. I think it also sort of builds on pillar -- we are trying to cut clean lines between this session and the next couple of sessions. But I think, in part, the point is to not. This builds on the first pillar that Marco was talking about earlier today: is there a way to bake into here some awareness about whether or not this person can be reverse-triaged, should they need to be?

And I don't know that the literature is going to support anything getting through the process yet, but it does seem like we could start to increase some sort of awareness about our inpatient triage system.

CO-CHAIR PITTS: Terry?

MEMBER ADIRIM: Yes, what I like
about this measure, I agree with my colleagues, is that it is an integrative measure. I think it puts accountability, just so that everybody is concerned about it away from the ER when it is not the ER’s fault.

So, one of the recommendations I would suggest or somehow put in here, that measure developers should move towards developing measures that are integrative. Because each one of the issues that you have brought up, you know, really not just at the isolated ED part, the one thing that I have always been concerned about is like, why are we, the ED, always considered separate from the rest of the healthcare system?

CO-CHAIR PITTS: Manish?

MEMBER SHAH: This may be a little bit on a tangent to this, but one of the things that always concerned me, and I think is also a potential to measure how the region is doing, is the EMS offload time. I mean, that has been a big issue in our area, you
know, with EMS having to stay at the hospital for an hour and a half trying to offload the patients. And that is something that is measured within them, so it can potentially give a fair amount of information. And so, that might be useful to somewhere integrate, whether it is, whether it is expanding this recommendation a little to include maybe input for in-flow measures also.

CO-CHAIR PITTS: Arjun?

MEMBER VENKATASH: I was just going to say, for the EMS offload time, I think that is a good measure, and it is one that in the previous phase of this project, when we were speaking to David Cone, he spent some time in Australia and did a lot of assessment of their EMS systems. And that is standardly reported in Australia. So, it would be worth reconnecting with him in thinking about that.

CO-CHAIR PITTS: Oh, I am sorry.

I didn't see you there.
MEMBER LEVINE: I just wanted to add that I like this one as well, but this is one that we could also add in, if we wanted to or felt the need to have a risk-adjustment piece in it to really get at the acuity of the patients in the hospital, that there is a lot of established risk-adjustment methodology already out there versus the iffy-ness that we had for the ED patients. So, that is an opportunity to throw that in as we make this more systemwide or hospital-based.

MR. PINES: Great. So, I do want to take about a 15-minute break in a minute here.

One of the things I did want to have people think about is other recommendations. We have 10 recommendations here. So, maybe what we could do is, before the next session starts on accountability and regionalization, we can just do a quick around the room and see if anyone has any other recommendations that they want to bring up for
the crowding and boarding. Or if you came up
with something from this morning, please let
us know then.

So, let's reconvene here at 3:15.

(Whereupon, the foregoing matter
went off the record at 2:56 p.m. and went back
on the record at 3:18 p.m.)

CO-CHAIR PITTS: We are just
slightly out of sequence. No problem.

For this next step here, we will
talk about accountability and regionalization.
Since we have already touched on a lot of
issues on that topic -- and, subsequently,
recommendations, right?

MR. PINES: Yes, why don't we just
do a round robin of recommendations?

CO-CHAIR STONE-GRiffITH: Are
there any other recommendations we should add?

CO-CHAIR PITTS: So, we are
thinking of round-robining that. I know Mike
has got some stuff to say. After that, we
will go on to Arjun's presentation.
MEMBER STOTO: So, I have got a couple of suggestions about the preparedness area. There is some text in the draft report, but not made specifically into recommendations.

Angela, did you get the email yet? I did it on my iPad, so it may not look that great when it comes up.

So, anyway, I thought there really were three things that encapsulate some of the discussion that we had this morning.

One is to identify some of the capabilities that are important on a daily basis in small-scale emergencies and large-scale emergencies, basically across the whole spectrum of things, and then find ways to measure those things.

I am saying that because I think that we don't have many opportunities to measure in large-scale emergencies. But if we can find capabilities that are useful across the board, then as much as possible measure
that in these small-scale events, and so on.

The second one is to develop approaches to measures these capabilities at the regional and the system level for two different opportunities. One is in actual emergencies, and two is in exercises and simulations. I think both of those cases, it would involve developing -- it is the last one on there -- a protocol and a measurement tool like an instrument that would be used to do this. And, of course, you would need to assess its validity and reliability through standard methods.

And the third one, which is not on there yet for some reason, is to consider sets of measures. I think that a lot of these measures really don't work all that well by themselves, but we really need to think about, if we had a set of these things, we maybe could get a good picture.

Typically, NQF endorses one measure at a time. But I think we are really
talking to developers at this stage. If we
talk about developing a set of measures, they
could be put forward in a package, and that
would be useful.

MS. FRANKLIN: Yes, we do endorse
one measure at a time, but we do, also, look
at measures that are expected to be reported
together as well and composites.

MEMBER STOTO: Okay. That is the
concept I have in mind, yes. So, that is the
idea.

CO-CHAIR PITTS: You have,
hopefully, had a chance to think a little bit
about the topic. We are going to go around
the room, hoping that somebody will have an
idea on the topic of accountability and
regionalization.

MR. PINES: So, essentially, for
the next section, if we could just go around
the room, and if you have any recommendations
that we didn't think of that haven't been
mentioned, sort of other ideas that came to
your mind during a break or that you wanted to
mention that we could integrate into the
document as draft recommendations when we send
it back out to the group?

CO-CHAIR PITTS: Yes, go ahead,

Ellen. You are first.

MEMBER WEBER: I am going to throw
out three things. Just getting back to the
one that was the England experience, one of
the other measures they had was a maximum
length-of-stay measure. So, that might be
worth considering, where sort of all the roof
falls in on you, and the health minister comes
and sees you. They have what was called the
12-hour trolley wait. If you were admitted
for 12 hours and still in the department,
literally, that was like a failed institution.

(Laughter.)

So, it is kind of one of those
things where I would say most emergency
physicians would agree that 12 hours is beyond
the pale of what you really should be doing in
the ED. A lot of times that is a service issue where no one will take the patient, and so forth, or there is no bed. So, that might be just another, when you are talking about time targets, we are going into sort of the semi-non-controversial.

The other thing, because we have talked about this, actually, the experience of the patient and the experience of the staff. This is not my area, but just to think about whether there ought to be measures of that because so much of the crowding issue is a stress issue, a burnout issue.

And then, depending on how it goes for the patient and what your ED is like, do they feel whatever -- you know, some assessment of their experience in the ED I think is important. That sort of goes beyond "Was the doctor nice to you?"

CO-CHAIR PITTS: Do you want to add anything, Mike?

MEMBER STOTO: Yes, I have a
comment, and I am not sure it is quite a recommendation, about accountability. This is something we are struggling with on this -- I mentioned this National Health Security Preparedness Index that I am working on.

The issue is this: when you look at the Institute of Medicine model about preparedness, it talks about what the whole system, broadly defined, has to do together for a community to be prepared.

On the other hand, CDC gives money to health departments and ASPR gives money to mostly state health departments or other organizations at the state level. You really can't hold those organizations who have received the money accountable for what the others do or don't do in the community.

So, although we really would like to think about the contributions of all these different organizations, I think, quite appropriately, the groups that receive the federal funds say, "Well, we can't be
accountable for them. We can only be accountable for what we do."

Coming up with a way to deal with that I think is a big challenge. So, maybe I have articulated the challenge, as opposed to solving it. But, hopefully, that is helpful.

CO-CHAIR PITTS: Wes, any additional insights?

MEMBER FIELDS: I just want to quickly reiterate the two things that I think would be most valuable to the system. The first is somewhere within the first three recommendations. And it is just that, even if it is reported as something which allows the hospital to opt out of a door-to-admit time, I think beginning to track what is happening with observation services, both in the emergency department and in the hospital, is tremendously valuable in terms of its potential to show how we could reduce both the number of hospitalizations and have a positive impact on length of stay. So, that is one.
And then, the other, which I really think we probably are pretty close to on nine, is this concept that, if you really want to get to the population orientation of responses at the community level, I think Recommendation 9 looks more like structural measures that demonstrate how hospitals and hospital systems interact to improve the capacity at the community level in terms of potentially having a way to manage or predict surge capacity for disaster response.

CO-CHAIR PITTS: Gregg, would you like to contribute something?

MEMBER MARGOLIS: It is in a different line. So, does anyone have any response to the last comment?

(No response.)

I would like to ask the group to think about, why is it accountability and regionalization? Those two things are very different to me, I think. So, I am just curious if they make sense lumping them
MEMBER STOTO: Can I just jump in on that a minute? I don't think we need measures of accountability. I think that accountability is something we need to consider as we develop all these measures. So, I am uncomfortable with that phrasing as well.

MR. PINES: So, the thinking behind putting those together was based on basically taking a lot of these measures and actually aggregating them up higher levels. So, for example, the hospital across town, you would actually care if there length of stay was long if you were being measured on how they were doing. And essentially, you know, the whole notion of competition, that we know that places will continue to compete, but to have some incentive to cooperate. So, that was the thinking there. But if you wanted to separate those out, we could certainly do that.
MEMBER MARGOLIS: For the record, I have no advocacy position on it. I am just curious what everybody's thought is about it. To me, that might be an issue of how do we incentivize cooperation or "coopetition" in regionalization and that sort of stuff as opposed to who is accountable for it. But I am just curious what everybody else's thought is about the notion of accountability being lumped together with regionalization, and I have an open mind.

MEMBER STOTO: Jesse, hearing what you said, it strikes me that that is just a different way of articulating the point that I made just a moment ago. Is that it? Okay.

CO-CHAIR PITTS: Manish, have you got anything you would want to mention?

MEMBER SHAH: So, I will just reiterate kind of the question I brought up earlier, which is where EMS metrics and EMS falls within this, whether it should be a separate recommendation, talking about whether...
it is offload timing at the regional level or
even potentially at the individual hospital
level. And are there other things, as much as
I hate attestation-type things, are there
cooperative-type protocols/policies in place
to address various levels of issues around
preparedness, around crowding?

I hate diversion. We have gotten
rid of it in our system. Is that something
that should be in that list of things we
consider?

And the other thing, over break,
just kind of we were talking about it a little
bit, I don't know where to go with this. But
sometimes within regions, however you define
that, or within groups, you are going to have
a wide variation, right? You are going to
have those hospitals that are massively
crowded, boarding, running at 110-percent
capacity, and then you have the other ones
that are running at 60-percent capacity or 70
percent and don't have boarding.
Somewhere around there, is it worth thinking about that as a metric of how the region is working together to optimize the care of all the patients in the community? I don't know where to go with that, but it was just something that flew in, and maybe it should just fly right back out.

