Institute *for* Healthcare Improvement

WHITE PAPER

Whole System Measures 2.0

A Compass for Health System Leaders

AN IHI RESOURCE

20 University Road, Cambridge, MA 02138 • ihi.org

How to Cite This Paper: Martin L, Nelson E, Rakover J, Chase A. *Whole System Measures 2.0: A Compass for Health System Leaders.* Cambridge, Massachusetts: Institute for Healthcare Improvement; 2016. (Available at ihi.org)

AUTHORS:

Lindsay Martin, MSPH: *Executive Director, IHI*

Eugene Nelson, DSc, MPH: *Professor of Community and Family Medicine, The Dartmouth Institute*

Jeff Rakover, MPP: *Research Associate, IHI*

Alide Chase, MS: *Senior Fellow, IHI*

Acknowledgements:

The authors thank Derek Feeley, Rocco Perla, George W. Bo-Linn, the Delphi method participants, the IHI Scientific Advisory Group, the team at HealthPartners, Don Goldmann, and several others who provided invaluable feedback to inform this work. We also thank Jane Roessner and Val Weber of IHI for their support in developing and editing this white paper.

Contents

Executive Summary

In the last decade, the health care policy and economic landscapes in the United States have changed substantially. The passage of the Affordable Care Act in 2010 stimulated the growth of accountable care organizations (ACOs) and other value-based payment arrangements in the US, and data that focused on outcomes and patient experience gained greater visibility. As health care spending increased, many providers began to look for new ways to cut costs and curb perceived runaway spending. US health care systems grew, merged, and acquired new assets to provide services across the full continuum of care.

In line with new payment programs and a focus on value and efficiency, health care systems faced a heavy burden of complex measures to collect, report, and review. One analysis published in 2014 found more than 500 distinct measures in use in 48 state and regional measure sets.[1] This complexity is compounded by the proliferation of new care and payment models at the same time. The innovative ideas currently being tested are exciting, and they add a new layer of complexity as we test across and within systems to find the best models for improving outcomes, improving the experience of care, and decreasing cost.

The burden of measurement from payers, accreditors, and other sources, such as consumer transparency initiatives, increases the need for measure rationalization. It is important to know which measures should be tracked, by whom (e.g., health systems, communities, payers, and individual providers), and how often — so that the right measures are tracked, by the right people, with the right frequency.

The Institute for Healthcare Improvement (IHI) developed Whole System Measures 2.0 (WSM 2.0) to provide specific guidance to health care system leaders and boards on how to measure current overall system performance and use this data to inform organizational strategy. WSM 2.0 is a set of 15 measures that help leaders better understand their organization's current (and desired) state across three domains: health, experience of care, and per capita cost. This work builds on the original *Whole System Measures* IHI White Paper,[2] published in 2007, and ongoing efforts to advance the Triple Aim (improving population health, improving the individual patient experience of care, and decreasing the per capita cost of care). While directive, this small measure set creates the opportunity for health care system leaders, managers, clinicians, and staff to drill down further to understand specific performance challenges or successes, and to identify strategic opportunities for improvement.

To develop WSM 2.0, IHI used a version of the RAND Corporation's Delphi method — a technique for reaching consensus under conditions of disagreement and uncertainty, while trying to avoid giving undue weight to the loudest voices in the room. The Delphi method helped us to narrow the potential set of measures to 15 measures that health care system boards and executives can use to gauge the performance of their system as a whole.

Introduction

IHI developed the original set of Whole System Measures (see Appendix A) to measure performance against the Institute of Medicine (IOM) six dimensions of quality articulated in the 2001 *Crossing the Quality Chasm* report — to push health care systems beyond hospital-based, condition-specific measures, aiming to include broader measures that span inpatient and

outpatient care.[3] Our belief was that health system leaders needed a small set of measures that reflected a health system's overall performance on core dimensions of quality guided by the Triple Aim: the health of the population, the individual experience of care, and per capita cost.

Some leading health care systems embraced the original Whole System Measures to guide their measurement framework and adapted the measures to suit their contexts and organizational strategies. The IHI White Paper, *A Guide to Measuring the Triple Aim*,[4] was published in 2010, informed by the experience of more than 100 sites worldwide that participated in IHI's Triple Aim prototyping initiative.[5] That paper offered a menu of options for measuring the Triple Aim, based on several underlying frameworks associated with each element of the Triple Aim.

Many efforts to develop core measures of quality have emerged in recent years. For example, the National Quality Strategy, established as part of the Affordable Care Act, led to the development of a core measure set by the US Department of Health and Human Services Measurement Policy Council.[6] The IOM's (now National Academy of Medicine's) 2015 *Vital Signs* report offers another core measure set, organized around the Triple Aim.[7] In early 2016, the Centers for Medicare & Medicaid Services (CMS) released a set of core measures intended to measure physician performance, with a central focus on the specialties, as well as patient-centered medical homes, accountable care organizations, and primary care.[8]

While all of these efforts are important and substantive in their own right, they also contribute to health care measurement complexity, highlighting the need for clarity and parsimony to enable senior leaders to understand the overall performance of their systems. The individual measures that comprise WSM 2.0 are not new; pulling them together to gain the appropriate level of understanding of quality across the system is new. While we do need to reduce measurement burden, we also need to rationalize the measures that exist. WSM 2.0 is intended to provide specific guidance to health care system leaders and boards on how to do just that: measure overall system performance and use this data to inform organizational strategy. As organizations test this measure set, we hope that this approach proves helpful and leads to greater measure rationalization at different levels and by different constituents, including public and private payers.

WSM 2.0 is based on the following principles:

- **Balance:** A measure set must address each of the three elements of the Triple Aim (health, care, cost). In addition, a measure set should balance the current need of a system and the possible future direction of the system.
- **Parsimony:** To maintain a systems perspective, a small set of measures is required. If there are too few measures, significant dimensions will be overlooked; if there are too many, the measures cease to have targeted value in providing strategic guidance to system leaders.
- **Alignment:** WSM 2.0 builds on the IOM Core Metrics[9] and other existing measure sets so as not to duplicate efforts. Table 1 below highlights the alignment of WSM 2.0 with select additional measure sets.
- **Immediate usefulness:** The measures need to be useful to health system leaders and boards to drive improved performance. The measures, as a set, must signal stability, improvement, or decline. In addition, we sought to include measures that had a track record of use. We did not include measures that, while novel or innovative, might be too early in the vetting process to fully understand their usefulness. We looked for measures that met high standards for vetting including, but not limited to, professional bodies like the National Quality Forum (NQF).

- **Consensus:** Health system leaders, quality improvement professionals, measurement experts, national advisors, and experts in patient- and family-centered care have knowledge in this area to be shared. To achieve the strongest set of Whole System Measures, we sought to leverage experience from all of these groups through constructive dialogue. To this end, IHI used a modified Delphi method to arrive at WSM 2.0, as described in the Methods section below.
- **Adaptability:** Empirical validity is essential. We need measures that work in the present, and we need to account for the ever-changing health care landscape. To that end, WSM 2.0 will need to be periodically revisited to ensure that the measures serve the intended purpose. As conditions change, some measures may become less viable or relevant and others may need to come to the forefront. By both pilot testing WSM 2.0 across multiple organizations and ensuring that leaders within an organization verify internal relevance, WSM 2.0 can and should be adapted when appropriate.

Methods Used to Develop Whole System Measures 2.0

Development of WSM 2.0 included four main inputs:

1. The development of subdomains, using the Triple Aim as a guiding framework, to support the identification of candidate measures.
2. A review of existing measure sets developed to quantify system performance and/or reduce measurement burden, including a review of academic and grey literature to identify potential measures.
3. A modified Delphi method through which health care system leaders, health care quality and measurement experts, leaders from professional medical societies, and other stakeholders contributed to iterative rounds of voting, measure discussion, and recommendation.
4. Additional consultations with IHI Senior Fellows, IHI senior leaders, and the IHI Scientific Advisory Group[10] to provide guidance on and, for measures related to health equity, further refine the selected measures.

