Boston Children's Hospital
Until every child is well℠

September 8, 2014

Ms. Lois Johnson
General Counsel
Health Policy Commission
Two Boylston Street, 6th Floor
Boston, MA 02116

Dear Ms. Johnson,

Attached, please find the testimony of Boston Children's Hospital, signed under pains and penalties of perjury, in response to questions provided by the Health Policy Commission and the Office of the Attorney General.

As the President and Chief Executive Officer of Boston Children's Hospital, I am legally authorized and empowered to represent the organization for the purposes of this testimony.

If you have any questions, please contact Joshua Greenberg, Vice President of Government Relations, at (617) 919-3055.

Sincerely,

Sandra L. Fenwick
President and CEO
Boston Children's Hospital

Exhibit A: Notice of Public Hearing

Pursuant to M.G.L. c. 6D, § 8, the Health Policy Commission, in collaboration with the Office of the Attorney General and the Center for Health Information and Analysis, will hold a public hearing on health care cost trends. The hearing will examine health care provider, provider organization and private and public health care payer costs, prices and cost trends, with particular attention to factors that contribute to cost growth within the Commonwealth's health care system.

Scheduled hearing dates and location:

Monday, October 6, 2014, 9:00 AM
Tuesday, October 7, 2014, 9:00 AM
Suffolk University Law School
First Floor Function Room
120 Tremont Street, Boston, MA 02108

Time-permitting, the HPC will accept oral testimony from members of the public beginning at 4:00 PM on Tuesday, October 7. Any person who wishes to testify may sign up to offer brief comments on a first-come, first-served basis when the hearing commences on October 6.

Members of the public may also submit written testimony. Written comments will be accepted until October 16, 2014 and should be submitted electronically to HPC-Testimony@state.ma.us, or, if comments cannot be submitted electronically, sent by mail, post-marked no later than October 16, 2014, to the Health Policy Commission, Two Boylston Street, 6th floor, Boston, MA 02116, attention Lois H. Johnson.

Please note that all written and oral testimony provided by witnesses or the public may be posted on the HPC's website.

The HPC encourages all interested parties to attend the hearing. Visit the Suffolk Law School website for driving and public transportation directions. Suffolk Law School is located diagonally across from the Park Street MBTA station (Red and Green lines). Parking is not available at the law school but information about nearby garages is listed at the link provided.

If you require disability-related accommodations for this hearing, please contact Kelly Mercer at (617) 979-1420 or by email Kelly.A.Mercer@state.ma.us a minimum of two weeks prior to the hearing so that we can accommodate your request.

For more information, including details about the agenda, expert and market participant panelists, testimony and presentations, please check the Annual Cost Trends Hearing section of the HPC's website. Materials will be posted regularly as the hearing dates approach.

Exhibit B: Instructions and HPC Questions for Written Testimony

Instructions:

On or before the close of business on September 8, 2014, electronically submit, **using the provided template,** written testimony signed under the pains and penalties of perjury to: HPC-Testimony@state.ma.us. **You may expect to receive the template for submission of responses as an attachment received from HPC-Testimony@state.ma.us.** If you have any difficulty with the template or did not receive it, please contact Kelly Mercer at Kelly.A.Mercer@state.ma.us or (617) 979-1420.

Please begin each response with a brief summary not to exceed 120 words. The provided template has character limits for responses to each question, but if necessary, you may include additional supporting testimony or documentation in an Appendix. Please submit any data tables included in your response in **Microsoft Excel or Access format**.

The testimony must contain a statement that the signatory is legally authorized and empowered to represent the named organization for the purposes of this testimony, and that the testimony is signed under the pains and penalties of perjury. An electronic signature will be sufficient for this submission.

If you have any other questions regarding this process or regarding the following questions, please contact: Lois Johnson at Lois.Johnson@state.ma.us or (617) 979-1405.

[Remainder of page intentionally left blank]

Questions:

We encourage you to refer to and build upon your organization's 2013 Pre-Filed Testimony responses, if applicable. Additionally, if there is a point that is relevant to more than one question (including Exhibit C questions from the Attorney General), please state it only once and make an internal reference.

1. Chapter 224 of the Acts of 2012 (c. 224) sets a health care cost growth benchmark for the Commonwealth based on the long-term growth in the state's economy. The benchmark for growth between CY2012-CY2013 and CY2013-CY2014 is 3.6%.
   SUMMARY: Boston Children's Hospital has met the state's cost growth target in each of the last two years, and has met below-inflation cost targets for several years prior to the passage of c. 224. We have reduced our costs by more than $200M since 2009 on a unit cost basis. As a unique pediatric subspecialty organization with a national and international reputation for delivering the highest quality care, 30% of our revenue is derived from non-Massachusetts patients. Combined with our efforts to support local care delivery in community hospitals and other lower-cost settings, we are experiencing overall increases in acuity, length of stay and concommitant resource needs of our patients.
   a. What trends has your organization experienced in revenue, utilization, and operating expenses from CY 2010-CY2013 and year-to-date 2014? Please comment on the factors driving these trends.
      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."
   b. What actions has your organization undertaken since January 1, 2013 to ensure the Commonwealth will meet the benchmark, and what have been the results of these actions?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

   c. What actions does your organization plan to undertake between now and October 1, 2015 (including but not limited to innovative care delivery approaches, use of technology and error reduction) to ensure the Commonwealth will meet the benchmark?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

d. What systematic or policy changes would encourage or enable your organization to operate more efficiently without reducing quality?

   This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

2. C. 224 requires health plans to reduce the use of fee-for-service payment mechanisms to the maximum extent feasible in order to promote high-quality, efficient care delivery. SUMMARY: Boston Children's Hospital has participated in the Blue Cross Blue Shield Alternative Quality Contract since 2012 as the only specialty pediatric hospital/providers with such an arrangement. In calendar year 2013, for our attributed population in the AQC, we managed cost growth to a trend more efficient than the network comparison. We are committed to providing high-value care to all patients regardless of their insurance type and our programs related to quality improvement and cost-effectiveness are generally applied on an all-payor basis. There is extensive benchmarking data available through the Pediatric Health Information System database overseen by the Children's Hospital Association. We are typically a top tier performer in terms of quality, and have below median costs and receive below median payments relative to other children's hospitals on a case mix adjusted basis.
   a. How have alternative payment methods (APMs) (payment methods used by a payer to reimburse health care providers that are not solely based on the fee-for-service basis, e.g., global budget, limited budget, bundled payment, and other non-fee-for-service models, but not including pay-for-performance incentives accompanying fee-for-service payments) affected your organization's overall quality performance, care delivery practices, referral patterns, and operations?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

   b. Attach and discuss any analyses your organization has conducted on the implementation of APMs and resulting effects on your non-clinical operations (e.g., administrative expenses, resources and burdens).

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

c. Please include the results of any analyses your organization has conducted on this issue, including both for your patients paid for under APMs and for your overall patient population.

   This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

3. Please comment on the adequacy or insufficiency of health status risk adjustment measures used in establishing risk contracts and other APM contracts with payers.
   SUMMARY: Boston Children's Hospital has evaluated several risk adjustment models with Milliman (an actuarial firm). We have found that, in general, traditional systems are not as accurate in projecting pediatric utilization as systems optimized for children. Further, risk adjustment systems do not generally account for socio-economic status, which can be a very significant problem for an organization that treats many low income children enrolled in the Medicaid program. (See Attachment 3.) It would be helpful if the payors and/or the state could produce a comparative schedule of risk adjustment systems utilized throughout Massachusetts to better understand the relative risk of patients utilizing our system across payors.
   a. In your organization's experience, do health status risk adjustment measures sufficiently account for changes in patient population acuity, including in particular sub-populations (e.g., pediatric) or those with behavioral health conditions?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

   b. How do the health status risk adjustment measures used by different payers compare?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

   c. How does the interaction between risk adjustment measures and other risk contract elements (e.g., risk share, availability of quality or performance-based incentives) affect your organization?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

4. A theme heard repeatedly at the 2013 Annual Cost Trends Hearing was the need for more timely, reliable, and actionable data and information to facilitate high-value care and performance under APMs. What types of data are or would be most valuable to your organization in this regard? In your response, please address (i) real time data to manage patient care and (ii) historic data or population-level data that would be helpful for population health management and/or financial modeling.
SUMMARY: In order to design accountable care arrangements and potential financial models, we believe that providers will need access to patient level claims data that includes information on both utilization and spending. As APM arrangements are evolving across payors, we have identified opportunities to improve quality and cost-effectiveness that relies on payor data most often provided under risk arrangements. With respect to real time data, we are especially interested in data that helps us manage patients' care, including chronic diseases. This would include medical, surgical, pharmacy, and behavioral health data. As we work to marry clinical and financial data, it has become clear that providers and others would be well-served by standardized data extract formats across payors and that it is costly to tailor systems to the format and data of specific payors.

   ANSWER: This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

5. C. 224 requires health plans to attribute all members to a primary care provider, to the maximum extent feasible.
SUMMARY: We believe that there is no single, "best" attribution method. Rather, this will depend on the nature of the care delivery and financial model under consideration. In general, we are comfortable with an approach that seeks to empanel patients to primary care providers for the development of models relying on medical home infrastructure. However, 90% of the patients seen at Boston Children's have pediatricians that are not members of our primary care network; these patients have substantially higher risk scores, than other local AMCs. A different attribution approach is necessary if you want to compare hospitals or subspecilists on a risk-adjusted basis. As noted in your reports, this is where much of the spending occurs.
   a. Which attribution methodologies most accurately account for patients you care for?

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

   b. What suggestions does your organization have for how best to formulate and implement attribution methodologies, especially those used for payment?

This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

6. Please discuss the level of effort required to report required quality measures to public and private payers, the extent to which quality measures vary across payers, and the resulting impact(s) on your organization.
   SUMMARY: Our investment in quality and safety is substantial and we have made significant contributions to the field of pediatric quality measurement and improvement. We have worked closely with our clinical and quality leaders to present an increasingly well-rounded view of quality performance and improvement across inpatient and ambulatory care. As a specialty organization, we often work to educate payors about pediatric-specific quality and safety measures and available benchmarks. While we have had success in harmonizing our quality portfolio across payors, we have not made as much progress in establishing a common set of pediatric specialty measures across local pediatric providers (i.e. the other providers are not being held to the same standards for the care they deliver in their contracts).

   ANSWER: This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

7. An issue addressed both at the 2013 Annual Cost Trends Hearing and in the Commission's July 2014 Cost Trends Report Supplement is the Commonwealth's higher than average utilization of inpatient care and its reliance on academic medical centers.
   SUMMARY: It is our belief that care of pediatric patients is best delivered in a family-centered way by those with pediatric expertise, whether they are part of the BCH enterprise or not. Currently, we are working in a variety of ways to support affiliated community hospitals in their care of pediatric patients. Note, however, that for children, much of the subspecialty capacity only exists in academic medical centers. While others under risk arrangements have worked to manage "leakage," we have taken a patient and family centered position that relies on patients, working with their pediatricians, to make the right decisions for their care. We do have concerns that patients who fall under the risk arrangements of non-pediatric systems face limited access to pediatric expertise, whether at BCH or elsewhere.

   a. Please attach any analyses you have conducted on inpatient utilization trends and the flow of your patients to AMCs or other higher cost care settings.

      This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the

Attorney General Questionnaire."

b. Please describe your organization's efforts to address these trends, including, in particular, actions your organization is taking to ensure that patients receive care in lower-cost community settings, to the extent clinically feasible, and the results of these efforts.

This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

8. The Commission found in its July 2014 Cost Trends Report Supplement that the use of post-acute care is higher in Massachusetts than elsewhere in the nation and that the use of post-acute care varies substantially depending upon the discharging hospital. SUMMARY: Our strong preference is that children return home after inpatient admission. At times, this is facilitated through the use of home-based services and supports (visiting nurses, care managers, etc.) This is not always possible for children in state custody or requiring habilitative care. It is relatively common that hospitalized children with behavioral health needs step down to residential treatment programs. In this circumstance, we frequently must negotiate the transition with relevant payors (insurers and/or school systems).

a. Please describe and attach any analyses your organization has conducted regarding levels of and variation in the utilization and site of post-acute care, as well as your efforts to ensure that patients are discharged to the most clinically appropriate, high-value setting.

This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

b. How does your organization ensure optimal use of post-acute care?

This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

9. C. 224 requires providers to provide patients and prospective patients with requested price for admissions, procedures and services. Please describe your organization's progress in this area, including available data regarding the number of individuals that seek this information (using the template below) and identify the top ten admissions,

procedures and services about which individuals have requested price information. Additionally, please discuss how patients use this information, any analyses you have conducted to assess the accuracy of estimates provided, and/or any qualitative observations of the value of this increased price transparency for patients.

SUMMARY: Beginning in January 2014, Boston Children's Hospital formalized a process for providing estimates to patients and/or families via a dedicated phone line, or via a Website. Such estimated include Hospital charges and professional fees (if known).

Health Care Service Price Inquiries				
Year		**Number of Inquiries via Website**	**Number of Inquiries via Telephone/In Person**	**Average (approximate) Response Time to Inquiries***
CY2014	**Q1**	4	19	24h
	Q2	3	27	24h
	Q3	5	20	24
	TOTAL:	12	66	

** Please indicate the unit of time reported.*

ANSWER: This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

10. Please describe the manner and extent to which tiered and limited network products affect your organization, including but not limited to any effects on contracting and/or referral practices, and attach any analyses your organization has conducted on this issue. Describe any actions your organization taken (e.g., pricing changes) in response to tier placement and any impacts on volume you have experienced based on tier placement.

SUMMARY: Most of the spending and utilization by children is concentrated in a small group of children with relatively complex conditions or acute needs. These children, in turn, tend to concentrate in a few pediatric centers locally. The National Association of Insurance Commissioners is in the process of updating its model statute governing network adequacy requirements. There was extensive testimony submitted during this process, with many of the commentators remarking on the special protections required for children. We have deep skepticism that payors are utilizing consistent approaches to the determination of which tier a given provider or physician is placed into, and we receive frequent complaints from physicians that they are not being allowed to send patients to Boston Children's Hospital when they believe it is in the best interest of the patient to do so.

ANSWER: This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

11. The Commission has identified that spending for patients with comorbid behavioral health and chronic medical conditions is 2-2.5 times as high as spending for patients with a chronic medical condition but no behavioral health condition. As reported in the July 2014 Cost Trends Report Supplement, higher spending for patients with behavioral health conditions is concentrated in emergency departments and inpatient care.
SUMMARY: Mental and behavioral health continue to be among the most arduous challenges facing the State, providers, and patients, and the systems to care for these patients are, at present, simply not working. For over 30 years, Boston Children's Hospital has been a leader in identifying and treating patients with mental and behavioral health conditions through a myriad of approaches, including clinical interventions, community programs run both in and out of the hospital, and partnerships with the Boston Public Schools, the Children's Mental Health Campaign (of which BCH was a founder), and the State. However, we continue to face systemic challenges as well as challenges with payors with respect to adequately identifying, classifying, and treating patients with clear, often highly complicated needs.
    a. Please describe ways that your organization is collaborating with other providers to integrate physical and behavioral health care services and provide care across a continuum to these high-cost, high-risk patients.

       This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

    b. Please discuss ways that your organization is addressing the needs of individuals to avoid unnecessary utilization of emergency room departments and psychiatric inpatient care.

       This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

    c. Please discuss successes and challenges your organization has experienced in providing care for these patients, including how to overcome any barriers to integration of services.

       This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the

Attorney General Questionnaire."

d. There has been increased statewide interest in data reporting across all services, inclusive of behavioral health. Please describe your organization's willingness and ability to report discharge data.

   This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

12. Describe your organization's efforts and experience with implementation of patient-centered medical home (PCMH) model.
    SUMMARY: With respect to the hospital-owned and operated practices , over the past two years Boston Children's Hospital has invested in patient-centered medical home transformation using the Change Concepts from the Safety Net Medical Home Project. The practices seek to deliver high-quality, proactive preventive and acute health care to all of our primary care patients through the use of care teams, population management techniques, evidence-based care for chronic illness, and approaches which empower our patients. The Pediatric Physicians Organization at Children's (PPOC) has implemented a PCMH program across its network of 80 practices to support integration of PCMH standards and principles of care across practices.

    a. What percentage of your organization's primary care providers (PCPs) or other providers are in practices that are recognized or accredited as PCMHs by one or more national organizations?

       This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

    b. What percentage of your organization's primary care patients receives care from those PCPs or other providers?

       This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

    c. Please discuss the results of any analyses your organization has conducted on the impact of PCMH recognition or accreditation, including on outcomes, quality, and costs of care.

This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

---

13. After reviewing the Commission's 2013 Cost Trends Report and the July 2014 Supplement to that report, please provide any commentary on the findings presented in light of your organization's experiences.
    SUMMARY: Given the high number of residents in the Commonwealth who are covered by public coverage programs, we continue to be both surprised and concerned at what seems to be a lack of focus and concern on the part of CHIA with respect to this particularly vulnerable population, especially with respect to the Medicaid program.
    ANSWER: This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

Exhibit C: Instructions and AGO Questions for Written Testimony

Please note that these pre-filed testimony questions are for hospitals. To the extent that a hospital submitting pre-filed testimony responses is affiliated with a provider system also submitting pre-filed testimony responses, each entity may reference the other's response as appropriate.

1. Please submit a summary table showing for each year 2010 to 2013 your total revenue under pay for performance arrangements, risk contracts, and other fee for service arrangements according to the format and parameters provided and attached as AGO Hospital Exhibit 1 with all applicable fields completed. Please attempt to provide complete answers. To the extent you are unable to provide complete answers for any category of revenue, please explain the reasons why.

   Completed in Attachment AGO Hospital Exhibit 1

   This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

2. For each year 2010 to present, please submit a summary table showing for each line of business (commercial, government, other, total) your inpatient and outpatient revenue and margin for each major service category according to the format and parameters provided and attached as AGO Hospital Exhibit 2 with all applicable fields completed. Please submit separate sheets for pediatric and adult populations, if necessary. If you are unable to provide complete answers, please provide the greatest level of detail possible and explain why your answers are not complete.

   Completed in Attachment AGO Hospital Exhibit 2

   This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

3. Please explain and submit supporting documents that show how you quantify, analyze and project your ability to manage risk under your risk contracts, including the per member per month costs associated with bearing risk (e.g., costs for human resources, reserves, stop-loss coverage), solvency standards, and projections and plans for deficit scenarios. Include in your response any analysis of whether you consider the risk you bear to be significant.

   This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

4. Please explain and submit supporting documents that show how you analyze and track the volume of inpatient and outpatient referrals to your hospital and the associated revenue from those referrals by particular physicians or provider groups. Please include a description and examples of how your organization uses this information.

This is a complex question and requires a more thorough response than the space allotted. Please refer to the attached document entitled "Boston Children's Hospital Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire."

Boston Children's Hospital

Responses to 2014 Health Policy Commission and Office of the Attorney General Questionnaire

September 8, 2014

Table of Contents

Health Policy Commission Written Questions

1. Chapter 224 of the Acts of 2012 (c. 224) sets a health care cost growth benchmark for the Commonwealth based on the long-term growth in the state's economy. The benchmark for growth between CY2012-CY2013 and CY2013-CY2014 is 3.6%.

a. What trends has your organization experienced in revenue, utilization, and operating expenses from CY 2010-CY2013 and year-to-date 2014? Please comment on the factors driving these trends.

Boston Children's Hospital is a unique organization in Massachusetts due to our sole focus on the care of pediatric patients. We are consistently recognized as one of the top children's hospitals in the country, and draw a significant percentage of our patients (and our revenue) from regional, national, and international referrals. These patients tend to be more complex, utilize more resources when hospitalized, and have longer lengths of stay than typical patients, and are therefore by definition "more costly" than the average pediatric patient.

In the local market, we have aggressively pursued actions to reduce the total cost of care, frequently through the use of lower cost settings. In some cases, these settings are other community hospitals (many of which our physicians staff) and in some cases we have been able to replace inpatient care with outpatient services. We have also continued to have a significant focus on prevention-based strategies (to avoid the use of services altogether) including a very substantial commitment to primary care medical home development.

The results of these trends on the hospital are:

- Significant increases in the acuity and length of stay (LOS) of patients that are hospitalized on our Longwood Campus;
- Growth in the percentage of bed days/revenue from combined regional, national, and international patients;
- Significant increases in volume seen in lower-cost satellite outpatient facilities;
- Deepening investment in accountable care infrastructure and supports.

These results are illustrated in **Chart 1**.

However, we question whether the current reimbursement system established for all hospitals is designed to work well for pediatric hospitals such as Boston Children's, given both our unique safety-net role for high complexity pediatric services and our high Medicaid payor mix. The costs of unique ("stand by") pediatric services should be covered and shared proportionately by all payors and in a manner that does not impose unrealistic cost-sharing burdens on families with special pediatric needs. Our response to the HPC last year laid out much of the data regarding the regionalization of higher complexity pediatric care; this is a national, not a local, phenomenon that arises as a result of the heterogeneity of rare conditions in children, the high level of specialization required, and the need for scale to adequately address these conditions.

b. What actions has your organization undertaken since January 1, 2013 to ensure the Commonwealth will meet the benchmark, and what have been the results of these actions?

Pediatric spending is generally a very small proportion of health care spending. The Commonwealth is thus unlikely to meet its cost targets through a focus on pediatric spending (at least in the short term); as a corollary, efforts at Boston Children's Hospital are likely to make only a small contribution to the state meeting its cost targets.

In the course of negotiations with third party payors, we have established modest rates of payment increase in support of the state's goals, even as we do not anticipate being a driver of increased costs. Our rates of inflation fall below medical inflation increases and often below general inflation, even as we experience substantial increases in pharmacy and other operating costs.

In order to achieve this reduced level of cost growth, we have implemented several initiatives that are intended to address unnecessary utilization of specialty care, including Standardized Clinical Assessment and Management Protocols (SCAMPs), a rapid-cycle quality improvement effort focused (in part) on the elimination of unnecessary utilization of services. We have worked with major pediatric primary care groups to develop shared care protocols in order to enable enhanced medical home capabilities. As noted in our answer to question 1c and question 4, we have invested heavily in accountable care capabilities that target high-needs patients through enhanced care coordination and case management. We also continue to explore enhanced telemedicine supports with an eye towards avoiding unnecessary visits and improving patient satisfaction, although the regulatory and financial environment does not optimally enable these efforts.

c. What actions does your organization plan to undertake between now and October 1, 2015 (including but not limited to innovative care delivery approaches, use of technology and error reduction) to ensure the Commonwealth will meet the benchmark?

Boston Children's Hospital has been aggressively pursuing cost reduction strategies since 2009, prior to the passage of c. 224 and the establishment of specific cost targets. In each of the years since that time, we have managed to keep unit cost growth well below CPI-M (for years prior to c. 224) and below the Massachusetts cost growth targets in the most recent period. We intend to continue our successful and sustainable strategies designed to eliminate unnecessary utilization, and to deliver care in the most appropriate setting.

We think about cost containment from three perspectives: 1) how can we reduce our internal costs; 2) how can we reduce costs to the overall healthcare system through prevention-based strategies; 3) how can we attempt to ensure that these reductions are passed along to employers and consumers through price reductions?

Boston Children's Hospital has been working on efforts to reduce costs from every angle, including unit price, efficiency and utilization. The hospital has decreased our overall per unit cost, volume adjusted, each year for the last five years. We have taken over $125M of expenses out of our system. In FY2013 we implemented $76M in clinical cost savings and identified an

additional $24M in cost savings for FY2014. If successful, we will surpass reductions in costs of over $200M over the last several years. **Chart 2** provides a depiction in our unit costs relative to CPI and CPI-M benchmarks since 2009. Early end-of-fiscal year projections suggest that the trend will be better than budgeted for FY14.

In addition to internal cost reductions, the enterprise has worked to reduce the total cost of care to the health care system, including: reducing lengths of stay, reducing utilization, reducing admissions altogether, and transitioning care to lower-priced settings. As part of our participation in the Blue Cross Blue Shield Alternative Quality Contract (AQC) through the Children's Hospital Integrated Care Organization (CHICO), we have worked across the continuum to support appropriate utilization of services and focused on cost-effective settings. Some examples include:

- Reducing Admissions: Our home ventilation program for technology-dependent children sends teams comprised of ICU physicians, nurses, and respiratory therapists on home visits to prevent admissions to the ICU through better at-home management of these very complex patients. This program was featured by Blue Cross Blue Shield in its annual report this year (2014).
- Reducing ED Utilization: Asthma is the most common cause of ED admission for children in Massachusetts. Our community asthma initiative has reduced asthma ED visit rates by 68% for children with uncontrolled asthma through intense patient education and environmental mitigation efforts and is more fully described in a journal article from Pediatrics. (**See abstract as Attachment 1**.) Within CHICO, we have worked with our primary care practices to implement evidence-based strategies such as extended office hours, nurse call line offerings, and follow-up to sick visits.
- Lower-Priced Settings: Our physicians staff the pediatric services in a number of lower cost community hospitals in eastern Massachusetts. Our general experience is that these staffing arrangements result in fewer admissions overall and fewer transfers of low complexity care to our Longwood campus. Similarly, we have increasingly used our outpatient satellites to deliver less-complex care in lower costs settings; nearly a third of all outpatient care is now delivered in these settings.
- Reducing Variability in Care Delivery: In addition to the SCAMPs efforts cited in Question 1b, we will be undertaking a major cross-institutional initiative with the Healthcare Performance Initiative to continue to enhance our work around creating a culture of safety in the organization. The focus of this work is harm reduction (as opposed to cost containment/meeting state cost targets). We do expect it to lead to a reduction in medical errors consistent with your question. In general, we agree with the hypothesis that error reduction leads to enhanced value and reduced cost, though this can be difficult to prove with respect to individual initiatives.
- Further Development of Accountable Care Capabilities: We continue to make significant investments in our integrated care organization and will begin to roll out enhanced care management capabilities for medically complex patients. We are also working hard to enhance the capacity of our primary care network to address the needs of children with behavioral health concerns (see answer to Question 11).

d. What systematic or policy changes would encourage or enable your organization to operate more efficiently without reducing quality?

The state should encourage the use of telemedicine capabilities and payors (including Medicaid) should define clear pathways for approval of its use. We should streamline requirements for credentialing providers including eliminating the requirement that individual hospitals need to separately credential providers offering "remote" services.

The state should recognize that intermediate levels of care are frequently important mechanisms for improving patient flow and should assess whether existing licensure statutes and regulations enable hospitals to flex beds and/or services at times of high need.

Medicaid underpayment significantly constrains the overall ability to meet cost targets. Boston Children's Hospital receives one third of its reimbursement from the Medicaid program, has a negative margin on these patients and routinely experiences rate reductions (as opposed to reasonable inflationary increases) from Medicaid. **(See Chart 3.)** This experience should be factored in to the establishment of cost targets for individual institutions and for the overall system.

Pediatric quality measures lag behind adult measures. If the state wants to assure that cost reductions are not impacting quality, it will need some mechanism for assessing the quality side of the equation. It is likely, for hospitals, that these measures have been best developed and deployed in major national children's hospitals (for children) and not in smaller pediatric settings. The relevant quality measures will need to address more complex patients (for example: there are excellent, NQF-validated measures on congenital heart outcomes) and not just primary-care based well-child care.

Lastly, we should assure that patients with behavioral health needs receive the care they require on a timely basis and in a well-coordinated manner. There is a growing body of evidence that patients with co-morbid behavioral health conditions are some of the least well-managed and most costly patients in terms of their medical (i.e., non-behavioral health) needs. It is our frequent experience that the children we treat with behavioral health concerns experience by far the most bureaucratic hurdles in accessing the care they need. It is not a good clinical outcome or a good use of resources to have a child boarded on our medical floor for two weeks awaiting placement in a behavioral health hospital; this occurs all too frequently. We should absolutely assure that mental health parity protections are fully implemented, that we are closely monitoring the performance of payors in delivering behavioral health services, that we have adequate clinical capacity across all levels of care in the state to serve patients, and that we eliminate as many unnecessary bureaucratic barriers as possible to accessing necessary care. The state and its regulatory agencies must play a lead role in assuring this occurs.

2. C. 224 requires health plans to reduce the use of fee-for-service payment mechanisms to the maximum extent feasible in order to promote high-quality, efficient care delivery.

a. How have alternative payment methods (APMs) (payment methods used by a payer to reimburse health care providers that are not solely based on the fee-for service basis, e.g., global budget, limited budget, bundled payment, and other non-fee-for-service models, but not including pay-for-performance incentives accompanying fee-for-service payments) affected your organization's overall quality performance, care delivery practices, referral patterns, and operations?

Our integrated care organization is the entity responsible for developing and implementing alternative payment models within our enterprise (hospital, specialists, and primary care network). We have participated in the Blue Cross Blue Shield Alternative Quality Contract since 2012 as the only specialty pediatric hospital/providers with such an arrangement. In addition, we have worked with HPHC to establish bundled payments for certain surgical care. In both examples, we have worked to include quality measures along with cost-effectiveness considerations, including tracking TME and directed care to high-value settings, such as our satellite locations.

We have been effective in delivering on areas of accountability related to quality and cost. This requires continued coordination and integration across the continuum reflected in our integrated care organization. In calendar year 2013, for our attributed population in the AQC, we managed cost growth to a trend more efficient than the network comparison. Our bundled payments have been associated with continued growth in use of our lower-cost satellite settings.

As a matter of principle, we have not focused on managing leakage within our integrated care organization. That is, we seek to direct patients to care in cost-effective, high-value settings, whether they are within our own integrated care organization or not. Our focus on keeping care local leads our providers to establish care coordination and information exchange relationships across a variety of providers and systems. It has been our position that patients and families should have access to high-quality care close to home, whether delivered by a BCH provider or otherwise. It has been our experience as a hospital that this may not be true for other networks establishing APMs. We hear anecdotally from patients and families that they are steered to stay within certain systems for care, even when BCH may be their preferred or the most appropriate setting. While some providers may seek to manage out-of-system care for the purpose of assuring tight integration and coordination of care, we expect some of this steerage is in service to business goals not necessarily in the best interest of patients and families.

We are committed to providing high-value to all patients and families regardless of their insurance type and our programs related to quality improvement and cost-effectiveness are generally applied on an all-payor basis. A few exceptions exist where action is taken on the basis of data made available under APM arrangements. While this serves patients and families well and allows us to advance our mission, it does create a lack of parity among payors, some of whom are investing with us in improvement efforts (through infrastructure and incentive payments). For those not offering such incentives or resources, their patients will benefit from programming funded by others. This is not a sustainable approach.

b. Attach and discuss any analyses your organization has conducted on the implementation of APMs and resulting effects on your non-clinical operations (e.g., administrative expenses, resources and burdens).