CO-CHAIR PITTS: I think it is really important. I am not sure how it fits into the NQF measures, but there is a huge difference that has not been investigated scientifically because it is really almost impossible to get ED-level information either in HCUP or in NHAMCS. It is possible, but you have to go through a bunch of steps. So, you just don't see much research on that topic. I think it is really important.

Brent?

MEMBER ASPLIN: Can I just highlight things that I and others have mentioned earlier? One is the rollup of institutional-directed measures to a region.
I think that is an important concept that we really should stress. And then, taking the opportunity to move beyond crowding and really just purposely call that we want to change the nomenclature.

CO-CHAIR PITTS: Arjun, I'm sorry, I skipped over you. Did you have something to say?

MEMBER VENKATASH: I guess it is another measure concept. I don't know where that fits. Maybe it was just in the previous concept.

But, as we were talking about this, a lot of what we have talked about considers surge on top of a system that is already crowded as a problem. But I actually think that if we want to get at some of the issues we were discussing before, which is how much flex the system has or knowing what a high-performing system, perhaps the measure should be surge recovery. Maybe the measure is something along the lines of number of days
or number of hours of sustained boarding at whatever -- I am not going obviously raise the discussion of whether that is two hours, four hours, or three hours.

But, to me, it seems like if we want to understand how systems respond to these types of things, then, actually, their ability to recover in minor surge, be it the bus or the day at 105 percent, the 106 percent, whatever it is. And whichever systems -- and that could be hospital or regionally measured -- recover the quickest are probably high performance from a perspective of preparedness. That is something you could measure at the hospital level and aggregate up at the regional level, and get an idea for preparedness out of that.

CO-CHAIR PITTS: Brendan?

MEMBER CARR: I have two. The first we touched on, but I just want to make sure that it doesn't get lost. I think that there is an ability to use that inpatient
triate measurement or to sort of make that a firmer piece, both at the hospital level to relieve boarding, but also at the aggregated level to get a sense of what proportion of patients, how many beds you could create, given the inpatient -- I am using the words -- "inpatient triage". We have called it reverse-triage, sending them home. I don't know what the criteria are.

But if everybody was flagged in some way, it strikes me that you would then know something about the hospital's ability to absorb a punch, and you also know, then, something about the region's ability to absorb one.

And the second piece is I think that there is a lot of story that could be told in transfers, because that speaks to how efficiently you get people out. So, I don't know what happens in boarding-speak when someone gets transferred to another hospital. Do we follow it? Do you know?
CO-CHAIR PITTS: Yes, I can tell you that NHAMCS, which has time intervals, does not consider psych transfers boarding. You can try to get at that by identifying a psych diagnosis and seeing whether they were transferred or not. And you can actually do interesting analysis that way. But census field reps have not called that "boarding".

MEMBER CARR: And what about the way that they get reported to CMS? What do we do with them? We just throw them out? Because they are the window, I think, to the region or the regional-ness or to the coalition-ness, I think, right? There is something. The transfers are telling us the story about how well I can offload my patients I can't take care of.

CO-CHAIR STONE-GRIFFITH: Yes, in the measure there are just strata, but if they are transferred, I mean, they are essentially a discharge.

MEMBER CARR: What happens in
hospital two, presuming they go to a floor?

I am not talking about psych. I am talking about medical patients.

CO-CHAIR PITTS: Yes, but there is no transfer criterion in the NQF criteria, I don't think.

CO-CHAIR STONE-GRIFFITH: No, no.

CO-CHAIR PITTS: Transfer doesn't come up.

CO-CHAIR STONE-GRIFFITH: Right.

MEMBER CARR: I am talking specifically about boarding metrics that have been accepted by CMS.

CO-CHAIR STONE-GRIFFITH: If you are discharged from the hospital through a transfer mechanism, you would not be in that boarding --

MEMBER CARR: I understand. Sorry. I get it now. Because you are not considered admitted.

CO-CHAIR STONE-GRIFFITH: Correct.

MEMBER CARR: You are being sent
to my hospital to be admitted.

CO-CHAIR STONE-GRIFFITH: Whoever receives them will likely admit them, but they are not in that measure yet.

MEMBER CARR: So, yes, I don't know what the metric is, but I do think that that tells the story of how well-connected my hospital is.

MR. PINES: So, just to clarify a recommendation for measure developers, how would you frame it for measure developers?

MEMBER CARR: I thought the "I don't know" part was clear.

(Laughter.)

Yes, I will get back to you. I will write something and email it to you.

CO-CHAIR PITTS: All right.

Manish, Brent; that leaves Terry.

MEMBER ADIRIM: Thank you.

I have three points that I wanted to make, the first one with regard to any recommendation that you have with regard to
regionalization. I would encourage you to incorporate the concepts or actually just straight-out within your recommendation to measure developers to consider the capability for special populations.

I am here because I am a pediatrician, but my colleague, Dr. Shah, points out that in Katrina most of the people affected were elderly. But I think considering those populations is a very important capability to have. So, I would like that to be included.

The second is that kind of goes to what Jesse was talking about with structural measures. I think any recommendation with regard to structural measures that may at least indirectly go to throughput would be with regard to designation of the institution, what type of trauma center designation, whether it is pediatric designated. So, I kind of thought that may be -- and I could be wrong -- but that may be something to
consider, as well as the training of the
staff.

So, in some departments you have
to be a Board-certified general emergency
physician. If it is a pediatric institution,
you have to -- so, that kind of indirectly
goes to quality -- well, not indirectly; I
think directly -- but it may have an impact on
patient outcomes.

And No. 3, we were discussing
earlier amongst ourselves that we would
encourage measure developers to look at the
existing care-coordination measures, to look
to adapt them for applicability to ED. I
think it is already done; you already have
them. And you could either respecify them or
whatever you can do to them to make them
applicable to regionalization or any of the
other issues we have discussed, would be good.

CO-CHAIR PITTS: Emily?

MEMBER CARRIER: So, are these
only recommendations to measure developers or
can some of them be recommendations about a research agenda or overall larger policy changes that could facilitate measure development?

I just wanted to say I feel like talking about regionalization and accountability, it is really important to note that certainly currently, and even most likely under future generations of Meaningful Use, as I am aware of it, it may still be impossible to track the clinical course of a patient from EHR to EHR if they are transferred from one hospital to another. So, you couldn't follow the patient who arrives in the ED and is transferred for cath, for example.

And in many markets, those patients will stay within a hospital system with a single EHR, a single shared EHR. So, a unified record may be possible in those cases. But not all markets are going to have those systems.

Understanding regional dynamics,
like in some markets it gets really complicated how a community works together versus surge. Like in some of the market, in one of the markets we study, it is quite possible that one hospital's ED could be overflowing and a neighboring hospital could be accepting transfers from out of state for a high-cost surgical procedure. And there are a lot of other markets like that.

So, the more that we can really follow individual patients from system to system, I think that would help get a good understanding, a better understanding of these three quality measures.

CO-CHAIR PITTS: Great.

AnnMarie?

MEMBER PAPA: I will echo what Terry said about the continuum in the care coordination. We did have that conversation.

I think, to your point, Ellen, with regard to the experience of the patient, absolutely. But the experience of the
caregivers. And I know in ICU there is a lot of work done with moral distress and futility and things like that. So, I don't know if there are other measures out there that would be similar or transferrable for us as well.


MEMBER ROBINSON: Thank you.

I just want to support some of the comments that have been previously made about there needing to be a recommendation in regards to system evaluation. We have talked a little bit about diversion and offload times. But if there isn't any coordination between hospitals and the individuals that you expect to transport patients to other locations, such as long-term care facilities, such as rehab facilities, to move them from one clinic to another, or whatever that is, ambulance services don't have unlimited resources. So, to staff up for that from a vehicle or manpower standpoint is something
that they need to plan for.

So, I guess I would like to advocate for that kind of consideration. When you are talking about regionalization, it doesn't just mean hospitals.

CO-CHAIR PITTS: Wes, did you want to go back?

MEMBER FIELDS: I just want to elaborate on what Brendan raised and what others have sort of touched on. I really think there are more and more reasons, as you regionalize care, you should expect more transfers. I think you need to begin to measure the difference between a move inside of a hospital system and a transfer across hospital systems because I think both of those things are likely to occur more and more.

And then, I just wanted to make sure I qualified something I said about observation services. I feel the same way about what is increasingly understood to be a transition of care. There, too, moving a
patient from independent living to assisted living is just as significant a transaction or a transfer as any other. And especially, that is true in terms of population management and trying to improve outcomes.

So, I just feel like, as we let go of the idea that crowding is an emergency department problem, I think we need to embrace the idea that there are many transitions of care and/or transfers which potentially can add value to the system, improve the patient experience, and reduce cost. But if we are not finding ways to track them, we will lose our ability to really understand them.

CO-CHAIR PITTS: Ryan?

MEMBER MUTTER: Just a few points. First, to Emily's point about transfers moving in and out of system and being lost, administrative data like HCUP increasingly has the capacity to track a patient across time and settings with an encrypted unique patient identifier. And that capacity has started to
be used and the development of some
experimental indicators. So, that could
potentially get at some of the issue that you
raised.

Second, just thinking about this
sort of "coopetition" and the incentives and
all that, I think it is potentially important
to think about. I don't quite know what the
solution is. On the one hand, developing,
say, coalition-level measures is interesting.
We could see variation. We could get sort of
a sense of capacity and capability and all
that. But what if you see bad performance?
Then, what? What do you do? Say, "Do
something about it."

You know, basically, we represent
Hospital A and Hospital B and we are
competitors, and our area looks bad. I mean,
our incentive is to say, "Do something about
it" to them, and theirs is the same thing.
And we have every little incentive to do
something that is going to make them look
better.

So, there is something to think about. I don't know what the solution is. But if we are going to invest in measure development, we might want to be thinking about, well, to what end and can we incentivize improvement?

And then, I guess third and finally, I think the observation services angle is really interesting. I mean, observation services is increasingly used, although there is variation in its use and how it is used.

In the event of a disaster, those are beds. In some facilities, for example, there are a lot of prolonged observation stays which CMS seems increasingly concerned about. That could be something potentially to consider as something that is going on that is sort of in daily operations, but could potentially impact disaster preparedness.

So, I don't quite know what the
angle is, but it has sort of come up and it is interesting to think about.

CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: There was one thing that popped in my head as you were talking, Ryan. What measures does the VA use? I mean, they have that one system and there are things that they may already have in place that we might be able to piggyback off of. I think Terry said you are working with them on some things, right?

CO-CHAIR PITTS: Wes?