Defining Subdomains of the Triple Aim for WSM 2.0

For the three primary domains of the Triple Aim (health, care, cost), we defined a set of subdomains to provide a more refined framework for WSM 2.0. Starting with the IOM's *Vital Signs* subdomains, we made modifications to more closely match the needs of a health care system. The subdomains are as follows:

- Population Health Domain
  - Individual Health
  - Healthy Behaviors
  - Community Wellbeing and Health Equity
  - Workforce Wellbeing (i.e., for the health care workforce)

- Experience of Care Domain
  - Access
  - Prevention
  - Safety
  - Appropriateness and Effectiveness
  - Patient-Centeredness
- Per Capita Cost of Care Domain
  - Affordability
  - Societal Footprint

Appendix B includes additional discussion of the subdomains and related measures.

Aligning WSM 2.0 with Existing Measures

In an effort to align WSM 2.0 with existing measure sets, IHI used, as a starting point, the Core Metrics released by the Institute of Medicine (now known as the National Academy of Medicine) in the 2015 *Vital Signs* report.[11] In addition to considering the Core Metrics as potential measure candidates, WSM 2.0 also drew on other sources. For example, the "HHS Measurement Policy Council Core Measure Sets" addressed areas such as hypertension control, smoking cessation, depression screening, and care coordination.[12] We also considered IHI's previous research on measure alignment, including measures suggested in the IHI White Paper, *A Guide to Measuring the Triple Aim*, and the original Whole System Measures.[13] Literature searches and the Delphi method surfaced other candidate measures.

Refining WSM 2.0 Using a Modified Delphi Method

Multiple measures fit within the WSM 2.0 framework, and the proliferation of measures and sets of measures speaks to reasonable disagreement regarding the best measures to assess a health care system's performance in the subdomains.

The Delphi method, pioneered by the RAND Corporation, offers an approach to reach consensus under conditions of disagreement and uncertainty.[14] The Delphi method has been used in health care contexts to come to agreement on best measure sets, and generally involves a combination of surveys and discussions.[15] The traditional Delphi method, however, does not include discussion. To narrow the larger set of potential measures, IHI used a modified Delphi method to identify a smaller set of 15 measures that assess health care systems' progress against the domains and subdomains in WSM 2.0.

IHI recruited participants (see Appendix D) for the modified Delphi method with the intent of achieving balanced input from health care system leaders (e.g., CEOs, CMOs, CNOs, board members), representatives from government and major payers (e.g., CMS, Veterans Affairs, Blue Cross Blue Shield Associations), professional societies (e.g., Institute of Medicine, National Quality Forum, American Hospital Association), measurement experts, academic experts, and patient advocates. Of the initial 81 contacts selected through a judgment sample, 41 agreed to participate.

In the modified Delphi method IHI used to select WSM 2.0, three rounds of anonymous voting via surveys occurred, each followed by a 45-minute conference call to discuss the survey results and offer participants an opportunity to contribute alternative measures, make arguments to support their measure selections, and elaborate on survey comments if desired. Notes from these discussions were included in the next cycle of survey voting to inform all participants of the discussion. The three rounds of surveys had 30, 29, and 27 respondents, respectively.[16]

Survey participants were asked to rate each potential measure based on its importance as part of a balanced, parsimonious set of Triple Aim measures for health care system leaders and boards. The rating scale was a 1 to 5 Likert Scale (1 = Extremely Important; 2 = Very Important; 3 = Moderately Important; 4 = Slightly Important; 5 = Not at All Important).

During the second and third survey rounds, participants viewed the anonymous high-level results of previous rounds, including both numerical ratings for each measure and qualitative comments submitted on the surveys and during conference calls. Each round resulted in the exclusion of measures with limited support (i.e., those with low importance ratings) and the addition of measures suggested by survey participants, ultimately resulting in a set of 15 measures.

Measures Included in the Initial Survey

The first survey included the *Vital Signs* measures, along with a selected set of additional measures to supplement the measures for subdomains not addressed by *Vital Signs* (e.g., workforce wellbeing). Several criteria guided the addition of measures beyond the *Vital Signs* set. Based on these criteria, the initial survey included only:

- Defined measures, rather than ideal or aspirational measures not currently in use;
- Measures already used relatively widely, rather than measures not yet well tested; and
- Measures in the public domain, not commercial measures.

WSM 2.0 Resulting from the Modified Delphi Method

All subdomains included at least one measure with an average rating between 1 (Extremely Important) and 2 (Very Important), with the exception of workforce wellbeing.[17] Some subdomains had more than one measure with an average rating of at least 2 (Very Important); in these cases, WSM 2.0 includes the additional measures.

The inclusion of two additional measures for the community wellbeing and health equity subdomain resulted from consultation with leaders who voiced a preference for including a measure of equity or disparities that went beyond the typically recommended stratification of the selected measures by demographic variables like race, ethnicity, and gender (which IHI suggested in the original *Whole System Measures* white paper and continues to advocate). These measures were added after the modified Delphi method had concluded and, therefore, did not receive consideration through the Delphi process. We offer further discussion of the inclusion of these measures below.

Table 1 details the final set of WSM 2.0 — 15 measures derived using the modified Delphi method. The measures are organized by the three primary domains (health, care, cost) and the subdomains. See Appendix C for WSM 2.0 operational definitions and additional information.

Table 1. Whole System Measures 2.0: Measures to Assess Health System Performance on the Triple Aim

Subdomain	Measure	Definition	Notes: Measure Alignment, Additional Rationale
Population Health Domain Measures			
Individual Health	General health	Self-rated general health (Excellent, Very Good, Good, Fair, Poor)	*Vital Signs*[18] measure
Healthy Behaviors	Overweight/obesity	Percentage of overweight or obese adults	*Vital Signs*; NQF measure[19]; CMS measure[20]
	Optimal lifestyle metric	Percentage of adults who do not use tobacco, are physically active, eat five fruits and vegetables daily, and have limited use of alcohol	
Community Wellbeing and Health Equity	Social support	Self-reported extent to which people have the social and emotional support they need (Always, Usually, Sometimes, Rarely, Never)	*Vital Signs*
	Disparities in infant mortality rate	Difference in death rate for infants under age of 1 year between white, non-Hispanic women and: non-Hispanic black women; non-Hispanic Puerto Rican women; non-Hispanic American Indian or Alaskan Native women	Added as a measure of health equity based on IHI senior leadership vetting
	Disparities in high school graduation rate	Difference in percentage of high school students graduating in four years between: 1) students who do not have disabilities and students with disabilities; 2) students with limited English proficiency and students without limited English proficiency; 3) students from low-income families and students not from low-income families; 4) white, non-Hispanic students and black and Hispanic students; and 5) white, non-Hispanic students and Hispanic students	Added as a measure of health equity based on IHI senior leadership vetting
Workforce Wellbeing	Job satisfaction	Percentage of respondents who respond "Agree," on average, with select indicators of job satisfaction	Widely used, publicly available measure of job satisfaction
Experience of Care Domain Measures			
Access	Timely ambulatory care	Percentage of patients who answer "Always" to CG-CAHPS questions on their ability to get urgent care, routine care, or needed information from a physician's office	*Vital Signs*; NQF measure[21]; CMS measure[22]

Prevention	Childhood immunizations	Percentage of children receiving recommended vaccines by age 3	*Vital Signs*; NQF measure[23]; CMS measure[24]
Safety	Hospital-acquired conditions	Rates of select conditions in acute care hospitals (e.g., select infections)	*Vital Signs*; NQF measure (select subsets of infection)[25]; CMS measure[26]
	Serious reportable events (SREs)	Number of SREs (categories include surgical or invasive procedures, product or device events, patient protection, care management, environmental, radiologic, potential criminal)	
Appropriateness and Effectiveness	Preventable hospitalizations	Rate of hospital admissions for certain acute illnesses or chronic conditions that are preventable through effective ambulatory care (e.g., diabetes, dehydration)	*Vital Signs*; NQF measure[27]; CMS measure[28]
Patient-Centeredness	Patient-clinician communication satisfaction	Percentage of patients reporting the highest level of satisfaction with their provider's communication	*Vital Signs*; NQF measure[29]; CMS measure[30]
Per Capita Cost of Care Domain Measures			
Affordability	Unmet health care needs	Percentage of patients who either did not receive care due to cost in the past 12 months or delayed care due to cost in the past 12 months	*Vital Signs*
Societal Footprint	Health care cost per capita	Ideal measure: Sum of public and private health care expenditures divided by size of population Or, recommended surrogate measures: • Medicare: Medicare reimbursement per enrollee per year • Total Cost of Care: Primary care providers' risk-adjusted cost effectiveness at managing the populations they care for, including nearly all associated expenditures (hospitalizations, office visits, pharmacy, etc.)	*Vital Signs*; NQF measure[31]; CMS measure[32]; IHI 2007 Whole System Measures[33]

Implications of WSM 2.0

The modified Delphi method narrowed the potential measures to a set of 15 measures that health care system boards and executives can use to gauge their progress in meeting the Triple Aim of better population health, better individual experience of care, and lower per capita cost of care.