We have seen growth in our investment in our integrated care organization each year. With increasingly complex, numerous, and risk-shifting APMs, we have invested in human resources (analytics, care management, administration) and technology to support our work. **Chart 4** shows the growth in annual budget for our integrated care organization from FY12-FY15.

c. Please include the results of any analyses your organization has conducted on this issue, including both for your patients paid for under APMs and for your overall patient population.

As noted above, we have been generally successful in our performance under the AQC (our most comprehensive APM opportunity) for both quality and TME outcomes. Our next largest patient population is Medicaid. We are very interested in new payment models to support innovation with this patient population.

For our overall patient population, we consistently benchmark our quality and costs against other children's hospitals. There is extensive benchmarking data available through the Pediatric Health Information System (PHIS) database overseen by the Children's Hospital Association. We are typically a top tier performer in terms of quality, and have below median costs and receive below median payments relative to other children's hospitals on a case mix adjusted basis. None of the other local pediatric providers contribute information to this database, so it is not possible to benchmark their performance.

3. Please comment on the adequacy or insufficiency of health status risk adjustment measures used in establishing risk contracts and other APM contracts with payers.

a. In your organization's experience, do health status risk adjustment measures sufficiently account for changes in patient population acuity, including in particular sub-populations (e.g., pediatric) or those with behavioral health conditions?

Risk adjustment approaches for cost predictions are improving but have some significant issues. We have completed substantial risk adjustment work with Milliman (an actuarial firm), who have found that, in general, traditional systems are not as accurate in projecting pediatric utilization as systems optimized for children. There are significant issues projecting costs for the most expensive patients because these tend to be unpredictable relative to an adult population that is larger, more frequently utilizes hospital care, and tends to suffer in general from acquired (rather than congenital) medical conditions. **(See Attachment 2.)**

In most populations, it is well-known that predicting cost is most challenging at the extremes, especially patients who are expected to incur high-costs. In a recent analysis, we clarified that for an attributed population, standard risk adjusters (such as DxCG) were poorly predictive among the most complex patients and as a specialty hospital we have a disproportionate share of such

patients. We have addressed these types of issues through reinsurance and truncation. While these are tools for complementing risk adjusters in APMs, they can be costly, making their utility limiting at times.

As noted by many others, risk adjustment systems do not generally account for socio-economic status (SES). This can be a very significant problem for an organization that treats many low income children enrolled in the Medicaid program. **(See Attachment 3.)**

b. How do the health status risk adjustment measures used by different payers compare?

To the extent major payors in Massachusetts utilize risk adjustment systems, they appear to have generally settled on DxCGs as the basis of their analysis. We do not have an optimal understanding of the relative risk of the patients utilizing our system across payors. We have better insight into patients that have been attributed to our primary care practices, but this is a small subset of our overall patient mix. It would be helpful if the payors and/or the state could actually produce a comparative schedule.

More recent versions of DxCG have worked to consider key characteristics of pediatric patients more specifically. While still inadequate, this presents some improvement in the utility of DxCG in pediatric populations.

c. How does the interaction between risk adjustment measures and other risk contract elements (e.g., risk share, availability of quality or performance-based incentives) affect your organization?

We believe that risk adjustment is crucial when comparing things like quality performance across provider organizations and seek to embed such an approach in both our internal and external measurement and benchmarking. To the extent that our contracts contain performance-based incentives, we seek to have them risk adjusted. As an example, we developed a portfolio of quality performance measures for our Blue Cross Blue Shield Alternative Quality Contract. These are risk adjusted and in general both we and Blue Cross Blue Shield are satisfied that they are working as intended.

We would note that in some cases, the solution to the "outlier" problem may be the appropriate use of truncation methods by payors. At this point, it appears that individual payors utilize different truncation approaches. It might be helpful to explore this issue further in order to assess whether some level of standardization across payors would be helpful.

4. A theme heard repeatedly at the 2013 Annual Cost Trends Hearing was the need for more timely, reliable, and actionable data and information to facilitate high-value care and performance under APMs. What types of data are or would be most valuable to your organization in this regard? In your response, please address (i) real time data to manage patient care and (ii) historic data or population-level data that would be helpful for population health management and/or financial modeling.

In order to effectively model accountable care arrangements and potential financial models, we believe that providers will need access to patient level claims data that includes information on both utilization and spending. In some cases, this information could be de-identified. For benchmarking purposes, providers will need the ability to compare their performance against network performance (e.g., in the Blue Cross Blue Shield Alternative Quality Contract, the quality and TME components are relative to a comparison group). The level of data may be specific to the population attributed or for which the provider is being held accountable.

As mentioned previously, we aim to deliver high-value care on an all-payor basis. As APM arrangements are evolving across payors, we have identified opportunities to improve quality and cost-effectiveness that relies on payor data most often provided under risk arrangements. Inasmuch as a payor relationship does not include specific APMs relative to a target population of patients, there may still be utility to sharing key information that supports care management. Current rules and regulations make sharing such data in the absence of an APM difficult.

As an example, Boston Children's has worked with Milliman to better understand the relative risk profile of patients seen here for inpatient care versus those seen at other academic medical centers serving children in Massachusetts. This is important because the vast majority of patients seen at Boston Children’s are not cared for by primary care physicians in our affiliated network. Milliman created an attribution model that assigned patients to a hospital provider (as distinguished from a primary care provider) based on where that child received most of his/her inpatient and/or sub-specialty care. This analysis found that the relative risk of patients attributed to Boston Children's was 1.7 times greater than those seen at other pediatric AMC providers. This will become an important consideration if payors and/or public policy begin to utilize performance indicators for quality or outcomes system-wide.

With respect to real time data, we are especially interested in data that helps us manage patients' care, including chronic diseases. This would include medical, surgical, pharmacy, and behavioral health data. For example, a key factor in good asthma care management is the use of preventive inhaled steroids; the ability to determine in real time whether patients are filling their prescriptions is therefore critical. We have invested substantially in our EMR infrastructure and are considered a national leader among children's hospitals in this regard. Further, we have worked to develop an enterprise data warehouse that allows us to bring together clinical and financial data. In this work, we have found substantial variation in the format and content of data downloads from various sources, including payors. As we work to marry clinical and financial data, it has become clear that providers and others would be well-served by standardized data extract formats across payors and that it is costly to tailor systems to the format and data of specific payors. There may be opportunity to leverage work done in building of the APCD.

5. C. 224 requires health plans to attribute all members to a primary care provider, to the maximum extent feasible.

a. Which attribution methodologies most accurately account for patients you care for?

See answer to Question 5b.

b. What suggestions does your organization have for how best to formulate and implement attribution methodologies, especially those used for payment?

We believe that there is no single, "best" attribution method. Rather, this will depend on the nature of the care delivery and financial model under consideration.

- In general, we are comfortable with an approach that seeks to empanel patients to primary care providers for the development of models relying on medical home infrastructure. If this is the model, then we believe more work needs to be done to keep empanelment lists accurate and timely and to require non-managed care members to nevertheless designate a primary care clinician (we recognize some of the ERISA issues involved in this).
- It may be the case that children with very high levels of medical need should be attributed to a subspecialist for a period of time or permanently on a condition-specific basis. For example, children with End Stage Renal Disease (ESRD) tend to be "managed" by our renal program. There are similar issues for adults with chronic behavioral health needs. We think that any model seeking to utilize a primary care medical home should address whether there are specific circumstances that warrant carving individuals out from empanelment requirements and/or creating empanelment opportunities with subspecialists.
- That said, 90% of the inpatients seen at Boston Children's Hospital are referred from outside our affiliated primary care network. Part of the value of APMs is to present different provider types with aligned incentives. Primary Care Clinicians situated in a primary care-based risk arrangement would benefit from APMs that align the incentives of sub-specialists to whom they refer. In the absence of such incentives, PCPs are referring to providers who are under FFS arrangements. While BCH providers seek to optimize resource utilization even in the absence of APMs, the interests of referring providers and BCH providers may be at odds without an increased focus on sub-specialty focused APMs. It is in the best interest of patients and families for PCPs and the specialists to whom they refer to have aligned incentives such that they are able to focus on the care and treatment of patients.
- As such, if the question is, "for whom are we, as a delivery system, responsible," the state clearly needs to develop an alternative approach to attribution. See the answer to Question 3 regarding work conducted by Milliman regarding the relative risk of patients seen at Boston Children's Hospital.

6. Please discuss the level of effort required to report required quality measures to public and private payers, the extent to which quality measures vary across payers, and the resulting impact(s) on your organization.

Quality measurement and improvement is a key component of our strategy and operations. Our investment in quality and safety is substantial and we have made significant contributions to the field of pediatric quality measurement and improvement, in addition to demonstrating leading performance across a number of safety and quality areas of priority.

In our current arrangements with private payors, we have worked closely with our clinical and quality leaders to present an increasingly well-rounded view of quality performance and improvement across inpatient and ambulatory care. In the last four years, we have successfully negotiated a portfolio of measures that are evidence-based, risk adjusted, and externally benchmarked. Given our focus and prioritization on quality measurement and improvement, our approach has been to draw from our extensive internal reporting and monitoring to reflect a picture in our quality contracts with payors, often in alignment with their priorities.

As a specialty organization, we often work to educate payors about pediatric-specific quality and safety measures and available benchmarks. This frequently entails going beyond HEDIS and other measures most commonly tracked for adults. We are able to utilize the deep expertise of our specialty quality and safety leaders in constructing measures appropriate for risk contracting. Because of the more limited set of available measures, we typically use a phased approach to implementing new quality performance standards: initial use of pay-for-reporting arrangements to establish benchmarks and targets followed by implementation of those metrics considered mutually meaningful through pay-for-performance/APM contracts.

In general, our specialty measures are compared to national benchmarks, which may include local institutions. While we have had success in harmonizing our quality portfolio across payors, we have not made as much progress in establishing a common set of pediatric specialty measures across local pediatric providers (i.e. the other providers are not being held to the same standards for the care they deliver in their contracts). Nor have we made material progress in incorporating pediatric quality measurement into tiering methodologies. (See answer to Question 10).

We have robust efforts within BCH to develop and advance pediatric measurement not specific to our payor contracts, and we are considered a national leader among pediatric systems in this regard. Our efforts include the CHIPRA Center of Excellence in Pediatric Measure Development, led by Dr. Mark Schuster and the Center for Patient Safety and Quality Research, led by Drs. Kathy Jenkins and Al Ozonoff. These efforts have begun to produce measures that have achieved NQF endorsement and will be well suited for payor contracts, including, for example, a pediatric version of the HCAHPs survey.

7. An issue addressed both at the 2013 Annual Cost Trends Hearing and in the Commission's July 2014 Cost Trends Report Supplement is the Commonwealth's higher than average utilization of inpatient care and its reliance on academic medical centers.

Please attach any analyses you have conducted on inpatient utilization trends and the flow of your patients to AMCs or other higher cost care settings.

As noted in our answers to Questions 1 and 2, we have worked extensively to use community-based resources wherever possible. Recent data from our affiliated primary care network indicates that 75% of all inpatient admissions are to hospitals other than BCH. It should be noted, however, that pediatric subspecialists are relatively rare nationally and generally not available in community settings. This is unlike the adult delivery system. Please see our response to last

year's HPC questions for an extensive review of the data in this regard. In addition, we believe that the admissions rate overall is declining for the pediatric population. We do not have comprehensive statewide-data to prove this, but understand that to be the trend among major pediatric providers locally.

b. Please describe your organization's efforts to address these trends, including, in particular, actions your organization is taking to ensure that patients receive care in lower-cost community settings, to the extent clinically feasible, and the results of these efforts.

As an AMC, we have established informal clinical relationships and formal clinical affiliations with a variety of care providers, including other AMCs, community hospitals, community specialists, and PCPs. It is our belief that care of pediatric patients is best delivered in a family-centered way by those with pediatric expertise, whether they are part of the BCH enterprise or not. Currently, we are working in a variety of ways to support affiliated community hospitals in their care of pediatric patients. This ranges from staffing inpatient, ED, and neonatal settings to simulation education and quality programming. We also implement joint goals with our community hospital partners to improve quality and monitor utilization. We monitor our progress on these monthly. Further, we have established telemedicine linkages to ED settings to support our ability to avoid transfers to a higher intensity setting by enabling our physicians in the community to communicate with pediatric sub-specialists at BCH and to prepare better for transfers when they are necessary for pediatric patients admitted to our affiliated community hospitals.

In order to maximize the availability of services in Longwood for the most complex patients, we have built ambulatory capacity in our satellites in Eastern Massachusetts. While these settings allow patients to receive care closer to home, they also offer a lower-cost setting for surgical procedures and imaging services.

In addition to these system-based approaches, we have worked to assure that there are not systematic barriers to our patients under risk-arrangements accessing other AMCs and non-affiliated community hospitals. While others under risk arrangements have worked to manage "leakage," we have taken a patient and family centered position that relies on patients, working with their pediatricians, to make the right decisions for their care. Our integrated care organization has worked to establish clinical bridges and information exchange with hospitals most frequently used by our patient population, irrespective of corporate or contracting affiliation.

We do have concerns that patients who fall under the risk arrangements of non-pediatric systems face limited access to pediatric expertise, whether at BCH or elsewhere. Provider consolidation, coupled with increased prevalence of risk arrangements, has led systems to limit access for patients outside of tightly managed networks. Given the need for pediatric patients to receive appropriate care, such "leakage" efforts could be harmful to patients and families and will likely drive up costs in the long-run.

8. The Commission found in its July 2014 Cost Trends Report Supplement that the use of post-acute care is higher in Massachusetts than elsewhere in the nation and that the use of post-acute care varies substantially depending upon the discharging hospital.

a. Please describe and attach any analyses your organization has conducted regarding levels of and variation in the utilization and site of post-acute care, as well as your efforts to ensure that patients are discharged to the most clinically appropriate, high-value setting.

See answer to Question 8b.

b. How does your organization ensure optimal use of post-acute care?

Our strong preference is that children return home after inpatient admission. At times, this is facilitated through the use of home based services and supports (visiting nurses, care managers, etc.) This is not always possible for children in state custody or requiring habilitative care.

In general, the availability of facility-based post-acute care for children may be more limited than for adults, but is also likely to require more specialized settings and services. The options available are few and far between. As a result, we tend to have strong and well-coordinated communication between our clinical staff and the post acute care settings, as we frequently need to "customize" a service for the unique needs of the child in question.

At present, the only pediatric-specific post-acute inpatient medical facility in the Commonwealth is Franciscan Hospital for Children. They are a unique organization that provides a range of services from medical care, to behavioral health, to residential educational programming. Their capabilities are more comprehensive than those available through home health agencies or pediatric skilled nursing facilities. We believe that appropriate use of these services can reduce length of stay at BCH and reduce overall costs to the health care system because the FHC cost structure is lower than that at a tertiary/quaternary hospital like BCH. We have an ongoing joint clinical operations committee designed to plan and trouble shoot individual patient cases.

It is worth noting that services for many pediatric patients tends to be habilitative as distinguished from rehabilitative (e.g., we are seeking to optimize a child's potential function, not to restore it to a prior baseline). For example, a three year old has not fully developed gross and fine motor skills, so a three year old receiving physical therapy services requires a different approach than a post-stroke elder. Like Boston Children's, Franciscan's cost structure differs from that of its adult-serving counterparts due to the relatively resource-intensive needs of children.

For medically complex children or those with chronic illness, we are increasingly developing models of care (including telemedicine, home visiting, and the use of community health workers). These are frequently not reimbursed by payors.

It is relatively common that hospitalized children with behavioral health needs step down to residential treatment programs. In this circumstance, we frequently must negotiate the transition with relevant payors (insurers and/or school systems). Given the limited capacity in

Massachusetts and payor-specific issues, it would be inaccurate to say we choose where a child goes. We attempt to assure that children are placed in the best care setting given their clinical needs and profile, but frequently spend an enormous amount of time negotiating these placements.

For home based services for children with behavioral health needs, we primarily rely on the matrix of services available through the Children's Behavioral Health Initiative (CBHI). These services are not available to privately insured patients, but arguably should be given the intermediate care requirements of the parity law.

9. C. 224 requires providers to provide patients and prospective patients with requested price for admissions, procedures and services. Please describe your organization's progress in this area, including available data regarding the number of individuals that seek this information (using the template below) and identify the top ten admissions, procedures and services about which individuals have requested price information. Additionally, please discuss how patients use this information, any analyses you have conducted to assess the accuracy of estimates provided, and/or any qualitative observations of the value of this increased price transparency for patients.

Boston Children's Hospital, beginning January 2014, formalized a process for providing estimates to patients and/or families via a dedicated phone line, or via a Website. Estimates being provided include Hospital charges and professional fees (if known). In addition, Boston Children's has provided some payers with internal contacts to assist and expedite estimates that are requested from the payer by members. The vast majority of estimates are being requested by patients/families that have appointments or services scheduled at Boston Children's or with its physicians. Only a limited number of "shoppers" have requested information (less than five YTD). Types of services for which estimates have been requested are:

Outpatient Surgery: 26
Diagnostic Testing: 24
Consults/Office Visits: 26
Inpatient Admissions: 2

Of the above estimates provided, only three were for patients/families with no insurance coverage. The remaining estimates were requested to determine their Insurance Out-of-Pocket expense. We have not yet performed retrospective analytics on the accuracy of estimates. We plan to do so after the end of our current fiscal year.

Health Care Service Price Inquiries				
Year		Number of Inquiries via Website	Number of Inquiries via Telephone/In-Person	Average (approximate) Response Time to Inquiries *
CY2014	Q1	4	19	24 hours
	Q2	3	27	24 hours
	Q3	5	20	24 hours
	Total	12	66	-

** Please indicate unit of time reported*

10. Please describe the manner and extent to which tiered and limited network products affect your organization, including but not limited to any effects on contracting and/or referral practices, and attach any analyses your organization has conducted on this issue. Describe any actions your organization taken (e.g., pricing changes) in response to tier placement and any impacts on volume you have experienced based on tier placement.

Most of the spending and utilization by children is concentrated in a small group of children with relatively complex conditions or acute needs. These children, in turn, tend to concentrate in a few pediatric centers locally. For example, there are only six hospitals in all of Massachusetts that maintain a pediatric ICU. This pattern is true throughout the country. We are generally concerned that limited and tiered network products are developed for the adult population and do not recognize the separate needs of children or the more limited care options available to them.

The Division of Insurance was required to investigate the special needs of children as a result of Chapter 61 of the Acts of 2012. They have not, as yet, issued their report. In the meantime, the National Association of Insurance Commissioners is in the process of updating its model statute governing network adequacy requirements. There was extensive testimony submitted during this process, with many of the commentators remarking on the special protections required for children. The attached consensus statement, with a growing list of endorsing organizations, summarizes many of the key concerns. **(See Attachment 4.)**

In addition, we have deep skepticism that payors are utilizing consistent approaches to the determination of which tier a given provider or physician is placed into. We have found ourselves placed all over the map in terms of tiers, and there seems to be no rhyme or reason to the placement. We think that very little emphasis is placed on quality differentials (and especially quality differentials related to the care of more complex patients). We also have found that the cost analysis used tends to focus on relative price for individual services and does not account for differences in utilization or TME.

We receive frequent complaints from physicians that they are not being allowed to send patients to Boston Children's when they believe it is in the best interest of the patient to do so. We have also had occasional experiences in which complex patients have been redirected elsewhere based

primarily on limited network designs (which seem to have been most aggressively pursued by health care systems for their own employees).

11. The Commission has identified that spending for patients with co-morbid behavioral health and chronic medical conditions is 2-2.5 times as high as spending for patients with a chronic medical condition but no behavioral health condition. As reported in the July 2014 Cost Trends Report Supplement, higher spending for patients with behavioral health conditions is concentrated in emergency departments and inpatient care.

a. Please describe ways that your organization is collaborating with other providers to integrate physical and behavioral health care services and provide care across a continuum to these high-cost, high-risk patients.

For over 30 years, Boston Children's Hospital has cared for patients with co-morbid physical and behavioral health care needs through provision of mental health clinicians integrated into inpatient medical/surgical services. These collaborations range from straight-forward co-location care to full integrated care models. Some examples of these collaborations include:

- Social work, psychology, and psychiatry clinicians are available for consultation to medical/surgical specialty physicians throughout all medical/surgical inpatient units. These clinicians are supported by resource specialists and care coordinators, whose role is to access community resources for the patient and family in a timely manner, in order to decrease the need for more acute and costly services.
- Social work, psychology, and psychiatry clinicians are highly integrated into a number of specialty care programs where specific disorders are targeted including pain disorders, neurodevelopmental disorders, solid organ transplants, epilepsy, brain disorders, deafness/hearing loss, and pediatric cancers. These collaborations involve longitudinal care across both inpatient and outpatient settings.
- After many years of providing co-located care in primary care, BCH began in Spring 2013 to implement a stepped-care integrated healthcare model into its primary care clinics and into the Pediatric Physicians Organization at Children's (PPOC), comprising 300 community pediatricians across Eastern Massachusetts.

b. Please discuss ways that your organization is addressing the needs of individuals to avoid unnecessary utilization of emergency room departments and psychiatric inpatient care.

There are a number of interventions that have been undertaken to avoid unnecessary utilization of emergency room departments and psychiatric inpatient care. They include:

- **24-7 Behavioral Health Emergency Coverage:** BCH provides 24-7 ED psychiatry coverage by clinicians well-trained in behavioral health emergencies including assessment, crisis intervention, and triage/disposition. Behavioral health clinicians respond immediately to all providers both within the hospital and in the community in order to help providers/families to access crisis teams and diversionary levels of care when clinically appropriate. These clinicians use a crisis intervention model in order to

avoid unnecessary hospitalizations for patients who can be stabilized and discharged to a less restrictive level of psychiatric care. The Psychiatry Department is exploring the establishment of an urgent care system within Outpatient Psychiatry, in order to divert patients from the Emergency Department, where appropriate.

- **School Consultation:** The Boston Children's Hospital Neighborhood Partnerships provides school-based consultation and offers a prevention-oriented curriculum in the Boston Public Schools.
- **Inpatient Units/Outpatient Settings:** As described above, there is a psychosocial clinician integrated into the medical/surgical care teams across the hospital, with the goal of responding to behavioral health crises in those settings, thereby avoiding unnecessary utilization of ED and inpatient psychiatric care. Outpatient primary care clinics have the capacity to respond to behavioral health crises through the use of their internal behavioral health clinicians, thereby decreasing ED utilization.

c. Please discuss successes and challenges your organization has experienced in providing care for these patients, including how to overcome any barriers to integration of services.

Our successes in providing care have focused mainly on providing integrated BH models of care for highly complex physically ill patients seen in medical and surgical specialties consistent with the finding noted in your question. These have been programs that have either had external mandates for multispecialty care (i.e., transplant programs are required to offer these supports) and/or are embedded in services that care for high cost patients (i.e., cardiology). In these services, the core components of collaborative care (direct service, consultation, education, and collateral care) are provided. The need for integrated care that goes beyond to reach more patients at earlier stages of their illnesses is well-recognized in today's healthcare. The challenge is to move beyond fee-for-service models to more population based financial models that allow for integrating healthcare into primary care and school based settings.

Additional challenges include the following:

- The stigma of mental health disorders may result in families/providers seeking expensive medical answers to complex symptoms; education for both patient-families and providers about the positive impact of well integrated care might decrease the expenditures on expensive workups.
- At a time of critical shortages in the behavioral health clinician workforce, a particular challenge in behavioral health is the resistance among health plans to permitting trainees under appropriate oversight and supervision to provide reimbursed clinical services to their subscribers. While the intention is to assure that their subscribers get optimal care, the reality is that it creates uneven access to care; those patients on more restrictive health plans have to wait longer for care than those whose plans permit supervised trainee services.
- Behavioral health care exists across a continuum of care providers, some of whom lie outside the BCH enterprise (for example: residential care providers). These providers are essential to the care of children and adolescents with medical and psychiatric co-morbidities.

d. There has been increased statewide interest in data reporting across all services, inclusive of behavioral health. Please describe your organization's willingness and ability to report discharge data.

Boston Children's Hospital reports discharge data as required by the Department of Public Health, and is interested in continuing to report data across services, including behavioral health. However, collecting discharge data does not seem sufficient to address the issues of care integration and cost containment that are raised here. In order to develop a clearer and more data-driven understanding of the reasons that patients with medical and behavioral health conditions cost 2-2.5 times those with medical conditions alone, more comprehensive data collection and analysis is necessary.

Understanding the utilization of services over the entire course of an acute event and immediately thereafter would help better shape patient care and help to increase both quality and efficiency of care delivery. For these purposes, crisis evaluation, admission, service utilization, and discharge data might together provide a better lens for interpreting what factor(s) are driving the costs. We hypothesize that some of the patterns of utilization and drivers of cost differentials may differ for children.

We would caution against wholesale collection of "new" data without a thoughtful, multi-stakeholder discussion of the research questions we are trying to answer, the most effective means of answering them, and whether individual organizations have the ability to supply the data. It could be better to structure this as a "more traditional" research project.

12. Describe your organization's efforts and experience with implementation of patient-centered medical home (PCMH) model.

Our primary care network includes both large, hospital-operated practices and a community-based network of pediatricians.

With respect to the hospital owned and operated practices (Boston Children's Primary Care at Longwood and Martha Eliot Healthcare Center), over the past two years Boston Children's Hospital has invested in patient-centered medical home transformation using the Change Concepts from the Safety Net Medical Home Project. The practices seek to deliver high-quality, proactive preventive and acute health care to all of our primary care patients through the use of care teams, population management techniques, evidence-based care for chronic illness, and approaches which empower our patients.

Empanelment: In July 2013 the practices achieved the goal of assigning 95% of patients to a primary care provider and team and are currently maintaining that level. By August 2013 over 90% of patient families queried were able to correctly identify their primary care physician.

Team-Based Care: In January 2013 the practices rolled out care teams that included nurses, clinical assistants, physicians, social workers, administrative staff, and patient navigators. All

care is delivered through these patient care teams. By June 2014, 85% of staff reported that the teams were highly functioning.

Loop Closure: Electronic communication strategies have been implemented for communication with specialists and loop closure is in place for referrals, labs, and tests.

Population Management/Chronic Disease Management: Registries were established for patients with complex care needs, asthma, and obesity. Care coordination was instituted and patients' progress in the medical system continues to be tracked. Patients in need of well child care are also tracked and supported using automated reports. Electronic decision-making support has been developed but is not yet in place for obese children.

Patient Empowerment: A parent advisory board was established and there is a parent representative on the transformation team. A monthly newsletter provides seasonal advice. Group-based visits have been developed for obese children and their families, which include nutritional and direct exercise components and have been wildly successful. Group visits for children with asthma are now under development. Cultural competence training is mandatory for all staff.

Mental Health Integration: The practices are now actively working on mental health integration with practice-wide screening, secondary assessment, stepped care and coordination.

With respect to the community-based affiliated practices, see the answers to 12a-c.

a. What percentage of your organization's primary care providers (PCPs) or other providers are in practices that are recognized or accredited as PCMHs by one or more national organizations?

The Pediatric Physicians Organization at Children's (PPOC) has implemented a PCMH program across its network of 80 practices to support integration of PCMH standards and principles of care across practices. The hospital-based primary care practices at BCH also provide care in a PCMH model.

At present, there is only one practice that is an accredited PCMH: the South Cove Community Health Center. All the other PPOC practices have adopted and passed (scored by the PPOC) the 10 Must-Pass Elements of PCMH 2008. Thus, they operate as medical homes but are not formally accredited PCMHs.

b. What percentage of your organization's primary care patients receives care from those PCPs or other providers?

The practices apply the PPOC PCMH Model to all their patients.

c. Please discuss the results of any analyses your organization has conducted on the impact of PCMH recognition or accreditation, including on outcomes, quality, and costs of care.

The PPOC Medical home framework consists of three components:

1. NCQA Standards: These are based on the National Committee on Quality Assurance 2008 Patient Centered Medical Home 10 Must-Pass Core Elements. These elements include increased access and communication with patients and families, organizing clinical information, identifying important diagnoses and conditions, adopting and implementing evidence-based guidelines, test and referral tracking, population management, performance reporting and improvement. Practices are required to examine their current processes and make the necessary changes to meet these standards. As part of this effort, practices strive to implement processes that can be maintained as well as measured over time to assess effectiveness.

2. Care Coordination: Practices create a care coordination plan and hire a care coordinator. Care coordination is aimed at improving the transfer of patient information and facilitating transitions in care. With care coordination, there is an emphasis on greater teamwork among the practice and a commitment to following a patient's care before, during, and after their visit to pediatrician's office.

3. Family Engagement and Partnership: Many practices are engaging the families in partnership with the clinical staff in the care of the child and in creating stronger medical homes within the practices. The practices are expanding and acting upon opportunities by adding families as a voice in patient care.

Benefits of Medical Home in PPOC Practices include:

- Improved access and communication with patients and families;
- Increased pre-visit information collection resulting in improved office efficiency and higher quality office visits;
- Improved test tracking and prompt follow up with results;
- Improved inventory of community resources for families;
- Stronger connections with community agencies and schools resulting in improved coordination of services;
- Better identification and management of children with chronic conditions;
- Increased patient and family involvement in self- management;
- Increased teamwork in the practices.

13. After reviewing the Commission's 2013 Cost Trends Report and the July 2014 Supplement to that report, please provide any commentary on the findings presented in light of your organization's experiences.

In answering this question, we also reviewed the recent FY14 report release.

- It is not clear that the Commission has provided much assistance in helping to think about or measure reductions and/or differentials in service utilization between providers. We believe that reduction of unnecessary utilization is likely to be one of the most important levers in effective cost reduction over the long term. Price reductions alone will have a

relatively finite, "one time" impact. In order to incent utilization-related activities, the Commission should make the measurement and understanding of unnecessary utilization a focus of its analytic work.

- We found the finding on APM uptake somewhat disingenuous. The Report acknowledges that the majority of Massachusetts residents are enrolled in public coverage. Many of the initial APM experiments have been in the Medicare program. In our experience, the clinical redesign necessary to achieve APM objectives is considerably "payor agnostic."
- Consistent with the preceding point, if 60% of Massachusetts residents are covered in public coverage programs, one would think that more attention would be paid in the report and in the Commission's work to public coverage data, analytics, opportunities and issues. Focus on the Medicaid program is particularly lacking. There were also quite clear directives in c.224 with respect to APM adoption in Medicaid and there is virtually no reporting on progress in this area contained in the report.