MEMBER FIELDS: Yes, I actually forgot the most important one of all. This is a joke, but it is that time of the day. Every emergency department should have different color of socks for the patients they discharge home. This would allow you to get beyond the Meaningful Use problem because you would be able to know where the patient most recently was treated by the color of their socks.
that?

(Laughter.)

CO-CHAIR PITTS: Ellen? I'm sorry.

MEMBER WEBER: I just want to reiterate what a couple of people have said. I just want to make sure. To me, the more I think about it, the diversion measure would be really important, both in terms of crowding and in terms of regionalization. Because, basically, there has to be some level of agreement. Well, I can't say that for sure. But if there is no diversion in an area, that generally suggests people have become enlightened in some way, and perhaps there has been some work around how do we avoid overloading one hospital versus another hospital. Is there better communication as a result of it?

And time on diversion would be sort of like a bad thing. So, that could be another push towards either cooperation or at...
least improving your flow.

So, I see it as kind of bridging both of the areas. Actually, the more I think about it, I think it is worth looking at.

CO-CHAIR PITTS: Is there any formal diversion national-level sort of criteria? I mean, I remember there used to be in the old days certainly the nurse could decide to go on diversion for the next five patients and then go back. In some places, diversion is a formal process involving multiple layers of the hospital, and in others it is sort of an ad-hoc kind of a decision. I wonder if there is any standardization at all. I am sort of not up on that.

No? The answer is no? Okay.

MEMBER WEBER: I thought there was something I just read in here, that each hospital is supposed to have a plan in place, you know, criteria for going on diversion.

MR. PINES: Right. I think that is part of the Joint Commission Flow Standard.
MEMBER WEBER: Yes.

MR. PINES: But I don't think there is any like national criteria. Essentially, the hospital has to have their own criteria.

CO-CHAIR PITTS: Manish?

MEMBER SHAH: I was going to say, I mean, I think diversion the way we are talking about it, because you are crowded, because you are boarding a lot of patients, or whatever, my sense of at least the environment is people are moving away from it in Massachusetts, in San Diego, and we have done it in Rochester.

Diversion because, you know, the plane just crashed into your hospital is a completely different thing. I think that is usually what most of us write into our JCAHO or the Joint Commission requirements, that there are going to be instances where you have to divert.

CO-CHAIR PITTS: Were you going to
say something, Brent?

MEMBER ASPLIN: Real quick, there is no national standard, but some regions have done pretty clear quantitative criteria before you can go and divert. I don't know how much enforcement there is.

CO-CHAIR PITTS: All right. Ryan, are you still vertical? Okay. I am sorry I haven't got to you guys.

DR. HUNT: A couple of quick observations from the discussion. The discussion surrounding the gross measures versus granular measures of crowding, specifically around the admit order piece, a long discussion about that. As I listened to that, I flashed back to the discussions a long time ago about an EMS measure of dispatch time, and it took a long time to figure out about how many calls there were to make that happen, when the wheels of the ambulance actually moved. I mean, there were a lot of
cuts to that.

And I thought about it and said, well, you know, the system, the system itself, that just needs a gross measure, but to really do problem-solving, not just at a facility level, not just at an EMS level, but also at a system level, at some juncture this will need to evolve to a much more granular level than this sort of like gross measure. So, I would encourage getting to the granularity sooner than later.

I am really sensitive that, while we don't have definitions around that, start making marks in the sand. That would be my encouragement about measure.

And then, the other one that I had a sidebar conversation about, I was charged a while back with being the Chair of the Crowding and Surge Committee for the hospital. The most surprising thing, the aha moment, when we got down to granularity of data, it wasn't discharge from the hospital orders
written; it was the time from discharge orders written by the resident, and the nurse signed off, to actually having the bed vacated, cleaned, and staffed by staff. So, that interval from discharge orders, hospital discharge orders, to actually vacating the bed and having it cleaned, et cetera, et cetera, we were just stunned. That was the bottleneck. It wasn't the discharge order issue.

So, again, that is granular, but I think at some juncture this is going to have to move toward granularity to do the problem-solving to improve it. That may mean you have got to gross first, but you have got to go granular to be able to do the problem-solving.

CO-CHAIR PITTS: Okay. Brendan?

MEMBER CARR: So, I wanted to respond what Ryan said because I think that he was asking about why people wouldn't cooperate within a coalition, within a region. I mean, I guess I am wondering if we need to be more
explicit in the report, then, to suggest that
the reason to develop metrics at that level is
to create incentives and/or the opposite of
incentives, so that people do cooperate. I
wonder if we need to be more explicit in the
report, if that is not clear. I mean, I think
that is the reason that we are having this
conversation, is to be able to benchmark my
region versus yours versus someone else's.

And then, with respect to the
transfer thing, I wanted to offer this: I
think that time from decision to transfer to
leaving the department might be an okay
benchmark, but decision to transfer is going
to be very difficult. I wonder if there isn't
some utility in just knowing time from
presentation to the emergency department or
triage or doc, or whatever you pick, to time
to be transferred for all transfer patients.

On some level, shouldn't there be
some awareness that we can't take care of this
patient? Right? Arjun sort of said to me,
maybe you pick a couple of diseases, and if it is intracranial hemorrhage and you are transferred to a neurosurgical service, it is time from CT scan. So, it is essentially when you got the diagnosis.

But, on some levels, I can't manage this patient. How long should it take me to figure out that I can't manage this patient? Isn't just time from presentation to time from leaving the door telling?

CO-CHAIR PITTS: I think it is telling. It certainly is telling with respect to behavioral problems. That time clearly distinguishes psych illness from other illness.

So, you are saying that maybe we could limit ourselves to time-in versus time-out plus a marker for transfer or not transfer, essentially.

MEMBER CARR: Not as a marker. Just within transferred patients, for transferred patients.
CO-CHAIR PITTS: So, if you can identify who is transferred, all you are going to do is time-in and time-out, basically, if I understand you.

MEMBER CARR: You do understand me, and I think those are going to be short numbers in places that have a game plan ahead of time and long numbers in places that have a hard time offloading patients, either because their neighbors are overwhelmed or because they didn't belong to a coalition or they didn't participate in something.

CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: But to that point, Brendan, what about the extenuating circumstances? Suppose you are a small hospital on an island, a critical-access hospital on an island, and the only way you get a patient off is by boat or by helicopter, and you can't fly and you can't get the boat out. I mean, that is going to affect your transfer time. So, there are things that have
to go into play when you are looking at that, I think.

In our area, in Philadelphia, it shouldn't be a big issue. But, you know, sometimes we can't fly them by PennSTAR. So, we have to go by land. You can't go by land on some of these places. So, just something to consider. I mean, I am not saying that it is a bad measure. I just think it is something we have to consider.

CO-CHAIR PITTS: Kathy?

MEMBER ROBINSON: The comments that Dr. Hunt made really resounded with me from the standpoint that we have been talking about crowding and boarding and preparedness, and the need to perhaps consider some of the greater detail in that regard.

And I just think of the example of EMS and fire agencies that might be charged to evacuate a community at the same time that a hospital is implementing their plan to discharge patients, and those personnel, those
resources can't be in two places at the same time.

Without those sorts of discussions ahead of time, or someone suggesting to them that that might be a consideration, that is going to get lost.

CO-CHAIR PITTS: Jay, do you have something to say?

MEMBER SCHUUR: Sure, around transfers. One thing comes to mind around that idea, which is there is an NQF measure in the chest pain set around time to transfer for patients with ACS. And so, that might be a good model for the sort of disease-specific measures.

And I think it probably is worth, if we are going to really hold people accountable for times, thinking about diseases where there is a clear time-to-outcomes relationship, whether it is sepsis, ACS, or something like that.

I think there is also a role for
transfer measures in the capacity piece,
looking at what institutions are transferring.
Because one aspect is the sick patient you
can't care for. The other type of transfer is
the patient that you should be able to
transfer, the hand injury, but you are
transferring them because your orthopedist
won't come in after hours, insurance issues,
this, that, or the other thing. That also is
a measure of system capacity resilience. I
would recommend exploring that also.

CO-CHAIR PITTS: I guess, Arjun,
you are next. Nothing to add? Okay.

Jay, do you want to say anything
more general that you have been thinking
about? Okay.

David? You're good?

Linda? No?

Emily? No?

Melissa?

MEMBER McCARTHY: Just that I love
the idea of the sunsetting with crowding.
(Laughter.)

It is brilliant, Brent.

And the idea, too, of just daily operations kind of at a facility level or a healthcare system level to me is very different from a regional level. Because if we just even take maybe length of stay, I mean, it works at the facility level or maybe even within a healthcare system. But what does it mean at a regional level? You have eight different hospitals. So, their average either length of stay or their median or their 90th, or whatever, is so variable within those eight hospitals. I don't know what we get from a summary.

But once you start thinking about them as separate concepts, what are measures of regional systems of care, then I think they have to be thought of differently? So, that is one recommendation I would make.

CO-CHAIR PITTS: Anthony?

MEMBER MacINTYRE: Yes, I would
just echo that. I am still struggling with
the conversation this morning about whether
this is really two ends of the spectrum or two
different things with interrelated components.
I am still in the camp of two different
things.

Two specific comments. Some of
the recommendations that are very specific
under the crowding piece really seemed to
arise from the facilities-specific world. I
don't know what the validity is when they are
suddenly scaled-up to a regional level.

Now, if their intent is to compare
facilities within a region, maybe there is
something there. But, as Melissa said, at a
regional level, I don't know what we are
looking at with some of those numbers.

The second comment with respect to
the preparedness side, I like where Mike was
headed. He is trying to give you some
specificity to the recommendations there.

I would just reiterate two points.
One is preparedness and response are two different beasts. If you are going to have measures, you are going to have measures for both.

And the second is that surge does not exist in isolation. I mean, that is an antiquated thought that many clinicians, primarily, still hang their hat on. Surge exists within the construct of how you manage your organization during a time of duress. And those management systems apply whether it is surge, a resiliency issue, or a safety and security issue.

The way in which I communicate with my staff during a surge event, an evacuation event, an active shooter in my hospital should be very similar. And if they are not, then you are missing the boat with all-hazards emergency preparedness and response.

So, I think examining surge in isolation is a bit shortsighted.
CO-CHAIR PITTS: Emily?

MEMBER CARRIER: I am sorry, I did remember one thing. I think it would be great, I guess going back to what you are saying about examining surgeon isolation, it would be great not to examine what we no longer call crowding in isolation.

(Laughter.)

It would be great to see suites of measures that look not only at performance in length of stay or meeting time targets, but also paired them with things like rates of 72-hour revisits, rates of discharge within 24 hours from inpatient admission, to get a whole picture of how the system is functioning, and not just how this one track is doing.

CO-CHAIR PITTS: So, Melissa, what are we going to call the crowding interest group? I thought about operations maybe.