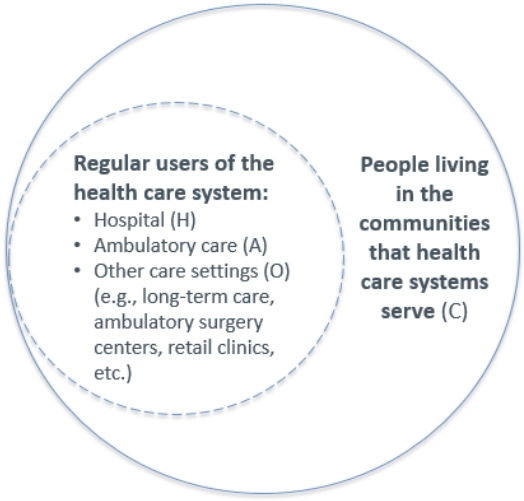
As discussed previously, the need to reflect current thinking in the shift from a focus on disease to a broader focus on health resulted in significant changes in the "health" measures as compared to the original Whole System Measures. The Population Health Domain in WSM 2.0 contains seven suggested measures: general health, overweight/obesity, optimal lifestyle metric, social support, disparities in infant mortality rate, disparities in high school graduation rate, and job satisfaction.

By focusing on these measures of health, health systems can further engage around patients' social needs and address a wider array of the determinants of health.

The Experience of Care Domain measures also include many changes from the original Whole System Measures, reflecting a desire to align with the *Vital Signs* measures and move toward composite measures like serious reportable events. The Per Capita Cost of Care Domain was also broadened to include new measures for unmet health care needs and per capita expenditures, embracing a view that is more inclusive of the communities in which health care systems reside.

Figure 1 depicts a broader WSM 2.0 framework, illustrating the suggested scope for the measures — including regular users of a health care system as well as the people living in the communities that health care systems serve; the WSM 2.0 domains and subdomains; and the specific measures and likely scope for each (i.e., hospital [H], ambulatory care [A], other care settings [O], and communities [C]).

Figure 1. Whole System Measures 2.0 Framework



Domains and Subdomains	Measures (and Likely Scope)
Population Health Domain	
Individual Health	General health (H, A, O, C)
Healthy Behaviors	Overweight/obesity (H, A, O, C)
	Optimal lifestyle metric (H, A, O)
Community Wellbeing and Health Equity	Social support (H, A, O, C)
	Disparities in infant mortality rate (H, A, C)
	Disparities in high school graduation rate (A, C)
Workforce Wellbeing	Job satisfaction (H, A, O)
Experience of Care Domain	
Access	Timely ambulatory care (A)
Prevention	Childhood immunizations (H, A, O, C)
Safety	Hospital-acquired infections (H)
	Serious reportable events (H, A, O)
Appropriateness and Effectiveness	Preventable hospitalizations (H, A, O)
Patient-Centeredness	Patient-clinician communication satisfaction (H, A, O)
Per Capita Cost of Care Domain	
Affordability	Unmet health care needs (H, A, O, C)
Societal Footprint	Per capita health care expenditures (H, A, O, C)

Challenges and Potential Solutions

The process of identifying WSM 2.0 also unearthed a rich set of challenges and solutions. The major challenges and IHI's approach to addressing them are briefly summarized below.

1. **Lack of adequate measures in some subdomains:** The Delphi participants were quick to acknowledge frustration with the existing state of tested, available outcome measures, specifically for joy in work/healthy workforce, and community wellbeing and health equity. IHI researched known and emerging measures and proposed existing, tested measures. In some cases, the proposed measures were less than ideal; as measures continue to evolve, WSM 2.0 will need to be updated and refined.
2. **Desire for clear, readily understandable composite measures:** In general, we were advised to use composite measures, incorporating numerous clinical processes and outcomes into one well-understood measure. Furthermore, measures that focused on specific clinical conditions were excluded after discussion as the initial intent was to move away from these measures toward a broader understanding of system performance for senior leaders.
3. **Resistance to including measures that are outside of the health care delivery system's control:** The Delphi panel and other experts who informed WSM 2.0 had mixed opinions about including measures that went beyond the immediate scope of the health system. Several measures (e.g., crime rate and available housing) were excluded during the modified Delphi method because leaders felt the health system had limited ability to directly impact those measures. However, given the desire to create a set of measures that is both realistic and aspirational, some broader community measures (such as high school graduation rate disparities) are included in WSM 2.0.

   The measures included in WSM 2.0 create one way for health systems to think about their broader role in addressing the multiple determinants of health. We recognize this is a start and that some health systems have moved further along by screening for and then playing a role in addressing food insecurity, utility insecurity, homelessness, lack of transportation to medical appointments, and more. We also realize that some leaders believe that they should not track, at the highest levels, measures they cannot directly impact. Taking into account both ends of the spectrum, the Delphi participants, the authors, and others consulted believe it is necessary for health systems to have a stronger understanding of the communities they serve. Tracking a few surrogate measures is important to advancing this understanding.
4. **Broadening the WSM 2.0 framework while retaining parsimony:** We sought to keep the number of measures small to decrease complexity and focus leaders. The majority of the Delphi panel participants favored identifying between 7 and 12 measures; the participants ultimately identified 15 measures to include in WSM 2.0.[34]
5. **Measuring joy in work:** The common theme of clinician burnout and the documented association between workforce wellbeing and clinical quality drove a strong desire to elevate the importance of the workforce wellbeing subdomain measure.[35] At the same time, this subdomain in particular highlights the need for development of more meaningful measures. Several commercial efforts have gained traction in recent years to help employers measure employee engagement (e.g., Gallup); however, there continues to be a lack of widely used, non-proprietary measures that gauge employee engagement and go beyond satisfaction.[36]

6. **Needed versus available measures:** Better measures are needed across all of the subdomains, and in particular in these areas:
    - Immediately actionable measure of community wellbeing for health care systems;
    - Composite measure of access that spans care settings (ambulatory, hospital, etc.);
    - Composite measure of prevention;
    - Composite measure of appropriateness and effectiveness of care; and
    - Measures of health care affordability that address both how health care impacts a patient's overall financial state, and the choices that patients make about how to allocate their own funds to health care relative to other needs.
7. **Frequency of data collection:** Data for many of the measures included in WSM 2.0 are collected annually, although some data can be and are routinely collected more frequently (e.g., hospital-acquired infections). Ideal measures for health care system boards and leaders should be available at least quarterly and electronically, if possible (to aid data gathering). The community wellbeing measures posed the greatest challenge with respect to frequency of typical data collection, with most captured annually. As we test the use of WSM 2.0, we will endeavor to understand the feasibility of updating some measures quarterly or more frequently; some may only be updated annually.
8. **Outcome versus process measures:** Following the advice of Don Berwick in his article, "Era 3 for Medicine and Health Care," and input from the Delphi panel participants, we included outcome measures whenever possible, while still acknowledging that some process measures with high relevance, such as prevention screening, are helpful.[37,38]
9. **Identifying impactful measures of health equity:** We struggled to identify appropriate measures of health equity that move beyond stratification of existing measures by demographic characteristics like race and ethnicity. While not perfect metrics, disparities in infant mortality rate and high school graduation rate were added upon the advice of senior IHI leaders (with additional discussion and vetting), in recognition of the importance of health equity in WSM 2.0. The US continues to lag behind other developed countries in infant mortality, with poor minority women in the US faring worse than similar women in other countries.[39]