Office of the Attorney General Written Questions

1. Please submit a summary table showing for each year 2010 to 2013 your total revenue under pay for performance arrangements, risk contracts, and other fee for service arrangements according to the format and parameters provided and attached as AGO Hospital Exhibit 1 with all applicable fields completed. Please attempt to provide complete answers. To the extent you are unable to provide complete answers for any category of revenue, please explain the reasons why.

Please refer to enclosed table (Exhibit 1).

Please see **Exhibit 1**. We have separately broken out in-state vs. out-of-state business for both commercial and Medicaid as a very substantial portion of our revenue is derived from out-of-state payors and patients. We do not separately track revenue by type of business (e.g. PPO vs. HMO).

2. For each year 2010 to present, please submit a summary table showing for each line of business (commercial, government, other, total) your inpatient and outpatient revenue and margin for each major service category according to the format and parameters provided and attached as AGO Hospital Exhibit 2 with all applicable fields completed. Please submit separate sheets for pediatric and adult populations, if necessary. If you are unable to provide complete answers, please provide the greatest level of detail possible and explain why your answers are not complete.

Please refer to enclosed table (Exhibit 2).

We do not capture information on margins by service line in this way and are skeptical that we would provide accurate information in response. Individual patients commonly receive services across multiple service lines during the course of treatment, but revenue for that patient may only be captured in one service line. In addition, accuracy of the answer assumes that internal cost allocation systems would adequately reflect costs for individual service lines. We cannot submit a response to this question signed under "pains and penalties of perjury" given that we do not analyze or aggregate information in this fashion.

In order to provide some insight into the relative scale of services provided at Boston Children's, we have created a summary table from our reported hospital cost reports by service category. Please see **Exhibit 2**.

3. Please explain and submit supporting documents that show how you quantify, analyze and project your ability to manage risk under your risk contracts, including the per member per month costs associated with bearing risk (e.g., costs for human resources, reserves, stop-loss coverage), solvency standards, and projections and plans for deficit

scenarios. Include in your response any analysis of whether you consider the risk you bear to be significant.

The hospital's participation in the AQC risk contract is via CHICO, which provides key administrative and management functions to the Hospital and its other member organizations. The hospital's share of gains or losses through the risk contract is allocated based on a formula agreed to within CHICO. Each quarter, CHICO provides the hospital an estimate of financial impact based on risk contract performance. To date, specific reserves have not been established beyond estimates due to/from third parties. Stop-loss coverage has been obtained through the payor with whom we have our risk agreement. Costs of administering the contract is by and large accounted for in the dues paid to CHICO. It is not possible to calculate the per member per month costs since CHICO's efforts to improve quality, reduce fragmentation, and assure access to cost-effective care are applied in a payor-blind manner. To the extent that we entered into agreements that include substantial exposure to downside risk, we would need to revisit our current approach and would work through CHICO to develop a policy across all members. We await guidance on the requirements of risk-bearing provider organizations to inform such a policy. Similarly, additional risk agreements would likely require further investments in analytical, IT, care management, and contract management functions commensurate with access to additional data and likely additional performance measures.

4. Please explain and submit supporting documents that show how you analyze and track the volume of inpatient and outpatient referrals to your hospital and the associated revenue from those referrals by particular physicians or provider groups. Please include a description and examples of how your organization uses this information.

As a subspecialty hospital that serves a majority of patients that are referred from outside our own primary care network, we make some attempt to track overall volume by individual referring provider and by their network relationship. We are not, however, confident, that we have accurately assigned individual physicians to the correct networks given the ongoing consolidation and realignment of individual physicians and groups in the marketplace. It is our understanding that the state does not have such a list, either.

For those referring provider groups with whom we have formalized contracted relationships, we have limited information on referral patters for their risk population. We often rely on this information augmented by EMR data to identify opportunities to improve and coordinate clinical care and information exchange. For example, we have worked closely with a group of referring providers to identify the extent to which our Community of Care hospitals are used by their patients. Similarly, we streamlined the process for notifying the BCH Longwood ED about patients to expect from practices whose patients are proximal to our Longwood location.

Finally, for our own primary care network with whom we have accountable care arrangements, the location of services is used primarily to identify entities with whom we target improved coordination and communication. For example, we identified that a significant number of patients receive care at several of our Community of Care hospitals. We have worked with these hospitals to establish automated daily reports including census and overnight ED visits. In

addition, we have identified those pediatric sub-specialties where a substantial number of outpatient visits occur with BCH sub-specialists and work to develop cost-effective care models. For example, we identified that BCH orthopedists represent a significant proportion of orthopedic care for our at-risk population. On this basis, we developed a Learning Community to train pediatric practices in the PPOC on orthopedic care that can be provided in the medical home. We have developed rapid response access through text and phone to support those practices in keeping care within the medical home. In addition to these efforts, we have defined population based outcomes, such as asthma control and rates of teen pregnancy, and drawn providers into improvement efforts targeted at these disparities. While many peers use the information on location of care to manage "leakage," this is not a primary strategy for us and we do not track or report such data on a regular basis.

APPENDIX

Attachments

Attachment 1: "Community Asthma Initiative: Evaluation of a Quality Improvement Program for Comprehensive Asthma Care"

Attachment 2: "Risk Adjustment for Pediatric Populations"

Attachment 3: **"**Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors"

Attachment 4: "Safeguarding Access to Medically Complex Care for Children by Requiring Health Plans and Exchanges to Develop Adequate Provider Networks"

Charts

Chart 1: Recent Hospital Trends

Chart 2: Cumulative Savings and Cost Containment

Chart 3: Medicaid Underpayment

Chart 4: Integrated Care Organization Budget

Exhibits

Exhibit 1: Revenue Summary (In-State and Out-of-State)

Exhibit 2: Summary of Hospital Costs

Attachment 1:

"Community Asthma Initiative: Evaluation of a Quality Improvement Program for Comprehensive Asthma Care"

Community Asthma Initiative: Evaluation of a Quality Improvement Program for Comprehensive Asthma Care

AUTHORS: Elizabeth R. Woods, MD, MPH,[a] Urmi Bhaumik, MBBS, MS, DSc,[b] Susan J. Sommer, MSN, RNC, AE-C,[a] Sonja I. Ziniel, PhD,[c] Alaina J. Kessler, BS,[a] Elaine Chan, BA,[a] Ronald B. Wilkinson, MA, MS,[d] Maria N. Sesma, BS,[a] Amy B. Burack, RN, MA, AE-C,[b] Elizabeth M. Klements, MS, PNP-BC, AE-C,[f] Lisa M. Queenin, BA,[b,g] Deborah U. Dickerson, BA,[b] and Shari Nethersole, MD[b,h]

[a]Division of Adolescent/Young Adult Medicine, [b]Office of Child Advocacy, [c]Clinical Research Program, and [d]Information Services, Children's Hospital Boston, Boston, Massachusetts; [e]ESAC Boston Asthma Initiative, Jamaica Plain, Massachusetts; [f]Medicine Patient Services, [g]General Pediatrics, and [h]Office of Government Relations, Children's Hospital Boston, Boston, Massachusetts

KEY WORDS
asthma, cost analysis, community health worker, emergency department visits, health disparities, health outcomes, hospitalizations, nurse case management, pediatrics, return on investment

ABBREVIATIONS
AAP—Asthma Action Plan
CAI—Community Asthma Initiative
CHW—Community Health Worker
CI—confidence interval
ED—emergency department
FTE—full-time equivalent
FY—fiscal year
GEE—generalized estimating equation
IPM—Integrated Pest Management
NAEPP—National Asthma Education Prevention Program
QI—quality improvement
ROI—return on investment

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or Health Resources and Services Administration.

www.pediatrics.org/cgi/doi/10.1542/peds.2010-3472

doi:10.1542/peds.2010-3472

Accepted for publication Nov 11, 2011

This work was presented in part at the Pediatric Academic Society meeting; May 2, 2010; Vancouver, BC, Canada.

Address correspondence to Elizabeth R. Woods, MD, MPH, Division of Adolescent/Young Adult Medicine, Children's Hospital Boston, 300 Longwood Ave, Boston, LO-306 MA 02115. E-mail: elizabeth.woods@childrens.harvard.edu

(Continued on last page)

WHAT'S KNOWN ON THIS SUBJECT: Comprehensive home visits conducted by Community Health Workers including environmental remediation and office-based nurse case management improve asthma outcomes.

WHAT THIS STUDY ADDS: Implementation of a comprehensive quality improvement program as part of enhanced care of pediatric asthma patients with a history of hospitalizations or emergency department visits can improve health outcomes and be cost-effective as well as reduce health disparities.

OBJECTIVE: The objective of this study was to assess the cost-effectiveness of a quality improvement (QI) program in reducing asthma emergency department (ED) visits, hospitalizations, limitation of physical activity, patient missed school, and parent missed work.

METHODS: Urban, low-income patients with asthma from 4 zip codes were identified through logs of ED visits or hospitalizations, and offered enhanced care including nurse case management and home visits. QI evaluation focused on parent-completed interviews at enrollment, and at 6- and 12-month contacts. Hospital administrative data were used to assess ED visits and hospitalizations at enrollment, and 1 and 2 years after enrollment. Hospital costs of the program were compared with the hospital costs of a neighboring community with similar demographics.

RESULTS: The program provided services to 283 children. Participants were 55.1% male; 39.6% African American, 52.3% Latino; 72.7% had Medicaid; 70.8% had a household income <$25 000. Twelve-month data show a significant decrease in any (≥1) asthma ED visits (68.0%) and hospitalizations (84.8%), and any days of limitation of physical activity (42.6%), patient missed school (41.0%), and parent missed work (49.7%) (all $P < .0001$). Patients with greatest functional impairment from ED visits, limitation of activity, and missed school were more likely to have any nurse home visit and greater number of home visits. There was a significant reduction in hospital costs compared with the comparison community ($P < .0001$), and a return on investment of 1.46.

CONCLUSIONS: The program showed improved health outcomes and cost-effectiveness and generated information to guide advocacy efforts to finance comprehensive asthma care. *Pediatrics* 2012;129:465–472

Developing and testing the effectiveness of new chronic care models is essential for cost savings under health care reform. Asthma is 1 of the most common chronic illnesses for children in the United States, and rates have reached historically high levels nationally with large racial/ethnic health disparities.[1,2] For children <18 years, asthma rates had increased to 9.6% in 2009.[3] At the time of planning for this project (2003–2005), the asthma prevalence rate was 9.5% overall in Massachusetts, but the average prevalence reported in the urban Boston Public Schools was 16% with 5 schools reporting rates >24%. In addition, asthma was the leading cause of hospitalization at Children's Hospital Boston (hereafter referred to as "Children's") with 70% of children hospitalized with asthma from urban, low-income neighborhoods in Boston. There were substantial health disparities with rates of asthma-related hospitalizations 5 times higher for black (14.2 per 1000) and Latino (14.1 per 1000) compared with white children (2.9 per 1000).[4] Care for children with poorly controlled asthma provides an important opportunity for the development of novel models of care and new payment systems under health care reform.

The National Asthma Education Prevention Program (NAEPP)[5] provides guidelines for asthma management that have been effective in improving health outcomes including decreasing hospitalization and emergency department (ED) visits.[5,6] Previous studies, including several randomized clinical trials, have demonstrated that multifaceted community-based environmental interventions for children with asthma that follow the NAEPP guidelines are particularly successful.[7–16] Effective interventions incorporating trained community health workers (CHWs) who provide asthma education and environmental materials (such as bedding encasements, cleaning materials, and HEPA vacuums) reduce household antigens, improve quality of life and symptom-free days, and decrease hospitalizations beyond office-based nurse case management.[8–11]

This program was modeled after several community-based comprehensive programs that address health disparities. The "Yes We Can Urban Asthma Partnership" and other culturally sensitive community-based programs provide a road map for comprehensive approaches that improve asthma-symptom-free days and reduce ED visits.[12–14] The Harlem Children's Zone Asthma Initiative focused on a specific geographic community so that health outcomes could be tracked, and demonstrated reduced ED and urgent-care office visits through a combination of care coordination and CHW home visits.[15,16] The combination of CHW asthma education and office-based nurse case management have demonstrated cost-effectiveness[17] and improved quality of life.[18] However, none of these programs have incorporated nurse home visits to address medication issues and compliance as well as environmental support. In addition, limited cost analyses of comprehensive programs are available.[19]

As we implemented the current program, we used the health outcomes presented by Lieu and colleagues, who demonstrated that asthma quality improvement (QI) indicators can be tracked including ED visits and hospitalizations.[20] QI efforts to develop individualized care plans with an up-to-date Asthma Action Plan (AAP) have been shown to reduce acute care visits.[21] The Community Asthma Initiative (CAI) was developed to address health disparities in the Boston neighborhoods most impacted by asthma by providing an enhanced model of care for children previously seen in the ED or hospitalized because of asthma. The objective of this article is to evaluate the CAI health and quality-of-life outcomes, to compare cost data with a similar community, and to calculate the return on investment (ROI) to society for this QI initiative.

METHODS

Setting

CAI was designed to reduce health disparities by addressing asthma issues at multiple levels of the socioecological model.[22,23] A community-based participatory approach involving active Community and Family Advisory Boards[24] and evidence from previous programs were used in the design. The model was developed for children 2 to 18 years old living in 4 urban zip codes showing a high prevalence of asthma and encompassing diverse underserved communities neighboring a major pediatric urban hospital and the hospital's community health center. The model includes (1) nurse case management and coordination of care with primary care and referral services, (2) nurse (bilingual) or nurse-supervised CHW (bilingual/bicultural in Spanish) home visits for asthma education, environmental assessment, and remediation materials (HEPA vacuum, bedding encasements, and Integrated Pest Management (IPM) materials tailored to the needs of the family), and connection to community resources; (3) referral to an IPM exterminator or Inspectional Services (http://www.cityofboston.gov/isd/housing/bmc/) when indicated.

Population

CAI services were offered to children from the 4 zip codes who had a recent ED visit or hospitalization. The nurse case manager (hereafter referred to as "nurse") reviewed daily, weekly, and monthly admission and ED logs for patients with the diagnosis codes for asthma. Patients were prioritized at greatest need for services because of a hospitalization or multiple ED visits in

the past year. Patients with intake from October 1, 2005 to June 30, 2008 had sufficient follow-up time to be included in this study. Services and follow-up care were provided for 1 year.

Patients were contacted by the nurse through face-to-face visits during hospitalizations or through telephone contact, and were offered case management services and home visits. Clinical releases were obtained to allow communication with providers and home visitors contracted through a community agency. Baseline, 6-, and 12-month standardized interviews were completed as part of clinical care to assess asthma symptoms and control, number of ED visits and hospitalizations, days of limitation of physical activity, child missed school days, parent/guardian missed work days, insurance access, up-to-date AAP (updated within the past year), environmental issues, and medication adherence. Asthma severity scores were obtained from AAPs, modified through clinical assessment by the nurses and discussions with primary care providers, and categorized as intermittent, or mild, moderate, or severe persistent asthma according to the NAEPP guidelines.[5]

For the cost analyses, CAI patients were compared with children from 4 similar zip code neighborhoods (not statistically different): similar diverse low-income communities (41.2% vs 59.2% black; 46.1% vs 34.6% Hispanic), male gender (53.9% vs 59.8%), mean age (7.9 ± 4.4 years vs 7.· ± 5.4 years), and socioeconomic status (77.5% vs 73.3% Medicaid) with ED visits or hospitalization during the same study period. From hospital administrative data for the CAI and comparison community, the number of hospitalizations and ED visits, and costs were assessed the year before the baseline visit, and 1 and 2 years of follow-up. Children's Internal Review Board waived the need for consent for the enhanced clinical care program, and approved access to case management data and hospital administrative databases for intervention and comparison groups with waiver of informed consent for the evaluation.

Statistical Analyses

Data were analyzed with the use of Stata version 10.1. Outcomes obtained by parental report included whether patients in 6-month time intervals had ED visits or hospitalizations (events), or limitation of physical activity, missed school or parent/guardian missed work (days) because of asthma, and if the patient had an up-to-date AAP. The events/days were analyzed both as dichotomous variables of the percentage of patients with ≥1 events/days versus none, and continuous variables of the number of events/days. Demographic characteristics such as age, gender, race/ethnicity (black/African American versus others, Hispanic versus others), insurance status (private versus public), household income (<$25 000 versus higher income), and asthma severity scores were collected. For the trichotomous variable for asthma severity (severe, moderate, others), indicator variables were developed for moderate versus others and severe versus others for the multivariate analyses. The number of home visits and any (≥1) nurse home visits were tracked. Analyses evaluated changes from baseline to 6 or 12 months, or the combined follow-up variable (with the use of the latest follow-up visit available).

For the intervention group, attrition analysis for demographic and asthma characteristics was performed with the use of χ^2 tests for categorical variables and unpaired t tests for continuous variables comparing baseline values for initial and follow-up time points. Paired analyses used the McNemar test to assess differences in dichotomous outcomes between the baseline and follow-up measurements. Paired t tests were applied for comparisons of continuous variables at 2 time points. Dichotomous outcomes across 3 time points were compared by using unadjusted and adjusted repeated-measures random intercept logistic regression models (displayed with odds ratios with their 95% confidence interval [CI]). Generalized estimating equation (GEE) repeated measures random intercept Poisson regression analyses tested differences for the counts of number of events/days for outcome variables (displayed with the change in number of events/days and 95% CI). Because of a small increase in all outcomes at 12 months, a quadratic term was inserted in the equation for multivariate models to correct for seasonal variation.

Hospital administrative data were used to compare the admissions, ED visits, and hospital cost for the intervention and comparison populations for Fiscal Year (FY) 2006. Cost of the ED visits and hospitalizations for each patient was calculated with the baseline event included in the previous year and assessing events at · and 2 years of follow-up.[19] A comparison group was identified for those with an ED visit or hospitalization from demographically similar neighborhoods; the first visit in the time period was used as the baseline visit. Hospital charges were adjusted with the appropriate Medicare modified rate (□0.42) to estimate hospital costs and brought to net present value (current dollar amounts). The ROI was calculated for the CAI patients, comparing the cost savings for society (due to the reduction in ED visits and hospitalizations) over the cost of the clinical program (ROI = difference in hospital costs of baseline from year 1 and year 2 for CAI patients divided by the cost of the program). The clinical cost of the program in FY2006 for 102 new families included 1.0 full-time equivalent (FTE) nurse, 1.0 FTE subcontracted CHW, 0.25 FTE program

coordinator, 0.1 program director, 0.1 FTE evaluator, IPM materials, and IPM exterminator services (including $194 246 personnel, $58 712 materials, and $5000 exterminator services).

RESULTS

During the study period, 562 children were identified and 283 (50.4%) children's families agreed to participate. The participating children were 55.1% male; 39.6% black, 52.3% Latino; 72.7% Medicaid; 70.8% household income <$25 000 (Table 1). One hundred twenty (42.9%) were scored as having moderate or severe persistent asthma; the remainder of the children had intermittent (24.3%) or mild persistent asthma (32.9%) with exacerbations resulting in ED visits or hospitalizations. One hundred fourteen (40.3%) were enrolled face-to-face by the nurse during the hospitalization and the rest by phone. A total of 203 (71.7%) families had a mean of 1.28 home visits (±1.27 SD), including 176 nurse visits (performed by nurse, or CHW and nurse) and 145 CHW visits, and 40 IPM exterminator visits. The retention rate was 68% at 6 months and 60% at 1 year, and 78% of participants had follow-up at 1 or both time points (follow-up). Attrition analyses showed minimal differences for baseline values of variables for the population compared with those cared for at 6 or 12 months of follow-up, with the exception of fewer low-income patients at 6 months, and fewer Hispanic patients at 12 months. Demographic variables were controlled for in the final models.

There were highly significant (all $P <$ .0001) reductions in any (≥1) ED visits (66.5% at 6 months, 68.0% at 12 months, and 56.0% with any follow-up), hospitalizations (79.7%, 84.8%, 82.6%), days of limitation of physical activity (50.4%, 42.6%, 38.7%), patient missed school days (44.9%, 41.0%, 42.3%), and parent missed work days (53.2%, 49.7%, 47.7%) (Fig 1). There was a large improvement in having an up-to-date AAP at follow-up (59.1%, 55.3%, 55.6%; $P <$ .0001). Also, for the continuous variables, there were similarly highly significant reductions in the number of events/days at 6 and 12 months (all $P <$ .000[illegible]) (Table 2).

TABLE 1 Baseline Demographic Information and Asthma Characteristics for Community Asthma Initiative Participants

	Baseline $N = 283$ *n* (%)
Age, mean in years (SD) ($n = 283$)	7.9 (4.6)
Gender (male) ($n = 283$)	156 (55.1)
Insurance (private) ($n = 282$)	66 (23.4)
Household income (<$25 000) ($n = 257$)	82 (70.8)
Race/ethnicity ($n = 283$)	
Hispanic	148 (52.3)
Black/African American	112 (39.6)
Other	23 (8.1)
Asthma Severity Score ($n = 280$)	
Intermittent	68 (24.3)
Mild persistent	92 (32.9)
Moderate persistent	99 (35.4)
Severe persistent	21 (7.5)
Enrollment ($n = 283$)	114 (40.3)
Face-to-face during hospitalization	114 (40.3)
Number of families receiving home visits	203 (71.7)
Mean number of home visits/family (SD)	1.28 (1.27)
Total number of nurse or CHW home visits	321
Number nurse or CHW and nurse visits (nurse visits)	176 (54.8)
Number of CHW-performed visits	145 (45.2)
Number of families receiving IPM extermination services	30 (14.7)
Number of IPM extermination service visits	40

Multivariate logistic regression models for dichotomous outcomes (controlling for demographic variables, asthma severity, number of home visits, any nurse home visits, and a quadratic term) showed that there were greatly reduced odds of having any ED visits, hospitalizations, days of limitation of activity, patient missed school days, parent/guardian missed work days, and increased odds of an up-to-date AAP at follow-up (Table 3). Patients with greatest functional impairment from ED visits and missed school were more likely to have any nurse home visits and greater number of home visits, respectively.

GEE for continuous variables, controlling for the same variables, showed significantly decreased number of ED visits (−2.84 events; 95% CI −3.98 to −1.71), hospitalizations (−3.16; −5.06 to −1.26), days of limitation of activity (−2.11 days; 95% CI −2.68 to −1.53), missed school days (−0.75; −1.11 to −0.40), and missed parent/guardian work days (−1.31; −1.87 to −0.74). Those with more home visits and any nurse visits were associated with more days of limitation of physical activity (0.06; 0.02–0.1[illegible]) and (0.14; 0.01–0.27), and missed school (0.04; 0.00–0.07) or (0.23; 0.13–0.32), respectively.

The cost of ED visits and hospitalizations for FY2006 CAI patients and a comparison population 1 year back and 2 years forward by using hospital administrative data showed remarkable differences (Fig 2). CAI patients started out with higher average cost per patient in the 1 year before entering the program compared with the comparison community, had similar costs at 1 year (with a greater decline from baseline for CAI patients), and had further reduction in costs at 2 years (repeated-measures analysis comparing intervention and comparison groups was $P <$ .001). Services were provided for 1 year with □10% of patients needing care after the first year. The cost of the clinical program was $2529/child and the savings for the intervention group was $3827/child over 2 years of follow-up yielding a ROI of 1.46. In other words, for every dollar spent on the program, 1.46 dollars were saved to society because of reduced ED visits and hospitalizations.

DISCUSSION

CAI augmented traditional asthma care by providing nurse case management, nurse and/or CHW home visits, asthma

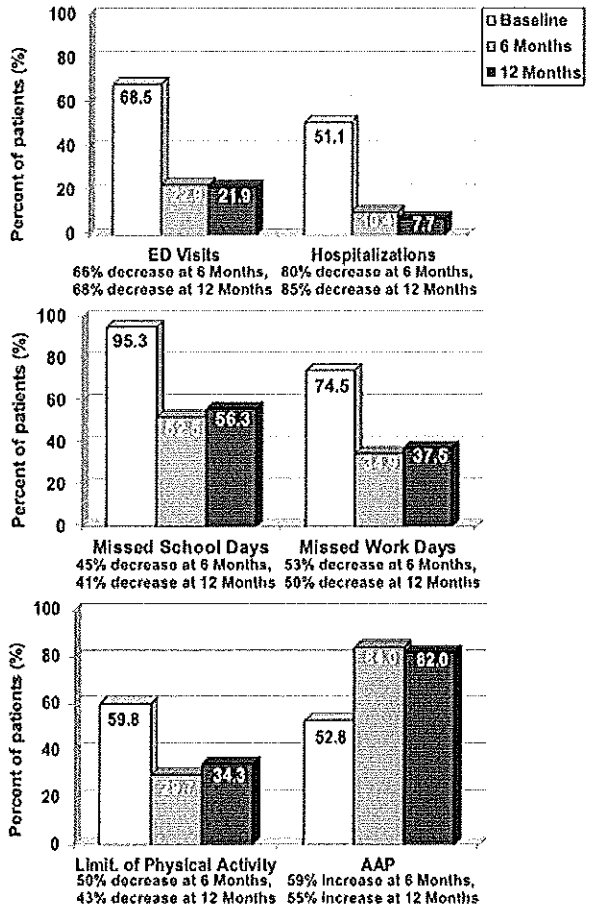


FIGURE 1

Community Asthma Initiative dichotomous outcomes at baseline, 6 months, and 12 months. Percentage of patients who experienced any (≥1 versus none) ED visits, hospitalizations, missed school days, missed work days (parents/caregivers), limitation of physical activity, and AAP for 283 children (all $P \leq .0001$).

education, and environmental assessment and remediation based on the evidence of previous national programs. The evaluation of this QI initiative has demonstrated a significant reduction in asthma ED visits, hospitalizations, limitation of physical activity, missed school, and parent/guardian missed work at 6 and 12 months of follow-up. The improvements in hospital costs were particularly remarkable when compared with the demographically similar diverse, low-income neighborhoods that had not received services during the study period. The continued reduction in cost at 2 years may indicate the ongoing improvement due to reduction in allergens through IPM and continued use of controller medications.

Home visits by CHWs and nurses allow the care system to reach families in their own homes to provide asthma education and can address the obstacles to good asthma control.[5] The CAI patients identified from these low-income communities were primarily black and Hispanic, and the program was developed to reduce health disparities for these populations. A comprehensive treatment plan needs to address social determinants of health, such as exposure to high levels of asthma triggers in the form of pests, mold, and dust found in poor housing and deteriorating schools, and chronic stress due to community violence.[25] The home environmental issues that families face required more aggressive services to reduce common asthma triggers than we originally anticipated. CAI provided all patients with HEPA vacuums and bedding encasements, environmental materials tailored to their needs, and IPM extermination on a case-by-case basis. Culturally sensitive communication about asthma treatment and medications also helped to address the personal beliefs of patients and their families and to identify barriers to adherence. Not surprisingly, nurse home visits were provided to patients with more ED visits and addressed the medication issues in greater detail. Nurse home visits and CHW visits closely supervised by nurses have not been reported in previous published initiatives, and their added value should be investigated further.

The changes in the comparison community over time may have reflected some degree of "regression to the mean," because some patients may not require subsequent ED visits or hospitalizations after 1 episode. Also, the comparison neighborhoods controlled for the impact of community-wide

TABLE 2 The Community Asthma Initiative Continuous Outcomes

	Continuous Outcomes (3 Time Points)			
	Baseline	6 mo	12 mo	*P* (Repeated Measures)
ED visits	1.0	0.3	0.3	<.0001
Hospitalizations	0.5	0.1	0.1	<.0001
Days of limitation of physical activity	2.7	1.2	1.2	<.0001
Missed school days	5.1	3.1	2.4	<.0001
Missed work days	2.1	1.1	1.1	<.0001

Number of events or days at baseline, 6 months, and 12 months using GEE (unadjusted) repeated-measures analyses for continuous outcomes, including number of ED visits, hospitalizations, days of limitation of physical activity, child missed school days and parental missed work days ($N = 283$).

TABLE 3 Logistic Regression Models for Dichotomous Outcomes (≥1 Versus None) Adjusted for Quadratic Term, Age, Gender, Race/Ethnicity, Number of Home Visits, Any Nurse Visits, Income, Insurance, and Asthma Severity

	Days of Limitation of Physical Activity OR (CI)	ED Visits OR (CI)	Admissions OR (CI)	Missed School Days OR (CI)	Missed Work Days OR (CI)	Up-to-date AAP OR (CI)
χ^2 for model	46.2[a]	117.3[a]	92.5[a]	87.05[a]	84.9[a]	76.1[a]
P value	<.0001[a]	<.0001[a]	<.0001[a]	<.0001[a]	<.0001[a]	<.0001[a]
Follow-up	0.09 (0.02–0.40)[a]	0.01 (0.00–0.04)[a]	0.01 (0.00–0.08)[a]	0.00 (0.00–0.01)[a]	0.01 (0.00–0.03)[a]	63.81 (11.85–343.60)[a]
Quadratic term	1.66 (1.16–2.39)[a]	2.68 (1.75–4.11)[a]	2.38 (1.34–4.21)[a]	3.72 (2.43–5.70)[a]	2.98 (1.98–4.49)[a]	0.42 (0.27–0.64)[a]
Age	0.99 (0.94–1.03)	0.95 (0.91–0.99)[a]	0.96 (0.91–1.01)	0.96 (0.91–1.02)	0.89 (0.83–0.95)[a]	1.01 (0.95–1.06)
Male	0.88 (0.59–1.30)	0.82 (0.56–1.19)	1.06 (0.68–1.66)	0.61 (0.39–0.96)[a]	0.79 (0.46–1.34)	1.24 (0.78–1.96)
Hispanic	1.48 (0.53–3.70)	1.45 (0.65–3.24)	2.87 (1.03–7.98)[a]	1.30 (0.51–3.28)	2.81 (0.95–8.29)	0.51 (0.19–1.37)
African American	2.09 (0.84–5.17)	1.06 (0.48–2.36)	3.4 (1.24–9.37)[a]	1.47 (0.58–3.68)	2.48 (0.86–7.16)	0.63 (0.23–1.68)
Number of home visits	1.07 (0.91–1.26)	1.00 (0.86–1.17)	1.06 (0.88–1.28)	1.36 (1.11–1.67)[a]	1.12 (0.86–1.46)	1.33 (1.06–1.68)[a]
Any nurse visit	1.39 (0.93–2.10)	1.59 (1.07–2.35)[a]	1.28 (0.80–2.04)	1.48 (0.93–2.37)	1.30 (0.74–2.28)	1.28 (0.78–2.10)
Low income	1.00 (0.58–1.71)	1.10 (0.64–1.89)	0.70 (0.38–1.29)	0.74 (0.39–1.41)	0.81 (0.43–1.55)	0.70 (0.37–1.32)
Private insurance	0.87 (0.46–1.65)	1.19 (0.65–2.18)	1.41 (0.70–2.82)	0.85 (0.42–1.71)	1.44 (0.68–3.07)	1.12 (0.54–2.30)
Moderate persistent asthma[b]	2.01 (1.32–3.07)[a]	1.17 (0.78–1.75)	0.74 (0.45–1.22)	1.19 (0.74–1.91)	1.00 (0.56–1.79)	3.39(2.01–5.73)[a]
Severe persistent asthma[c]	3.71 (1.73–7.98)[a]	2.56 (1.24–5.28)[a]	2.16 (0.96–4.89)	2.17 (0.83–5.66)	2.59 (0.86–7.79)	4.52 (1.52–13.45)[a]

Quadratic term was added because of the small increase in outcomes at 1 year compared with 6 months (reflecting a similar time of year as enrollment). Continuous variables included age and number of home visits.