(Laughter.)

Suzanne?

CO-CHAIR STONE-GRIFFITH: Well, as
I have been listening to several of the comments, one of the things that has resonated with me suddenly is sort of, what about freestanding emergency departments? When you brought up the movement transfer in both the context of transfers and in the context of movements from one department to another, and the proliferation of freestanding emergency departments that is occurring in our communities, what role do they play in terms of -- I don't know -- just the population that they serve, the surge, the EMS? They become part of it.

I think whether that is a strata or a way to look at those in aggregate or separated, I would like to sort of put that one in the mix.

CO-CHAIR PITTS: Okay. Helen, did you want to say anything in particular about this whole business? No?

All right. Jesse?

MR. PINES: So, I just also wanted
to make sure that I think we have had a pretty robust discussion about regionalization and accountability, but I wanted to make sure that if there are any other comments on that topic. Essentially, the way it is going to be framed in the report is, basically, what we talked about before, which essentially is taking these current measures and essentially aggregating them to different levels, regional levels, hospital and coalition level.

Is there any other discussion around that or other comments that we should get into that section of the paper?

CO-CHAIR PITTS: I'm sorry? Yes?

Peggy?

MS. SPARR: One that hasn't really been discussed very much, but it would seem appropriate if you are talking about it at a regional level, is just revenue. I think that it is unrealistic just to take the numbers from individual facilities and work their way up. I think there has to be, at least from
our experience in big disasters, there has to be some way to make sure that people are made whole by cooperating with each other and by doing what they need to do.

I think that within some sort of a context that that can be done. I don't know that it can all be done under existing authorities, under CMS, or different insurer plans, or whatever. But, unless there is some way to make sure that people are made whole, it is a lot less likelihood that they are going to be working together.

Whereas, I think people are very motivated during disasters and during emergency events, but they also need to know that, when they do this, at the end of it all there is going to be some way that things will go back to normal, and they are going to start working their way back down to normal amounts of delivery of care.

It is not something that has been really said very much. And yet, I think it
does affect the ability of these measures to actually work or not.

CO-CHAIR PITTS: Sure, Terry?

MEMBER ADIRIM: Yes, I just wanted to ask a question, to go to what Jesse was saying about kind of tying up this section. It may just be the time of day, but did you clearly hone-in on what you are going to write about accountability?

MR. PINES: So, we do have a draft of that section in the report. We can take a look at that. But, essentially, what is currently in there is that the measures are going to be aggregated at different levels. So, essentially, a lot of these crowding measures and preparedness measures would be basically taken from the facility level to the regional level or to the hospital/coalition level to promote the "coopetition".

So, next, we are going to go ahead and move to the discussion of the NQF criteria. Arjun Venkatash had a short
presentation he was going to give for us.

MEMBER VENKATASH: It is not as much of a presentation, I think, as it is a valuable exercise for this process, which is that a good handful of people in the room have served on NQF Steering Committees in the past through a consensus-development process, but some haven't as well. I think there is probably value for those who have not served to look at the NQF measure evaluation criteria when thinking about these concepts, because that helps set the bar of understanding what it would take for any of these concepts to actually turn into something that would be endorsed.

And for those who have been through the process before, I think it would be valuable to add those insights to this discussion, because the criteria have not necessarily evolved over time, but have been specified over time. And I think probably the two things that happened, most importantly,
were two reports about a year ago on the

evidence expectations regarding measure
evaluation, and then the second being around
testing.

I think probably the easiest thing
to do is -- oh, you have got them right there?

Good.

Helen presented this morning the
four general categories by which measures are
evaluated. I think where this group can
provide some value here in terms of what the
final report looks like is thinking about
places where you think either an exception
needs to be made to the current NQF measure
evaluation criteria in order for a measure
around crowding, boarding, and preparedness to
make it through the process or, secondly,
places where there are clear inadequacies that
measure developers need to be aware of before
they go through the measure development
process.

So, if we start at the very top of
the measure evaluation process, No. 1 is a
must-past criteria that is the importance of
the measure.

Angela, do you have that other
table?

So, for the purposes of this, I
don't know what everybody here thinks, but my
idea was, if we use the IBA measure that we
alluded to this morning as kind of a frame of
reference to the back of our head, I think it
will be valuable because it is a preparedness
measure. I think it is the preparedness
measures really that are going to have the
most trouble when you look at the way
measurement evaluation criteria are set up and
the degree to which they are outcome- and
patient-measurement-focused. So, I think if
we think of that measure in the back of our
head, and then go through these criteria, that
is probably the most valuable.

So, we will start with importance.
I don't have it in front of me, either, I
don't think.

Within importance, if you think of that IBA measure as it currently stands, I think that a lot of the preparedness measures, importance is kind of subdivided by categories. That IBA measure would be able to -- one of the first criteria is, does this fall within a national strategy around quality or improvement? I think that the preparedness concepts will often fall within national strategies, either by fiat or whatever it is. But that part wouldn't be very challenging.

The second step of that would often be, is there an evidence base to suggest that there is a performance gap, would come under importance. So, I guess a question for the group is, what does performance gap data look like? What does variation data look like for a proposed measure, be it at the hospital level or regional level, around preparedness? Because if the outcome is rare, we may not have that type of gap data.
MEMBER CARRIER: May I ask a question?

MEMBER VENKATASH: Yes.

MEMBER CARRIER: When you are talking about importance, is it enough to say that in a disaster it is important that hospitals be able to move beds? Or do we have to say in a disaster we have demonstrated that it is important that hospitals be able to move 20 percent of their acute beds within four hours, which are two different questions I think?

MEMBER VENKATASH: So, I mean, the latter. But we are almost even a step up from that, in the sense that that would get more at the focus of measurement and the actual specifications of a measure. But you were to be a measure developer and make a measure around IBA, part one of importance would be that it is high impact. So, we would argue that preparedness in the setting of either surge or large-scale massive disaster is
important to the National Quality Strategy, things like that.

And then, when you describe the opportunity for improvement in the current performance gap, you would have to show data that says, okay, in the setting of disasters, currently, there is variation amongst hospitals, and 50 percent of hospitals are unable to create 20-percent capacity.

So, the question I would have is, if we think about this in terms of the IBA measure, right, the IBA measure asks, can you increase capacity by 20 percent in four hours? What data would a measure developer have to show you for that measure for you to believe that there is a performance gap, that right now, at either the hospital level or the regional level, that there is currently a gap between being able to do that or not?

DR. BURSTIN: Just one clarification. Actually, it doesn't have to be a gap in care. It could also be variation
across providers or across entities, whatever
the case may be. But that would be adequate
as well.

CO-CHAIR PITTS: Mike, you have
got something to say?

MEMBER STOTO: Yes. I think that
criteria is a really important one. But,
given the state of development of this field,
where no one has any data at all about this,
we just can't do that right now.

So, I think that in this case the
importance argument has to rely on the kind of
arguments that Dave was making this morning
about, if you can't do this, you can't do
anything else, is basically what it comes down
to.

I think that, given the state of
development, we just have to think about that
differently now.

MEMBER ADIRIM: You could say it
has face validity.

MEMBER STOTO: I'm sorry?
MEMBER ADIRIM: You could say it has face validity.

CO-CHAIR PITTS: Face validity.

MEMBER STOTO: Yes, that's right. Right. And I think of that in terms of a logic model that you really think through how is this going to lead to the outcomes that we want. It really is on the order of face validity and kind of logical thinking, as opposed to data, at this stage.

MR. PINES: Right. Right. So, essentially, what we are trying to do here is basically take the objective criteria, NQF criteria, that are applied to all measures, and given the unique nature of preparedness data/evidence, basically, modify those standards a little bit. And we don't want to say "reduce the standards," but modify those standards, so that measures that are important can potentially get through.

MEMBER VENKATASH: I guess a related question here is, traditionally, when
thinking about the performance gap, right, it helps us identify those areas which require performance improvement. So, in the case of preparedness, part of this is justifying that this is worth measuring and that it needs improvement.

So, does it need simulated exercise to demonstrate performance is inadequate? Or is this kind of face concept that we believe in general consensus that current performance with respect to that measure is inadequate? I think that is probably a question that this group can provide guidance on because, if that is not enough, if we are going to dedicate a lot of resources to the improvement, should the standard be that at least simulated exercise has demonstrated that current performance is inadequate before you continue through the rest of measure development?

CO-CHAIR PITTS: Is there not any empirical data on performance being inadequate
in response to some disaster? There must be.
You hear about disasters, and this and this
was done, but was it adequate?

MEMBER CARR: What sources do they
use? Do they need peer-reviewed literature?
Does anybody know this?

MEMBER STOTO: I don't think that
people have looked at this in a systematic
way. I think people have looked at one
disaster after another and have said, "Oh,
gee, we didn't have enough capacity in the
hospital to do this." And that is the kind of
reasoning that this is based on, but I don't
think it is the kind of statistical analysis
that we often see in NQF.

MEMBER CARR: But does anybody
know the threshold? I mean, does that work?
There is a series of six disasters, right? We
just heard that Aurora's emergency department
was crowded before they got 20-some-odd --

DR. BURSTIN: For the performance
gap, yes. For the evidence for the measure
focus, no. Those are different. So, you
could certainly use, I think, the cumulative
experience to assess that there are issues.

I think the question that is
raised up here on the top one is really
evidence of measure focus. And again, I think
there is going to be plenty of data, I would
think, to suggest that better-organized --
that the availability of beds makes a
difference. No? Okay.

MEMBER STOTO: But, you know, I
think that we have to -- I want to be careful
about not saying we lower standards, but just
think about it differently. Because the state
of this field is just very early. The fact
that we don't have evidence doesn't mean that
it is not important.

DR. BURSTIN: So, two things to
add to that, the first of which is, if it is
an outcome -- and I guess the question is, is
this an outcome? I am not sure this really
is. No, I guess it is a process measure. So,
that doesn't really count.

But we actually do have an exception as well which we put in there specifically for areas that just aren't as further along in terms of evidence, which is that if there is no empirical evidence, expert opinion is systematically assessed with agreement that the benefits to patients greatly outweigh potential harms. I mean, to me, this seems like a logical area for where the exception could potentially be invoked. But I think it is important to note that it probably will need to be invoked for these kinds of measures.

CO-CHAIR PITTS: So, I am speaking for Suzanne here, who wonders if there is not data internationally in places where they have disasters every couple of weeks, Israel or someplace like that.

MEMBER STOTO: You know, there is another issue here, too. Where did the 20 percent come from? It may be that 20-percent
increase in capacity doesn't do you any good at all in most kind of exercises. Maybe you need to double or triple your capacity in this regard. But that, to me, is a bigger challenge than the idea that you need to increase it.