    High school graduation rate emerged early in the Delphi method as a key indicator of social and community wellbeing, and was included in the *Vital Signs* measure set. The development of WSM 2.0 underscores an urgent need for the development of broad but informative equity measures. Until such measures are developed, we believe these two measures included in WSM 2.0 are suitable surrogates.
10. **Application in specialty health care systems:** We considered the applicability of WSM 2.0 to specialized health systems during the Delphi method. For example, concern was raised that WSM 2.0 would provide the pediatric community with few applicable measures. The Agency for Healthcare Research and Quality (AHRQ) has funded an effort to develop and test high-priority pediatric measures, upon which we hope to build.

Limitations of the Methodology

While the Delphi method offered an opportunity for a group of national health care leaders to identify a consensus-driven set of 15 measures, several methodological limitations exist.

1. **Potential bias in measures:** We worked to robustly survey available measures in the academic and grey literature. Yet inevitably, some strong candidate measures may not have been included in the original set of measures assessed using the Delphi method. The ample opportunity for the participants to advocate for additional measures during the Delphi method itself helped to address this concern, but some strong measures may not have been considered as carefully as a result. Pilot testing of these measures will help to address this bias.
2. **Potential bias in participants:** The Delphi participants were selected via a judgment sample. While they represent some of the leading voices in US health care — among providers, payers, government agencies, academia, and other organizations — this process inevitably results in potential bias in the selection process. A different group of participants may have resulted in a different set of measures. In addition, some of those invited elected not to participate, further introducing the potential for bias. However, analysis of those invited versus those who participated showed few differences between the groups in terms of composition in role (e.g., board member versus executive) and in type of organization (e.g., health system versus think tank). Overall, the Delphi participants included a slight overrepresentation of academic participants compared to the group invited, and a slight underrepresentation of health system participants.
3. **Potential for undue influence between Delphi rounds:** The traditional Delphi method includes a higher degree of anonymity than the modified Delphi method used to identify WSM 2.0. This creates potential for bias due to advocacy by national leaders, with understandable deference to their credentials. At the same time, all Delphi participants were leaders in their disciplines, likely reducing the potential for such undue influence. Participants were able to voice their beliefs either anonymously through the written survey or by sharing their reflections verbally in the discussion sessions. To further reduce bias, written comments reflecting the prior round of survey results and the discussion were included in the next round of survey results.
4. **Need for further development of measure specifications:** Some measures, especially in the Population Health Domain, will require further development to ensure smooth adoption in a health care system context. For example, we will need to learn how to introduce measures of general health and social support into regular use. We will need to determine how to apply the optimal lifestyle metric at a health care system level for regular review. At present, not all measures have been vetted according to standards of evaluation articulated by bodies like the National Quality Forum.[40]

   At the same time, the Delphi participants selected these measures as representative of the areas they see as crucial to understanding their progress in meeting Triple Aim goals. Some areas, such as community wellbeing and equity, lack adequate examples of measures that meet the highest standards of evaluation, but we do not believe that should slow the pursuit of measures such as these. Thus, Whole System Measures 2.0 serves as both a practical and aspirational set that should help advance both health system performance and measure development. A start for any health system is to consider the measures from the lens of patients and families, and to consider what components of the items being measured can be influenced by the health care system.

Recommendations for Use of WSM 2.0

WSM 2.0 can serve as a basis for immediate testing by a diverse set of health care systems. IHI is looking to join with others to set up a series of testing cycles to support initial data collection and measurement, and to hardwire measures as appropriate into regular measurement strategy. Opportunities also exist to use WSM 2.0 to guide shared improvement work by health care systems, whether around particular topic areas or in programs that address the Triple Aim overall. In addition, programs led by IHI or its partners may include WSM 2.0 as part of their own measurement strategy.

We acknowledge the need to develop a deep understanding of how the measures are used for the purpose of governance at the system level. While we wait for community measures to emerge and the frequency of reporting to shorten, IHI would like to understand the impact of introducing WSM 2.0 to health care boards and executives for their immediate use. Does WSM 2.0 encourage system leaders to advance their thinking around achieving the Triple Aim? To this end, IHI hopes to engage a small number of diverse health care systems to help us test and refine the measures.

Moving Beyond WSM 2.0

The journey from today's available and useful measures to tomorrow's needed and more meaningful measures presents challenges, but the destination is becoming clearer. We interpreted the degree of excitement and positive energy among the Delphi participants as signals of willingness to move forward, embrace complexity, and address the measurement challenges outlined in this white paper.

IHI recognizes that, inevitably, health care systems will select sets of measures that align with their unique missions, populations, and strategic initiatives. We see WSM 2.0 as a "yes and" solution. Health care systems will augment WSM 2.0 with additional measures that are most applicable and appropriate for their organizations. The following principles emerged during the Delphi method as guidance for inclusion and/or modification of additional measures:

1. **Where possible, select measures that address multiple requirements by payers, accreditors, and other regulatory bodies.** Consider programs, initiatives, and contracts in which your health care system actively participates and those in which it may hope to participate in the future.
2. **Select measures with room for improvement.** Including measures with little room for growth in your organization adds limited value for a health care system and may no longer support ambitious goals.
3. **Where possible, select measures with data available monthly or quarterly, and electronically.** As noted, we struggled to identify measures for which data can be collected at least quarterly and that were still oriented toward outcomes rather than processes. Electronically available data reduces the data-gathering burden.
4. **Prioritize measures that are not overly complex.** The measures selected by the Delphi participants are relatively simple to understand, measure, and obtain in a timely way; the measures should have a clear line of sight to the desired aim and corresponding outcome.

5. **Select measures that capture multiple services and sites of care.** Given the complexity of services provided today and the need for parsimony, we should move beyond individual measures (e.g., hospital-acquired infections) to composite measures for broader areas like safety that can provide more detailed data drill-down to identify root cause problems in specific service areas.
6. **Select measures that are important in driving toward the Triple Aim**. By continuously keeping the Triple Aim at the forefront, the intention of WSM 2.0 is to present a balanced set of measures that address population health, individual experience of care, and per capita cost.

Conclusion

Increasingly, health care systems face numerous measurement demands from payers, accreditors, and others.[41] IHI believes that WSM 2.0 can serve as a core set of 15 measures that health care systems can adapt as they seek to better understand their overall system performance using measures that are meaningful. We hope that health care systems willing to test and learn together can use this small set of shared measures to create momentum for measure rationalization — reducing the noise from extraneous measurement and providing the signals for improvement initiatives that lead to better health for the population, improved experience of care for the individual, and decreased per capita cost.

Appendix A: Original Whole System Measures

In 2007, the Institute for Healthcare Improvement published the white paper, *Whole System Measures,*[42] for health care system leaders (see Table 2). The aim was to 1) develop a balanced set of measures based on the Institute of Medicine's six dimensions of quality (safe, effective, patient-centered, timely, efficient, and equitable); 2) specify the measures and possible aims at a high enough level to provide useful insight on a health care system's performance; and 3) limit the number of measures to enable senior leaders to understand the overall performance of their health care systems and set strategic direction to make improvements.