[a] Significant results.

[b] Moderate persistent asthma = moderate persistent versus all others (indicator variable).

[c] Severe persistent asthma = severe persistent versus all others (indicator variable).

changes in asthma care, case management, and education. CAI patients were selected to be at greatest need of services by the nurse case manager and therefore showed higher initial cost, but ended up with costs similar to the comparison community at the end of the first year, and even lower costs at the end of the second year, which resulted in significant cost savings. Identification of an ideal comparison group is challenging, and our program was able to compare costs with demographically similar zip code neighborhoods. The use of the nonenrolled population in the same neighborhoods as a comparison group would reflect additional biases, because nonrespondents may have higher risks of poorer outcomes owing to the inability to be contacted and the refusal of enhanced care. Future matching strategies or risk adjustment for patients with initial hospitalizations might help correct the differential baseline cost of the 2 populations.

There were strengths and limitations to this study, because CAI was not a randomized clinical trial. The comparison data were drawn from hospital administrative data, but similar case management information was not available for the comparison group. The retention rate was lower than ideal, but reasonable for a voluntary QI study with no evidence of differential attrition. Additional initiatives may need to be developed to reach the "unreachable" populations not served by the program. The regression analyses indicated that patients with greater functional impairment had nurse visits and more home visits. However, the analyses could not separate out the impact of specific services. Because administrative data were used for the cost evaluation of CAI and for the comparison population, there were no biases due to lack of follow-up for the cost analyses.

Hospital administrative data cannot identify care at other hospitals; however, parental reports contain information across institutions as well as quality-of-life information, but they may lack accuracy. Parent and hospital administrative data were remarkably similar in this study and complemented each other. The cost estimate is conservative, because some of the staff time included in the analyses was used for CHW training and supervision, community meetings and collaborations, program planning, and evaluation in

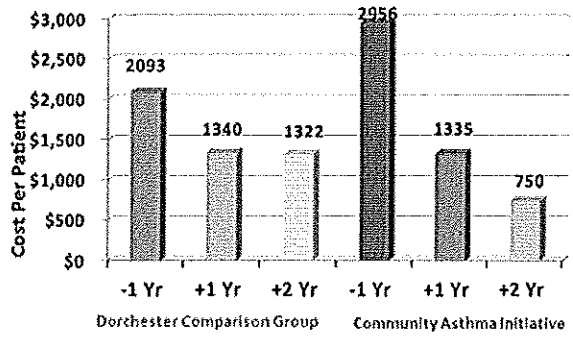


FIGURE 2
Cost of ED visits and hospitalizations for Community Asthma Initiative patients (N = 102) and comparison population (Dorchester Comparison Group) (N = 559) 1 year back (−1 Yr) and 2 years forward (+1 Yr and +2 Yr) (FY2006) (repeated-measures analysis comparing intervention and comparison groups, P < .001).

addition to patient care. The cost analysis did not include physician fees or financial estimates of impact on quality of life, and so the ROI underestimates the true cost savings. Future cost analyses should consider merging program information with insurance company data to include the costs of urgent care visits and medications that may increase when asthma is in better control and patients have more connection to their primary care providers.[2]

These remarkable results provide a model of effective care for high-risk asthma patients with substantial cost savings. The initial ROI calculation of 1.46 exceeds the break-even threshold of 1.0. Case management and home visits combined have helped patients who previously needed a higher level of care to have better control of their asthma. CAI incorporates a culturally sensitive, family-centered approach through home visits and care coordination, and is based in the community, as recommended by the Institute of Medicine's chronic care model.[26,27] Cost-effectiveness calculations support the business case for payers to cover these chronic care services and materials that are not reimbursed in a fee-for-service system.[28] The program has partnered with asthma policy organizations in Massachusetts to develop the "Investing in Best Practices for Asthma: A Business Case"[29] that moved ahead policy changes for care of children with asthma.

CAI provides an effective enhanced-care model that could be included in a bundled or global payment system to reduce the cost of asthma care to society and improve the health and the quality of the life of children living with asthma. The CAI model can be used to respond to the health care reform call for "accountable care organizations" and expansion of care under the medical homes for patients with chronic illnesses.[30] Accountable care organizations are responsible for the quality of care, as measured by standard outcome metrics,[31] and would receive bundled or global payments for care with potential shared savings for providers and payers.[32,33] CAI has started working with Medicaid and other stakeholders to develop and implement a bundled payment pilot.

CONCLUSIONS

CAI was developed to address health disparities for urban low-income children, and the cost-effectiveness of the program has generated information to guide advocacy efforts to finance comprehensive asthma care for children.

ACKNOWLEDGMENTS

We are grateful for the evaluation and quality improvement expertise of Sion Kim Harris, PhD and Gareth Parry, PhD, editorial comments of S. Jean Emans, MD, and early efforts of research assistant Lauren Ebe, BA, as well as the Community Asthma Initiative collaborators and participants.

REFERENCES

1. Akinbami LJ. The State of Childhood Asthma, United States 1980-2005. Hyattsville MD: National Center for Health Statistics, 2006. Advance data from Vital and Health Statistics 381. Available at: www.cdc.gov/nchs/data/ad/ad381.pdf. Accessed May 17, 2009
2. Barnett SBL, Nurmagambetov TA. Costs of asthma in the United States: 2002-2007. *J Allergy Clin Immunol.* 2011;127(1):145–152
3. Centers for Disease Control and Prevention. Vital Signs: asthma prevalence, disease characteristics, and self-management education — United States, 2001-2009. *MMWR Morb Mortal Wkly Rep* . 2011;60(17):547–552
4. The Health of Boston 2008. Boston, MA: Boston Public Health Commission, Research Office; 2008. Available at: www.bphc.org/about/Documents/Health%20of%20Boston%202008.pdf. Accessed May 17, 2009
5. National Heart, Lung, and Blood Institute. *National Asthma Education and Prevention Program, Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma.* Washington, DC: US Department of Health and Human Services; 2007
6. Cloutier MM, Hall CB, Wakefield DB, Bailit H. Use of asthma guidelines by primary care providers to reduce hospitalizations and emergency department visits in poor, minority, urban children. *J Pediatr.* 2005;146(5):591–597
7. Krieger J, Jacobs DE, Ashley PJ, et al. Housing interventions and control of asthma-related indoor biologic agents: a review of the evidence. *J Public Health Manag Pract.* 2010;16(suppl 5):S11–S20
8. Evans R, III, Gergen PJ, Mitchell H, et al. A randomized clinical trial to reduce asthma morbidity among inner-city children: results of the National Cooperative Inner-City Asthma Study. *J Pediatr.* 1999;135(3):332–338
9. Morgan WJ, Crain EF, Gruchalla RS, et al; Inner-City Asthma Study Group. Results of a home-based environmental intervention among urban children with asthma. *N Engl J Med.* 2004;351(11):1068–1080
10. Bryant-Stephens T, Kurian C, Guo R, Zhao H. Impact of a household environmental intervention delivered by lay health workers on asthma symptom control in urban, disadvantaged children with asthma. *Am J Public Health.* 2009;99(suppl 3):S657–S665
11. Krieger J, Takaro TK, Song L, Beaudet N, Edwards K. A randomized controlled trial of asthma self-management support comparing clinic-based nurses and in-home community health workers: the Seattle-King County Healthy Homes II Project. *Arch Pediatr Adolesc Med.* 2009;163(2):141–149
12. La Roche MJ, Koinis-Mitchell D, Gualdron L. A culturally competent asthma management intervention: a randomized controlled pilot study. *Ann Allergy Asthma Immunol.* 2006;96(1):80–85
13. Thyne SM, Rising JP, Legion V, Love MB. The Yes We Can Urban Asthma Partnership: a medical/social model for childhood asthma management. *J Asthma.* 2006;43(9):667–673
14. Flores G, Abreu M, Tomany-Korman S, Meurer J. Keeping children with asthma out of hospitals: parents' and physicians' perspectives on how pediatric asthma hospitalizations can be prevented. *Pediatrics.* 2005;116(4):957–965

15. Center for Disease Control and Prevention. Reducing childhood asthma through community-based service Delivery—New York City, 2001-2004. *MMWR Morb Mortal Wkly Rep.* 2005;54(01):11–14
16. Nicholas SW, Jean-Louis B, Ortiz B, et al. Addressing the childhood asthma crisis in Harlem: the Harlem Children's Zone Asthma Initiative. *Am J Public Health.* 2005;95(2): 245–249
17. Karnick P, Margellos-Anast H, Seals G, Whitman S, Aljadeff G, Johnson D. The pediatric asthma intervention: a comprehensive cost-effective approach to asthma management in a disadvantaged inner-city community. *J Asthma.* 2007;44(1):39–44
18. Clougherty JE, Levy JI, Hynes HP, Spengler JD. A longitudinal analysis of the efficacy of environmental interventions on asthma-related quality of life and symptoms among children in urban public housing. *J Asthma.* 2006;43(5):335–343
19. Bhaumik U, Walker SP, Sommer SJ, et al. Social return on investment from an asthma community-based care management intervention program. American Public Health Association Annual Meeting, Denver, CO, November 8, 2010
20. Lieu TA, Quesenberry CP, Jr, Capra AM, Sorel ME, Martin KE, Mendoza GR. Outpatient management practices associated with reduced risk of pediatric asthma hospitalization and emergency department visits. *Pediatrics.* 1997;100(3 pt 1): 334–341
21. Zemek RL, Bhogal SK, Ducharme FM. Systematic review of randomized controlled trials examining written action plans in children: what is the plan? *Arch Pediatr Adolesc Med.* 2008;162(2):157–163
22. Institute of Medicine. *The Future of the Public's Health in the 21st Century.* Washington, DC: National Academies Press; 2003
23. Dahlgren G, Whitehead M. *Policies and Strategies to Promote Social Equity in Health.* Stockholm, Sweden: Institute for Futures Studies; 1991
24. Israel BA, Eng E, Schulz AJ, Parker EA. In: Israel BA, Eng E, Parker EA, eds. Introduction to methods in community-based participatory research for health. In: *Methods in Community-based Participatory Research for Health.* San Francisco, CA: John Wiley & Sons; 2005:3–29
25. Wright RJ, Mitchell H, Visness CM, et al. Community violence and asthma morbidity: the Inner-City Asthma Study. *Am J Public Health.* 2004;94(4):625–632
26. Institute of Medicine. *Crossing the Quality Chasm: A New Health Care System for the 21st Century.* Washington, DC: National Academy Press; 2001
27. Wessel L, Spain J. The Chronic Care Model: A collaborative approach to preventing and treating asthma in infants and young children. Zero to Three; 2005. Available at: www.zerotothree.org and www.clinicians.org/images/upload/Asthma.pdf. Accessed September 12, 2010
28. Sommer SJ, Queenin LM, Nethersole S, et al. Children's Hospital Boston Community Asthma Initiative: Using Outcomes to Drive Policy Change. *Prog Community Health Partnersh.* 2011;5(3):327–335
29. Hoppin P, Jacobs M, Stillman L. Investing in best practices for asthma: a business case – August 2010 update. New England Asthma Regional Council. 2010. Available at: www.asthmaregionalcouncil.org. Accessed September 19, 2010
30. Rittenhouse DR, Shortell SM, Fisher ES. Primary care and accountable care—two essential elements of delivery-system reform. *N Engl J Med.* 2009;361(24):2301–2303
31. Porter ME. A strategy for health care reform —toward a value-based system. *N Engl J Med.* 2009;361(2):109–112
32. Commission on a High Performance Health System. Developing innovative payment approaches: finding the path to high performance. New York, NY: The Commonwealth Fund; 2010. Available at: www.commonwealthfund.org/□/media/Files/Publications/Issue%20Brief/2010/Jun/1401_Guterman_developing_innovative_payment_approaches_ib.pdf. Accessed November 7, 2010
33. Massachusetts Executive Office for Administration and Finance. Recommendations of the Special commission on the Health Care Payment System, July 16, 2009. Available at: www.mass.gov/eohhs/docs/dhcfp/pc/final-report/final-report.pdf. Accessed January 24, 2012

(Continued from first page)

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

FINANCIAL DISCLOSURE: *The authors have indicated that they have no financial relationships relevant to this article to disclose.*

FUNDING: Supported in part by Centers for Disease Control and Prevention REACH US Cooperative Agreement grant 1U58DP001055, Healthy Tomorrows Partnership for Children Program H17MC06705 from Health Resources and Services Administration, Leadership Education in Adolescent Health training grant T71MC00009 from the Maternal and Child Health Bureau, Health Resources and Services Administration, and the Ludcke and Thoracic Foundations.

Community Asthma Initiative: Evaluation of a Quality Improvement Program for Comprehensive Asthma Care

Elizabeth R. Woods, Urmi Bhaumik, Susan J. Sommer, Sonja I. Ziniel, Alaina J. Kessler, Elaine Chan, Ronald B. Wilkinson, Maria N. Sesma, Amy B. Burack, Elizabeth M. Klements, Lisa M. Queenin, Deborah U. Dickerson and Shari Nethersole

Pediatrics; originally published online February 20, 2012;
DOI: 10.1542/peds.2010-3472

Attachment 2:

"Risk Adjustment for Pediatric Populations"

Risk adjustment for pediatric populations

Howard Kahn, FSA, MAAA
Rob Parke, FIA, ASA, MAAA
Rong Yi, PhD

The use of risk adjustment in provider reimbursement arrangements has increased as alternative payment arrangements are becoming more widespread in health insurance. Risk adjustment has been used by Medicare Advantage and managed Medicaid programs to reimburse health plans for the unique risks and populations in their care. More recently, as carriers have transferred utilization risk to providers through alternative payment arrangements such as global budgets and bundled payments, risk adjustment has been used to reflect a provider's patient's severity. Also, under the Patient Protection and Affordable Care Act (ACA), beginning in 2014 risk adjustment will be used to transfer payments among all fully insured individual and small group plans.

However, many commercial risk adjustment methodologies applied were developed using a standard population representing a combination of adults and children. Adults comprise a larger proportion of the average population, and as a consequence, the disease states recognized in these methodologies were optimized with greater emphasis on adults. Because a chosen risk adjustment methodology should reflect the characteristics of the underlying patient population, organizations such as children's hospitals, pediatric provider groups, and health plans that enroll a large proportion of children have begun to question these standard risk adjustment models. These groups argue that there are fundamental differences in clinical profiles, patient mix, treatment options, and patient management needs between the pediatric population and the general population.

Risk adjustment is commonly used to recognize the relationship between the cost of delivering care and patient severity in order to redistribute the total amount of available funds among groups by relative risk scores.

In this paper, we compare results from a model we optimized for a pediatric population with a control model we developed for a standard population. This model is similar to many commercially available versions developed from the open source hierarchical condition categories (HCC) system.[1]

Risk adjustment modeling background

We begin by describing the construction of a typical HCC risk adjustment model.

A general risk adjustment formula for n defined conditions is represented as:

$$Y = I + C_1X_1 + C_2X_2 + \ldots C_iX_i + \ldots + C_nX_n$$

where Y = risk-adjusted expected claims cost (or risk score) for member x

I = intercept equivalent to the minimum cost (or risk score) assigned to a member

C_i = coefficient (risk weight) for the i^{th} clinical classification

X_i = member's value for the i^{th} clinical classification, such as asthma, diabetes, COPD, etc.

The first step in creating a risk adjustment model is to determine the number and definitions of the clinical classifications required. These classifications can represent any driver of healthcare cost found in claims data, but typically represent a collection of diagnoses and member demographics.[2] Careful consideration should be given to creating the clinical classifications, as they must have clinical face validity, not be so specific that they lose statistical credibility, be robust to coding pattern differences, and accurately predict average costs for all members in the population.

1 The HCCs are used in Medicare Advantage and Part D plans, in the federally administered risk-adjustment model for commercial individual and small groups starting in 2014, and in several states' Medicaid and subsidized insurance programs. The HCCs used in all of these systems have not been calibrated for a pediatric population.

2 There are also pharmacy-based risk adjustment models, which are typically used when the quality of medical diagnosis coding is questionable, e.g., due to capitation.

For example, a very simple risk adjustment formula might consist of only two classifications: age and diabetes. The formula is then represented as:

$$Y = I + C_{diabetes}X_{diabetes} + \sum_{i \in A_{ge}} C_iX_i$$

where $X_{diabetes}$ = 1 if the member has diabetes, 0 if the member does not (as defined by a set of ICD-9 diagnosis codes)

X_i = 1 if the member's age equals i, 0 otherwise

Statistical techniques are then used to estimate the value of the classification coefficients (risk weights). The resulting formula for a 40-year-old could be:

$$Y = \$100 + \$50X_{40} + \$2{,}000X_{diabetes}$$

This risk adjustment formula would then predict that a 40-year-old member who has diabetes would cost $2,150 for the year. However, a 40-year-old member who does not have diabetes is predicted to incur only $150 for the year, because that is the value of the age component (CageXage) added to the intercept ($100) in the formula. The intercept and risk weights (the resulting costs for each of the classifications, 50 and 2,000 in this example) are important aspects of the risk adjustment model, along with the definitions of the classifications. For instance, a refinement to the above simple model could be that diabetes is split into Type I and Type II diabetes, with and without complications.

Existing risk adjustment methodologies have been developed and used on populations that include a mix of adults and children. Does this type of methodology accurately capture the different characteristics of a pediatric-only population?

Is there a better alternative?

Adjustments to the general risk adjustment model

Before applying an existing risk adjustment model to a specific population, special consideration must be taken to ensure the model is a good fit. There are many reasons why a risk adjustment model would need to be adjusted:

1. **Unique population** – nuances about the population included in the claims data used to develop the risk weights.
2. **Unique contract** – the claims data does not represent the total cost of care but rather a component of the total (e.g., mental health and behavioral health carve-outs).
3. **Secular changes** – the risk weights were developed using data from a few years back and need to be updated to reflect current practice and treatment patterns.
4. **Coding convention changes** – starting in October 2014, diagnosis coding will be converted to ICD-10-CM. Both the classifications and the risk weights will need to be revised and updated. Classifications need to be ICD-10 ready before the official conversion date. Risk weights recalibrated on ICD-10 claims data will need to wait until an adequate volume of claims is available.

In this paper, we explore the consequences of item 1 above by measuring the effectiveness of a standard risk adjustment model on pediatric-only populations. To this end, we built a control model for a standard commercial population from the Truven Health Analytics MarketScan[®][3] database. We limited our focus to New England States[4] and developed a concurrent[5] risk adjustment model with 184 disease classifications based on the HCC system. We note that this model is not a Milliman Advanced Risk Adjusters™ (MARA) risk adjustment model, but instead a control model for this specific analysis.[6]

The R-squared value[7] in our control model is 58% on the standard population. This is very similar to the reported R-squared values for many commercially available concurrent risk adjusters.[8] However, if we remove the adults from this population, our model's R-squared reduces significantly to 45% because the model's disease classifications and coefficients were optimized for a population that includes both adults and children.

3 Truven Health Analytics MarketScan® is a large and nationally representative commercial claims database. It is used to develop risk adjustment tools by many vendors of commercial risk adjustment tools.
4 We only used claims in New England states–Maine, Massachusetts, Connecticut, New Hampshire, Rhode Island, and Vermont–for model development.
5 A concurrent model uses the current year's data to risk adjust total cost of care within the year. We chose to develop a concurrent model because many recent global risk contracts retrospectively use risk adjustment at settlement.
6 For more information, go to http://us.milliman.com/Solutions/Products/Milliman-Advanced-Risk-Adjusters.
7 The R-squared statistic measures the amount of variability a model is capable of explaining in a population and is often used to evaluate the effectiveness of a risk adjustment model. A more accurate model results in a higher R-squared value.
8 See TABLE IV.7 of the 2007 SOA risk adjuster comparison study: www.soa.org/files/research/projects/risk-assessmentc.pdf.

Pediatric risk adjustment model

To improve the control model's R-squared of 45% for pediatric-only populations, we developed a pediatric-only model. We achieved this result through an iterative process using only the pediatric population included in our MarketScan database sample.

The detailed work flow of the model development process is summarized below:

1. We began the modeling at the DxGroup level that underlies the HCCs. There are 784 DxGroups in the original HCC classification system.

2. We modeled DxGroups with more than 30 patients separately and left those with fewer than 30 patients in their original HCCs.

3. We created two-way and three-way disease interactions for inclusion in the model (e.g., diabetes and chronic obstructive pulmonary disorder [COPD] would be included as an additional explanatory variable, in addition to diabetes alone and COPD alone). We calculated the sample size of each and retained only those that had at least 30 patients in a cell.

4. We regrouped DxGroups and disease interaction terms with statistically insignificant coefficients (at a 5% significance threshold) with the other small-cell DxGroups in the same HCC and recalculated their coefficients (risk weights).

5. We reset the coefficients of DxGroups and disease interaction terms with statistically significant but negative coefficients to zero. Negative coefficients often imply a confounding variable; if left in the model, they will produce spurious relationships among conditions. From a payment perspective, negative coefficients result in reduction in payment for diagnosing or treating a condition, which does not have face validity either.

6. We repeated steps (4) and (5) until all variables left in the model had statistically significant and non-negative coefficients. This resulted in 570 DxGroups/HCC categories.

By way of an example, the control model has a category called "other infectious diseases." Using the control model, we would only have one risk weight associated for all diseases falling under this category. However, in the pediatric model we refined this classification by splitting out "other bacterial infections," "bacterial infection in other diseases," "other viral infections," "other infections," "Lyme disease," and "bacteremia." Table 1 summarizes the risk weights for the general HCC category "other infectious diseases" and compares it to the pediatric model calibration:

TABLE 1

	RISK WEIGHT	
	PEDIATRIC MODEL	CONTROL MODEL
OTHER BACTERIAL INFECTIONS	$4,045	$5,410
BACTERIAL INFECTION IN OTHER DISEASES	$2,207	$5,410
OTHER VIRAL INFECTIONS	$93	$5,410
OTHER INFECTIONS	$392	$5,410
LYME DISEASE	$355	$5,410
BACTEREMIA	$13,126	$5,410

In addition, illnesses that are more important in a pediatric population, such as developmental disability, were refined in our model. Table 2 below shows the risk weights for the HCC category "other developmental disability" used in the control model and compares it to the pediatric model calibration.

TABLE 2

	RISK WEIGHT	
	PEDIATRIC MODEL	CONTROL MODEL
EMOTIONAL DISORDERS OF CHILDHOOD/ADOLESCENCE	$931	$830
LEARNING/DEVELOPMENT DISORDERS	$1,061	$830
UNSPECIFIED CHROMOSOMAL ANOMALIES AND CONGENITAL MALFORMATION SYNDROMES	$4,119	$830
SEX CHROMOSOME ABNORMALITIES (E.G., KLINEFELTER'S/TURNER SYNDROMES)	$7,550	$830

Howard Kahn, Rob Parke, Rong Yi

Milliman

Results

This pediatric risk adjustment model has an R-squared of 58% on pediatric populations, which is a significant improvement from the control model's R-squared of 45%.

This increase in statistical fit will affect the financial results of organizations bearing financial risk for pediatric populations. For example, using the pediatric-only model on children in the data used to develop our model results in a risk score that is approximately 1.5% higher than the control model developed for a standard population.

Other considerations

The pediatric risk model we developed is intended for a commercially insured pediatric population and was designed to risk adjust total cost of care. As with any risk model, further fine-tuning to better reflect the business needs and the characteristics of a population under consideration is required. For example, in risk-based contracts where a subset of services is carved out, such as neonatal intensive care, the model may also need to be recalibrated to better reflect the scope of the global payment arrangement.

Conclusion

These results show that a risk model calibrated for a standard population has significantly lower predictive power if it is applied to a pediatric-only population. In alternative payment models that use risk adjustment to distribute payments to providers, this could also result in inequitable reimbursement to providers specializing in pediatric populations. As a result, providers specializing in serving pediatric populations should carefully review the risk models used in any alternative payment arrangement before participation.

Guidelines issued by the American Academy of Actuaries require actuaries to include their professional qualifications in all actuarial communications. Rob Parke and Howard Kahn are members of the American Academy of Actuaries, and meet the qualification standards for performing the analyses in this report.

Howard Kahn is a consulting actuary with Milliman's New York office. Contact him at howard.kahn@milliman.com.

Rob Parke is a principal and consulting actuary with Milliman's New York office. Contact him at rob.parke@milliman.com.

Rong Yi, is a senior consultant with Milliman's New York office. Contact her at rong.yi@milliman.com.

Attachment 3:

“Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors”

Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors

TECHNICAL REPORT

August 15, 2014

This report is funded by the Department of Health and Human Services under contract HHSM-500-2012-00009I task order 7.

NATIONAL QUALITY FORUM

Contents

Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors

TECHNICAL REPORT

Foreword

Convened by the National Quality Forum (NQF), the Expert Panel tackled the challenging issue of whether or not to adjust performance measures for socioeconomic status and other demographic factors, including income, education, primary language, health literacy, race and other factors. The Panel's report has implications for NQF policy.

NQF recommends that outcome measures be adjusted for clinical severity because it affects outcomes, but up until now we have not recommended adjustment for sociodemographic factors, in part because of their link to disparities. The question of whether to adjust measures for sociodemographic factors is being called now because there is a growing body of evidence that sociodemographic factors also affect patient outcomes. These outcome measures, increasingly used in accountability programs such as public reporting and pay-for-performance, are under more intense scrutiny. Getting the measures "right" is important given that they are being used to determine which providers to include in networks, how to determine financial rewards or penalties, where to go for healthcare services, and where to focus improvement efforts.

Whether to adjust measures for sociodemographic factors is of great interest to stakeholders who have passionate views and legitimate concerns on all sides of this issue – NQF received more public comments on this topic than any other project to date. At the heart of it though, people want performance measures to provide fair comparisons across those being measured, but also agree that we cannot lose sight of disparities in healthcare and health faced by disadvantaged patients or ignore the challenges of the providers and health plans that care for them.

In this report, the Expert Panel recommends that for comparative performance assessment sociodemographic adjustment is appropriate if certain conditions are met, and further that if a measure is adjusted for sociodemographic factors it must be specified for stratification so that any disparities are made visible. The report lays out the conceptual and methodological basis for this and other recommendations. The Panel also made specific recommendations for operationalizing potential sociodemographic adjustment, including guidelines for selecting risk factors and the kind of information to submit for measure review. Finally, the Panel recommended that NQF appoint a standing Disparities Committee.

In its deliberations on the report's policy implications, the Consensus Standards Approval Committee recommended, and the NQF Board of Directors approved, a trial period during which the NQF restriction against sociodemographic adjustment will be lifted; they approved a new standing Disparities Committee. The Board emphasized the need for a time-limited, robust trial period and strongly urged the field to develop and use sociodemographic-adjusted measures so that the data necessary to inform NQF's permanent policy in this area is generated. The procedures, guidance, and timeline for the trial period will be developed over the coming months and posted on the NQF web site.

Reaching consensus on a path forward for this consequential and controversial issue involving measurement science and strongly held views about adjustment was challenging, and I believe no other organization could have accomplished it. Please join me in thanking the members of the Expert Panel, NQF Members and others who submitted comments, the Consensus Standards Approval Committee, the Board, and the NQF staff for their contributions to helping improve measurement of quality.

Best,

Christine Cassel

President and CEO

Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors

TECHNICAL REPORT

Executive Summary

Introduction

There is a large body of evidence that various sociodemographic factors influence outcomes, and thus influence results on outcome performance measures. ***Sociodemographic Status (SDS)*** refers to a variety of socioeconomic (e.g., income, education, occupation) and demographic factors (e.g., age, race, ethnicity, primary language). There also is a large body of evidence that there are disparities in health and healthcare related to some sociodemographic factors. Given the evidence, the overarching question addressed in this project is, "What, if anything, should be done about sociodemographic factors in relation to outcome performance measurement?"

NQF endorses performance measures that are intended for use in both performance improvement and accountability applications such as public reporting and pay-for-performance. In this context, the overall performance measure score is used to make a conclusion about a healthcare unit's (a unit refers to an hospital, health plan, practice or other unit that is being assessed) quality in relation to other units or some other comparator such as average performance. The general question being addressed is: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?** That is, the measure scores are used to inform decisions of individuals seeking care; purchasers paying for care, including bonuses or penalties; or networks contracting for healthcare service. Such comparisons should be affected as little as possible by factors other than quality of care, such as patient characteristics already present at the start of care.

Because healthcare outcomes are a function of patient attributes (including SDS) as well as the care received, and patients are not randomly assigned to units for healthcare services so that all units have the same mix of patients, risk adjustment is essential to ensuring an "apples to apples" comparison when examining outcome performance in real-world settings. ***Risk adjustment (also known as case-mix adjustment)*** refers to statistical methods to control or account for patient-related factors when computing performance measure scores; methods include multivariable modeling, indirect standardization, or direct standardization. Risk adjusting outcome performance measures to account for differences in patient health status and clinical factors (e.g., comorbidities, severity of illness) that are present at the start of care is widely accepted. This report explores also adjusting performance measures for sociodemographic status (SDS) when appropriate.

Core Principles

The Expert Panel on Risk Adjustment for Sociodemographic Factors agreed on a set of core principles to ground its recommendations.

1. Outcome performance measurement is critical to the aims of the National Quality Strategy.
2. Performance measurement and risk adjustment must be based on sound measurement science.