CO-CHAIR PITTS: Wes?

MEMBER FIELDS: Well, I may be missing something here, but it seems to me that this provides the best rationale for fusing these two areas of activity. Because I think if you go back to sort of the calculus of this, there is lots and lots of peer-reviewed evidence that crowding translates into bad patient outcomes. If you aggregate all that, I think it provides a compelling case for why, if we are having this much trouble managing low- to mid-range surges in demand, then we can reasonably conclude that, whether it is 20 percent or 40 percent above the highest level of the tide in the nation's EDs, that additional capacity needs to be
created.

So, I actually think that is the reason why we are all here. It is ironic that crowding data that has been peer-reviewed can make the case for achieving a fairly high-level national strategy around surge capacity.

CO-CHAIR PITTS: Peggy? Yes, sorry.

MS. SPARR: I'm sorry. I am new to ASPR. As I have been aware, this standard of 20 percent, even for individual hospitals, it has been above their normal daily operating ability, not for IBA going 20 percent below it.

But it is tied, also, to the ability of when a community will essentially say, "We need help." So, if you can do that, somebody has come up with the number, and it predates me, but it is tied to at what point you are going to ask the state to come in to offer resources, and then the state to ask the feds to come in for resources. So, it is not
just a number; it is a number tied to how many
days that you can hold off, how many hours you
can hold off on your own without anybody else
coming in.

MR. PINES: And also, just to
clarify the discussion here, essentially, we
are talking about the specific criteria under
impact. I don't see a major issue with a lot
of preparedness measures being able to show,
certainly, a national impact here.

But, specifically, under
performance gap and evidence, where there may
be issues specifically tying specific
interventions to outcomes -- so, for example,
IBA is an idea right now. It is not something
that has been necessarily tested. How do we
modify our criteria for acceptance of
performance measures where there may not be a
lot of data on performance gap? And what
Helen had said we could potentially do is have
expert consensus.

And then, basically, for evidence,
in a lot of the preparedness literature there is not really any sort of tie to outcomes, for the reasons that Mike mentioned earlier. No counterfactual; you know, you don't really have a control group, and other issues. You don't really know what would have happened had you not done what you did.

So, essentially, how do we modify that language to make sure that measures like IBA can actually get through?

MEMBER STOTO: Well, you know, actually, in a way, I don't think we do, thinking about it that way, because this is a major feature of the HPP capacities and the PHEP capacities. Just the way that we would say that the Preventive Services Task Force recommends this, I think that has similar standing.

CO-CHAIR PITTS: Brent?

MEMBER ASPLIN: I think we could do that, but I thought that was the whole point of this meeting. You know what I mean?
So, if we were going to do that, why have this second phase of the whole process? We could have just been doing the measures today?

I am not suggesting that you are saying that is your preferred outcome. I would really like to avoid it, if we can, though, because, ultimately, I think it serves this body of work best if we meet the same standards for the whole consensus-development process as the other measures that go through NQF. I think it will help the body of work better.

MR. PINES: I think the issue that we are going to run into, especially with taking a look at the measures from the environmental scan and other measures that Mike has brought up, essentially, we would systematically make it very difficult for any measure to get through if we use the current NQF criteria. And essentially, not for crowding measures, but specifically for a lot of the preparedness measures that are HPP.
MEMBER VENKATASH: The last part I think that is worth a quick discussion within evidence is the consistency standard, especially because, as more measures have been getting reviewed recently, that has been one where there has been a lot more, I think, within Steering Committees discussion.

What that basically expects is that, for things that have been studied in multiple contexts -- and I can imagine the situation here being that, if we look at historical examples, achieving consistency is going to be difficult. It would be easy for somebody to say, "Well, that experience is a little different than this one," right there. That shooting in Aurora was different than that flood there.

And so, coming to some agreement around what evidence consistency means here for the measure focus I think is actually going to be something that would be fairly challenging. Since I don't know this
literature at all -- and maybe, Mike, this is looking to you and others in the room -- the idea around, how well agreed-upon is the idea that what we have known from historical examples, if that is what we are going to use to get at this idea of importance, how consistent is that across different examples?

MEMBER STOTO: This is a good question. I mean, I think that consistency in this context is also basically used the same way it is used in a systematic review. I mean, do you see not only that there is a strong effect in a meta-analysis, but there is relatively-little heterogeneity. So, we just don't know that because this kind of study hasn't been done.

But I do think, though, that if you think about it logically, there are some, I guess there are likely to be some kinds of emergencies where having surge capacity is important; some where it is not that important.
Now does that mean that we shouldn't have a measure of it, because it is only important in some situations but not in others? I don't think so.

What I would say is that we interpret this in terms of heterogeneity in a meta-analysis, and it just doesn't apply here. I don't want to seem like an apologist for this, but I think that the process, we need to really think carefully about how to apply the process to a very different situation for the situations we used before.

MEMBER VENKATASH: I mean, if that is the recognition that is made from the group, that is a big deal because right now consistency is a must-pass criteria. So, if there was a measure developer out there developing a measure --

MEMBER STOTO: I am not saying that it shouldn't apply. I am just saying that it doesn't -- we just don't know. We are not in a position to assess it because we
don't have the kind of statistical studies
that you would in others.

MR. PINES: But, also thinking
about consistency in a broader way, you know,
are there ways to measure consistency or other
potential standards that we could either
replace that with, that would kind of seem
like a similar bar for preparedness measures?
Because, essentially, I think what we don't
want is what seems like sort of measures that
are sort of experiments, that are sort of put
out there, saying that this is what we think
we should do. So, this is your measure of
performance.

So, essentially, that we do have
at least some sort of an expert panel, expert
consensus, some rigorous methodology that was
used. I know that for HPP there was a lot of
rigorous methodology, a lot of expert panels
that have gone into develop those measures.
That could potentially be used for this.

MEMBER STOTO: And maybe this is
the place where the simulation studies you
mentioned, you suggested, might come into
play, is that you can sort of do it, look at
it over a range of different kinds of
situations; does this consistently make a
difference?

CO-CHAIR PITTS: Okay. Brendan?

And then, Emily.

MEMBER CARR: I also wonder if
there is room to lean on crisis standards-of-
care document to talk about the fact that
there are times -- I don't know. I am
sensitive to Brent saying he doesn't want the
rules to be different. But, at the same time,
you know, we are being asked to play by
evidence-based rules in a world that doesn't
use evidence-based decisionmaking. This is
the intersection of public health and
healthcare, and it is tricky to do that.

So, I don't know. I wonder if we
sort of go through the crisis standards-of-
care conceptual framework and try to apply it
here, if we feel less conflicted, given that central entities have said it is okay.

MEMBER ASPLIN: Keep in mind I am coming at it from the crowding side more than -- so, I am spending more time in that world. So, I probably feel more strongly about that, those measures, than I would from preparedness.

MEMBER CARR: Sure. Sure. Yes.

DR. BURSTIN: I think it is actually quite different on the emergency preparedness side, where there just isn't a robust literature to draw on. We wouldn't, for example, let a lot of the care-coordination measures go forward, even though there is systems research that clearly showed evidence that was applicable.

In this case, again, I think we did this for some of our palliative care measures, a new, emerging field as well. There aren't a lot of studies on the benefits of spirituality for patients undergoing end-
of-life, but, boy, you would sure note, you
know, there is a lot of expert consensus that
says it is important. So, we are not going to
hold back a measure that is going to have
benefits to the nation just because it doesn't
emerge from that same level of database.

CO-CHAIR PITTS: So, Emily, and
then Rick.

MEMBER CARRIER: I was just
wondering if there was any benefit to thinking
about the frame of measures around never
events, particularly for the disaster
preparedness. I don't know much about how
those measures are developed, but no one is
going to do an RCT of wrong-side surgery
certainly. And yet, we are able to build a
measure that this should never happen.

Are there things in preparedness
or in crowding that are never events?

MEMBER ASPLIN: Apparently not in
crowding.

(Laughter.)
CO-CHAIR PITTS:  Rick?

DR. HUNT:  I am not sure if we have resolved the performance gap piece, but I want to get really concrete, really specific, and really practical for just a second.

Performance gap. After four days at the Madrid hospital that saw the majority of the patients, they said, "We did not perform very well. We performed very badly."

They saw 272 patients in 2.5 hours. For the ER docs like me around the room, that ought to make you really tachycardic and lose bodily functions. Okay? Two hundred and seventy-two, 2.5 hours. "We performed very badly." That is what he said. That was the surgeon in charge of the hospital that day. They performed badly in terms of they had a long list of these are the things we did not -- it was a long list.

And so, when you think about where there is a performance gap, I kept going,
well, how in the world did they do this at all? It was a 1600-bed hospital. Blocks were 1600 beds. And by the way, we got lucky because it was switch of shift. So, they had staffing for 3200 patients.

So, if there is a performance gap -- and that is real, that really happened. It is not theoretical. So, in terms of are there lists of kinds of events -- and people can do Auroras. The gaps definitely exist.

So, I am hopeful that that one we can get beyond.

CO-CHAIR PITTS: Anthony?

MEMBER MacINTYRE: Just the literature, the medical literature is scant and mainly anecdotal. There are recurring patterns of process issues that are repeated. But I wonder how looking at other bodies of literature might be approached, not to increase your workload, Jesse. But there is a tremendous amount of emergency management literature. There is a good body of business
crisis continuity literature. There is a whole lot of DoD literature out there, consultant-generated and otherwise. If you really want to look at organizational management, there could be some good pieces out there.

    Just as a specific example, you know, DoD has spent a lot of time looking at how to develop competencies for individuals to perform in unusual situations, and how do they maintain those skills, knowledge, and ability. I think there are some parallels there. Obviously, you are not going to develop a whole new NQF standard based on one article, but it might be worth looking at some of these other bodies of literature. We tend to stay siloed sometimes in our own realm.

    DR. BURSTIN: And NQF has experience doing this with our safe practices, for example, where it is perfectly reasonable to invoke other industries where safety has been so far ahead of us for quite some time.
So, that is very fair game, a great suggestion, Anthony.

CO-CHAIR PITTS: Jay?

MEMBER SCHUUR: I really like Emily's idea around the never events. I think that that is an interesting frame.

Maybe you guys can remind me.

When the Patient Safety Committee sort of went through never events again, as I remember, ACEP put one in which was death in the waiting room, as an idea of a boarding never event. And from what I remember of the form, there was a different sort of standard. It wasn't the typical full measure form. It was a shorter form.

And has that now changed with the new evidence development process? Or is there a different process that that goes through?