Table 2. Original Whole System Measures

Whole System Measure	IOM Dimension of Quality	Outpatient Care	Inpatient Care
1. Rate of Adverse Events	Safe	X	X
2. Incidence of Nonfatal Occupational Injuries and Illnesses	Safe	X	X
3. Hospital Standardized Mortality Ratio (HSMR)	Effective		X
4. Unadjusted Raw Mortality Percentage	Effective		X
5. Functional Health Outcomes Score	Effective	X	X
6. Hospital Readmission Percentage	Effective	X	X
7. Reliability of Core Measures	Effective	X	X
8. Patient Satisfaction with Care Score	Patient-Centered	X	X
9. Patient Experience Score	Patient-Centered	X	
10. Days to Third Next Available Appointment	Timely	X	
11. Hospital Days per Decedent During the Last Six Months of Life	Efficient	X	
12. Health Care Cost per Capita	Efficient	X	X
13. Equity (Stratification of Whole System Measures)	Equitable	X	X

Each of the original Whole System Measures had a corresponding "Toyota Specification," which set an ambitious goal that represented breakthrough performance in quality and appropriate cost (see Table 3). At the time, the specification provided either the top decile of performance or the best practice seen in another industry; if neither of those parameters demonstrated breakthrough performance, IHI assigned the specification.

Table 3. Toyota Specifications for the Original Whole System Measures

IOM Dimension of Quality	Whole System Measure	Toyota Specification
System Level		
Patient-Centered	Patient Experience Score: Response to the question in the How's Your Health database, "They give me exactly the help I want (and need) exactly when I want (and need) it."	72% of Patients Report, "They give me exactly the help I want (and need) exactly when I want (and need) it."[43]
Effective and Equitable	Functional Health Outcomes Score	5% of Adults Self-Rate Their Health Status as Fair or Poor[44] (*Self-rating will not differ by income)*[45]
Efficient	Health Care Cost per Capita Surrogate measure: Medicare Reimbursement per Enrollee per Year	$3,150 per Capita per Year[46] $5,026 per Enrollee per Year[47]
Component Level		
Safe	Rate of Adverse Events	5 Adverse Events per 1,000 Patient Days
Safe	Incidence of Nonfatal Occupational Injuries and Illnesses	0.2 Cases with Lost Work Days per 100 FTEs per Year[48]
Effective	Hospital Standardized Mortality Ratio (HSMR)	HSMR = 25 Points Below the National Average
Effective	Hospital Readmission Percentage	30-Day Hospital Readmission = 4.49%[49]
Effective	Reliability of Core Measures	10^{-2} Reliability Levels[50]
Patient-Centered	Patient Satisfaction with Care Score	60% of Patients Selected the Best Possible Score
Timely	Days to Third Next Available Appointment	Primary Care: Same-Day Access Specialty Care: Access Within 7 Days
Efficient	Hospital Days per Decedent During the Last Six Months of Life	7.24 Hospital Days per Decedent During the Last Six Months of Life[51]

Appendix B: WSM 2.0 Subdomain Measure Definitions

Population Health Domain: Subdomain Measures

- **Individual Health:** The extent to which individuals served by a health care system and in the communities served by the system can meet the World Health Organization's definition of health: "A state of complete physical, mental, and social wellbeing and not merely the absence of disease or infirmity."[52] This concept encompasses health status, functional status, and individual behaviors, as well as intermediate and non-intermediate health outcomes.
- **Healthy Behaviors:** The prevalence of "behavior patterns, actions and habits that relate to health maintenance… health restoration and to health improvement."[53] Healthy behaviors underlie both primary and secondary prevention efforts and are crucial to population health.
- **Community Wellbeing:** "How well a locality is functioning, how well the locality is governed, how well local services operate, and how safe, pleasant, and rewarding it feels to live in the locality."[54]
- **Health Equity:** "A state in which all people have the opportunity to attain their full health potential and no one is disadvantaged from achieving this potential because of their social position or other socially determined circumstance."[55,56,57]
- **Workforce Wellbeing:** Many definitions exist. For WSM 2.0, we focused on the definition advanced by Juniper: "The part of employees' wellbeing that they perceive to be determined primarily by work and can be influenced by workplace interventions."[58]

Experience of Care Domain: Subdomain Measures

- **Access:** Multiple definitions exist. We focused on the definition advanced by RAND: "The ease with which an individual can obtain needed medical care or services."[59]
- **Prevention:** The receipt of "interventions aimed at reducing the incidence of disease and disability or at slowing the progression and exacerbation of illnesses…"[60]
- **Safety:** Per the *Crossing the Quality Chasm* report, "…the avoidance of injuries to patients from the care that is intended to help them…"[61]
- **Appropriateness and Effectiveness:** Per the *Crossing the Quality Chasm* report, "…[based] on scientific knowledge, the provision of services to all who could benefit and refraining from providing services to those not likely to benefit…"[62]
- **Patient-Centeredness:** Per the *Crossing the Quality Chasm* report, "…the provision of care that is respectful and responsive to individual patient preferences, needs and values, and ensures that patient values guide all clinical decisions…"[63]

Per Capita Cost of Care Domain: Subdomain Measures

- **Affordability:** Per the National Quality Forum, the patient's out-of-pocket spending on health care compared to the annual household budget.[64]
- **Societal Footprint:** The extent to which health care utilization imposes an overall financial burden on society, displacing other investments.

Appendix C: WSM 2.0 Operational Definitions

Subdomain	Measure	Operational Definition	Steward	Data Sources	For Additional Information
Individual Health	General health	Percentage of survey respondents who report that their health is "Excellent," "Very Good," or "Good" in response to the question: "Would you say that in general your health is…?" Response scale: Excellent, Very Good, Good, Fair, Poor	Measure defined by BRFSS (Centers for Disease Control and Prevention)	BRFSS (states, select MSAs and counties); potentially health care system surveying	CDC, BRFSS. "The BRFSS Data User Guide." 2013. BRFSS. "2015 Behavioral Risk Factor Surveillance System Questionnaire."
Healthy Behaviors	Overweight/obesity	Percentage of patients aged 18 years and older with a documented body mass index (BMI) during the current encounter or during the previous six months AND when the BMI is outside of normal parameters, a follow-up plan is documented during the encounter or during the previous six months of the encounter Normal parameters: Ages 65 and older: BMI > or = 23 and < 30 Ages 18-64: BMI > or = 18.5 and < 25	Centers for Medicare & Medicaid Services (CMS)	Health care system data	National Quality Forum. "Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-up." (#3039)
Healthy Behaviors	Optimal lifestyle metric	Percentage of survey respondents who report that they adhere to three or four of the following behaviors: • Nonsmoking (current) • Physically active (30 minutes or more of physical activity on at least four days per week) • Healthy diet (five or more servings of fruits and vegetables per day) • Moderate alcohol use (no more than two drinks [men] or one drink [women] per day)	HealthPartners	Requires health care system surveying	Pronk N. An optimal lifestyle metric: Four simple behaviors that affect health, cost, and productivity. *ACSM Health & Fitness Journal.* 2012.
Community Wellbeing and Health Equity	Social support	Percentage of survey respondents who report that they "Always" or "Usually" have the support they need in response to the question: "How often do you get the social and emotional support you need (from any source)?" Response scale: Always, Usually, Sometimes, Rarely, Never	Measure defined by BRFSS (CDC)	BRFSS (states, select MSAs and counties); potentially health care system surveying	CDC, BRFSS. "The BRFSS Data User Guide." 2013. BRFSS. "2015 Behavioral Risk Factor Surveillance System Questionnaire."