3. Disparities in health and healthcare should be identified and reduced.
4. Performance measurement should not lead to increased disparities in health and healthcare.
5. Outcomes may be influenced by patient health status, clinical, and sociodemographic factors, in addition to the quality and effectiveness of healthcare services, treatments, and interventions.
6. When used in accountability applications, performance measures that are influenced by factors other than the care received, particularly outcomes, need to be adjusted for relevant differences in patient case mix to avoid incorrect inferences about performance.
7. Risk adjustment may be constrained by data limitations and data collection burden.
8. The methods, factors, and rationale for risk adjustment should be transparent.

Recommendations

The Expert Panel made ten recommendations. The recommendations may apply to outcome performance measures (including resource use and patient-reported outcomes) and some process performance measures. However, ***each performance measure must be assessed individually to determine appropriateness of SDS adjustment***. The recommendations may apply to any level of analysis including health plans, facilities, individual clinicians, accountable care organizations, etc.

Although the recommendations to adjust for sociodemographic factors when indicated are grounded in sound measurement science methods and principles, the Expert Panel addressed concerns raised in the public comment period about appropriateness of adjusting for SDS in three substantial ways:

- requiring measure specifications for stratification to identify disparities if a performance measure is SDS-adjusted;
- recommending a transition period during which a clinically-adjusted version of the performance measure would be specified and available only for comparison purposes to the SDS-adjusted score; and
- recommending an NQF standing disparities committee to monitor implementation of the revised policy as well as ensure continuing attention to disparities.

Recommendations Related to NQF Criteria and Processes Related to SDS Adjustment

Recommendation 1: When there is a conceptual relationship (i.e., logical rationale or theory) between sociodemographic factors and outcomes or processes of care and empirical evidence (e.g., statistical analysis) that sociodemographic factors affect an outcome or process of care reflected in a performance measure:

- those sociodemographic factors should be included in risk adjustment of the performance score (using accepted guidelines for selecting risk factors) unless there are conceptual reasons or empirical evidence indicating that adjustment is unnecessary or inappropriate;

AND

- the performance measure specifications must also include specifications for stratification of a clinically-adjusted version of the measure based on the sociodemographic factors used in risk adjustment.

Recommendation 2: NQF should define a transition period for implementation of the recommendations related to sociodemographic adjustment. During the transition period, if a performance measure is

adjusted for sociodemographic status, then it also will include specifications for a clinically-adjusted version of the measure only for purposes of comparison to the SDS-adjusted measure.

Recommendation 3: A new NQF standing committee focused on disparities should be established. A standing disparities committee would review implementation of the revised policy about sociodemographic adjustment as recommended in this report (including key decisions by developers and purchasers) and monitor for any unintended consequences of the revised policy.

Recommendation 4: The NQF criteria for endorsing performance measures used in **accountability applications** (e.g., public reporting, pay-for-performance) should be revised as follows to indicate that patient factors for risk adjustment include both clinical and sociodemographic factors:

> ***2b4.*** *For outcome measures and other measures when indicated (e.g., resource use, some process):*
> *an evidence-based risk-adjustment strategy ~~(e.g., risk models, risk stratification)~~ is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome ~~(but not factors related to disparities in care or the quality of care)~~ and are present at start of care;[14,~~15~~] and has demonstrated adequate discrimination and calibration* ***OR*** *rationale/data support no risk adjustment~~/ stratification~~.*
> *14. Risk factors that influence outcomes should not be specified as exclusions.*
> ~~*15. Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.*~~

Recommendation 5: The same guidelines for selecting clinical and health status risk factors for adjustment of performance measures may be applied to sociodemographic factors, and include the following:

- Clinical/conceptual relationship with the outcome of interest
- Empirical association with the outcome of interest
- Variation in prevalence of the factor across the measured entities
- Present at the start of care
- Is not an indicator or characteristic of the care provided (e.g., treatments, expertise of staff)
- Resistant to manipulation or gaming
- Accurate data that can be reliably and feasibly captured
- Contribution of unique variation in the outcome (i.e., not redundant)
- Potentially, improvement of the risk model (e.g., risk model metrics of discrimination, calibration)
- Potentially, face validity and acceptability

Recommendation 6: When there is a conceptual relationship and evidence that sociodemographic factors affect an outcome or process of care reflected in a performance measure submitted to NQF for endorsement, the following information should be included in the submission:

- A detailed discussion of the rationale and decisions for selecting or not selecting sociodemographic risk factors and methods of adjustment (including a conceptual description of

relationship to the outcome or process; empirical analyses; and limitations of available sociodemographic data and/or potential proxy data) should be submitted to demonstrate that adjustment incorporates relevant sociodemographic factors unless there are conceptual reasons or empirical evidence indicating that adjustment is unnecessary or inappropriate.
- In addition to identifying current and planned use of the performance measure, a discussion of the limitations and risks for misuse of the specified performance measure.

Recommendations Relevant to NQF Policy

Recommendation 7: NQF should consider expanding its role to include guidance on implementation of performance measures. Possibilities to explore include:

- guidance for each measure as part of the endorsement process;
- guidance for different accountability applications (e.g., use in pay-for-performance versus pay-for-improvement; innovative approaches to quality measurement explicitly designed to reduce disparities).

Recommendation 8: NQF should make explicit the existing policy that endorsement of a performance measure is for a specific context as specified and tested for a specific patient population (e.g., diagnosis, age), data source (e.g., claims, chart abstraction), care setting (e.g., hospital, ambulatory care), and level of analysis (e.g., health plan, facility, individual clinician). Endorsement should not be extended to expanded specifications without review and usually additional testing.

Recommendations about Broader Related Policy Issues

Recommendation 9: When performance measures are used for accountability applications such as public reporting and pay-for-performance, then purchasers, policymakers and other users of performance measures should assess the potential impact on disadvantaged patient populations and the providers/health plans serving them to identify unintended consequences and to ensure alignment with program and policy goals. Additional actions such as creating peer groups for comparison purposes could be applied.

Recommendation 10: NQF and others such as CMS, Office of the National Coordinator (ONC) for Health Information Technology, and the Agency for Healthcare Research and Quality (AHRQ) should develop strategies to identify a standard set of sociodemographic variables (patient and community-level) to be collected and made available for performance measurement and identifying disparities.

Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors

TECHNICAL REPORT

Section 1: Introduction

NQF endorses performance measures that are suitable for both performance improvement and "accountability applications" (e.g., pay-for-performance, public reporting), when those measures meet a standard set of criteria. Measures of outcomes of care are among those endorsed by NQF. Clinical health outcomes (e.g., survival, improvement or maintenance of function, relief of pain or distressing symptoms) are considered important for performance measurement because they often are the reasons for seeking and providing healthcare and can reflect the quality of care received. Other outcomes for which measures may be endorsed include cost or resource use, referred to broadly as economic outcomes.

Because outcomes can be influenced by many factors other than the healthcare services and interventions received, the current NQF criteria include risk adjustment or stratification for outcome performance measures on the basis of clinical factors like comorbidity or severity of illness. In general, more severe or more complex disease in a cohort of patients, all else being equal, is associated with poorer outcomes. Risk adjustment is designed to improve the ability to make comparative conclusions about quality. Avoiding incorrect conclusions or inferences about quality is important to consumers/patients and purchasers in making informed decisions about where to obtain care; to payers, health plans, and providers regarding rewards/penalties; and to providers and plans in terms of reputation and the ability to improve care for the various subpopulations that they serve.

Current NQF criteria for performance measures direct that some sociodemographic factors, for which disparities in quality of care have been documented in the past, such as socioeconomic status (SES) and race, should not be included in statistical risk models; the related current NQF guidance (provided in a footnote) indicates that stratification is the preferred approach for these factors. The main reason for this current position on sociodemographic factors was a concern that adjustment for variables like income, education, or English proficiency would "mask disparities," and essentially allow or create lower standards of performance for "disadvantaged"[a] populations. The current criterion and concern are examined in this report.

Risk adjusting outcome performance measures to account for differences in patient health status and clinical factors (e.g., comorbidities, severity of illness) that are present at the start of care is widely accepted. This report explores also adjusting performance measures for sociodemographic status (SDS) when appropriate. See Box 1 for examples of clinical and sociodemographic factors that affect complexity of condition, which can influence patient outcomes.

[a] In this report, "disadvantaged" is used to refer to social, economic, and/or environmental disadvantage. It could be related to a variety of sociodemographic factors such as income, race, and education.

Box 1. Clinical and Sociodemographic Complexity

Clinically Complex Patient	Sociodemographically Complex Patient
• Multiple Chronic Conditions • Severe Primary Condition (e.g., severe heart failure, metastatic cancer, end-stage renal disease) • Concurrent mental and physical health problems • Disease affects multiple organ systems • Disease causes significant functional deficit or disability • Condition requires treatment by multiple providers and/or specialized sites of care	• Poverty – Low income and/or no liquid assets • Low levels of formal education, literacy, or health literacy • Limited English proficiency • Minimal or no social support –not married, living alone, no help available for essential health-related tasks • Poor living conditions – homeless, no heat or air conditioning in home or apartment, unsanitary home environment, high risk of crime • No community resources – social support programs, public transportation, retail outlets

NQF also endorses process performance measures, which typically are not adjusted for clinical status or SDS. SDS adjustment of process performance measures also is addressed in this report.

Reason to Re-Examine the NQF Policy

The increased use of NQF-endorsed performance measures beyond public reporting and quality improvement to other accountability applications, such as payment rewards and penalties, has brought increased scrutiny to performance measures. The validity and fairness of some performance measures that do not account for patients' sociodemographic complexity when used to make comparative conclusions have been questioned. Consequently, reaching consensus on NQF endorsement of outcome performance measures for use in accountability applications has become increasingly controversial over the issue of adjusting outcome performance measures for SES or other sociodemographic factors. Recent examples are NQF #1789: Hospital-wide all-cause unplanned readmission (See the Readmissions Project, section titled Candidate Consensus Standards Review) and NQF #2158-Medicare Spending per Beneficiary Measure (MSBP) (See Cost and Resource Use Phase 1, section titled Pre-Meeting Member Comment, Phase 1).

The impact of sociodemographic factors on health and healthcare has been well documented.[14] In fact, most epidemiological and health services research studies that focus on quality commonly adjust for patient SES or other demographic factors. In contrast, SDS adjustment of quality measures has typically been avoided. There are at least two divergent views regarding adjustment for sociodemographic factors:

1) Adjusting performance measures for sociodemographic factors is essential to making fair comparative conclusions about quality and is important to consumers/patients, payers, and others making decisions about choice of providers or health plans or assigning rewards or penalties. Disadvantaged patients confront varying barriers, often lifelong, to health and healthcare, and failing to account for the sociodemographic factors when indicated creates an uneven playing field for performance measurement. For example, Satin[5] states "Asking clinics and physicians who work primarily with poor patient populations to achieve the same results as those working with wealthier populations is effectively asking for more, and in some cases,

impossibly more from these providers/plans. The results of such unrealistic demands may be fewer and fewer providers/plans willing to serve the already underserved."

2) Adjusting performance measures for sociodemographic factors should not be done because it obscures disparities and implies that differences in outcomes based on SDS are expected and accepted. For example, Iezzoni[6, p. 21] states: "For some purposes, ethical concerns raise questions about whether and how to risk-adjust. Such situations arise when persons with certain attributes (e.g., gender, race, SES) that might be potential risk factors for a given outcome simultaneously face the likelihood of receiving substandard care because of those attributes."

Interestingly, both of these positions are based in part on a shared concern about entrenching or worsening disparities in health or healthcare. In the first view, if performance measurement fails to recognize sociodemographic complexity, then it may create a disincentive for healthcare providers and health plans to serve disadvantaged patients, decreasing access to healthcare. In the second view, if performance measurement adjusts for sociodemographic factors, then it may create a disincentive for healthcare providers and plans to improve care to disadvantaged patients.

The issues and concerns about the potential unintended consequence of adjusting or not adjusting for sociodemographic factors on disparities for "disadvantaged" patient populations are addressed in more detail later. However, it is important to note that any recommendations about risk adjusting performance measures must be grounded in sound measurement science, which also is addressed in this report.

Terminology and Key Definitions

In this report, the following key terms are used.

- **Unit** will be used to signify the entity whose performance is being measured, which could be a hospital, health plan, clinician, etc. Performance measurement (and sociodemographic adjustment) can be applied to any setting and level of analysis.
- **Clinical adjustment** refers to adjustment for only clinical variables.
- **Sociodemographic or SDS adjustment** refers to adjustment for both clinical and sociodemographic variables.

The key concepts used in this report are defined as follows and also included in the glossary in Appendix B.

- **Confounding** refers to the distortion in the degree of association between an exposure (independent variable) and an outcome (dependent variable) due to a mixing of effects between the exposure and an incidental (confounding) factor. Confounding represents systematic error and threatens the internal validity of an epidemiologic study since it can lead to false conclusions regarding the true relationship between an exposure and outcome. (See the basics of confounding in Appendix D.)

- **Risk adjustment (also known as case-mix adjustment)** refers to statistical methods to control or account for patient-related factors when computing performance measure scores; methods include multivariable modeling, indirect standardization, or direct standardization. These

methods can be used to produce a ratio of observed-to-expected, a risk-adjusted rate, or other estimate of performance. (See the basics of risk adjustment in Appendix C.)

- **Stratification** refers to computing performance scores separately for different strata or groupings of patients based on some characteristics(s)—i.e., each healthcare unit has multiple performance scores (one for each stratum) rather than one overall performance score.

- **Peer groups for comparison** refers to creating peer groups of healthcare units caring for a similar mix of patients, within which to examine performance scores.

- **Sociodemographic Status (SDS)** refers to a variety of socioeconomic (e.g., income, education, occupation) and demographic factors (e.g., race, ethnicity, primary language).

- **Outcome** – the result of providing healthcare. The term, outcome, will be used to broadly include the following types of outcomes relevant to performance measurement: **quality outcomes** of health outcome (e.g., mortality), intermediate clinical outcome (e.g., BP < 140/90), patient-reported outcome (e.g., depression), and **economic outcomes** of cost and resource use.

Project Purpose, Scope, Approach

There is a large body of evidence that various sociodemographic factors influence outcomes, and thus influence results on outcome performance measures. There also is a large body of evidence that there are disparities in health and healthcare related to some of those sociodemographic factors. Given the evidence, the overarching question addressed in this project is "What, if anything, should be done about sociodemographic factors in relation to outcome performance measurement?"

The purpose of this project was to:

- Identify and examine the issues related to risk adjusting outcome performance measures for SDS (i.e., SES and/or other sociodemographic factors).
- Make recommendations regarding if, when, for what, and how outcome performance measures should be adjusted for SES or other sociodemographic factors.
- Make recommendations for NQF's endorsement criteria for outcome performance measures.

During the project, the Expert Panel identified that process performance measures also may need adjustment.

This project did not include recommendations for:

- specific performance measures;
- adjustment for determining payment for services provided, such as capitated payments;
- use of particular risk adjustment or statistical procedures; or
- structuring performance reward/penalty programs such as pay-for-performance.

A multistakeholder Expert Panel (Appendix A) with a variety of experiences related to outcome performance measurement and disparities reviewed the issues and made recommendations regarding the use of SES and other sociodemographic variables for adjusting outcome performance measures. The

Expert Panel's draft recommendations were presented for public comment. This report and the recommendations reflect the Expert Panel's modifications in response to public comments.

Core Principles

The Expert Panel agreed on a set of core principles to ground its recommendations. The principles were not intended to imply a particular direction for recommendations related to risk adjustment for SES and sociodemographic factors; rather, they represented a baseline of agreement on the key issues that must be considered in making recommendations.

1. Outcome performance measurement is critical to the aims of the National Quality Strategy.
2. Performance measurement and risk adjustment must be based on sound measurement science.
3. Disparities in health and healthcare should be identified and reduced.
4. Performance measurement should not lead to increased disparities in health and healthcare.
5. Outcomes may be influenced by patient health status, clinical, and sociodemographic factors, in addition to the quality and effectiveness of healthcare services, treatments, and interventions.
6. When used in accountability applications, performance measures that are influenced by factors other than the care received, particularly outcomes, need to be adjusted for relevant differences in patient case mix to avoid incorrect inferences about performance.
7. Risk adjustment may be constrained by data limitations and data collection burden.
8. The methods, factors, and rationale for risk adjustment should be transparent.

Section 2: Recommendations

The Expert Panel made the following ten recommendations. A brief rationale accompanies each recommendation in this section. However, an in-depth discussion of the methodological basis and other considerations that led the Panel to these recommendations is in the following sections.

Although the draft recommendations were supported by the great majority of the Expert Panel and the NQF member and public commenters, the purchaser stakeholders and some, but not all, of the consumer stakeholders remained concerned about the appropriateness of adjusting for SDS. The Expert Panel carefully considered these ongoing concerns and modified their draft recommendations in three substantial ways:

- requiring measure specifications for stratification to identify disparities if a performance measure is SDS-adjusted;
- recommending a transition period during which a clinically-adjusted version of the performance measure would be specified and available only for comparison purposes to the SDS-adjusted score; and
- recommending an NQF standing disparities committee to monitor implementation of the revised policy as well as ensure continuing attention to disparities.

In addition, the Expert Panel provided a more detailed methodological discussion (section 4) to facilitate better understanding of what risk adjustment does and does not do. See Appendix G for public comment themes and Panel responses.

Applicability of Recommendations

The recommendations may apply to outcome performance measures (including resource use and patient-reported outcomes) and some process performance measures used for comparative performance assessment. However, ***each performance measure must be assessed individually to determine appropriateness of sociodemographic adjustment***. The recommendations may apply to any level of analysis including health plans, facilities, individual clinicians, accountable care organizations, etc.

Recommendations Related to NQF Criteria and Processes Related to SDS Adjustment

Recommendation 1: When there is a conceptual relationship (i.e., logical rationale or theory) between sociodemographic factors and outcomes or processes of care and empirical evidence (e.g., statistical analysis) that sociodemographic factors affect an outcome or process of care reflected in a performance measure:

- those sociodemographic factors should be included in risk adjustment of the performance score (using accepted guidelines for selecting risk factors) unless there are conceptual reasons or empirical evidence indicating that adjustment is unnecessary or inappropriate;

AND

- the performance measure specifications must also include specifications for stratification of a clinically-adjusted version of the measure based on the sociodemographic factors used in risk adjustment.

Rationale: Patient characteristics that are present before care begins can influence patient outcomes or some processes of care. In order to avoid incorrect inferences or conclusions about quality in the context of comparative performance evaluation of various healthcare entities, some performance measures need to be adjusted for relevant patient characteristics when certain conditions are met.

Adjustment of performance measures for clinical complexity of the mix of patients served is widely accepted and the same principles and methods apply to sociodemographic characteristics. There are conceptual and statistical conditions for selecting risk factors that must be met and evaluated for each individual performance measure. Not all performance measures, or even all outcome performance measures, may need to be adjusted for sociodemographic factors. For example, the outcome of central line infection occurring during a hospital stay or the process of administering the correct medication at the correct time during a procedure would not have a conceptual basis for SDS adjustment, as there is no logical reason why these measures should be affected by variables like poverty, illiteracy, or limited English proficiency. However, if there is a conceptual relationship (i.e., logical rationale or theory, prior research) and empirical relationship (i.e., based on statistical analysis) with the outcome or process being measured and the guidance for selecting risk factors is followed, relevant SDS factors should be included in risk adjustment procedures to avoid incorrect inferences about quality based on an overall performance score. This approach is grounded in accepted methods and principles related to statistical inference and confounding discussed in section 4.

The recommendation acknowledges there may be situations where SDS adjustment is unnecessary or inappropriate based on conceptual reasons or empirical evidence. Important considerations include whether the key processes leading to an outcome are directly under the control of the healthcare unit and do not depend on active patient participation as in the examples noted above or whether the effect of an SDS variable on an outcome is due primarily to differences in the quality of care received. The information submitted with a performance measure considered for NQF endorsement should justify the approach taken as outlined in the recommendations. These topics are discussed in sections 4 and 6.

Identifying and reducing disparities in health and healthcare are important national priorities and require additional analysis of performance data by patient subgroups. If sociodemographic factors are included in a risk model, it indicates that the measure is disparities-sensitive and should also be stratified to identify differences by patient subgroups. Stratified performance data are most useful and most transparent as a means of identifying where disparities exist, which isn't possible in an overall score, whether only clinically-adjusted, or SDS-adjusted. Requiring that an SDS-adjusted measure also be specified for stratification is a continuation and strengthening of NQF's prior guidance to stratify disparities-sensitive performance measures. Performance data should be stratified on the basis of the sociodemographic factors used in risk adjustment so that clinically-adjusted scores are computed for each stratum (not one overall clinically-adjusted score). Specifications would include how the strata are constructed and how to compute the clinically-adjusted score for those strata. It is important to note a major limitation of stratified data by healthcare unit: small cell sizes decrease the reliability of the estimates, and they should not be used for comparative performance evaluation. Appropriate explanations about limitations or minimum cell sizes to be reported should accompany the stratified data.

Clearly, a concerted effort among providers, health plans, policymakers, researchers, and the public is needed to address healthcare disparities. For example, when sociodemographic factors influence a

performance measure, providers need to examine their own data to identify opportunities for improvement in serving disadvantaged patient populations. The Centers for Medicare & Medicaid Services (CMS) or other producers of performance reporting should make such stratified data easily available to interested parties, such as consumer advocates, researchers, health plans, and providers. Doing so could serve a dual purpose of providing finer grained data to interested parties and for assessing and addressing healthcare disparities.

Recommendation 2: NQF should define a transition period for implementation of the recommendations related to sociodemographic adjustment. During the transition period, if a performance measure is adjusted for sociodemographic status, then it also will include specifications for a clinically-adjusted version of the measure only for purposes of comparison to the SDS-adjusted measure.

Rationale: A defined transition period with specific evaluation parameters will facilitate a systematic collection of information about the change in policy, including additional information about the effects of sociodemographic adjustment and any unintended consequences. Additional guidance related to implementing stratification as outlined in recommendation 1 may need to be developed. Therefore, during the transition period, specifications for a clinically-adjusted version of the SDS-adjusted measure would be included within the endorsed SDS-adjusted measure submission and identified for comparison purposes only. "Comparison" here means comparison between overall scores of the clinically-adjusted and SDS-adjusted versions of a measure to understand the effects of SDS adjustment. It does not mean use of the clinically-adjusted measure for actual comparisons of health plans or providers in public reporting or pay-for-performance programs. The clinically-adjusted version of the SDS-adjusted measure is an essential step to stratification as recommended and also has been seen by some stakeholders as important to understanding the effect of the policy change. The second part of recommendation 1 indicates that an endorsed SDS-adjusted measure always includes specifications for stratification of the clinically-adjusted version of the measure; therefore, specifying a clinically-adjusted version of the measure is a required step toward stratification. The recommended Disparities Committee would be tasked with further detailing requirements for stratification.

Recommendation 3: A new NQF standing committee focused on disparities should be established.

Rationale: A standing disparities committee would review implementation of the revised policy about sociodemographic adjustment as recommended in this report (including key decisions by developers and purchasers) and monitor for any unintended consequences of the revised policy. It would also assess trends in disparities and review and provide guidance related to methodologies for adjustment and stratification such as use of community factors, and standard sociodemographic data collection. The membership of the committee should follow standard NQF policy about representation of diverse stakeholders and balance of perspectives.

Such a committee would also help ensure that social and demographic disparities in care do not get overlooked, but rather remain an integral part of quality measurement. The committee would be explicitly tasked with examining evidence for unintended consequences to patients across the full range of NQF-endorsed measures—including lowered expectations and incentives to improve care to

disadvantaged patients—by monitoring disparities both between and within providers. The committee would review decisions regarding when measures are adjusted for sociodemographic factors and how. It would assess the impact of the NQF policy changes on disadvantaged patients and on safety net providers. It would recommend the collection of additional sociodemographic data (individual- or community-level). The committee would suggest ways to better address and/or integrate healthcare equity and value. The committee could investigate how risk adjustment methodologies and stratification may influence our understanding of where and why disparities exist. It also could play a role in assisting developers and end users understand the role of risk adjustment and stratification in portraying and evaluating provider and health plan performance.

Because of the change to long-standing NQF policy proposed in the panel's recommendations, the disparities committee would be specifically tasked with preparation of an annual report, for at least the first five years of its existence, for public release, on the issues listed above. Its first task would involve a one-year look back at the consequences of the recommendations, both intended and unintended. This would help ensure that the recommendations were having the intended effect.

Recommendation 4: The NQF criteria for endorsing performance measures used in **accountability applications** (e.g., public reporting, pay-for-performance) should be revised as follows to indicate that patient factors for risk adjustment include both clinical and sociodemographic factors.
(*Note: additions are underlined; deletions are indicated with strikethrough*)

> ***2b4.*** *For outcome measures and other measures when indicated (e.g., resource use, some process measures):*
> *an evidence-based risk-adjustment strategy ~~(e.g., risk models, risk stratification)~~ is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome ~~(but not factors related to disparities in care or the quality of care)~~ and are present at start of care;[14,~~15~~] and has demonstrated adequate discrimination and calibration* ***OR*** *rationale/data support no risk adjustment~~/ stratification~~.*
> *14. Risk factors that influence outcomes generally should not be specified as exclusions.*
> ~~*15. Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.*~~

Rationale: This change in the NQF criteria removes the prohibition against adjusting for sociodemographic factors and is consistent with recommendation 1.

Recommendation 5: The same guidelines for selecting clinical and health status risk factors for adjustment of performance measures may be applied to sociodemographic factors, and include the following:

- Clinical/conceptual relationship with the outcome of interest
- Empirical association with the outcome of interest

- Variation in prevalence of the factor across the measured healthcare units
- Present at the start of care
- Is not an indicator or characteristic of the care provided (e.g., treatments, expertise of staff)
- Resistant to manipulation or gaming
- Accurate data that can be reliably and feasibly captured
- Contribution of unique variation in the outcome (i.e., not redundant)
- Potentially, improvement of the risk model (e.g., risk model metrics of discrimination, calibration)
- Potentially, face validity and acceptability

Rationale: The guidelines for selecting clinical risk factors apply equally well to sociodemographic factors. Selecting risk factors and developing a model is an iterative process, but is based first on a conceptual relationship and demonstration of an empirical relationship with the outcome or process of interest. A detailed discussion of selecting risk factors is provided in section 6.

Recommendation 6: When there is a conceptual relationship and evidence that sociodemographic factors affect an outcome or process of care reflected in a performance measure submitted to NQF for endorsement, the following information should be included in the submission:

- A detailed discussion of the rationale and decisions for selecting or not selecting sociodemographic risk factors and methods of adjustment (including a conceptual description of relationship to the outcome or process; empirical analyses; and limitations of available sociodemographic data and/or potential proxy data) should be submitted to demonstrate that adjustment incorporates relevant sociodemographic factors unless there are conceptual reasons or empirical evidence indicating that adjustment is unnecessary or inappropriate.
- In addition to identifying current and planned use of the performance measure, a discussion of the limitations and risks for misuse of the specified performance measure.

Rationale: NQF submission currently requires information on risk adjustment specifications, risk factor selection, assessment of the risk adjustment procedure, and current and planned use of the performance measure. The developer's decisions regarding sociodemographic factors, including use of proxy data, should be transparent and open to review and evaluation. See section 6 for a discussion about information to be submitted for evaluation.

Recommendations Relevant to NQF Policy

Recommendation 7: NQF should consider expanding its role to include guidance on implementation of performance measures. Possibilities to explore include:

- guidance for each measure as part of the endorsement process;
- guidance for different accountability applications (e.g., use in pay-for-performance versus pay-for-improvement; innovative approaches to quality measurement explicitly designed to reduce disparities).

Rationale: A measure that is ideal for one use may not be ideal for another. How a measure is implemented involves multiple decisions that could affect the validity of conclusions (inferences) made about quality of care and potential unintended consequences. The review of the detailed information about the performance measure for potential endorsement provides an opportunity to identify any specific considerations or limitations for use in specific accountability applications.

Recommendation 8: NQF should make explicit the existing policy that endorsement of a performance measure is for a specific context as specified and tested for a specific patient population (e.g., diagnosis, age), data source (e.g., claims, chart abstraction), care setting (e.g., hospital, ambulatory care), and level of analysis (e.g., health plan, facility, individual clinician). Endorsement should not be extended to expanded specifications without review and usually additional testing.

Rationale: This is implicit in the current NQF criteria and process for endorsing a performance measure as specified and tested. However, it should be clearly stated that expansions to additional patient populations, data sources, settings, or levels of analyses are not endorsed and would require an ad hoc review to expand endorsement.

Recommendations about Broader Related Policy Issues

Recommendation 9: When performance measures are used for accountability applications such as public reporting and pay-for-performance, then purchasers, policymakers, and other users of performance measures should assess the potential impact on disadvantaged patient populations and the providers/health plans serving them to identify unintended consequences and to ensure alignment with program and policy goals. Additional actions such as creating peer groups for comparison purposes could be applied.

Rationale: Even if a performance measure is adjusted using sociodemographic factors, this does not ensure protection of safety net providers and additional strategies may be needed. For example, SDS adjustment or stratification for patient-level factors does not address potential differences in community factors such as public funding or area healthcare resources, which may have a substantial impact on comparative performance results. Given that safety net providers are differentially funded (a function of local and state taxing jurisdictions), making comparisons even among safety net providers may be problematic. Accountability programs should consider if and how to incorporate this type of community factor into comparative evaluations for purposes of assigning rewards and penalties.

Although NQF does not control how measures are implemented, it is important to signal that the impact of program polices on providers or health plans caring for disadvantaged populations should be considered. These healthcare units may have fewer resources to improve the care they provide. The recent MedPAC recommendation regarding hospital readmissions is an example of creating peer groups for comparison as a way to lessen the impact of a performance penalty on safety-net hospitals.

Recommendation 10: NQF and others such as CMS, Office of the National Coordinator (ONC) for Health Information Technology, and the Agency for Healthcare Research and Quality (AHRQ) should develop strategies to identify a standard set of sociodemographic variables (patient and community-level) to be collected and made available for performance measurement and identifying disparities.

Rationale: Even when performance measures should be adjusted for sociodemographic factors, data limitations currently pose a substantial barrier. Although mandated data collection is beyond the scope of NQF, there is a need for a national effort to collect relevant sociodemographic information in a standardized way that allows for its valid use in adjustment models for performance measures that will be applied across states and regions. Most sociodemographic variables, particularly socioeconomic factors, that could conceivably be used in risk adjustment models are not currently collected in a standard way by health plans, doctors, hospitals, and other healthcare providers, and are not included in claims data bases that are often used to develop risk models. Data on sociodemographic factors also are important for providers when providing care and when reviewing their performance for quality improvement.