DR. BURSTIN: It is a different process for both serious reportable events and safe practices. But I think the example that was raised by Emily was the wrong-side
surgery. That actually is also in our Quality Indicator. So, I mean, sometimes they are measures and sometimes they are practices or serious reportable events. And I think you can go either way.

CO-CHAIR PITTS: Mike?

MEMBER STOTO: I could think of lots of potential issues, preparedness issues, and measures for which there is no good evidence, and people don't really agree. But does anybody really think that it doesn't matter whether you can increase the capacity of your healthcare system during an emergency? I mean, that just seems so obviously true, right?

MR. PINES: Well, I think the question would come particularly if a measure was submitted for IBA and the 20-percent number was used. We would, hopefully, want to see some sort of rigor around that 20-percent number or at least some expert consensus or at least some sort of a methodology that would
demonstrate that that number was vetted.

MEMBER STOTO: Yes. No, I agree about the 20 percent. But about the concept of being able to increase your capacity to treat acute cases, it is hard to imagine that not being true.

MR. PINES: Right. Right, and I think that would come under the importance to measure and report.

MEMBER MacINTYRE: You also need to have assumptions, though, with that. And the assumptions are that the sky is blue, Metro is running, the bridges haven't been closed for the Inauguration, on and on. Because, unfortunately, most disasters do occur with those other issues, and that impacts your ability to surge 20 percent. So, having some assumptions in there is pretty important, even if you do scope out the number well.

I mean, around here, Hurricane Isabelle, Metro shuts down; there goes your 20
percent.

CO-CHAIR PITTS: Did you cancel your comment? Okay.

All right. Arjun, do you want to continue on?

MEMBER VENKATASH: I think it is probably, given how much time we have, worth going on to scientific acceptability.

MEMBER STOTO: I haven't heard actually how it is defined, this measure.

MEMBER VENKATASH: Well, it is not specified. Right now, it really sits at a measure concept level. So, that is why I say think of it as, if that were the concept, what would your expectations be at each of these junctures? Especially that applies primarily to Category 1, importance, and Category 2.

MEMBER STOTO: Well, I mean, how can we talk about validity and reliability if we haven't defined it?

MEMBER VENKATASH: That is a good question. I think in some of the sub-
questions, though, within validity and reliability, we can get it without even the definition. And that is that scientific acceptability as a whole contains two categories: first, validity, and then, second is reliability.

Part of this does get at how specific, how well it is specified, but a lot of the validity question actually gets to how well the specifications are supported by evidence.

So, now, with a lot of the preparedness measures, for the same reason that we may not have evidence around the measure focus above, it would be the same reason why we wouldn't have evidence, potentially around specific measure specifications.

And the classic example I like to use here is denominator exclusions within this. And that is thinking about, when you apply this measure and we think about who is
in and who is out, which hospitals you
measure, which hospitals you don't measure,
which patients within a hospital within a
region would be measured, which ones wouldn't,
the expectation is often that, if the patient
is going to be excluded, that is based on an
evidence base that lives below that.

It may be that for these measures
this doesn't become a huge issue because we
say everybody is in, which I think is probably
the clear concept and gets to some of what we
were saying before.

But if there are going to be
exclusions, there is not going to be evidence
base around any of that. I don't know if that
applies well, but it is certainly worth a
little discussion.

MR. PINES: I think that that kind
of gets too deep into the weeds. We can't
even think about that.

So, I am thinking there are a
couple of ways I can imagine of assessing
this. One is you can go around to whatever these regions are and say, "Can you bump up your capacity in four hours by 20 percent?"

And they say, "Oh, yeah, yeah, we can do that." I mean, I wouldn't regard that as evidence that is valid and reliable.

Well, then, I might say, "Well, how about if you did it in an exercise where you actually called the people in and did it and moved them around, and so on?" Well, then, I would think a lot more highly of that.

And maybe you guys have thought this already, but I think until we hear whether it is that, one or the other or something else, we can't even begin to think about these other issues.

MR. PINES: And I think we are also thinking about the reliability and validity in the context of what instrument is used to actually measure what you are trying to measure. I think there are a lot of ways to do that.
I know that for emergency preparedness, particularly preparedness, on the preparedness side there are tons of instruments out there with variable reliability and validity, some with a lot more testing than others. That could potentially go into the NQF submission. You know, really linking preparedness to actually may be a little bit more tenuous.

And essentially, one of the only things we may have in that part of the application would be the reliability and validity of the instrument itself.

MEMBER STOTO: But this is a particular measure we are talking about, your ability to increase your capacity for acute cases by 20 percent in four hours. And so, the validity or reliability of some other issue is not relevant here. This is a specific thing for which you can actually get some evidence on this specific thing. But you have to specify what you mean and how you are
going to measure it first, before you can even
think about this.

MR. PINES: Right, right. So, I
mean, it could be done through a tabletop
exercise. And essentially, you would
basically do repeated tabletop exercise to
demonstrate. I mean, that may be what we are
talking about.

CO-CHAIR PITTS: Oh, I'm sorry, go
ahead, Peggy.

MS. SPARR: I just wanted to make
sure because I am not clear. I don't think I
heard Marco say that this morning. So, I just
want to make sure, when people think about
IBA, that we are not talking about surge
above; we are talking actual offloading of
people within four hours. I just want to make
that within four hours you will have beds
available already versus having to pull them
out of your storage area, call in more staff,
because they are already there. And that is
his novel concept.
CO-CHAIR PITTS: Okay. Okay,
Arjun, go ahead.

MEMBER VENKATASH: So, I guess
within some of this reliability and within the
validity question, then, I guess, a reasonable
question for guidance is, would an attestation
measure be something that would be considered
for endorsement? Or is the general group
consensus that, listen, that bar is way too
low?

To demonstrate that this is worth
measuring, not worth measuring, but that this
is valid, at least there has to be some data
that suggests that it is done via exercise or
something else. Is that reasonable?

MEMBER STOTO: To me, it is.

MEMBER ADIRIM: There is a
systematic way of assessing.

MR. PINES: Yes, I think that is
what we are looking for, is basically to see
how much the group really agrees with this
language. You know, we will certainly send
this draft document out to the group. But, essentially, to make sure that in the context of the NQF standards that it does not appear like we are going to basically modify the standard so that, in order to get measures through these campy ideas, there has to be some scientific basis behind them.

CO-CHAIR PITTS: It sounds like we all agrees attestation is not good enough. Okay.

MEMBER VENKATASH: I think it is sort of related in some way, you know, in the specified measure, but the usability and feasibility side of this. If a measure is developed in this space, I think on the crowding side this seems much clearer because a lot of these things you can take a group of hospitals, you can see what it took to measure it, and then demonstrate how understandable those findings are to various people, right, be it boarding time, waiting time, any of those? That kind of makes sense to us.
On the preparedness side, what is a demonstration of usability? Does a measure developer have to demonstrate that this information is meaningful at the policy level, patient level? How do you demonstrate that the findings from this are meaningful, outside of saying that we really wanted to measure preparedness, I guess?

MEMBER STOTO: Oh, in the public health preparedness world, what it usually comes down to is these are one of the things that CDC has required that we have been reporting, and people have been looking at it to make judgments, and so on.

DR. BURSTIN: I think the read of what we have under usability and use fits the extent to which potential audiences, which in this case could be ASPR, CDC, and others, are using or could use the performance results for both accountability and performance improvement. I think that is actually a pretty easy fit on this one. Arjun, I am not
too worried.

MEMBER VENKATASH: And then, if you look down at feasibility, about the same general consensus, that if somebody has collected some of this data in order to be able to apply and demonstrate it, that that is kind of the demonstration of feasibility?

CO-CHAIR PITTS: Yes, Wes?

MEMBER FIELDS: I think where we may have to demonstrate feasibility is if we really get to the point of demonstrating how hospital systems cooperate and collaborate in this mode. And again, that might be a little challenging, but I think the concept of scalability has come up several times and it makes sense to me. But whether or not that will meet the test for a standard measure, you guys will have to tell me.

MEMBER VENKATASH: That is actually a good question. If a measure is developed and the level of measurement is specified as coalition or region, should data
already be present at that level? Or is hospital-level data from multiple different regions but not all within one existing coalition sufficient?

DR. BURSTIN: It depends on the level of analysis. So, I think it will work either way.

MEMBER CARR: You think it will work, even though we don't know what the region is or what a coalition looks like?

DR. BURSTIN: I think you have to have preciseness --

MEMBER CARR: Okay. Yes, right. It does strike me that the geographic unit here --

DR. BURSTIN: Yes.

MEMBER CARR: -- is a really big gap. I mean, we all could sum the data, but we do have to know what we are summing it to. So, will they say, "We are not letting anything through until you tell us what the unit is?" or will they allow us to have a
nebulous unit?

DR. BURSTIN: I mean, one of the
most important issues here is precision of
specifications. You have got to be able to
compare apples to apples. So, that is the
goal. So, the precision is important, but we
do require measure testing at the level of
analysis in which it is intended. So, in some
ways, having a regional measure makes it
easier because it is a whole lot easier to be
potentially, with data already collected, to
do some of the signal-to-noise analysis at a
higher level of analysis than it is to take
the deeper dive on the reliability of the
provider-level data aggregated up.

So, there are different approaches
here. I think both of them are workable.

MEMBER STOTO: I think that might
be true for this particular measure about
capacity, because this really is summing up,
I think. But some of the other preparedness
measures we want to do really have to do with
how the different parts of the system work

together. And that is very different from

summing up and that would be a different kind

of situation.

CO-CHAIR PITTS: Jesse? Did you

have something? Oh, I'm sorry.

MEMBER VENKATASH: Actually,

Brendan just raised a good question, which is,

I mean, currently, as measures would be

specified, you would have to specify, if you

say this is a regional measure, what the

region is.

DR. BURSTIN: You have to define

what a region is, right.

MEMBER VENKATASH: That probably

needs big, red underlining in terms of

expectations of developers.

DR. BURSTIN: And I think there

are some good examples. I mean, for example,

if you look at the AHRQ Prevention Quality

Indicators, which are defined at a community-

level, a population-level, those are examples.
You recently endorsed a measure in population health, one of the only, looking at late-stage presentation for HIV, again, at a more regional approach. So, there are some examples.

There are plenty of HRSA regions and HHS regions and lots of ways to cut and paste these things, as appropriate.

MEMBER ADIRIM: Does it have to be so specific that it is saying that it is a state being a region versus just coming up with a definition, so that there is flexibility for local communities? So that, for example, a region could be multiple states versus a local community. Tiered. I mean, I could see various ways to define a region.

MEMBER VENKATASH: Would it raise a harmonization issue? If you have multiple measures with differently-specified regions, then if you endorsed multiple measures, the same agents are functioning within preparedness under different collaboratives or
different collections of what a region is.