Subdomain	Measure	Operational Definition	Steward	Data Sources	For Additional Information
Community Wellbeing and Health Equity	Disparities in infant mortality rate	For patients or the local population, the relative risk of infant mortality (death before age of one year) for non-Hispanic black women, non-Hispanic Puerto Rican women, and non-Hispanic American Indian or Alaskan Native women, compared to white women	Measured defined by CDC	Estimates (of rates) for states, counties, and select urban areas available by race from National Vital Statistics System (CDC)	For example: CDC. "Defining and measuring disparities, inequities, and inequalities in the Healthy People initiative." (2010 presentation)
Community Wellbeing and Health Equity	Disparities in high school graduation rate	The difference in the four-year adjusted cohort graduation rate between: 1) students who have disabilities and students who do not have disabilities; 2) students with limited English proficiency and students without limited English proficiency; 3) students from low-income families and students not from low-income families; 4) white, non-Hispanic students and black, non-Hispanic students; and 5) white, non-Hispanic students and Hispanic students The four-year cohort adjusted graduation rate is the number of students who graduate in four years with a regular high school diploma divided by the number of students from the adjusted cohort for the graduating class Low-income is often defined as those earning less than twice the federal poverty level	Measure defined in federal regulation	Four-year adjusted cohort graduation rate available by county; information by race, ethnicity, English-language proficiency, and income may be available from local education departments and/or schools	For example: National Center for Education Statistics. "Public High School Graduation Rates." May 2016.
Workforce Wellbeing (i.e., for the health care workforce)	Job satisfaction	Percentage of survey respondents who respond "Agree," on average, with the following indicators of job satisfaction, as graded on a Likert scale (1-5): • This hospital/facility is a good place to work • I am proud to work at this hospital/facility • Working in this hospital/facility is like being part of a large family • Morale in this unit is high • I like my job "Agree" includes responses for both 4 and 5 on the Likert scale (1-5): 1. Disagree strongly, 2. Disagree	Center for Healthcare Quality and Safety — University of Texas at Houston, and Memorial Hermann Healthcare System	Requires health care system surveying	Sexton JB, et al. The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. *BMC Health Services Research*. 2006.

Subdomain	Measure	Operational Definition	Steward	Data Sources	For Additional Information
		slightly, 3. Neutral, 4. Agree slightly, 5. Agree strongly			
Access	Timely ambulatory care	The percentage who affirmatively respond "Always" to CG-CAHPS survey items on getting timely appointments, care, and information from a physician's office. These three items include: 1. In the last six months, when you contacted this provider's office to get an appointment for care you needed right away, how often did you get an appointment as soon as you needed? 2. In the last six months, when you made an appointment for a check-up or routine care with this provider, how often did you get an appointment as soon as you needed? 3. In the last six months, when you contacted this provider's office during regular office hours, how often did you get an answer to your medical question that same day? Response scale: Never, Sometimes, Usually, Always	Agency for Healthcare Research and Quality (AHRQ)	CG-CAHPS reports and health care system data	For example: Robert Wood Johnson Foundation. "How to Report Results of the CAHPS Clinician & Group Survey." 2010.
Prevention	Childhood immunizations	Percentage of children 2 years of age who had four diphtheria, tetanus and acellular pertussis (DtaP); three polio (IPV); one measles, mumps and rubella (MMR); three H influenza type B (HiB); three hepatitis B (HepB); one chicken pox (VZV); four pneumococcal conjugate (PCV); one hepatitis A (HepA); two or three rotavirus (RV); and two influenza (flu) vaccines by their second birthday The measure calculates a rate for each vaccine and nine combination rates.	National Committee for Quality Assurance (NCQA)	Health care system data	NCQA. "Childhood immunization status." 2009.

Subdomain	Measure	Operational Definition	Steward	Data Sources	For Additional Information
Safety	Hospital-acquired conditions	The rates for acute care hospitals of the following conditions: • Foreign object retained after surgery • Air embolism • Blood incompatibility • Falls and trauma • Manifestations of poor glycemic control • Catheter-associated urinary tract infection • Vascular catheter-associated infection • Surgical site infection, mediastinitis, following coronary artery bypass graft (CABG) • Surgical site infection following certain orthopedic procedures • Surgical site infection following cardiac implantable electronic device • Deep vein thrombosis/pulmonary embolism following certain orthopedic procedures • Iatrogenic pneumothorax with venous catheterization	Measures defined by CMS hospital-acquired conditions program	CMS and health care system data	Various resources; see, for example: CMS. "Hospital-Acquired Conditions." 2015.
Safety	Serious reportable events (SREs)	The total number SREs during the reporting period. SREs include: • Surgical or invasive procedure events • Product or device events • Patient protection events • Care management events • Environmental events • Radiologic events • Potential criminal events	National Quality Forum (NQF)	Health care system records and data	NQF. "Serious Reportable Events in Healthcare—2011 Update: A Consensus Report." 2011.
Appropriateness and Effectiveness	Preventable hospitalizations	Rate of hospital admissions for certain acute illnesses or chronic conditions preventable through effective ambulatory care (e.g., diabetes, dehydration) PQI #90: Overall composite per 100,000 population, ages 18 years and older, including admissions for: diabetes with short-term complications, diabetes with long-term complications, uncontrolled diabetes	AHRQ	Hospital-level data reported by CMS	AHRQ. "Prevention Quality Overall Composite: Technical Specifications." 2016.

Subdomain	Measure	Operational Definition	Steward	Data Sources	For Additional Information
		without complications, diabetes with lower-extremity amputation, chronic obstructive pulmonary disease, asthma, hypertension, heart failure, dehydration, bacterial pneumonia, or urinary tract infection Admissions captured by diagnosis ICD diagnosis code AHRQ methodology for composite weighs each of the above conditions using a weighting number			
Patient-Centeredness	Patient-clinician communication satisfaction	Percentage of patients who respond "Always" to CG-CAHPS survey items on provider communication quality. These two items include: 1. In the last six months, how often did this provider explain things in a way that was easy to understand? 2. In the last six months, how often did this provider listen carefully to you? Response scale: Never, Sometimes, Usually, Always	AHRQ	CG-CAHPS reports and health care system data	For example: Robert Wood Johnson Foundation. "How to Report Results of the CAHPS Clinician and Group Survey." 2010.
Affordability	Unmet health care needs	Percentage of survey respondents who either did not receive care due to cost in the past 12 months or delayed care due to cost in the past 12 months *May create a composite percentage based on affirmative answers to the following questions:* During the past 12 months, was there any time when you needed any of the following, but didn't get it because you couldn't afford it? *[Note: Each of these is a separate survey question.]* • Prescription medicines • Mental health care or counseling • Dental care • Eyeglasses • To see a specialist • Follow-up care AND	Measure defined by CDC (National Health Interview Survey)	National Health Interview Survey (state-level estimates possible); potentially health care system surveying	CDC. National Center for Health Statistics. "About the National Health Interview Survey." 2016.

Subdomain	Measure	Operational Definition	Steward	Data Sources	For Additional Information
		During the past 12 months, did you delay filling a prescription to save money? *(For those adults age 19 and older who were prescribed medication in the past 12 months)* Response scale for all questions: Yes, No, Refused, Don't know			
Societal Footprint	Health care cost per capita: Total cost of care	Includes all costs associated with treating health plan members (attributed to a primary care provider), including professional, facility inpatient and outpatient, pharmacy, lab, radiology, ancillary, and behavioral health services Attributes members to the provider that provides the largest percentage of primary care office visits as determined by the specialty of the servicing physician during the performance period	HealthPartners	Health care system data, data from commercial payers (and Medicaid data, if applicable)	HealthPartners. "Total Cost of Care and Resource Use: Frequently Asked Questions (FAQ)." 2014.
Societal Footprint	Health care cost per capita: Medicare reimbursement per enrollee per year	Total spending per annum for Medicare beneficiaries ages 65-99 enrolled in both Medicare Parts A and B, typically excluding patients enrolled in risk-bearing health maintenance organizations	CMS	CMS data; Dartmouth Atlas (regional data by Hospital Referral Region)	For example: The Dartmouth Institute for Health Policy and Clinical Practice. "A New Series of Medicare Expenditure Measures by Hospital Referral Region: 2003-2008." 2011.