Section 3: Background

Context of Comparative Performance Assessment

NQF endorses performance measures that are intended for use in accountability applications such as public reporting and pay-for-performance. In this context, the overall performance measure score is used to make a conclusion about a unit's quality in relation to other units or some other comparator such as average performance. The general question being addressed is: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?** That is, the measure scores are used to identify which units have better quality in order to inform decisions of an individual to seek care, a purchaser to pay for care or give a bonus or penalty, a network to award contracts, etc. Such comparisons should be affected as little as possible by factors other than quality of care, such as patient characteristics already present at the start of care.

Because healthcare outcomes are a function of patient attributes (including SDS) as well as the care received; and patients are not randomly assigned to units for healthcare services so that all units have the same mix of patients, risk adjustment is essential to examining outcome performance in real-world settings.[6] Thus, when comparing outcomes, the purpose of risk adjustment is to ensure like-to-like comparisons. [6] Without appropriate risk adjustment, units can be misclassified based on incorrect conclusions about comparative performance. (See the basics of risk adjustment in Appendix C.) Depending on the specific program in which the performance measures are used, misclassification can create disincentives to care for more complex patients (clinically or sociodemographically complex) and potentially decrease resources to those units with large shares of complex patients.

Although NQF does not control the structure of various accountability programs, NQF's primary role is to ensure that an endorsed performance measure is suitable for use in comparative accountability applications. An appropriately adjusted performance measure alone will not solve other issues or problems that could be present in various accountability programs or formulas for determining base payment for services to more complex patients, which are outside the role of NQF.

Evidence-Based Risk Adjustment Strategy

NQF measure evaluation criteria call for an "evidence-based" risk adjustment strategy. Identifying potential risk factors may be informed by prior studies, but is not required. Ultimately the final risk adjustment strategy requires empirical evidence from statistical analyses regarding the relationship of the potential factors to the outcome using actual data. The relationship to the outcome is assessed first individually and then in the context of other risk factors. Risk factors and their strength of association are unique to each individual performance measure. This requirement for an evidence-based risk adjustment strategy is different from the NQF requirement for clinical evidence that supports performance measures of structure, processes, and intermediate outcomes, which calls for a systematic assessment and grading of the body of clinical evidence that supports their link to desired outcomes.

Sociodemographic Factors and Outcomes

The term *sociodemographic* will be used to include a variety of socioeconomic (e.g., income, education, occupation) and demographic factors (e.g., race, ethnicity, primary language) that are often associated with disadvantage among affected populations. Although age is a demographic factor, it also is considered a clinical factor and already included in many risk adjustment procedures. A large body of evidence shows an association between various sociodemographic variables and outcomes.[1-4] In

general, sociodemographic "disadvantage" (e.g., low income, low education, homelessness) is often associated with poorer patient outcomes (e.g., higher morbidity, mortality, or readmissions). Low SES and social disadvantage tend to be associated with greater morbidity, disease severity, and worse quality of life.[7, 8]

The mechanism(s) for the association between sociodemographic factors and health status and outcomes is often complex and is not always clear.[9, 10] Depending on the specific SDS factor and outcome, it can involve the effect of mediators such as financial resources, community resources, or patient understanding on the ability to access healthcare services or follow through with treatments. These factors contribute to healthcare disparities.[1] Sociodemographic factors operate in the present but also may have a cumulative effect on health outcomes across the life course through a variety of mechanisms including early effects on sensitive periods during development and epigenetic effects. Historical and current discrimination impact the patient, ranging from biological stress levels to social confidence when interacting with the healthcare system. Another potential and simultaneous mechanism may be the implicit biases or assumptions on the part of healthcare providers that influence their interactions with, and the care options given to, patients with different characteristics (e.g., race/ethnicity), thus increasing the likelihood of providing substandard care. Disadvantaged patients also may be concentrated in areas of poorly resourced or lower quality healthcare services.

The characteristics associated with being disadvantaged (e.g., low SES) generally are associated with less than optimal clinical outcomes. However, for resource use and cost outcomes, the relationship could vary. Depending on timing and the population included, cost and resource use could be less in disadvantaged patients because of inability to access and use healthcare services, or more because of higher severity due to lack of preventive and early diagnostic services.

Essentially, the evidence of a relationship to SDS will vary depending on the specific outcome or process being measured. As will be discussed in the report, potential risk factors need to be assessed empirically with actual data for the proposed risk factors and the outcome being measured.

Process Performance Measures

Most of the same issues regarding the relationship between sociodemographic factors and outcomes might also apply to processes, especially processes that are not directly under the control of the healthcare provider and require some action by the patient (e.g., getting prescription filled). As with the outcome performance measures, adjusting process measures should be guided first and foremost by a causal theory. Many processes are primarily under the control of healthcare providers (e.g., administering the correct antibiotic to prevent surgical site infection), and adjustment for sociodemographic factors would not be appropriate because the relevant clinical guideline generally would make no exception for sociodemographic factors, and there is no plausible, acceptable causal path through which a sociodemographic factor would affect performance of the clinical process. Some processes, though, are not as strongly under the control of the provider (e.g., adherence with medications, receipt of screening colonoscopy), and adjustment for sociodemographic factors might be called for if the general criteria for selecting risk factors are met. Therefore, the recommendations regarding sociodemographic adjustment also apply to some process performance measures.

Perspectives on Adjusting for Sociodemographic Factors

The reasons for and against adjusting performance measures for SDS were identified during the Panel's deliberations, and they also were raised during the comment period. The reasons for opposing or

supporting SDS adjustment are based on a combination of evidence, logical arguments, and some assumptions about drivers of behavior.

The Expert Panel carefully considered all perspectives and ultimately recommended that performance measures be adjusted for SDS under certain conditions. The concerns about negative consequences attributed to SDS adjustment discussed below are specifically addressed in the recommendations to include specifications to stratify any performance measures that are SDS-adjusted, establish a standing disparities committee, and define a transition period. The concerns are also addressed in the methodological discussion.

During the review and comment process, some questioned the role of evidence in weighing the arguments for and against SDS adjustment. Ultimately, the recommendation for SDS adjustment could only be made if it was based on sound measurement science, which is discussed in section 4.

Concerns and Unintended Consequences about Adjusting Performance Measures for Sociodemographic Factors

The first and most important concern about adjustment for sociodemographic factors is that disadvantaged patient groups, on average, might receive worse quality of care. In other words, differences in observed performance, either across units or by patient group within units, reflect actual differences in the processes of care for disadvantaged versus other patients that would be "adjusted away". In a study of quality of care related to sociodemographic factors, Asch et al.[11] found small, but statistically significant differences in quality of care provided by the income level of patients; unexpectedly, blacks and Hispanics received slightly better quality care than whites. However, the differences in quality among sociodemographic subgroups were small in comparison to the difference between observed and optimal care.

There are three mechanisms through which healthcare units might provide worse care processes for disadvantaged patients. First, it could reflect bias in care by providers in general based on the sociodemographic characteristics of the patients (e.g., poverty, race, language).[12, 13] Second, it could reflect reduced resources and funding in places where patients receive care. If disadvantaged patients cluster within poorly resourced units or within units that provide worse care, then disadvantaged patients will on average, receive worse care.[14-16] A third mechanism involves attempts by the clinician to tailor care to perceived constraints of the patient. Such decisions might be appropriate, i.e., when they are collaboratively made in partnership with the patient, or inappropriate, i.e., when the physician unilaterally decides what the patient wants and/or can afford. Findings from empirical studies often differ depending on the performance measure and provider type.

A second concern is that adjustment will mask meaningful differences in quality or performance—that is, the adjustment will have a strong enough effect that meaningful differences in performance will not be detectable in adjusted performance scores. The concern reflects a belief that differences in unit outcome performance reflect the degree to which units implement interventions to mitigate the effects that sociodemographic factors have on those outcomes (e.g., instructions in multiple languages, interpreters, prescribing low-cost generic drugs, hospital discharge follow-up), rather than the effect of those factors on patient outcomes.

A third concern is that adjustment implies that worse outcomes are "expected" for certain patient groups such as those with low income, creating a double standard, and no expectation that healthcare

units try to mitigate the effect of such factors on outcomes. Some worry that if adjusting for sociodemographic differences results in performance being labeled "average" or "as expected" despite worse outcomes for disadvantaged patients, it will blunt the motivation to provide optimal care for disadvantaged patients. In other words, if the effect of sociodemographic factors is "adjusted away," one cannot or will not do something about them. Empirical evidence supporting or refuting this concern is lacking.

Concerns and Unintended Consequences When Performance Measures Are Not Adjusted for Sociodemographic Factors

The association between SDS and outcomes has been demonstrated with outcome performance measures for physicians, hospitals, and other healthcare providers.[10, 16-21] In general, caring for sociodemographically "disadvantaged" populations is associated with poorer performance (based on current performance measures) on average, although there are some noteworthy exceptions to the general pattern.[22]

As discussed in the methodological basis for adjustment (section 4), SDS is a potential confounder and can lead to incorrect comparative inferences or conclusions about quality. An alternative explanation to the conclusion of poor quality for poorer performance scores is that the unit is caring for a disproportionate share of disadvantaged patients, who all else being equal, have worse outcomes (just as do clinically complex patients).

In addition to hindering informed decisionmaking by patients, use of performance measures that fail to account for sociodemographic factors when indicated, could lead to harm of patients through other mechanisms. As healthcare moves toward increasing use of financial rewards for better quality and financial penalties for worse quality, use of measures that result in incorrect conclusions about quality poses a substantial risk for penalizing healthcare organizations and providers who serve more disadvantaged populations.[23-27]

Units serving "low-sociodemographic" populations and communities are more likely to be identified as "poor performers" and either be less likely to receive financial rewards, or be more likely to face financial penalties, in pay-for-performance programs. Joynt and Jha,[28] for example, found that safety-net hospitals were more than twice as likely as other hospitals to have high penalties in the first year of the Medicare Hospital Readmission Reduction Program. In another example, Young, et al., found a strong association between socioeconomic characteristics of members of Medicare Part D drug plans and the performance ratings of those plans.[29] Zaslavksy and Epstein found associations between sociodemographic characteristics and some HEDIS performance measures and noted that sociodemographic adjustment would have a meaningful impact on health plan comparison for some plans and some measures. [30]

In the context of public quality reporting and pay-for-performance, failing to account for the greater difficulty in achieving good outcomes in socially and economically disadvantaged populations could set up a series of adverse feedback loops that result in a "downward spiral" of access and quality for those populations. The net effect could worsen rather than ameliorate healthcare disparities. There are at least three potential adverse consequences, each of which could have the eventual effect of undermining the quality of care for disadvantaged patients, thus exacerbating disparities in health and healthcare.

First, with public reporting of performance, healthcare units will have a strong incentive to avoid serving disadvantaged populations, so as to avoid being labeled as a "bad performer." This could happen based on where physicians and other individual providers choose to work, where facilities are opened or closed, or expanded or contracted, where health plans operate, and through more subtle ways of "cherry picking." A study on public reporting of surgeon mortality rates for coronary artery bypass graft (CABG) that did not adjust for race resulted in decreased access to surgery for racial/ethnic minorities.[31] Second, with some pay-for-performance programs, substantial funding will be shifted away from organizations serving disadvantaged populations and communities and to organizations serving more affluent, less vulnerable, populations, and communities. Third, individual consumers, private and public payers, and others choosing among providers or plans whose performance will be publicly reported will tend to avoid units serving disadvantaged patients and communities based on performance scores that may not provide a valid comparative performance assessment.

In the context that measurement science supports adjustment for sociodemographic factors when certain conditions are met, most of the NQF Expert Panel members were concerned that the potential negative impact on care to disadvantaged patients resulting from the three phenomena listed above were, on balance, even more detrimental than the concerns about "masking disparities" or "setting a lower bar for performance." Therefore, the current position against sociodemographic adjustment should be reconsidered. However, the expressed concerns about masking disparities and lower standards were also addressed in the final recommendations.

Mitigation of Effect of Sociodemographic Factors

Adjustment for SDS does not mean that providers cannot take steps to mitigate the effects of some sociodemographic factors. Just as care is adjusted based on clinical severity and complexity, care should be adjusted to address specific needs related to sociodemographic factors. Some examples include providing interpreters, instructions in different languages, discharge clinics, prescribing generic drugs, outreach to homeless patients in community settings, etc. Strategies to mitigate the effects of sociodemographic factors are often resource-intensive and raise payment policy issues that are outside the scope of this project but discussed briefly later in this report (section 8).

Section 4: Methodological Basis for Risk Adjustment

As already mentioned, when performance measures are used for comparative assessments as with pubic reporting and pay-for-performance, risk adjustment is essential to avoid making incorrect inferences or conclusions about quality or performance. The goal of risk adjustment is to answer the question: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?**

The need for risk adjustment is based on accepted and foundational statistical theory and epidemiologic principles involving causal inference and confounding. This section describes the key foundation for risk adjustment and also provides responses to some of the concerns about sociodemographic status (SDS) adjustment from a methodological perspective. Other considerations for selecting risk factors are discussed in section 6.

Key Definitions

- **Confounding** refers to the distortion in the degree of association between an exposure (independent variable) and an outcome (dependent variable) due to a mixing of effects between the exposure and an incidental (confounding) factor. Confounding represents systematic error and threatens the internal validity of an epidemiologic study since it can lead to false conclusions regarding the true relationship between an exposure and outcome. (See the basics of confounding in Appendix D.)
- **Risk adjustment (also known as case-mix adjustment)** refers to statistical methods to control or account for patient-related factors when computing performance measure scores; methods include multivariable modeling, indirect standardization, or direct standardization. These methods can be used to produce a ratio of observed-to-expected, a risk-adjusted rate, or other estimate of performance. (See the basics of risk adjustment in Appendix C.)

Terms

- **Unit** will be used to signify the entity whose performance is being measured, which could be a hospital, health plan, clinician, etc.
- **Clinical adjustment** refers to adjustment for only clinical variables.
- **Sociodemographic or SDS adjustment** refers to adjustment for both clinical and sociodemographic variables.

Conceptual Basis for Risk Adjustment

In clinical comparative effectiveness studies, researchers often ask whether one treatment is better than another for reducing morbidity or improving survival. For example, in a randomized controlled trial comparing the success rate of two treatments, "A" and "B," the average "effect" of treatment A versus B can be estimated by the difference in the proportion of patients receiving treatment A who have a successful outcome and the proportion of patients receiving treatment B who have a successful outcome, $\hat{s}_A - \hat{s}_B$. Randomization ensures that patients receiving the two treatments are comparable. When treatments are not randomly assigned— as in a nonrandomized observational study— the observed difference $\hat{s}_A - \hat{s}_B$ may be biased (i.e., systematic deviation from the true value). To the extent that patients in each treatment group differ in ways that affect outcomes (e.g., they are sicker, frailer, etc.), the observed differences in outcomes may reflect different patient characteristics rather than the treatment effect of interest. In other words, the effect of treatment is confounded by differences in

pretreatment patient factors. An identical confounding issue arises in studies comparing outcomes of healthcare units, in which patients are not randomized to units (just as in the real-world environment of performance measurement in which patients are not randomized to units).

Table 1 illustrates potential confounding in performance measurement. In this example, clinical severity is associated with mortality in the national patient population with a difference of 1 percentage point between patients with low to average clinical severity vs. high severity (2% vs. 3%). Across the units, the proportion of high severity patients varies—unit A has exactly the same proportion as nationally (20%) and unit B has a higher proportion (60%). Unit B has a higher unadjusted mortality rate than the national average (2.6% vs. 2.2%). Because severity is a potential confounder, an alternative explanation for unit B's higher overall mortality rate is its substantially higher proportion of high-severity patients, rather than it delivers worse quality of care. The data stratified by clinical severity group indicate that for unit B, the higher overall mortality is a function of serving a larger proportion of high-severity patients because the mortality within each group is exactly the same as the national averages for those groups.

Table 1. Example of Confounding

	All Patients in National Population		**Unit A**		**Unit B**	
Clinical Category	Patient mix N/Percent	**Unadjusted Mortality** N/Percent	n	**Unadjusted Mortality** n/Percent	n	**Unadjusted Mortality** n/Percent
All Patients	1,000,000 100%	22,000 **2.2%**	1000	22 **2.2%**	1000	26 **2.6%**
Low-Average Severity	800,000 80%	16,000 2%	800 80%	16 2%	400 40%	8 2%
High Severity	200,000 20%	6,000 3%	200 20%	6 3%	600 60%	18 3%

Risk (or case-mix) adjustment refers to a collection of techniques for reducing the effect of confounding factors in studies where patients are not randomly assigned to different treatments. In performance evaluation, the "treatments" are different healthcare units. The "treatment effect" may be conceived as the difference between a patient's actual outcome and the outcome that would have occurred had the patient been treated by another unit. Risk adjustment aims to control for patient factors (e.g., morbidity or sociodemographic factors) that could affect the outcome so that residual differences in outcomes reflect the treatment effect of interest.[32]

The statistical and epidemiologic literature describes conditions in which valid inferences about treatment effects based on observational data are possible. In general, unbiased estimation (i.e., without systematic deviation from the true value) requires the assumption that outcome differences are unconfounded, conditional on a set of pretreatment covariates. This unconfoundedness assumption means that blocks of patients having identical values of pretreatment covariates and who receive each treatment, are like a random sample from a common population. Although the unconfoundedness assumption is unlikely to be literally true in a nonrandomized observational study, the risk of encountering large violations of the assumption can be minimized by careful consideration of all

potential confounders. The assumption becomes more plausible if the set of covariates is expanded to include all factors that may predict the outcome or the choice of healthcare unit or both. Data availability may be a practical constraint on the factors that can be considered.

What types of variables are appropriate for risk adjustment?

Risk adjustment involves an attempt to compare only patients who are alike with respect to a set of pretreatment covariates. When we say that a variable was "adjusted" or "included in the risk adjustment," we mean that the analysis aimed to compare outcomes of patients at different healthcare units who were similar with respect to that covariate. In general, covariates appropriate for risk adjustment are those factors that are hypothesized to remain the same if the patient were to be reassigned to a different unit.[33] By this rule, any patient characteristic that is present prior to treatment and is a known or suspected confounder of the treatment effect may be included. Variables of sociodemographic complexity could also cause confounding in the same way as severity of illness in the example in Table 1 and therefore, also eligible for adjustment.

Although it is generally desirable to adjust for all important confounding factors, theory dictates that we should not adjust for components of the treatment being evaluated.[34] Doing so may "adjust away" differences in outcomes that result from the adoption of more or less effective care practices by different units. For example, one would generally not adjust for the frequency of hand washing when comparing infection rates across hospitals because assiduous hand washing is one of the ways in which a hospital may seek to achieve a lower infection rate.[33] SDS factors are not treatment variables in the way that whether a specific surgical intervention was provided is a treatment variable, and therefore, they do not "adjust away" treatment effects. However, additional concerns about how SDS factors may influence treatment and "adjust away" the unit treatment effect are discussed in later sections below.

Does adjusting for sociodemographic factors mask disparities in outcomes for disadvantaged patients?

In terms of revealing disparities, a single SDS-adjusted score is no different than a single clinically-adjusted score. SDS adjustment may change the unit score if the proportion of disadvantaged patients is larger or smaller than average. However, ***a single score alone*** (e.g., 80% of patients improved in function) without additional information on case mix ***cannot reveal potential disparities in outcomes across population subgroups, regardless of whether the score is only clinically adjusted*** (the current practice) ***or adjusted for both clinical and SDS factors*** (when appropriate).

As recommended by the Expert Panel and in prior NQF projects, identifying disparities requires additional information and analysis by the relevant sociodemographic factors (e.g., stratification) if the question is: *how do the outcomes of patients with different characteristics compare (either within units or in the population across units)*? Therefore, risk adjustment that includes sociodemographic factors does not change the fact that additional methods (e.g., stratification) as recommended are needed to identify disparities. However, a by-product of adjusting for sociodemographic factors is information about whether within-unit disparities exist. *For example, with multivariate modeling approaches, whether or not and to what degree an SDS factor contributes to the variability in outcomes can be determined.*

Does risk adjustment for sociodemographic factors set a lower expectation for outcomes of disadvantaged populations?

With some risk adjustment procedures, observed counts for an outcome are compared to "expected" counts, which are based on the average experience for patients with similar characteristics. When the average outcome rate for patients with certain characteristics is worse than that for other patients and this value is used to adjust performance scores, some are concerned that it sets a lower standard for the group of patients who experience worse outcomes.

In probability theory, the term "expectation" has a specific technical meaning that differs from its usage in everyday discourse. Generally speaking, it is the value of a random variable that would be observed on average in a large series of repeated trials or random samples. In the context of indirect standardization, the term "expected rate" has a similar technical meaning. It may be loosely translated as describing the "average" or "typical" outcomes for a given case mix. Importantly, the term "expected" is not intended to convey a judgment that "average" or "typical" outcomes are morally acceptable.

Although statisticians use the term "expectation" in this narrow technical sense, it is important to ask what (implicit or explicit) value judgments are reflected in the various accountability initiatives that make use of risk-adjusted performance measures. Policy concerns about accepting or institutionalizing the status quo would not necessarily impact the choice of statistical methodology for risk adjustment and performance measurement, but might reasonably impact decisions about the design of accountability initiatives and the allocation of pay-for-performance (and other) incentives. Nevertheless, risk adjustment does change performance scores (depending on the mix of patients) because the intent is to answer the question: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?** If the interest is in the question: ***how do the outcomes of patients with different characteristics compare (either within an individual unit or at the population level)***, then performance data stratified by the relevant factors are needed.

Does risk adjustment for sociodemographic factors reduce the incentive to improve care for disadvantaged patients?

There is an expressed concern that sociodemographic adjustment will raise the performance status of some units with a large share of disadvantaged patients from "substandard" to "average" or "average" to "good" and this will lessen the incentive to improve care at those units. There is a parallel concern, sometimes expressed as "masking disparities," that the poor outcomes of such units will be labeled as average — the same label as for a unit producing better outcomes but for a less disadvantaged group. It is unknown whether such a change in labeling will have an impact on motivation to improve, but there is, of course, still an opportunity for such a unit to raise itself to a "superior" level by implementing solutions to problems that affect outcomes for its disadvantaged patients. Motivation to improve is also influenced by the structure of formal (e.g., financial) and informal (e.g., reputational) incentives. We do not know which of these incentives is more motivating.

In general, when there are different categories of patients, the largest group will have the greatest impact on a unit's performance score. Therefore, any improvement in performance in the largest group will improve the overall performance score by a greater amount than a similar improvement in the smaller group. Units interested in improving their overall performance score likely will focus on

improvements affecting the largest number of their overall population of patients. Sociodemographic risk adjustment (or not) does not change this tendency.

If a performance measure is SDS-adjusted, it means that there is a difference in outcomes for one or more specific sociodemographic factors. This signals a need to review data for sociodemographic subgroups to identify opportunities for improvement. The Panel recommended that specifications for SDS-adjusted measures must also include instructions for stratification of a clinically-adjusted version of the measure based on the sociodemographic factors used in risk adjustment. This stratification will allow identification of, and facilitate reduction of, sociodemographic disparities. If a unit's case mix includes a high proportion of disadvantaged patients, it will need to address the special needs of that population in order to improve its overall performance score.

When measures that are adjusted for SDS are implemented, the risk adjustment coefficients should be updated on a periodic basis. Thus, improvements in equity of outcomes or processes (that is, reductions in average within-unit quality differences) will be reflected in updated model coefficients and the effects of adjustment would diminish.

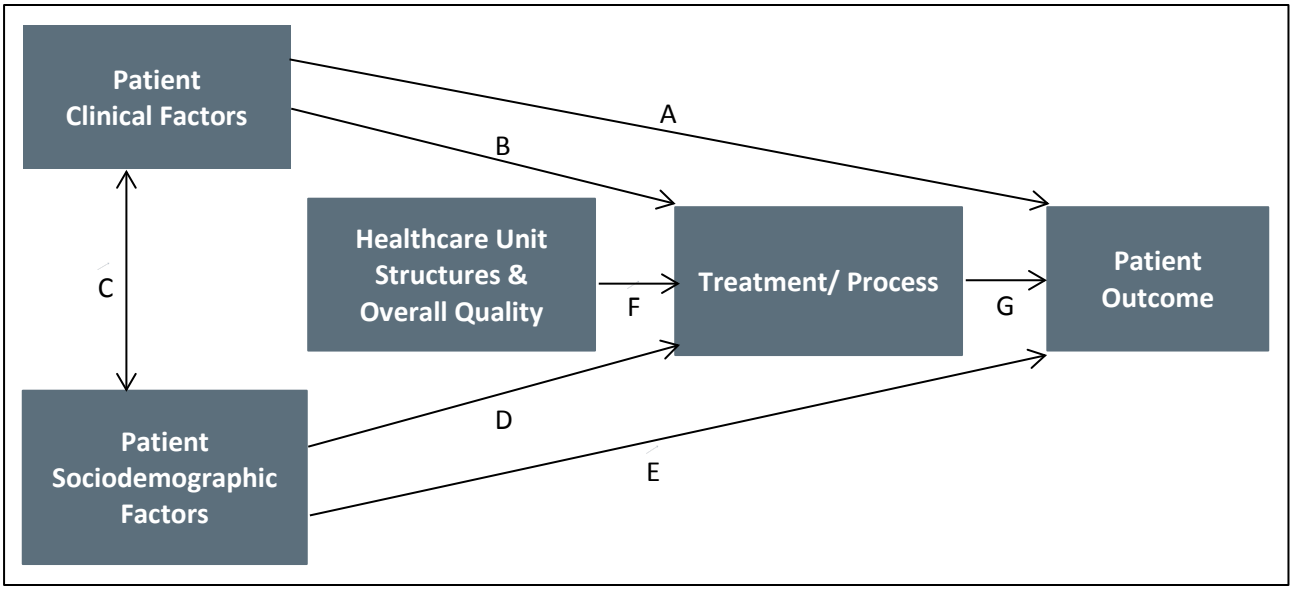
Finally, by appropriately risk-adjusting performance measures, units that have a high proportion of disadvantaged patients and are achieving better outcomes with those patients can be identified as examples for what can be achieved and a source of information about best practices. Their better outcomes might otherwise be masked by the absence of sociodemographic adjustment. Likewise, units achieving good outcomes, but with a low proportion of disadvantaged patients, are less likely to be identified as the best performers.

Does risk adjustment for sociodemographic factors mask disparities in quality if the reason sociodemographic factors affect an outcome is because of the care received?

There may be multiple and complex relationships between sociodemographic factors and outcomes. Following is a simplified path diagram for the effects of patient factors on outcomes.

The objective of performance measurement is to assess overall unit quality through its effects on measurable treatments and processes, as well as its effects on outcomes (path F to G). However, the inference about quality may be confounded by clinical characteristics of patients that affect success in implementing treatments/processes (path B) and/or directly affect outcomes, or through mechanisms not involving the healthcare unit (path A). ***Exactly the same causal relationships hold for sociodemographic characteristics*** of patients if they affect treatment/process (path D) or outcomes (path E).

Figure 1. Causal Paths



Not only should treatment variables be excluded from risk adjustment as discussed above, variables in the causal pathway between treatment and outcome (path G) should also be excluded because they can distort differences in outcomes by "adjusting away" the treatment effect of interest. For example, one would not adjust for a complication that arises after treatment begins. Even if a variable that occurs after beginning care with the unit does not directly cause outcomes, adjusting for it may cause bias (i.e., systematic deviation from the true value) if the variable is caused by the treatment and is correlated with the outcome. ***These concerns about variables that occur after treatment begins do not apply to sociodemographic factors that are present prior to treatment*** because such factors logically cannot be affected by the healthcare unit — i.e., the healthcare unit cannot affect the patient's level of income or education.

Healthcare unit structures (e.g., staffing ratio and expertise, financial health, performance on other quality measures) reflect their capacity to deliver quality treatments and processes. External factors can also affect the healthcare unit's capacity to deliver quality care (e.g., area pool of healthcare workers, public funding). These are unit characteristics, not individual patient characteristics, and are not used in risk adjustment procedures to account for differences in patient case mix across units.

In general, the path or mechanism of action for a patient factor's effect on an outcome does not need to be known in order to consider it a potential confounder to be assessed for risk adjustment. Adjustment for a variable might make sense if it is a direct cause, an indirect cause, or serves as a surrogate for a cause for which data are lacking. Inferences about comparative quality of healthcare units can be made only IF the potential confounding effects of the relevant factors are controlled (i.e., adjusted), regardless of the path or mechanism.

However, a concern remains that the reason for poorer outcomes on average for patients with particular SDS factors is that disadvantaged patients systematically receive poorer quality care than other patients because either:

- all or most healthcare units provide worse quality care to disadvantaged patients compared to other patients within the same units; or
- all or most disadvantaged patients primarily receive care from poorer quality healthcare units.

This concern is addressed in the following section.

Does risk adjustment for sociodemographic factors set a different standard if disadvantaged patients are concentrated in lower quality units?

The above concern about accounting for sociodemographic factors in a risk adjustment procedure if the factors are also related to poor quality can be further elucidated by distinguishing two sources of variation in outcomes across subgroups. Disparities in outcomes for disadvantaged patients can be caused by differences within and between units. Disparities in outcomes are a combination of two components:

- "disparities within": members of disadvantaged groups have worse outcomes than other patients within the same unit (could be due to a variety of reasons);
- "disparities between": members of disadvantaged groups receive care from units where a group of patients would experience inferior care (measured by other processes or outcomes) compared to a group of patients with similar clinical and sociodemographic characteristics receiving care at other units (some refer to this as a contextual effect[35]).

Either or both of these mechanisms can be at work in any dimension of quality, and their relative importance varies across measures and population subgroups. However, ***only the within-unit effects are adjusted for in a risk adjustment procedure because these are the ones that are related specifically to patient characteristics rather than differences across units***.

Adjustment for sociodemographic factors will not mask disparities in quality of care between units, provided that the risk-adjustment variables are measured at the patient level. The effect of those variables on individual patient outcomes then can be estimated, as long as there is variation in patient characteristics within units. While adjusting for sociodemographic variables will result in a different probability of an outcome, it just reflects the patient-based risk reality like any other patient comorbidity. The end result is that each unit's performance score will be compensated for the estimated effects of the sociodemographic factors in proportion to the number of patients in the sociodemographic categories, where those estimated effects are based on the experience of all units in the model. This is appropriate in the context of comparative performance assessment when addressing the question: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?** An illustration of the effect of risk adjustment appears in section 5.