MEMBER ADIRIM: Yes, but for
preparedness you would have the same elements
within a region. The capabilities would be
similar, right? So, you could compare a local
region -- I may be hallucinating -- but can
compare a local region, because it would have
certain elements that you would need to handle
disaster versus a multi-state region maybe?
I don't know.

CO-CHAIR PITTS: Go ahead, Gregg.

MEMBER MARGOLIS: Well, I don't
know if I have the answer, but we have
certainly had a lot of conversations about
this, and more questions than answers, I am
sure.

First of all, I think at least in
terms of the regions, there is some element of
geopolitical borders here that probably do
make sense, although we know that patients
don't necessarily follow those geopolitical
borders in terms of referral patterns. But
there are elected officials, whether at a county level or a municipal level, that bear certain responsibility and interest in some of these variables for the communities in which they have been elected to serve.

So, there probably is, at least in terms of preparedness and some of these other variables, some value in being able to make a comment to elected officials that the preparedness of your county or city or state is this. And that kind of gets a little bit to the accountability piece that was brought up a little bit earlier.

But I don't think that that is the whole piece of it. One of the things that we have been talking about and kicking around a little bit is what might some of the Dartmouth Atlas work in terms of geographic variations in care be able to offer to some of this conversation? And we were just having a little watercooler conversation. You know, what might be the analog of the hospital
referral regions for emergency care? Is it possible -- and Ryan might be able to help us with this -- to define that people that have emergency care issues within this geographic area have a 90-percent or 95-percent probability of staying within the healthcare resources of this area and, therefore, helping to define maybe not health or hospital referral regions, but maybe emergency care referral region? And do those, in fact, line up with some of these other ways to look at geographic variations in care, such as the Dartmouth Atlas?

CO-CHAIR PITTS: So, Ryan?

MEMBER MUTTER: So, yes, there have been a lot of these sort of conversations. We have kicked around a lot of ideas and at this point probably generated more heat than light.

But there are a few ideas that sort of come to mind. One is sort of what I will regard of as a positive approach, which
is we sort of come up with some kind of
construct and sort of empirically build a
region.

The other is sort of what I regard
as a bit of a negative/testing approach, which
is to look at an existing construct, whether
it is a geopolitical boundary, say a county,
or something like that, or sort of more of a
scientifically-based measure, an HRR, or
something, and then look at unplanned critical
illness and what is going on there.

So, there is an idea that it is
sort of from the economics literature. It is
Elzinga and Hogarty. Some of you all have
talked to me about this. You have heard me
talk about them.

And they came up with a measure
basically for alcohol. What they did is they
looked at an area and said, okay, so there are
people in this area and there are breweries in
this area. What we are interested in is how
much beer is coming into this area to be
consumed by these people and how much of the locally-produced beer is going out.

Okay. So, let's do the analog for healthcare, for unplanned critical illness. Let's look at a region and there are hospitals in there, and there are injured patients in there and all around. So, what percent of the injury or the unplanned critical illnesses in the HRR, let's say, are being treated in there? What percent are going out? And if that is a high number, then that suggests that maybe HRRs aren't so good for this. And what percent of unplanned critical illness is outside the HRR flowing into it? If that number is high, if that percent is high, maybe these things aren't so good.

So, basically, it is sort of an assessment. We could do an assessment of this sort for different constructs, be it HRRs, HSAs, counties, whatever. So, that is one sort of aspect of it.

The other aspect of it is to look
at variation in some quality-of-care metrics.

And there are a few that we could use within these entities, HRRs, HSAs, whatever, and sort of see, is there variation across these entities in these quality measures that we are interested in? And if so, if there is variation in quality, well, that suggests there is an opportunity for improvement.

Maybe these things are interesting.

And so, anyway, those are some of the concepts that in various conversations, and whatnot, we have sort of kicked around, because this is like a very big issue. It is, what is this area that we want to assess? Is it connected to the idea of coalition-building? Is it exactly the same as the coalition? Is it related? So, there is sort of the definition for coalition-building purposes, and then there is sort of the definition of area for sort of quality measurement purposes. And there may be an extent to which those two are related.
CO-CHAIR PITTS: Brendan?

MEMBER CARR: So, yes, Ryan and I spoke, and Gregg, about this at length. But, then, the last piece is, if we are going to talk about incentivizing geographies, regions, coalitions to do something, to cooperate, what happens to all the shared space? I mean, these are not going to be crystal-clear lines in the sand. There is going to be, you know, as you get further and further away from one coalition, you are going to bleed into the other coalitions. So, there is going to be lots and lots of -- I talked this morning about white space, Marco, because I worry about white space all the time, people that aren't part of the team.

And sort of going back to my trauma system roots, people that are not part of an inclusive system, they are sort of on their own, but I also worry about the shared space and the redundancies because we don't know which coalition. We don't know how to
share credit or responsibility or blame for poor outcomes for those places that are shared by multiple coalitions. You can imagine how miserable this becomes in urban areas where there is a lot of overlap.

CO-CHAIR PITTS: Jesse?

MR. PINES: Just a quick comment. I just want to make sure, you know, I think this is a really important discussion. I think we are going to be integrating a lot of the sort of need to empirically define regional boundaries as a recommendation.

But I also just want, just sort of getting back to this specific question about the precision of specification, regardless of whatever region it is, whether it is HRR or whether it is these newly-defined empirical regions, it has got to be precise when it is submitted to NQF.

CO-CHAIR PITTS: Mike?

MEMBER STOTO: You know, I think that this is a critically-important issue for
ASPR in implementing this approach. But I don't think it is so much a measurement issue or a harmonization issue. I think the real critical issue is how do you define these regions.

We did some work here in this region that we are sitting in at this moment now. We came up with five different views of what the region was. I heard two different ones, in addition, already just in the last few minutes.

So, I think coming up with regions that make sense is really hard, but that is the hard part. It is not a measurement issue.

MEMBER MARGOLIS: If I could actually just respond to that, they are not regionally-exclusive, either. I mean, it is entirely possible that you might want to be able to, for different purposes, if you are an elected official, you are only concerned about the entities within the borders of your community.
But if you want to look at it differently, I think it would be really cool if you are able to define the regions in different ways based on different needs.

MEMBER STOTO: Cool, but hard.

And the other dimension is that in this region here we cross state boundaries, maybe as many as five, depending on how you call it. You know, the Governor of Virginia doesn't care much about Maryland.

CO-CHAIR PITTS: Wes?

MEMBER FIELDS: Yes, actually, I think this sounds an awful lot like a management problem. I think if we assume that, from now on, it is a mutual problem, then, to me, what that means is that the usual chain of command that hospitals respond to are the ones that are going to make the most sense because of the improvement opportunities won't happen from disaster to disaster. They will happen from, you know, micro-event to micro-event.
So, I think, especially if the launch point is metrics and measures, which in theory changes the way providers behave and how hospitals deliver services, to me, what that implies is you may need the geographical management scheme for the big blowoff, but to actually make the hospitals work better, and for them to cooperate more across systems, you are going to have to use fairly familiar-looking management structures that aren't geopolitical, and they are much more about how the money flows.

CO-CHAIR PITTS: Arjun, are you still proceeding onward or are we done?

MEMBER VENKATASH: We can be done.

CO-CHAIR PITTS: Okay.

MR. PINES: So, one thing I did, I know we are sort of ending our time in the next couple of minutes here. I did want to just have a very brief discussion with the broader group. Essentially, we were asked to set up a runway for measure development, and
I just wanted to have a quick discussion about
the who and how, and to do a little
brainstorming about who are the measure
developers for this, how do we communicate
with them now to let them know that there may
be a consensus-development process in the
coming years, and what sort of tools do they
need? And what sort of mechanisms can we use
to make the highest likelihood of good
performance measures in this area?

MEMBER MUTTER: So, I will mention
that AHRQ has an IDIQ mechanism that we use to
facilitate measure development. That is the
means by which we are doing our current
emergency department Quality Indicator work
that we are just beginning right now.

Other agencies who want to use
this mechanism, basically, through an IAA, can
do so. For example, the current work we are
doing, as I mentioned before, has got a mental
health component to it that is funded by
SAMHSA. So, if there were other agencies that
wanted to support this kind of work, our IDIQ is a possibility to get that done.

CO-CHAIR PITTS: Does HCUP -- what kind of identifiers or at what level do you identify these types of data, all the way down to the hospital, the region, the address? What is available?

MEMBER MUTTER: So, I mean, I think some of that sort of depends on the scope of work that is envisioned. I mean, a lot of AHRQ's Quality Indicator work has been built around data that is coded like HCUP because we have HCUP data.

That doesn't mean that measure development work has to be done on HCUP-looking data. Basically, the IDIQ is a mechanism. If someone were to want to come in and do work on a different type of data, I mean I think that is a conversation. It is certainly not by nature limited to admin data that looks like HCUP.

CO-CHAIR PITTS: Arjun?
MEMBER VENKATASH: I think one obvious potential operator is the Joint Commission, right? They have already got flow standards. They already have a lot of standards around preparedness to some degree, and they have experience in the measure development space.

CO-CHAIR PITTS: AnnMarie?

MEMBER PAPA: I don't know, but what about IHI? They have a lot of measures and they do a lot of -- I don't know if they actually develop the measures or they use other people's measures. So, I can't answer that. But just something to think about.

MR. PINES: Jay, any thoughts on ACEP potentially being a measure developer in this area?

MEMBER SCHUUR: I think probably saying that the organization is going to be discussing over the next six months, but at this point I would say, no, I think not by themselves, but in collaboration with other
entities, whether governmental or academic
groups.

DR. BURSTIN: There is actually a

nice example in the world of stroke
measurement, where it is a combination of the
American Stroke Association, the Joint
Commission, the CDC, who actually came
together as a coalition to develop a set of
guidelines that everybody is comfortable with,
and then the Get with the Guidelines folks are
involved as well.

So, one thought might be kind of
thinking about how are the right people at the
table and seeing if there is a way to kind of
have some collaborative development, so you
don't wind up with measures in different
spaces that are really conflicting with each
other.

MEMBER CARR: For the less-
initiated, could we get like the 90-second
primer on how you go and who goes to engage
them? I mean, does NQF do advocacy in that
realm, to build coalitions? Is that something that feds usually do? Or we just hope for the best after we put the paper out?

DR. BURSTIN: It happens in a whole variety of ways. I think in many of the clinical areas those coalitions sort of already exist. They are sort of more natural. I think the question is if we could help play any matchmaking role here. It sounds like some of this is just who is ready around the table. It might be a good, logical place to start here.

But we do put out the measurement gaps. So, a lot of what this report will indicate is where the gaps or where measures need to be developed.