Appendix D: WSM 2.0 Delphi Method Participants

Name	Title	Organization
James Anderson	Advisor to President	Cincinnati Children's Hospital Medical Center
Suzanne Anderson	President and Executive Vice President	Virginia Mason Medical Center and Virginia Mason Health System
Katherine Lewis Apton, MPH	Program Officer	Leadership Consortium for Value and Science-Driven Health Care, National Academy of Medicine
Mary Barton, MD, MPP	Vice President, Performance Measurement	National Committee for Quality Assurance
Doug Bonacum, MBA	Vice President, Quality, Safety and Resource Management (*retired*)	Kaiser Permanente
Helen Burstin, MD, MPH	Chief Scientific Officer	National Quality Forum
Christine K. Cassel, MD	Planning Dean	Kaiser Permanente School of Medicine
Carolyn M. Clancy, MD	Deputy Under Secretary for Health, Organizational Excellence	Veterans Health Administration
Paul Cleary, PhD	Anna M. R. Lauder Professor of Public Health (Health Policy), and Dean	Yale School of Public Health
Marshall Chin, MD, MPH	Richard Parrillo Family Professor of Healthcare Ethics in the Department of Medicine Associate Chief and Director of Research, Section of General Internal Medicine Director, Chicago Center for Diabetes Translation Research Director, Robert Wood Johnson Foundation Finding Answers: Solving Disparities Through Payment and Delivery System Reform Program Office, The University of Chicago	University of Chicago School of Medicine
James Conway, MS	Board Member Board Member	Lahey Health Winchester Hospital
Patrick Conway, MD	Acting Principal Deputy Administrator Deputy Administrator for Innovation and Quality Chief Medical Officer	Centers for Medicare & Medicaid Services

Name	Title	Organization
Janet Corrigan, PhD, MBA	Chief Program Officer, Patient Care	Gordon and Betty Moore Foundation
Helen Darling, MA	Interim President and CEO	National Quality Forum
Victor J. Dzau, MD	President	National Academy of Medicine
Susan Edgman-Levitan, PA	Executive Director	John D. Stoeckle Center for Primary Care Innovation, Massachusetts General Hospital
Rick Foster, MD	Executive Director	Catalyst for Health, South Carolina Hospital Association
Susan Frampton, PhD	President	Planetree
John A. Gillean, MD, MHA	Executive Vice President and Chief Clinical Officer	CHRISTUS Health
Judith Hibbard, DrPH	Research Professor, Health Policy Research Group	University of Oregon
George Isham, MD, MS	Senior Advisor Senior Fellow	HealthPartners HealthPartners Institute
Brent James, MD, MStat	Chief Quality Officer	Intermountain Healthcare Institute for Healthcare Delivery Research, and Intermountain Healthcare
Karen Joynt, MD, MPH	Associate Physician, Cardiovascular Division Instructor, Department of Health Policy and Management	Brigham and Women's Hospital Harvard T.H. Chan School of Public Health
Thomas Kottke, MD, MSPH	Medical Director for Well-being	HealthPartners
Susan Knudson	Senior Vice President, Health Informatics	HealthPartners
Uma Kotagal, MBBS, MSc	Senior Executive Leader, Population and Community Health Senior Fellow	Cincinnati Children's Hospital Medical Center Cincinnati Children's Hospital Medical Center
Kevin Larsen, MD, FACP		Center for Medicare and Medicaid Innovation (CMMI) Office of the Administrator – CMS
Michael Leonard, MD	Principal	Safe and Reliable Healthcare
Sally Okun, RN, MMHS	Vice President for Advocacy, Policy and Patient Safety	PatientsLikeMe

Name	Title	Organization
David Pryor, MD	Executive Vice President and Chief Clinical Officer	Ascension Health
Michelle Schreiber, MD	Senior Vice President and Chief Quality Officer	Henry Ford Health System
Thomas Sequist, MD	Chief Quality and Safety Officer Associate Professor of Medicine and Health Policy	Partners HealthCare Harvard Medical School
Matthew Stiefel, MPA, MS	Senior Director, Center for Population Health Fellow	Care Management Institute, Kaiser Permanente Institute for Healthcare Improvement
MaryAnn Stump, RN, MBA, CPHQ	Care Model Innovation Consultant	
Paul Wallace, MD	Senior Scholar in Residence	AcademyHealth
Beth Waterman, RN, MBA	Chief Improvement Officer	HealthPartners
Gary R. Yates, MD	Partner, Strategic Consulting	Press Ganey Associates
Additional participants: Vice president of a major hospital association; senior vice president of a large not-for-profit health care system; and director of a state-level health care improvement organization		

References

[1] Cassel CK, Conway PH, Delbanco SF, Jha AK, Saunders RS, Lee TH. Getting more performance from performance measurement. *New England Journal of Medicine*. 2014 Dec 10;371:2145-2147.

[2] Martin LA, Nelson EC, Lloyd RC, Nolan TW. *Whole System Measures*. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement; 2007. www.ihi.org/resources/Pages/IHIWhitePapers/WholeSystemMeasuresWhitePaper.aspx

[3] Institute of Medicine. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press; 2001.

[4] Stiefel M, Nolan K. *A Guide to Measuring the Triple Aim: Population Health, Experience of Care, and Per Capita Cost*. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement; 2012. www.ihi.org/resources/Pages/IHIWhitePapers/AGuidetoMeasuringTripleAim.aspx

[5] From 2007 to 2012, IHI worked with more than 100 teams to test the Triple Aim framework in the US and internationally.

[6] These measures, while helpful, are too great in number and too specific to fit the intent of Whole System Measures — a set of 15 measures that aim to help leaders better understand their organization's current state (and desired state) across three domains: health, experience of care, and cost per capita.

[7] Institute of Medicine. *Vital Signs: Core Metrics for Health and Health Care Progress*. Washington, DC: The National Academies Press; 2015.

[8] Other efforts include the National Quality Forum's (NQF) Aligning Forces for Quality measure alignment work and tool development; and the NQF's Measure Applications Partnership, a public-private partnership convened to provide input to the US Department of Health and Human Services on selecting performance measures for its programs. [See: Bipartisan Policy Center. *Transitioning from Volume to Value: Consolidation and Alignment of Quality Measures*. April 2015.] In addition, a number of regional efforts at measurement alignment have taken off in recent years, such as the efforts of Minnesota Community Measurement, a statewide organization that both develops and harmonizes measures across Minnesota payers and providers. [See, for example: Chase J. Will Medicare Core Measures make MNCM obsolete? *MNSM News*. 2016.]

[9] Institute of Medicine. *Vital Signs: Core Metrics for Health and Health Care Progress*. Washington, DC: The National Academies Press; 2015.

[10] IHI's Scientific Advisory Group is an independent group of advisors, chartered by IHI's Board of Directors, which provides expert review and guidance to improve the rigor and credibility of IHI's results-oriented programs and activities. www.ihi.org/about/Pages/ScientificAdvisoryGroup.aspx

[11] Institute of Medicine. *Vital Signs: Core Metrics for Health and Health Care Progress*. Washington, DC: The National Academies Press; 2015.

[12] US Department of Health and Human Services (HHS). "HHS Measurement Policy Council Core Measure Sets (as of June 2014)." www.ahrq.gov/workingforquality/mpcmeasures.pdf

[13] Stiefel M, Nolan K. *A Guide to Measuring the Triple Aim: Population Health, Experience of Care, and Per Capita Cost*. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement; 2012. www.ihi.org/resources/Pages/IHIWhitePapers/AGuidetoMeasuringTripleAim.aspx

[14] The Delphi method was originally designed by the RAND Corporation as a means to make predictions in the face of a high degree of uncertainty (such as future innovations in fields like defense or consumer goods). The Delphi method in its purest form takes place without any face-to-face communication between participants, to avoid overweighting the influence of "strong personalities." Individuals, over the course of several rounds, have the opportunity to review the responses of other anonymous participants from previous rounds, as well as the arguments supporting their responses.

[15] See, for example: Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting health care quality indicators: A systematic review. *PLoS One*. 2011;6(6):e20476.

[16] Each round had different participants; some only participated in a single round. All who participated in a survey or conference call were considered participants in the process.

[17] In this subdomain, participants voiced concerns regarding the quality of existing measures. In the final set of measures, the measure selected for workforce wellbeing had the highest rating among those included.

[18] The *Vital Signs* report includes "best current measures" for a set of recommended measurement concepts.

[19] The NQF endorsed the measure "Preventive Care and Screening: Body Mass Index (BMI) and Follow-Up."