While patients with certain characteristics may tend to concentrate differently across units (this establishes a fundamental requirement for risk adjustment), generally there is enough overlap of factor types across units, so that a model correctly estimates the necessary compensation for the disproportionate concentrations. If units differ in quality after adjusting for globally-estimated factor effects, this will be reflected in the profiling results of performance scores.

On the other hand, if members of disadvantaged subgroups tend to be concentrated within units that are overall of lower quality, then methods that ignore such systematic between-unit (contextual) differences can produce biased (i.e., systematic deviation from the true value) unit comparisons. For example, patients without insurance may have poorer outcomes, but this may be in large part because units that treat large numbers of uninsured patients have correspondingly fewer resources, leading to lower quality care for all patients treated by the unit, not just the uninsured. If concentration in low-quality units is a concern, methods exist to appropriately evaluate and address this source of confounding. The possibility of such clustering of disadvantaged subgroups within lower-quality units should be addressed by developers. Examples of methods that can address these issues are given below and in Appendix E .

- The between-unit differences can be controlled for or analyzed using various statistical methods such as including dummy variables for each unit or a unit-level variable that represents the same factor (e.g., percentage of low income patients). If a unit-level factor has an effect that is substantial relative to the patient-level effect, including only a patient-level covariate may result in adjustment for differences in quality of treatment.

  In theory, a patient-level factor could have a strong association with an outcome when between-unit effects are excluded from the model, but a negligible association after adding unit variables for each unit (contextual variables). This would occur, for example, if care for poor and nonpoor patients is similar within each unit but the poor receive care at lower-quality units. In that case, the sociodemographic factor is not a confounder when comparing outcomes across units and efforts to adjust for this factor when comparing outcomes across units may not be needed.

  It is important to distinguish *controlling* for unit effects when estimating within-unit (individual-level) effects, as discussed in the preceding paragraph, from *adjusting* for effects of unit characteristics when reporting quality. The latter is not the intent of risk adjustment, the goal of which is to control for confounders in order to identify the treatment effect of the unit. When unit-level variables are used, they must be used appropriately so as not to adjust performance scores for between-unit differences in quality, which is what you are trying to identify. It is beyond the scope of this paper to provide detailed guidance on statistical methods; however, if unit-level variables are included as described here, procedures for computing the estimated performance score would be different than when only patient-level characteristics are used.

- In the illustration of indirect standardization in section 5 of this report, if there is concern about concentration of low-income patients in low-quality units, direct standardization would produce valid estimates of healthcare unit performance despite low-income patients being concentrated among units of lower overall quality (see Appendix F). However, if concentration in low-quality units is not a concern, indirect standardization would produce valid estimates of performance. The key point is that just as estimates of unit effects should be controlled for possible confounding by patient characteristics, estimates of the direct effect of patient characteristics (observed within unit) should be controlled for unit effects.

These methods are only mentioned as an indication of the kinds of approaches that are relevant and are not a replacement for more technical discussions of various methods. The analyses of within- vs.

between-unit effects can be reported and discussed in the measure submission so that reviewers understand these relationships for the specific performance measure. This also is an example, where the analyses for SDS adjustment could potentially be used to reduce disparities by identifying the ways in which SDS affects outcomes.

Limitations of Risk Adjustment

Risk adjustment procedures are not perfect even with attention to rigorous methods and principles. Risk-adjusted scores may give a false sense of security and the details warrant close review. The following limitations are acknowledged.

- Data for a potential SDS risk factor with a strong conceptual relationship to the outcome or process being measured, even when based on prior research, may not be available for adjustment. This is not unique to SDS and also occurs with some clinical factors (e.g., stroke severity).

- Even if analyses can identify that an SDS factor exerts its effect on outcomes primarily due to between-unit differences, the specific reason(s) for the between-unit differences cannot be determined without additional study. For example, it would not be known whether the presumed quality differences between units were due to direct action or inaction on the part of the healthcare teams or influenced by lack of public support of safety net providers and insufficient resources to address the increased complexity of disadvantaged patients or to recruit healthcare workers, etc. These unit characteristics or factors would not be included in risk adjustment procedures used to account for differences in patient case mix as discussed above. However, some community-level factors such as public funding could be critical for policy considerations.

- Risk adjustment can only account for measurable and reportable factors. If unmeasured factors are not randomly distributed across units, the risk adjustment procedure may not adequately mitigate the impact of these unmeasured factors on the performance score for certain units. For example, if homelessness is not accounted for in the risk adjustment procedure and some units care for high proportions of homeless patients, adjusted performance scores may not compensate for differences in outcomes for patients with those factors.

Conclusion and Implications

- With appropriate selection of risk factors and risk adjustment methods, sociodemographic-adjusted scores do not mask disparities or differences in quality.

- Based on epidemiologic principles related to confounding and statistical theory of causal inference, the specific path or mechanism for the effect of an SDS factor does not need to be known. However, the requirement for a conceptual and empirical relationship to the outcome (or process) of interest, as well as the other guidelines for selecting risk factors discussed in section 6, will determine whether a sociodemographic factor should be included.

- When considering sociodemographic adjustment, the concern of disadvantaged patients being concentrated in overall lower quality units can and should be empirically tested and if necessary, addressed in the method used for adjustment.

- Risk adjustment does change performance scores if the proportion of patients with various characteristics is different from the average. This is appropriate if the intent is to answer the question: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?** Regardless of whether the risk adjustment procedure only includes clinical factors or includes both clinical and sociodemographic factors, an adjusted overall score is not designed to answer the question ***how do the outcomes of patients with different characteristics compare (either within an individual unit or at the population level)*****?** If the interest is in the second question, then data stratified by the relevant factors are needed.

The recommendation regarding sociodemographic adjustment includes the requirement for a conceptual and empirical relationship to the outcome (or process) being measured. Conceptual considerations may include whether the effect of SDS is primarily mediated through quality of care and questions such as whether there is any reason to think that a central line infection acquired during a hospitalization is influenced by race or income. This is discussed in section 6.

Although it may be possible to provide some rare but real or simulated examples illustrating some level of presumed failure (to prevent incorrect inferences about quality), that would be the exception rather than the statistical rule. It is not possible to create rules that would accommodate all possible scenarios regarding the use of sociodemographic risk factors. The guidelines for selecting risk factors, beginning with a conceptual and empirical basis, along with statistical and epidemiological theory and practices, provide a sound basis for making those determinations.

Section 5: Effect of Risk Adjustment

Risk adjustment refers to statistical methods to control or account for patient-related factors when computing performance measure scores, including methods such as multivariable models, indirect standardization, or direct standardization. The methodological basis for risk adjustment is presented in section 4. The result of the statistical procedure is an adjusted overall performance score that takes into account the presence of patient-related factors. Generally, healthcare units serving a disproportionate share of higher-than-average-risk patients will have adjusted scores that look better than their raw scores; the reverse will be true for units serving a disproportionate share of lower-than-average-risk patients.

An important goal of risk adjustment is to "level the playing field" when making conclusions about quality of care or performance. That is, the performance scores should not simply be due to differences in the severity or complexity of the patients served. As noted above, the guidelines for selecting clinical and health status risk factors also apply to sociodemographic factors. Therefore, without controlling for sociodemographic factors that have a conceptual and empirical relationship to the outcome or process, the inference from the performance score would be incorrect in the context of comparative performance assessment where the central question is: ***how would the performance of various units compare if hypothetically they had the same mix of patients*****?** Sociodemographic factors can contribute to the severity and complexity of the patient population served. Healthcare units with a disproportionate share of disadvantaged patients will appear to provide lower quality care than they actually do, and vice-versa simply as a function of their case mix.

The following illustration is based on one approach to adjustment — indirect standardization. (See another illustration for direct standardization in Appendix F.) With indirect standardization, an expected number of outcomes is determined by applying stratum-specific rates based on all patients in the reference population to a unit's number of cases in each stratum.[6] An observed-to-expected ratio is then used to compute a standardized or risk-adjusted rate. Multivariable statistical models are an extension of indirect standardization based on the same concepts.

The table that follows illustrates risk adjustment using indirect standardization. This hypothetical illustration does not use actual data and is simplified with just two levels for a sociodemographic factor and numbers chosen for easy computation. ***For purposes of this illustration, one should assume that the sociodemographic risk factor meets the guidelines for selecting risk factors presented in section 6 and accepted principles regarding confounding discussed in section 4***. The key points are illustrated in the top of the table — rows 1-6; details about the calculations are provided in rows 7-10.

- The initial scores (row 3) are already adjusted for clinical factors. We will call the performance measure "mortality rate," but it could represent any relatively rare adverse event.
- In this hypothetical example, the national mix of patients is 80% average-high income and 20% low income. The national average experience for mortality is 2% for average-high income patients vs. 3% for low-income patients. Assume that this rate is already adjusted for relevant clinical factors.
- Comparing the overall computed mortality rates that are only clinically adjusted (row 3 labeled "All Patients," unit A has the lowest rate, followed by units B and C (2.2%, 2.6%, and 2.9% respectively). This is an example of the current situation for performance measures, in which clinical adjustment is done, but SDS adjustment is not done.

Table 2. Illustration of Risk Adjustment Using Indirect Standardization

1		All Patients in National Population		Unit A		Unit B		Unit C	
2	**SDS Strata**	Patient Mix N/ Percent	**Clinically-Adjusted Deaths** N/Percent	Patient mix n/ Percent	**Clinically-Adjusted Deaths** n/Percent	Patient Mix n/ Percent	**Clinically-Adjusted Deaths** n/Percent	Patient Mix n/ Percent	**Clinically-Adjusted Deaths** n/ Percent
3	**All Patients**	1,000,000 100%	22,000 **2.20%**	1000 100%	22 **2.20%**	1000 100%	26 **2.60%**	1000 100%	29 **2.90%**
4	**Average to High Income**	800,000 80%	16,000 2%	800 80%	16 2%	400 40%	8 2%	400 40%	8 2%
5	**Low Income**	200,000 20%	6,000 3%	200 20%	6 3%	600 60%	18 3%	600 60%	21 3.5%
6	**Income-adjusted rate**				**2.20%**		**2.20%**		**2.45%**
7	**Calculation Details**								
8	Expected deaths Sum of: National stratum rate * unit number of patients in each category				2%*800 + 3%*200 = **22**		2%*400 + 3%*600 = **26**		2%*400 + 3%*600 = **26**
9	Standard ratio = clinically-adjusted/ expected deaths				22/22 = **1.0**		26/26 = **1.0**		29/21 = **1.115**
10	**Income-adjusted rate** Ratio *National rate				1.0* 2.2% = **2.20%**		1.0* 2.2% = **2.20%**		1.115* 2.2% = **2.45%**

- Unit A's sociodemographic case mix is the same as the national mix. Its performance is also the same as the national average for both the average-high and low-income categories (2% and 3% respectively). In both case mix and performance, then, it is exactly average. Adjustment for income using this method (row 6) does not change its rate (2.20%).

- Unit B's sociodemographic case mix (rows 4-5) has a higher proportion of low-income patients, but its performance is exactly the same as the national average as well as that of unit A for the two income categories (2% and 3% respectively). When its performance score is not adjusted for income (2.6%), its performance appears to be "worse" than unit A, but in fact it is not. When its rate is adjusted for clinical factors and income (row 6), its performance score is identical to that of Provider B (2.2%). This reflects that the question being addressed is: *how would the performance of various units compare if hypothetically they had the same mix of patients*?
- Unit C has the same sociodemographic mix of patients (rows 4-5) as unit B, but its performance is worse for the low-income group (3.5% vs. 3%). Its income-adjusted rate (row 6) is higher than unit B's income-adjusted rate, reflecting its poorer performance for its low income patients.

There are three important points to emphasize about this example.

- First, adjustment for income in this particular illustration does not "adjust away" the differences in results achieved between unit B and unit C. Unit C still has a worse performance score than either A or B after adjustment.
- Second, income disparities are clearly visible in the data for each stratum (rows 4-5), and they are actually a key part of the middle steps of the indirect standardization calculations. This is the data that would be available to identify disparities both across and within units with the recommended stratification.
- Finally, all three units in this scenario may have incentives to improve. In a "star system" of rankings, Units A and B might have "three-star" designation because their performance is just average. If rewards are given for four- or five-star performance, they both need to improve. Unit C may have a two-star designation depending on how cut points are set, but it also has a clear incentive to improve. It may be the case that both Unit B and Unit C find that their best opportunity for overall improvement is to improve care for their low-income patients because they comprise a substantial proportion of their population.

Neither the clinically adjusted nor the SDS-adjusted overall performance rates alone provide any information on disparities. Without the specific information on performance for income subgroups, the overall performance rates neither identify nor mask disparities. The subgroup scores that are included in this method do reveal the disparities, though. This particular adjustment method meets the Panel's general principles of transparency, attention to disparities, and validity and fairness of performance assessment.

Risk adjustment is not perfect, and the same limitation when adjusting for clinical factors applies to sociodemographic factors — that is, risk adjustment can only account for measurable and reportable factors. Additionally, risk adjustment procedures only address patient characteristics and there could be unit characteristics (e.g., funding of safety net providers, area healthcare workforce) that might have policy implications related to some accountability applications. Therefore, risk adjustment does not necessarily preclude using additional methods when comparing performance such as constructing peer groups for comparison as described below.

Alternatives to Risk Adjustment

Stratification

Stratification refers to computing performance scores separately for different strata or groupings of patients based on some characteristics(s) — i.e., each healthcare unit has multiple performance scores (one for each stratum) rather than one overall performance score. For example, strata could be constructed based on poverty level and performance scores computed for each stratum. Sometimes stratification is considered a type of risk adjustment as a means to making like comparisons; however, the Expert Panel thought statistical procedures such as multivariable models and stratification were so different, that they are considered separately. With stratification, performance is reported and can be compared for subgroups of patients with similar levels of risk or sociodemographic characteristics. It offers the advantage of allowing identification of disparities in healthcare for certain subgroups of patients because scores are associated with the particular SDS factor. In essence, stratification "unmasks" healthcare disparities by examining performance for groups who have been historically disadvantaged compared to groups who have not been disadvantaged.

An illustration of stratification appears in the table that follows. Note that stratification is essentially the first step in adjustment as illustrated in the example above. Stratification is most likely to be useful when examining performance for groups where substantive differences in performance have been observed. It is particularly useful for providing finer-grained information and most notably for assessing and addressing disparities.

The biggest barrier for using stratification alone for accountability applications is feasibility. Each healthcare unit's patient population is divided into the specified categories, thus reducing sample sizes available for analysis in each category. Sample size affects reliability and the ability to distinguish differences and make correct inferences. If there is more than one relevant sociodemographic factor (e.g., race, ethnicity, income, language, etc.) then stratification becomes much more complex, increasing the number of categories and further reducing sample size in each "cell" of the resulting matrix of stratification factors and levels. Combining individual factors into composites may address this problem to some degree, but stratification by itself does not address the problem of needing a single performance score for each unit for a given measure in order to use in either public reporting or pay-for-performance programs.

Table 3 shows a very simple example of stratification. A single sociodemographic variable (income, for example) is divided into three levels, and patients are assigned to one of the three levels. The size of the population at the national level may be in the hundreds of thousands or millions for each of the three strata, so performance rates are very reliable. In this illustration a higher performance rate is desired and indicates better quality. Unit A has reasonably large sample sizes in each stratum, and performance scores close to the national average in each. Its performance, for each stratum, would probably be identified as average. There is no direct way, in this example, to make a judgment about unit A's overall performance, although its higher proportion of patients in the "low" stratum would tend to make its overall performance without SDS adjustment appear to be worse than average, as shown earlier in Table 2. Disparities in performance across the three strata are evident, and are essentially the same as the disparities found at the national level. Unit B has a much smaller sample in each stratum, and also lower performance scores in each. It actually has lower disparities across strata than unit A, but its overall performance score without SDS adjustment would be worse. (Note, though, that the stratified report does not actually provide an overall score.) However, the small sample sizes in each cell may make it difficult to identify the performance as significantly worse than either unit A or the national average.

Table 3. Illustration of Stratification

	National		**Unit A**		**Unit B**	
Socio-demographic Stratum (e.g., income)	Percent of Patients	Observed Rate	n/ Percent	Clinically-Adjusted Rate	n/ Percent	Clinically-Adjusted Rate
Low	30%	63%	500 50%	65%	20 20%	60%
Moderate	50%	72%	400 40%	70%	50 50%	65%
High	20%	85%	100 10%	83%	30 30%	67%

Peer Groups for Comparison

Peer groups for comparison refers to creating peer groups of healthcare units caring for a similar mix of patients, within which to examine performance scores. It could facilitate comparisons of units with similar resources, e.g., VA sites with VA sites, or federally qualified health centers with each other. Depending on how the peer groups are constructed, it also tends to match patient populations, e.g. proportion of uninsured patients or those covered through Medicaid. In this approach, performance scores for individual units are neither adjusted nor stratified for sociodemographic factors (using the definition of stratification in this report). Constructing peer groups for comparison occurs after performance scores are computed. This approach avoids the issue of reducing sample sizes seen with stratification. Recently, MedPAC recommended using this approach with the readmission reduction program.[36] Peer groups can help ensure that use of a performance measure to apply rewards or penalties is consistent with program and policy goals. For example, if units caring for a disproportionate share of disadvantaged patients will be disproportionately penalized using non-SDS-adjusted performance scores, then that may not be a desirable result. Applying the penalty on the basis of performance within groups of "peer units" rather than on the basis of performance relative to the entire universe of units is one way to avoid a disproportionate share of penalties to safety-net units. Adjustment of a performance score for sociodemographic factors would not always or automatically exclude the possibility of also using peer groups for comparison.

Table 4 presents a simple example of use of peer groups to establish different reference points for different units, which then could be used to apply financial rewards or penalties or to identify providers as relatively better or worse within that peer group. In this example, units (e.g., hospitals) are grouped into "quintiles" based on the percent of their patients at or below 138% of the federal poverty level. Hospitals in quintile 1 have relatively few such patients; hospitals in quintile 5 have many. The

performance measure here is something for which "more is better" — for example, percent of acute stroke patients arriving at the emergency department within two hours of symptom onset. There are clear disparities in performance on this measure at the national level in this example.

Units A and B are both assigned to a quintile based on their specific percent of patients at or below 138% of the federal poverty level. Unit A is in the middle quintile and unit B is in the fifth quintile. Unit A's performance is a bit worse than the quintile average (62% vs., 65%), so it might be identified as a "below average" performer for its quintile. Unit B's performance is a bit better than average for its quintile (59% vs. 55%), so it might be identified as an "above average" performer for its quintile, even though its performance is worse than A's in absolute terms.

The Panel had generally favorable views of this approach as a method to more fairly apply financial rewards and penalties. However, determining appropriate peer groups can be challenging. The method does not, however, identify disparities in care within units, nor does it indicate whether unit A or B is better than the other if scores had been adjusted for patient income, either for specific subgroups of patients or overall. It is possible, depending on the exact distribution of patients across income strata for the two units, that unit B would have a better score with an adjustment approach like that illustrated in Table 2. Additionally, some view this approach as more explicitly setting different benchmarks for healthcare units based on the proportion of disadvantaged patients served.

Table 4. Illustration of Peer Groups for Comparison

Quintile Based on Percentage of Low Income Patients	Number of Units	Quintile Cut Point for Percentage of Patients at or Below 138% of Poverty	National Average Clinically-Adjusted Rate for Units in Quintile	Unit A Clinically-Adjusted Rate	Unit B Clinically-Adjusted Rate
National	3000		70%		
1st Quintile	600	10%	75%		
2nd Quintile	600	20%	70%		
3rd Quintile	600	30%	65%	62%	
4th Quintile	600	40%	60%		
5th Quintile	600	55%	55%		59%

The three general approaches described here — statistical risk adjustment, stratification, and peer groups for comparison — are not mutually exclusive. They could be used in various combinations or in all three ways for a given performance measure, with the specific analytic approach chosen for a specific analytic or program purpose. In an analysis focusing on the presence of sociodemographic disparities in care, for example, stratification would be the natural first-choice approach, as it provides the clearest and simplest information about performance in relation to a particular sociodemographic factor. For some program purposes, like application of a hospital readmission penalty, a peer-grouping approach might be simplest and most desirable. Each has different strengths and limitations. The Panel concluded that different approaches serve different purposes. A strong majority of Panel members did not think, however, that either stratification or creating peer groups would be adequate for all "accountability" measurement purposes. When single performance scores are interpreted as indications of underlying quality of care, the large majority of the Panel thought that statistical adjustment for relevant sociodemographic factors when indicated would be necessary to support valid inferences about quality and that stratification was needed to assess and address disparities.

Section 6: Guidelines for Selecting Risk Factors

The Expert Panel reviewed the widely accepted guidelines for selecting clinical or health status risk factors and their rationales (Table 5). The Panel determined that these same guidelines may also be applied to sociodemographic factors. As indicated in recommendation 1, several conditions must be met before a performance measure is adjusted for SDS. These conditions are consistent with selecting clinical risk factors. Each performance measure must be assessed individually.

Conceptual and Empirical Relationship for SDS Adjustment

The first condition for selecting risk factors is that a conceptual relationship and an empirical relationship exist between the specific risk factor and the outcome (or process) being measured. A conceptual relationship refers to a logical theory or rationale that explains the association. The conceptual basis may be informed by prior research and/or healthcare experience related to the outcome of interest, but does not require a direct causal relationship (i.e., it could be a direct cause, an indirect cause, or serve as a surrogate for a cause for which data are lacking). An empirical relationship means that there is a statistical association between variables for the risk factors and the variable for the outcome.

Not all outcomes or processes of care are affected by sociodemographic factors. For example, outcomes and processes such as the outcome of central line infection occurring during a hospital stay, or the process of administering the correct medication at the correct time during a procedure, would not have a conceptual reason for a relationship with sociodemographic factors. One would expect the same things to be done, and the same results obtained, for any and all sociodemographic subgroups. Further, not all sociodemographic factors may affect all outcomes. For example, improvement in ambulation has no conceptual relationship to race, but does to age.

The recommendation on SDS adjustment also allows that SDS adjustment might be unnecessary or inappropriate based on conceptual reasons or empirical evidence. Some examples include whether the influence of the SDS factor is exerted primarily through the quality of care delivered; empirical analyses that indicate the potential factor does not account for variation in the outcome being measured; or empirical analyses that indicate that the effect is through disadvantaged patients being clustered in poorer quality units (as discussed in section 4).

An assessment of a conceptual relationship between an SDS factor and an outcome of interest includes a consideration of whether the effect of the SDS is primarily mediated by the quality of care delivered. That is, does the SDS factor lead to the delivery of inferior care processes, which in turn affect the outcome? An obvious example is unequal treatment to patients with a particular characteristic such as race or homelessness where they are consistently skipped in routine screening for hypertension, which leads to higher rates of blood pressure greater than 140/90. If this was the general and pervasive practice for those patients, it could be reason enough to not consider those SDS factors for risk adjustment, even if they have an empirical association with the outcome. The underlying mechanisms for the effect of specific SDS factors on specific outcomes may be complex, involving multiple paths, or essentially be unknown without additional study. As discussed in the methodological basis for risk adjustment, the exact mechanism of the effect of a factor on an outcome does not need to be known in order to consider it a potential confounder. Adherence to the epidemiological and statistical methods and principles related to confounding as well as the guidelines for selecting risk factors are used in

conjunction with conceptual considerations to inform decisions about whether or not to adjust for SDS factors.

Some potential questions for identifying a conceptual basis for adjusting a performance measure for sociodemographic factors include:

- Does prior research indicate a relationship between SDS and the outcome?
- Is there a logical relationship or theory about the relationship between SDS and the outcome?
- Is there a significant passage of time between the healthcare unit intervention and measured outcome during which other factors may have an effect?
- Do patient actions or decisions influence the outcome or process and are the decisions affected by SDS (e.g., ability to purchase medications)?
- Does the patient community have an influence (e.g., distance to pharmacies, groceries, healthcare services)?

If a conceptual relationship exists between a patient-level sociodemographic factor and outcome, it should be tested empirically. The Panel did not specify, and does not recommend, any particular analytic approach with which to assess empirical associations between sociodemographic factors and outcomes, nor any specific cutoff or threshold value to use for declaring the presence of an association. A common method to identify an empirical relationship is to assess the correlation between the two variables. For example, as income increases, mortality decreases. If the basic conditions for conceptual and empirical relationship are met, then SDS factors will be assessed for inclusion in risk adjustment procedures following the remaining guidelines for selecting risk factors.

Some have advocated that sociodemographic factors affect clinical and health status and therefore, may already be accounted for through those risk factors. That is a possibility that can be tested empirically. It is also important to consider that if sociodemographic factors lead to less use of healthcare services, data on health status and clinical conditions prior to the start of care may not exist to the same degree for disadvantaged patients as for those who use healthcare services more frequently and result in underestimation of clinical severity.

Risk factors should meet the first two and most of the other guidelines listed in Table 5, but not necessarily all of them. Developing a risk model is an iterative process that at times requires weighing various trade-offs.

Table 5. Guidelines for Selecting Risk Factors for Adjustment

Guideline	Rationale	Clinical/ Health Status Factors[b]	SDS Factors[c]
Clinical/conceptual relationship with the outcome of interest	Begin with conceptual model informed by research and experience to ensure relationships are not unique to any one data set	✓yes	✓yes
Empirical association with the outcome of interest	To confirm conceptual relationship	✓yes	✓yes
Variation in prevalence of the factor across the measured healthcare units	If there is no variation in prevalence of the factor across healthcare units being measured, it will not bias performance results	✓yes	✓yes
Not confounded with the effect of the healthcare unit — risk factors should:	Trying to isolate effects of the healthcare unit — quality of care	✓yes	✓yes
• be present at the start of care and	Ensures not a result of care provided	✓yes	✓yes
• not be an indicator or characteristic of the care provided (e.g., treatments, interventions, expertise of staff)	Although these could explain variation in the outcome, in performance measurement the goal is to isolate differences in performance due to differences in the care provided	✓yes	✓yes
Resistant to manipulation or gaming — generally, a diagnosis or assessment data (e.g., functional status score) is considered less susceptible to manipulation than a clinical procedure or treatment (e.g., physical therapy).	Ensures validity of performance score as representing quality of care (vs. for example, upcoding)	✓yes	✓yes
Accurate data that can be reliably and feasibly captured	Data limitations often represent a practical constraint to what factors are included in risk models. Generally, measurement error will be reflected in weaker correlations.	✓yes	✓yes
Contribution of unique variation in the outcome (i.e., not redundant or highly correlated with another risk factor)	Prevent overfitting and unstable estimates, or coefficients that appear to be in the wrong direction; reduce data collection burden	✓yes	✓yes

[b] Examples of clinical and health status factors include comorbidity; severity of illness; patient-reported health status, etc.
[c] Examples of sociodemographic factors include income; education; English language proficiency, etc.

Guideline	Rationale	Clinical/ Health Status Factors[b]	SDS Factors[c]
Potentially, improvement of the risk model (e.g., risk model metrics of discrimination — i.e., sensitivity/specificity, calibration) and sustained with cross-validation	• Change in R-squared or C-statistic may not be significant, but calibration at different deciles of risk might improve • May not appear to be a big change but could represent meaningful differences in terms of the outcome (e.g., lives, dollars) • Order of entry into a model may influence this result	✓yes	✓yes
Potentially, face validity and acceptability	Some factors may not be indicated empirically, but could improve acceptability — need to weigh against negative impact on model, feasibility, and burden of data collection	✓yes	✓yes

Information Submitted for Review and Evaluation for Potential Endorsement

The Expert Panel recognized that developing adjustment strategies for performance measures is an iterative process involving a conceptual basis and empirical analyses resulting in multiple decisions to arrive at a final risk adjustment procedure. There is more than one appropriate way to accomplish adjustment. Therefore, NQF should not be prescriptive regarding methods for adjustment or specific SDS variables. However, steering committees and stakeholders need to have sufficient information to evaluate performance measures for endorsement. When a measure is submitted to NQF for potential endorsement, it is important that the developer's rationale regarding adjustment for sociodemographic factors be transparent and open to review and evaluation.

In addition to the adjustment methods, factors, and rationale, the developer should discuss the potential risk of misuse of the measure. NQF already requires information on current and planned use of measures. The developer has detailed knowledge about the limitations of the performance measure that could impact its use in accountability applications.

The Expert Panel identified the following as important information for reviewers to evaluate whether SDS adjustment is appropriate.

- Conceptual description (logical rationale or theory informed by literature and content experts) of the causal pathway between sociodemographic factors, clinical factors , quality of care, and outcome

- Patient-level sociodemographic variables that were available and analyzed, for example:
    - Patient-reported data (e.g., income, education, language)
    - Proxy variables when sociodemographic data are not collected from each patient (e.g., based on patient address and use of census tract data to assign individual patients to a category of income, education, etc.) and conceptual rationale for use

    - Patient community characteristics (e.g., crime rate, percent vacant housing, smoking rate, level of uninsurance) assigned to individual patients for the specific community where they live — see discussion of community variables in section 7

- Analyses and interpretation resulting in decision to include or not include SDS factors. For example:
    - Prevalence of the factor across measured entities
    - Empirical association with the outcome
    - Contribution of unique variation in the outcome
    - Assessment of between-unit effects vs. within-unit effects as discussed in the methodological discussion in section 4

- Current and planned use of the measure and a discussion of risks for misuse of the specified performance measure

Section 7: Specific Sociodemographic Factors to Consider for Adjustment

Adjustment of the performance score generally involves patient-level data for the risk factors — i.e., an individual patient's diagnosis, lab value, income, education, etc. Although the Expert Panel agreed that performance measures should be adjusted for sociodemographic factors when appropriate, it also recognized the data challenges that constrain adjustment. Data about patient sociodemographic factors other than age and sex often are not collected, or not standardized sufficiently for use in performance measurement.[37] Collection of race and language by healthcare units is growing but SES-related data are not widely collected. Therefore, data availability is a critical consideration. Besides overcoming prior assumptions, data constraints may be the biggest barrier to adjustment for sociodemographic factors and will require further initiatives to define standards and to implement data collection.

When sociodemographic data are not collected for each patient, other methods may be used to assign a value for each patient (e.g., based on census data for the patient's home address or ZIP Code). Just as whether sociodemographic variables are used in adjustment should be based on conceptual relationships, use of proxies for patient sociodemographic data should also have a conceptual basis. For example, data for the area where a patient lives could be assigned as a crude proxy for individual SES, or as Krieger[38] suggests, could characterize the patient's environment. That is, if one uses census data on income for a given patient's neighborhood, one can either be saying "I think you're probably poor because you live in this neighborhood" or "You live in a neighborhood with mainly poor people in it."