And frankly, just to be honest, measurement development costs money. So, the other question is, if there are dollars on the table, that people can put on the table to say here are the defined gaps that ASPR and others find to be the three-four most important
measures we need developed in the next year,
I don't think it would be that hard between
AHRQ and us and others to find the right
to help you with that.

CO-CHAIR PITTS: Wes?

MEMBER FIELDS: That just takes me
back to my first response this morning to the
pitch. If you have $350 million funded, it
makes a lot more sense to me to use it to do
the focused measure development we have been
talking about rather than trying to pile it on
an ACO for disasters.

MR. PINES: So, any final comments
for the group? We have a few administrative
things before we end, but any other final
comments?

(No response.)

Angela is going to be taking us
through the timeline, but, basically, we have
to get this done this fall. So, we will be on
a very tight timeline.

But I just wanted to go ahead and
thank everyone for all their attention today.
I know today was a very long day. There were
a lot of really great comments, and this is
very helpful as we develop this document.

MS. FRANKLIN: Okay. And before
we wrap up, Arnika, I wanted to check to see
if there is public comment out there.

THE OPERATOR: Once again, to ask
a question, please press *1.

(Pause.)

There are no questions at this
time.

MS. FRANKLIN: Is there any member
of the public in the room who wants to make a
comment?

(No response.)

There's none? None here. We are
good.

Actually, what I would say, I will
turn it over to Adeela. But we will be
getting back with you to set up a call in the
next two to three weeks to go over all of the
changes that we want to aggregate into the
document, based on your comments. Please
don't hesitate to email us comments and
suggestions as well in the interim. But you
will be getting a poll from NQF as to your
availability for a call to get together to
walk through the new document.

MR. PINES: And one thing I
forgot, I just wanted to thank our Co-Chairs,
Suzanne and Steve, for leading this meeting
and, also, for all the folks who put together
extra stuff for this meeting, extra
presentations: Mike Stoto, Dave Marcozzi, and
Arjun Venkatash.

MS. FRANKLIN: If there is nothing
else, I think we are adjourned.

(Whereupon, at 5:02 p.m., the
meeting was adjourned.)
meetings 9:19

Melissa 2:13 22:21

member 14:5,15

member's 14:5

metaphor 136:5

mentioned 38:21

message 209:8

met 1:17

metaphor 136:5

method 149:4

methodological 26:12

methodologies 54:17 185:1

methodology 185:15 205:3

migrate 169:16

Mike 2:21 5:11

miles 156:17

million 64:13,18

mild 149:1

mind 16:13 62:4

mindset 125:14

mindsets 123:14

mind's 281:16

mini 20:18

minimized 160:6

minister 326:13

Minneapolis 17:22

minor 248:18

minorities 248:18
<table>
<thead>
<tr>
<th>Word</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>poor</td>
<td>422:2</td>
</tr>
<tr>
<td>pop</td>
<td>114:21</td>
</tr>
<tr>
<td>popped</td>
<td>350:5</td>
</tr>
<tr>
<td>pops</td>
<td>224:18</td>
</tr>
<tr>
<td>popular</td>
<td>16:15</td>
</tr>
<tr>
<td>population</td>
<td>40:11</td>
</tr>
<tr>
<td>populations</td>
<td>137:3</td>
</tr>
<tr>
<td>population-based</td>
<td>168:13</td>
</tr>
<tr>
<td>population-level</td>
<td>413:22</td>
</tr>
<tr>
<td>porcelain</td>
<td>288:13</td>
</tr>
<tr>
<td>portal</td>
<td>282:3</td>
</tr>
<tr>
<td>portfolio</td>
<td>24:15</td>
</tr>
<tr>
<td>portion</td>
<td>13:7</td>
</tr>
<tr>
<td>posed</td>
<td>70:11</td>
</tr>
<tr>
<td>position</td>
<td>332:2</td>
</tr>
<tr>
<td>positive</td>
<td>329:21</td>
</tr>
<tr>
<td>possibly</td>
<td>427:2</td>
</tr>
<tr>
<td>possible</td>
<td>34:15</td>
</tr>
<tr>
<td>practical</td>
<td>304:14</td>
</tr>
<tr>
<td>practicality</td>
<td>91:16</td>
</tr>
<tr>
<td>practice</td>
<td>97:10</td>
</tr>
<tr>
<td>practices</td>
<td>85:18</td>
</tr>
<tr>
<td>practicing</td>
<td>21:1</td>
</tr>
<tr>
<td>pre</td>
<td>18:8</td>
</tr>
<tr>
<td>preceded</td>
<td>31:8</td>
</tr>
<tr>
<td>precise</td>
<td>33:3</td>
</tr>
<tr>
<td>potential</td>
<td>54:10</td>
</tr>
<tr>
<td>pot</td>
<td>116:14</td>
</tr>
<tr>
<td>potable</td>
<td>184:14</td>
</tr>
<tr>
<td>potenti</td>
<td>73:15</td>
</tr>
<tr>
<td>potentially</td>
<td>73:16</td>
</tr>
<tr>
<td>power</td>
<td>149:1</td>
</tr>
<tr>
<td>pounding</td>
<td>309:5</td>
</tr>
<tr>
<td>pouring</td>
<td>212:6</td>
</tr>
<tr>
<td>powerfullness</td>
<td>395:5</td>
</tr>
<tr>
<td>practicality</td>
<td>91:16</td>
</tr>
<tr>
<td>praktisch</td>
<td>29:7</td>
</tr>
<tr>
<td>practice</td>
<td>97:10</td>
</tr>
<tr>
<td>practiced</td>
<td>85:18</td>
</tr>
<tr>
<td>practiced</td>
<td>21:1</td>
</tr>
<tr>
<td>pre</td>
<td>18:8</td>
</tr>
<tr>
<td>preceded</td>
<td>31:8</td>
</tr>
<tr>
<td>precise</td>
<td>33:3</td>
</tr>
<tr>
<td>presentations</td>
<td>43:20</td>
</tr>
<tr>
<td>presented</td>
<td>372:8</td>
</tr>
<tr>
<td>presenting</td>
<td>52:16</td>
</tr>
<tr>
<td>present-day</td>
<td>114:9</td>
</tr>
<tr>
<td>President</td>
<td>4:23</td>
</tr>
<tr>
<td>press</td>
<td>47:8</td>
</tr>
<tr>
<td>pressure</td>
<td>219:12</td>
</tr>
<tr>
<td>presume</td>
<td>339:1</td>
</tr>
<tr>
<td>pretend</td>
<td>145:8</td>
</tr>
<tr>
<td>pretty</td>
<td>13:10</td>
</tr>
<tr>
<td>previous</td>
<td>304:14</td>
</tr>
<tr>
<td>press</td>
<td>116:14</td>
</tr>
<tr>
<td>presiding</td>
<td>1:22</td>
</tr>
<tr>
<td>presuming</td>
<td>3:33</td>
</tr>
<tr>
<td>prevent</td>
<td>92:17</td>
</tr>
<tr>
<td>preventive</td>
<td>93:1</td>
</tr>
<tr>
<td>preview</td>
<td>114:6</td>
</tr>
<tr>
<td>previous</td>
<td>13:20</td>
</tr>
<tr>
<td>pre</td>
<td>182:12</td>
</tr>
<tr>
<td>potential</td>
<td>54:10</td>
</tr>
<tr>
<td>pot</td>
<td>116:14</td>
</tr>
<tr>
<td>potable</td>
<td>184:14</td>
</tr>
</tbody>
</table>

Neal R. Gross & Co., Inc.
202-234-4433

Neal R. Gross & Co., Inc. 202-234-4433
2:52, 2:24, 196
197:5:20
198:6:135:5

2 22:113:19 129:11
129:11 132:17
170:15 220:14
291:8 311:15
401:17
2.5 395:11, 15

2:5 322:6
20 50:4, 51:10
61:7 68:3 69:16
71:16 72:5 73:21
74:2, 6 104:3
109:3, 6 110:11
113:18 119:17
129:11, 12 132:17
213:16 267:12
375:10 376:13
382:21 383:20
384:11, 13 400:3
400:17, 22 404:3
405:17
20,000-visits-per...
244:4
20-percent 52:9, 19
53:16 376:9
382:22 399:18, 20
20-some-odd
380:20
200 58:5 65:15 74:5
74:10
2003 111:22
2004 268:11
2006 53:3
2008 190:10 193:10
198:12 216:1
224:15 241:15
274:11
2009 101:8 161:21
2010 269:10
2011 46:6
2012 1:13 45:1
234:8
220 6:5, 7

23 116:17 274:3
24 4:16, 18 366:13
24-hour-plus 281:3
240 6:8
243 6:9
25 116:16 192:6
25-percent 161:20
267 6:10
268 6:14
27 4:20
272 395:11
285 6:15
288 6:17
297 6:18, 20

3 3:7 275:20 290:5
290:6, 7 291:8
311:15 342:10
3-percent 80:12, 19
3:15 322:4
3:18 322:7
30 4:22 103:6
275:20 291:16
300 6:21
301 6:23
315 6:24, 26
320 396:5
322 7:4
343 7:6
371 7:8
374 7:12

4 4:4 147:2
4:9 53:13, 18
4:30 117:17
40 103:6 246:20
295:17, 20 383:20
400 58:5
42 4:12
425 7:15
43 5:4, 6
431 7:19
432 7:22
44 5:7, 8
45 116:21
48 274:3

495 206:8 216:4
496 206:8
497 206:8 208:5
216:5, 16 225:17
228:15 284:20

5 5:1 80:12 268:16
268:19
5-percent 80:19
5:02 433:17
50 158:2 217:9, 9
376:8
50-car 83:10
55-minute 202:2
212:2
58 5:10

6 6:1
60 158:2 247:1
60,000 128:12
60-percent 333:21

7 7:6:15 7:1 183:5, 10
286:18
70 333:21
72 366:12

8 8:4, 7, 8 45:6 275:5
8:30 1:21
8:44 8:2
80 45:14 66:10, 20
69:17 70:20, 22
205:22 213:15
80-year-old 224:18
87 5:11

9 9:4:10 6:21 196:16
300:6, 9 304:19
307:16 330:6
9th 1:19
90 45:14 51:1
90th 274:13, 22
275:6 297:14

298:5, 6 299:9
363:13
90-percent 417:5
90-second 429:20
95 220:1 268:16
269:20 284:10
95-percent 417:5
97 220:1
98 268:11, 13
283:20
99th-percent 271:9

Neal R. Gross & Co., Inc.
202-234-4433
CERTIFICATE

This is to certify that the foregoing transcript

In the matter of: Regionalized Emergency Medical Services Expert Panel Meeting

Before: NQF

Date: 10-17-12

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

______________________________
Court Reporter