[20] For example, Medicare Physician Quality Reporting System.

[21] NQF has endorsed the CG-CAHPS survey, which includes questions regarding timely ambulatory care.

[22] For example, Medicare Physician Quality Reporting System.

[23] NQF endorsed "Childhood Immunization Status."

[24] For example, Medicare Physician Quality Reporting System.

[25] For example, National Healthcare Safety Network CAUTI, CLASBI, and CDI outcome measures.

[26] Hospital-Acquired Conditions Reduction program.

[27] The individual prevention quality indicators (e.g., diabetes with short-term complication admission rate, heart failure admission rate) are endorsed, with this exception: hypertension admission.

[28] For example, Medicare Value-Based Modifier Program includes a composite of preventable hospitalizations for chronic conditions.

[29] NQF has endorsed the CG-CAHPS survey, which includes questions regarding patient-clinician communication.

[30] For example, Medicare Physician Quality Reporting System.

[31] NQF has endorsed the Total Cost of Care measure.

[32] CMS makes available information on expenditures per enrollee (fee-for-service) for states and counties, via the Office of Enterprise Data and Analytics.

[33] The original Whole System Measures, published in 2007, also included this measure.

[34] This approach is consistent with Berwick's view: "Purveyors of measurement… should commit to reducing (by 50% in 3 years and by 75% in 6 years) the volume and total cost of measurement currently being used and enforced in health care. The aim should be to measure only what matters, and mainly for learning." [See: Berwick DM. Era 3 for medicine and health care. *Journal of the American Medical Association*. 2016 Apr 5;315(13):1329-1330.]

[35] See, for example: Koy V, Yunibhand J, Yupin A, Fisher ML. Relationship between nursing care quality, nurse staffing, nurse job satisfaction, nurse practice environment and burnout: Literature review. *International Journal of Research in Medical Sciences*. 2015;3(8):1825-1831.

[36] Sexton JB, Helmreich RL, Neilands TB, Rowan K, Vella K, Boyden J, Roberts PR, Homas EJ. The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. *BMC Health Services Research*. 2006 Apr 3;6:44.

[37] Berwick DM. Era 3 for medicine and health care. *Journal of the American Medical Association*. 2016 Apr 5;315(13):1329-1330.

[38] Berwick DM. Era 3 for medicine and health care. *Journal of the American Medical Association*. 2016 Apr 5;315(13):1329-1330.

[39] Chen A, Oster E, Williams H. *Why Is Infant Mortality Higher in the US Than in Europe?* Working Paper. Cambridge, MA: National Bureau of Economic Research; 2015.

[40] See, for example: *Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement*. Washington, DC: National Quality Forum; April 2015. www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=79434

[41] Conway PH. The Core Quality Measures Collaborative: A rationale and framework for public-private quality measure alignment. *Health Affairs Blog*. June 23, 2015. http://healthaffairs.org/blog/2015/06/23/the-core-quality-measures-collaborative-a-rationale-and-framework-for-public-private-quality-measure-alignment/

[42] Martin LA, Nelson EC, Lloyd RC, Nolan TW. *Whole System Measures*. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement; 2007. www.ihi.org/resources/Pages/IHIWhitePapers/WholeSystemMeasuresWhitePaper.aspx

[43] Email communication with John Wasson, MD. March 29, 2007.

[44] Zahran HS, Kobau R, Moraiarty D, Zack M, Holt J, Donehoo R. Health-related quality of life surveillance—United States, 1993-2002. *MMWR: Surveillance Summaries*. 2005 Oct 28;54(SS-4):1-35.

[45] Due to the lack of nationally available data using the Functional Health Survey-6+, IHI used self-reported health status data from the Centers for Disease Control and Prevention *Health-Related Quality of Life Surveillance* report. There is more information about the Functional Health Survey-6+ in the 2007 *Whole System Measures* IHI White Paper.

[46] Organisation for Economic Co-operation and Development. "OECD Health Statistics 2016." www.oecd.org/els/health-systems/health-data.htm

[47] The Dartmouth Atlas of Health Care. www.dartmouthatlas.org/index.shtm

[48] US Bureau of Labor Statistics. Table 1. Incidence rates of nonfatal occupational injuries and illnesses by industry and case types, 2005. Washington, DC: US Department of Labor; 2005.

[49] Belk K. Premier Perspective™ Database Readmission Data (Calendar Year 2006). Premier, Inc.; 2007.

[50] Nolan T, Resar R, Haraden C, Griffin F. *Improving the Reliability of Health Care*. IHI Innovation Series white paper. Boston, MA: Institute for Healthcare Improvement; 2004. www.ihi.org/IHI/Results/WhitePapers/ImprovingtheReliabilityofHealthCare.htm

[51] The Dartmouth Atlas of Health Care. www.dartmouthatlas.org/index.shtm

[52] Preamble to the Constitution of the World Health Organization. Geneva, Switzerland: World Health Organization; 1946.

[53] Gochman D. "Health Behavior: Plural Perspectives." In: Bahar S (editor). *Health Behavior: Emerging Research Perspectives*. Springer; 1988.

[54] Lee SJ, Kim Y. "Searching for the Meaning of Community Well-being." In: Lee SJ, Kim Y, Phillips R (editors). *Community Well-being and Community Development: Concepts and Applications*. Springer; 2015. Also important to the notion of community wellbeing is the concept of "community" — "a number of people who have some degree of common identity or concerns often related to a particular locality or conditions… a community… is… a number of people who have repeated dealing with each other." Numerous definitions of both "community" and "community wellbeing" exist. Varied definitions of community wellbeing can touch on factors spanning psychosocial, cultural, political, and economic dimensions. A significant literature exists that addresses the varied meanings of community, in multiple contexts. [See, for example: Neal ZP. "Community." Oxford Bibliographies. www.oxfordbibliographies.com/view/document/obo-9780199756384/obo-9780199756384-0080.xml]

[55] See, for example: Braveman PA. Monitoring equity in health and health care: A conceptual framework. *Journal of Health, Population, and Nutrition*. 2003;21(3):181-192.

[56] See, for example: Kawachi I, Subramanian SV, Almeida-Filho. A glossary for health inequalities. *Journal of Epidemiology and Community Health*. 2002;56:647-652.

[57] See, for example: Whitehead M. The concepts and principles of equity and health. *International Journal of Health Services*. 1992;22(3):429-445.

[58] Juniper B, White N, Bellamy P. Assessing employee well-being: Is there another way? *International Journal of Workplace Health Management*. 2009;2(3):220-230. Numerous definitions of "workforce wellbeing" exist. For example, one report on employee wellbeing published by Mercer defined it as "the active state of pursuing health and life skills with the aim of achieving physical and emotional health and financial security." This definition notably focuses more on a set of activities than on health outcomes. [See: *Employee Well-Being: A New Way to Define Organizational Success*. Mercer, LLC; 2015.] The concept of "wellbeing" IHI advances for this subdomain is one articulated by the CDC: "…at minimum, well-being includes the presence of positive emotions (e.g., contentment, happiness), the absence of negative emotions (e.g., depression, anxiety), satisfaction with life, fulfilling and positive functioning." [See: Centers for Disease Control. "Well-Being Concepts." www.cdc.gov/hrqol/wellbeing.htm]

[59] RAND Corporation. "Health Care Access." www.rand.org/topics/health-care-access.html

[60] Goetzel RZ. Do prevention or treatment services save money? The wrong debate. *Health Affairs*. 2009;28(1):37-41.

[61] See, for example: Institute of Medicine. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press; 2001.

[62] Institute of Medicine. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press; 2001.

[63] See, for example: Institute of Medicine. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press; 2001.

[64] See, for example: *Measuring Affordability from the Patient's Perspective*. Washington, DC: National Quality Forum; September 2014. www.qualityforum.org/Publications/2014/09/Measuring_Affordability_from_the_Patient_s_Perspective.aspx

Institute for Healthcare Improvement
20 University Road
Cambridge, MA 02138 USA