The Expert Panel identified potential sociodemographic factors that might be useful for adjustment and discussed some of the pros and cons when considering those factors for adjustment. However, the Panel did not recommend specific variables to be used — that will depend on applying the guidelines for selecting risk factors for a particular performance measure, as well as on data availability.

Age is also considered a clinical variable as well as a demographic variable. Physiologic changes accompany age and the probability of disease increases with age. Age is already included in many clinical risk adjustment procedures and should continue to be utilized as indicated by the conceptual and empirical relationships with the measured outcome.

Socioeconomic Status (SES)

SES arguably represents a fundamental determinant of health,[39] and access to and use of healthcare.[40] SES represents a multidimensional construct that has been traditionally measured based on income, education, and occupation (although much greater attention has been given to the first two dimensions).[41]

Income

Income is a key dimension of SES. It affects health over the life course and healthcare access and affordability. These effects have been extensively documented.[1, 42-44] Optimally, household income should be collected directly from patients. This is currently done in selected instances, (e.g., to assess eligibility for charity care, subsidies for health insurance on the exchanges), but it is not widely collected in healthcare. A key barrier is reluctance to ask all patients about their income (potentially resistance from both patients and healthcare units).[45] A second barrier is that income is difficult to measure because household income can come from multiple sources for each person within a household.[46] A full assessment requires multiple questions.

Income variables need to be considered in light of variations in cost of living and purchasing power across the U.S. For national use, consideration should be given to standardization by wage or cost-of-living indexes.

When individual or household income data are lacking, proxies based on residence may be used.[38, 47] Area-level data may be used either to impute characteristics of individuals or to characterize the areas in which people live, and there is a rich literature on pros and cons of either usage. Area-based measures of income can be based on patient addresses geocoded to the Census Tract, Block Group, or Block. ZIP Codes can be linked to census data; however, ZIP Codes are limited because of greater socioeconomic heterogeneity within the area.[38] Smaller, less heterogeneous areas may yield more valid results when used as a proxy for individual income. The Geocoding Project showed findings regarding the association of SES with mortality and with cancer incidence were most consistent when addresses were geocoded at the Census Tract than at the ZIP Code or Census Block Group. Recent developments have improved the matching of addresses to areas and have minimized failures to successfully geocode addresses.[48-50]

Medicaid eligibility or dual eligibility for Medicaid and Medicare often is used as an indicator of low income. Although there is significant heterogeneity in Medicaid eligibility, benefits, and payments between states, it is a verified indicator of low income and the information is widely available. Expanded eligibility for Medicaid through the Affordable Care Act (ACA) represents a verified measure of household poverty (i.e., <138% federal poverty). Currently, 25 states, in addition to the District of Columbia, have opted for expansion. This expansion presumably will grow over time providing for a standard measure of poverty across states with similar eligibility. However, some low-income people will not be eligible for Medicaid with the ACA expansion due to immigration status or other reasons.

Education

Education represents another dimension of SES. It is powerfully related to health, health behavior, and healthcare.[1, 51, 52] Like other measures of SES, patient/parent education level varies across health plans andhospitals.[53, 54] Nearly two decades ago, the National Committee on Vital and Health Statistics proposed that education (i.e., years of schooling) is a core health data element that should be standardized in healthcare and healthcare information fields.[55] Despite this recommendation, education is not widely collected in healthcare outside of patient experience of care surveys (e.g., CAHPS) and is inconsistently collected by clinical personnel as part of the social history of a patient that is included in the medical record. In contrast to household income, education may be easier to collect from patients with fewer refusals.[45] Regulations and promotional efforts have fostered collection of race, ethnicity, and language among hospitals[56] and health plans.[57] Similar approaches could be used to promote collection of individual patient educational attainment within structured data fields (that can be exported). Until these data become available, community-based measures (discussed in more detail below) may be used as crude proxies.[58, 59] Standardized collection of patient (or parental education) in healthcare would obviate use of imputed measures of patient education. This represents an important priority related to improved measurement of SES in healthcare. An IOM report on optimal social and behavioral measures for collection in EHRs recommended inclusion of patient educational level.[60]

Occupation/Employment

Occupation is the third dimension of SES. Employment status is more easily obtained than occupation and potentially relevant given its relationship to health insurance, health behavior, and mortality, and represents an additional potential adjustor.[61-64] Existing methods for classification of occupations have limitations.[41] Moreover, relatively little is known about its effect on outcomes independent of other

measures of SES and sociodemographic-related factors. Obtaining standardized occupational data from patients generally does not lend itself to single questions.[65]

Sociodemographic Factors Related to SES

Language

Limited English proficiency (including communicating through American Sign Language) contributes to suboptimal healthcare, inadequate informed decisionmaking, poor self-management, and healthcare disparities.[66-72] These barriers persist despite language assistance regulations[73] and the recognition of language differences as barriers to quality and safety by The Joint Commission.[74] A 2009 Institute of Medicine (IOM) report recommended standardized data collection for language in addition to race and ethnicity.[75] Subsequently, progress has been made by hospitals and health plans in the collection of these data using a combination of direct and indirect methods.[76-78]

Insurance

The uninsured disproportionately includes minorities, the poor, those with low education, and those with limited English proficiency.[79] Health insurance is strongly associated with healthcare use, improved preventive and chronic care management, and reduced mortality for children and adults.[1, 80-83] The presence or absence of insurance may be useful for adjusting quality performance measures. An important related measure is underinsurance.[84-86] Out-of-pocket payments not covered by health insurance affect patients' healthcare decisions, particularly among poorer patients.[85] Optimally, data on the quality of insurance analogous to the designations for insurance purchased on health exchanges, (i.e., bronze, silver, gold, and platinum) could be collected to assess patient underinsurance.

Race and Ethnicity

Race and ethnicity are not and should not be used as proxies for SES; rather, their effects are confounded by SES.[87] That is, income, education, and related factors (including language and insurance) represent key contributors to racial and ethnic disparities in healthcare.[1, 42] Potential mediators of the effect of race on outcomes include source of care,[19, 88] discrimination,[89] and potential differences in biology (including those that are environmentally- or stress-induced). Potential biological effects include high rates of preterm birth among African Americans[90] and differences in levels of glycated hemoglobin between blacks and whites.[91-93] For other outcomes, such as hypertension control, there is conflicting evidence as to whether factors such as discrimination, fear of side effects, and/or adherence to treatment plan fully account for disparities in blood pressure control or not.[94, 95] Although some see race/ethnicity technically as no different than other potential confounders, because of the concerns about bias and racism, careful thought, consideration, and a clear rationale should be used when adjusting performance measures for race and/ethnicity. At the same time, reporting of data stratified by race and ethnicity should be encouraged to assess and address disparities in healthcare. Collection of race and ethnicity data is improving, but gaps remain hindering use of these data.[96, 97]

Homelessness

Homelessness is associated with poor healthcare access and high levels of unmet healthcare needs, poor health, and hospital re-admission.[98-101] However, patients frequently are not asked about their housing status, even during hospitalization.[102] Standardized definitions for homelessness have been developed and are used by Housing and Urban Development.[103]

Marital Status

Marital status is strongly associated with household income. It is not only related to health behaviors, but to health and mortality, particularly following disruption through divorce or death.[104-109] It is easily and often collected along with other demographic factors in the process of hospital admission or clinic registration. Marital status is also strongly related to the of caregiver availability that is known to be related to health outcomes in post-acute settings.

Literacy and Health Literacy

Literacy (ability to effectively read and write), numeracy (ability to understand and use numbers in daily life),[110] and health literacy (capacity to obtain, process, and understand basic information and services needed to make appropriate decisions regarding health)[111] are associated with educational attainment.[112] Both general literacy (and numeracy) and the related construct of health literacy are strongly associated with healthcare use and outcomes.[111, 113, 114] Brief screening tools show promise for health literacy.[115]

Community Variables

Risk factors are considered patient-level characteristics, and in that context, "community" refers to the community where the patient resides, not the community where the healthcare unit is located.

Patient Living Environment

Community variables could be used at the individual level to characterize the environment in which the patient lives. Community variables include the geographic distance to pharmacies, availability of public transportation, types and availability of food outlets, neighbor and social support infrastructure, and availability of parks and recreation areas. These may be as, or more, important than individual SDS characteristics in terms of accounting for access to economic and social infrastructure and healthcare services, all so important to good health outcomes. In rural communities, this includes the geographic distance to healthcare providers. Other examples include rates of crime or percentage of blue collar or professionals residing in the area.[116, 117] Because multiple variables of social disadvantage by Census area are available, some researchers have used composite measures based on factor analysis.[47, 118] However, a single measure (poverty) may perform as well as composite measures.[38]

Proxy for Patient-Reported Data

Community-based measures of SES have been used to characterize SES of patients in health plans and quantify socioeconomic disparities in quality.[58, 119, 120] The specific variables selected and how they are used should be based on the conceptual model. While community-based measures potentially will misclassify some individuals when used to impute individual-level characteristics due to socioeconomic heterogeneity within the area being measured,[121] they offer the potential for capturing contextual effects beyond individual measures including insurance availability or public support for health care.[122-124]

Community Factors Affecting the Healthcare Unit

Some community characteristics are most relevant as characteristics of the healthcare unit not the patient, for example, funding for safety net providers (a function of local and state taxing jurisdictions and associated public funding or lack thereof) and the pool of available healthcare workers for employment. Because they are not characteristics of individual patients, they would not be included in

risk adjustment procedures as discussed in section 4. How these types of factors may have implications for policy responses to performance assessment and need to be further explored.

Potential Mediators of Sociodemographic Factors

There are a number of potential mediators between sociodemographic factors and outcomes. Examples include social support (and its converse, isolation and loneliness),[125, 126] and "patient activation" which refers to patient confidence and skills needed to assume shared responsibility for their health and healthcare.[127] A range of behavioral factors, including smoking, alcohol use, physical activity, and diet,[128] may be mediators of effects of some sociodemographic factors; however, these are more likely than SDS factors to be included in clinical risk adjustment models, along with self-reported health status.[129]

Consideration of Specific Sociodemographic Factors

As previously noted, selection of sociodemographic risk factors should first be guided by the conceptual relationships, but before any analyses of relationships with outcomes can be conducted, the data must be available. There may be several options for operationalizing a sociodemographic concept, and the Expert Panel identified some of the pros and cons of various variables to consider when selecting variables for sociodemographic adjustment (see Table 6).

Table 6. Sociodemographic Factors – PROs and CONs

Factors/Concepts (specific variables)	**PROs**	**CONs**	**Caveats**
Factors that should be considered, depending on: data availability and the specific outcome or process			
Income	• Allows for use of various ranges	• Hard to collect privately (e.g., in clinician office) • Not easily collected with a single question • May not be an acceptable question to all patients • Meaning is not geographically consistent due to difference in costs of living	• For national performance measures, need to consider standardization to account for area wage and cost of living differences
Income in relation to federal poverty level	• Definition is standard • Being used under ACA • Researchers are used to using it	• Doesn't include receipt of other benefits (e.g., food stamps) • Doesn't account for cost of living or community offsets	.
Household income	• May be more meaningful than individual income	• Requires assessment of household size	.
Medicaid status as proxy	• Relatively easy to collect in claims data	• Eligibility not consistent across states	• Potentially becomes more useful as more States expand Medicaid to 138% federal poverty level

Factors/Concepts (specific variables)	PROs	CONs	Caveats
Social Security Supplemental Income (SSI)		• Correlated with Medicaid status, but not consistently across states	• In many states, receipt of SSI automatically makes one eligible for Medicaid
Education	• Perceived to be valid (i.e., less misreporting than for income) • Definitions fairly consistent across various subgroups (e.g., answers from immigrants comparable to those from others) • Fairly stable across time, at least after a certain age	• Not widely collected by healthcare units • If collected (e.g., in EHR text fields) may not be easily retrievable	
Homelessness	• Strongly associated with health outcomes • Measures something "beyond" income • Current HUD definition	• Multiple other definitions • Data often not collected • Status can change	• Prevalence tends to cluster among safety net healthcare units
Housing instability	• May be better indicator than homelessness which can change	• More difficult to define than homelessness	
English Proficiency	• Standard definition exists • Tied to need for translation services/other resource needs and therefore should be collected • Increasingly being collected (required by “Meaningful Use” and some states)		
Insurance Status	• Readily available • Some indication of access and resources • Benefit coverage strongly related to affordability	• Wide variability in insurance coverage • Data for underinsurance not widely collected	
Medicaid status	• Readily available • Some indication of limited income and resources	• Not consistent across states	
No insurance	• Readily available • Standard meaning		• Difficult to capture information about these patients (particularly if using claims data)

Factors/Concepts (specific variables)	PROs	CONs	Caveats

Community/ Neighborhood-level data used as proxy for individual data or as contextual variable	• Many variables available from Census data
	  • Income
	  • Education
	  • Immigration status
	  • Language
	  • Unemployment
	  • Home ownership
	  • Single parents
	  • Others
		• Census data do not include all potentially important variables
		• Residential heterogeneity will affect whether it is a good proxy for data about individuals.
		• Heterogeneity may differ based on levels of socioeconomic segregation and potentially population density.
		• Requires geocoding for Census Tract and smaller areas.	

Contextual - Proportion vacant housing	• Seen as indicator for other related issues such as poverty, crime, lack of resources		
Contextual- Crime rate	• May be an indicator for other related issues such as poverty, lack of resources		
Other factors that could be considered			
Factors/Concepts (specific variables)	**PROs**	**CONs**	**Caveats**
Social Support	• Some brief items have been used in previous research • Captures something that other variables do not	• Multidimensional construct that typically requires multiple questions • Lack of agreement about how to measure • Not consistently measured	
Living alone	• Available in OASIS data for home health	• Directionality may not be consistent. In some situations such as frailty or impairment, it could be a risk factor. In other situations, it might be an indicator of ability to live alone due to good health and function.	
Marital status	• Often collected		

Factors/Concepts (specific variables)	**PROs**	**CONs**	**Caveats**
Occupation	• May capture other concepts (e.g., environmental exposures)	• Multiple definitions • Potentially large data collection burden due to the complexity of the concept • Marginal value (i.e., over and above that contributed through use of other variables) may be limited • Unclear how to handle certain population subgroups (e.g., retirees, students, homemakers)	
Employment Status	• Often collected	• Employment status does not reflect income or availability of insurance • Simple yes/no does not reflect desire/happiness with situation (e.g., retirees may be happy to be unemployed) • Subject to change requiring continuous updating	
Literacy	• This concept may also be able to partially capture health literacy	• No standardized definitions • May be easy to game	If the correlation with education is high, then education could be used.
Health Literacy	• Potentially more relevant to healthcare • Three-item and single-item validated questions exist	• Not consistently collected/ available	
Local/state funding for safety net providers (e.g., tax base)	• Affect resources available to safety net providers beyond insurance	• Data not easily collected/ available	• Not a patient characteristic • Risk for unintended consequences (setting a lower standard for poorly supported institutions might send the wrong messages to tax payers)
Race/ Ethnicity	• Correlated with SES and may be more available than other variables	• May be more correlated with bias	• Should not generally be used as proxy for SES

Section 8: Policy-Related Discussion

Use of Performance Measures in Accountability Applications

NQF-endorsed performance measures are expected to be used in accountability applications such as public reporting and pay-for-performance. The NQF criteria focus on endorsing measures that demonstrate reliability and validity and adequate risk adjustment so that correct conclusions about the quality of care can be made by patients and others. NQF does not set different reliability and validity standards for different accountability applications. As already noted, concerns have been expressed about the policy response to performance results. For example, if providers or health plans serving disadvantaged populations have poorer outcome performance and incur financial penalties, it could worsen disparities in health and healthcare by reducing resources available to care for their patients. Therefore, it is imperative that various accountability applications be assessed for the potential impact on providers and plans caring for disadvantaged populations to identify unintended consequences to patients and to ensure alignment with program and policy goals.

Even if a performance measure is adjusted for sociodemographic factors, it does not rule out the potential need for also creating peer groups for comparisons in various accountability applications. Even when risk adjustment includes relevant patient-level factors, it may not fully account for differences in risk across units when patient mix differs widely across healthcare units or due to data limitations.

When a measure is submitted to NQF for endorsement, information on current and planned use should be submitted. Currently, NQF criteria and endorsement do not include requirements for or evaluation of procedures for implementation and reporting of the computed performance measure score (e.g., reporting with or without confidence intervals or sample sizes; methods for determining rankings or ratings, statistically significant differences, or incentives and penalties). However, the way a measure is implemented involves multiple decisions that could affect the validity of conclusions (inferences) made about quality of care and create potential unintended consequences. For example, cut points based on rankings of performance scores without confidence intervals could result in different classifications (conclusions) about quality without any significant difference in performance for units above or below a cut point (i.e., confidence intervals for scores above and below a cut point may overlap). Review of the detailed information about the performance measure for potential endorsement provides an opportunity to identify any specific considerations or limitations for use in specific accountability applications.

Because of the above concerns, the Expert Panel recommended that NQF should consider expanding its role to include guidance on implementation of performance measures. Possibilities to explore include:

- guidance for each measure as part of the endorsement process; or
- standards for different accountability applications (e.g., use in pay-for-performance versus pay-for-improvement; innovative approaches to quality measurement explicitly designed to reduce disparities).

Some Panel members expressed concern about endorsed measures being used inappropriately, and the Expert Panel recommended that NQF should make explicit the existing policy that endorsement of a performance measure is for a specific context as specified and tested for a specific patient population (e.g., diagnosis, age), data source (e.g., claims, chart abstraction), care setting (e.g., hospital, ambulatory care), and level of analysis (e.g., health plan, facility, individual clinician). This policy is implicit in the

current NQF criteria and process for endorsing a measure as specified and tested, but the Panel expressed concerns about inappropriate application of modifications to endorsed measures.

Use of Performance Measures to Identify and Reduce Disparities

Recommendation 1 acknowledges that when a performance measure is SDS-adjusted, it is disparities sensitive. The second part of the recommendation states: the performance measure specifications must also include specifications for stratification of a clinically-adjusted version of the measure based on the sociodemographic factors used in risk adjustment.

A single performance score (whether adjusted or not adjusted for sociodemographic factors) neither identifies nor masks disparities—that requires the additional information about the characteristics of the patients served. In other words, the current system of performance measurement does not allow disparities to be identified so that they can be eliminated. Doing so requires analysis of performance measures that are stratified as recommended with NQF-endorsed disparity sensitive measures.[130] Hence, the Panel made this recommendation. This approach also helps address concerns about masking performance for disadvantaged groups and represents an important step for ensuring high-quality care for all.

A variety of analytic approaches potentially could be useful for identifying disparities. Performance on a measure could be analyzed by key sociodemographic variables at different levels of analysis such as clinician, facility, or population. As noted above, indirect standardization is based on identifying various categories that could be examined by population and healthcare unit. Multivariable statistical model analysis can provide information about the strength of association of specific factors and how much additional variation in an outcome is accounted for by the variable. However, the Expert Panel recommended stratification as defined in this report to identify disparities and opportunities to reduce disparities.

The Expert Panel did not identify how best to operationalize the use of stratified performance data to identify and reduce disparities. Stratum-specific rates for each unit could prove useful to providers, plans, policymakers, researchers, and the public. However, mechanisms for making detailed data available do not widely exist. How to move toward meaningful use of data and shared accountability for identifying and reducing disparities is a topic that a standing disparities committee could address.

Healthcare units need to know whether their performance differs between groups based on sociodemographic factors within their own patient population. Units also might want to know how their performance with certain groups compares to that of other units. Such data also could prove critical in designing and implementing policies, strategies, and/or programs to improve healthcare equity. Policymakers could use such stratified data to inform funding allocation decisions (e.g., payment rates based on the sociodemographic characteristics of the population). Stratified data could also inform funding for targeted programs such as patient navigators, community health workers, improved access to language services, and other programs designed to mitigate disparities.

The Expert Panel discussed the benefits of transparency with stratified results, but did not resolve how best to present the additional detail in addition to sociodemographic-adjusted scores. Some individuals might find stratified data useful to identify which healthcare units would be best for patients similar to themselves (e.g., income, language, race, ethnicity). CMS or other producers of performance reports

should make such stratified performance data available when feasible and relevant (e.g., through hyperlinks). At a minimum, it should be publicly available through a clear-cut process for interested parties to request such data. Alternatively, the underlying data needed to construct the stratified performance scores for healthcare units could be made available upon request. Some key issues to be resolved include:

- potential confusion if data are reported more than one way;
- cautions about reliability when cell sizes become quite small; and
- how to construct strata and make drill-down data useful given the potential for use of multiple SDS factors

This is clearly an area where more work needs to done and would benefit from a standing disparities committee. Given the direct relevance of stratified performance data to improved healthcare equity, this is an area where payers such as CMS, states, and health plans could take the lead (as some have done).

Payment and Responsibility for Mitigating Effects of Sociodemographic Factors

During its deliberations the Expert Panel identified two related policy concerns — adequate payment to reflect higher intensity of services to disadvantaged populations and responsibility for mitigating the effects of sociodemographic factors. These concerns, briefly described below, extended beyond the scope of this project but have substantial policy implications.

Disadvantaged populations may have needs that require greater resources. Current payment systems better align resources with clinical/medical needs of patients than services to mitigate the effects of sociodemographic factors. This failure to align payment with supportive patient services for disadvantaged patients creates a mismatch between healthcare unit capacity and the needs of the patient population, thereby creating a potential for worse performance. There are some examples of attempts to adjust payments for services provided to address higher need for resources related to sociodemographic factors. Some examples of this type of payment adjustment are 1) hospital payment adjustment disproportionate share (DSH) of certain low income patients (see overview of Medicare hospital payment); and 2) inclusion of Medicaid status in case-mix adjustment for Medicare Advantage plans (see overview of Medicare Advantage payment). It was beyond the scope of this project to address the adequacy of payment adjustments related to sociodemographic factors. Nonetheless, improved alignment between payments for services and the needs of the patient population served by that unit could potentially partly mitigate the negative effect on patient outcomes. Much of the debate about adjusting for sociodemographic factors relates to setting appropriate expectations for investment in care for disadvantaged patients and concerns about which entity should be incentivized to do so.

Some question whether greater payment to address the needs of sociodemographically complex patients would eliminate the need for SDS adjustment of performance measures. There is some parallel here to clinical factors, where current case-mix payments to healthcare units aim to account for patient morbidity and severity (and thus the need for more costly care); however, performance measures are still risk adjusted for clinical complexity. Similarly, if resources targeted to address the needs of sociodemographically complex patients eventually reduce disparities, the effects of SDS on patient outcome and performance measurement will be reduced, but it likely would still be necessary to risk adjust for SDS complexity.

Finally, an important related issue is identifying who is responsible for mitigating the effects of sociodemographic factors on health and healthcare and paying for those efforts. Where does healthcare responsibility end and community responsibility begin? Should the costs of language translation be covered by the community (e.g., multipayer consortium or borne by each healthcare unit, perhaps through enhanced payments)? There are notable examples of extraordinary efforts by healthcare units to address sociodemographic factors such as funding hospice beds for terminally ill homeless patients or providing translators for a large number of languages. These types of efforts require resources above and beyond typical healthcare reimbursement. Just as important a question as who is responsible is the question, what is the most effective and efficient approach to address social determinants of health?

Full discussion and resolution of the related issues of payment and responsibility for mitigating the effects were beyond the scope of the Expert Panel's charge, but the recommendations represent a widely-held view among Panel members that improving equity in outcomes will require greater investments.

References

1. Agency for Healthcare Research and Quality (AHRQ). *2012 National Health Care Disparities Report.* Washington, DC: AHRQ; 2013. Report No: 13-0003. Available at http://www.ahrq.gov/research/findings/nhqrdr/index.html.
2. IOM Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care,* Washington, DC: The National Academies Press;2003.
3. Institute of Medicine. How far have we come in reducing health disparities? Progress since 2000: Workshop summary. Washington, DC: The National Academies Press; 2012.
4. Schoen C, Radley D, Riley P, et al. *Health Care in the Two Americas: Findings From the Scorecard on State Health System Performance for Low-Income Populations, 2013.* New York, NY: The Commonwealth Fund; 2013. Report No: Pub No. 1700. Available at http://www.commonwealthfund.org/Publications/Fund-Reports/2013/Sep/Low-Income-Scorecard.aspx?page=all.
5. Satin DJ. Paying physicians and protecting the poor. *Minn Med.* 2006;89(4):42-44.
6. Iezzoni, LI ed. *Risk Adjustment for Measuring Health Care Outcomes.* 4th ed. Chicago, IL: Health Administration Press; 2013.
7. Muennig P, Fiscella K, Tancredi D, et al. The relative health burden of selected social and behavioral risk factors in the United States: implications for policy. *Am J Public Health.* 2010;100(9):1758-1764.
8. Meyer PA, Yoon PW, Kaufmann RB. Introduction: CDC Health Disparities and Inequalities Report - United States, 2013. *MMWR Surveill Summ.* 2013;62 Suppl 3:3-5.
9. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. *Health Aff (Millwood).* 2002;21(2):60-76.
10. Franks P, Fiscella K. Effect of patient socioeconomic status on physician profiles for prevention, disease management, and diagnostic testing costs. *Med Care.* 2002;40(8):717-724.
11. Asch SM, Kerr EA, Keesey J, et al. Who is at greatest risk for receiving poor-quality health care? *N Engl J Med.* 2006;354(11):1147-1156.
12. van Ryn M, Saha S. Exploring unconscious bias in disparities research and medical education. *JAMA.* 2011;306(9):995-996.
13. Dovidio JF, Fiske ST. Under the radar: how unexamined biases in decision-making processes in clinical interactions can contribute to health care disparities. *Am J Public Health.* 2012;102(5):945-952.
14. Bach PB, Pham HH, Schrag D, et al. Primary care physicians who treat blacks and whites. *N Engl J Med.* 2004;351(6):575-584.
15. Reschovsky JD, O'Malley AS. Do primary care physicians treating minority patients report problems delivering high-quality care? *Health Aff (Millwood).* 2008;27(3):w222-w231.
16. Skinner J, Chandra A, Staiger D, et al. Mortality after acute myocardial infarction in hospitals that disproportionately treat black patients. *Circulation.* 2005;112(17):2634-2641.
17. Chien AT, Wroblewski K, Damberg C, et al. Do physician organizations located in lower socioeconomic status areas score lower on pay-for-performance measures? *J Gen Intern Med.* 2012;27(5):548-554.
18. Fiscella K, Franks P. Influence of patient education on profiles of physician practices. *Ann Intern Med.* 1999;131(10):745-751.
19. Hasnain-Wynia R, Baker DW, Nerenz D, et al. Disparities in health care are driven by where minority patients seek care: examination of the hospital quality alliance measures. *Arch Intern Med.* 2007;167(12):1233-1239.
20. Hong CS, Atlas SJ, Chang Y, et al. Relationship between patient panel characteristics and primary care physician clinical performance rankings. *JAMA.* 2010;304(10):1107-1113.
21. Zaslavsky AM, Hochheimer JN, Schneider EC, et al. Impact of sociodemographic case mix on the HEDIS measures of health plan quality. *Med Care.* 2000;38(10):981-992.
22. Nuzum R, McCarthy D, Gauthier A, et al. *Denver Health: A High-Performance Public Health Care System.* New York: Commonwealth Fund; 2007. Available at

http://www.commonwealthfund.org/Publications/Fund-Reports/2007/Jul/Denver-Health--A-High-Performance-Public-Health-Care-System.aspx.
23. Casalino LP, Elster A, Eisenberg A, et al. Will pay-for-performance and quality reporting affect health care disparities? *Health Aff (Millwood).* 2007;26(3):w405-w414.
24. Friedberg MW, Safran DG, Coltin K, et al. Paying for performance in primary care: potential impact on practices and disparities. *Health Aff (Millwood).* 2010;29(5):926-932.
25. Ryan AM. Will value-based purchasing increase disparities in care? *N Engl J Med.* 2013;369(26):2472-2474.
26. Weinick RM, Chien AT, Rosenthal MB, et al. Hospital executives' perspectives on pay-for-performance and racial/ethnic disparities in care. *Med Care Res Rev.* 2010;67(5):574-589.
27. Werner RM, Goldman LE, Dudley RA. Comparison of change in quality of care between safety-net and non-safety-net hospitals. *JAMA.* 2008;299(18):2180-2187.
28. Joynt KE, Jha AK. Characteristics of hospitals receiving penalties under the Hospital Readmissions Reduction Program. *JAMA.* 2013;309(4):342-343.
29. Young GJ, Rickles NM, Chou CH, et al. Socioeconomic characteristics of enrollees appear to influence performance scores for medicare part D contractors. *Health Aff (Millwood).* 2014;33(1):140-146.
30. Zaslavsky AM, Epstein AM. How patients' sociodemographic characteristics affect comparisons of competing health plans in California on HEDIS quality measures. *Int J Qual Health Care.* 2005;17(1):67-74.
31. Werner RM, Asch DA, Polsky D. Racial profiling: the unintended consequences of coronary artery bypass graft report cards. *Circulation.* 2005;111(10):1257-1263.
32. Kuhlthau K, Ferris TG, Iezzoni LI. Risk adjustment for pediatric quality indicators. *Pediatrics.* 2004;113(1 Pt 2):210-216.
33. Zaslavsky AM. Statistical issues in reporting quality data: small samples and casemix variation. *Int J Qual Health Care.* 2001;13(6):481-488.
34. Rothman KG, Greenland S, Lash TL. *Modern Epidemiology,* Lippincott Williams & Wilkins;2008.
35. Feaster D, Brincks A, Robbins M, et al. Multilevel models to identify contextual effects on individual group member outcomes: a family example. *Fam Process.* 2011;50(2):167-183.
36. Medicare Payment Advisory Commission (MedPAC). *Refining the Hospital Readmissions Reduction Program.* Washington, DC: MedPAC; 2013. Available at http://www.medpac.gov/.
37. Weissman JS, Stern RS, Epstein AM. The impact of patient socioeconomic status and other social factors on readmission: a prospective study in four Massachusetts hospitals. *Inquiry.* 1994;31(2):163-172.
38. Krieger N, Chen JT, Waterman PD, et al. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. *Am J Epidemiol.* 2002;156(5):471-482.
39. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav.* 1995;Spec No:80-94.
40. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J Health Soc Behav.* 2010;51 Suppl:S28-S40.
41. Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc.* 2007;99(9):1013-1023.
42. Dubay LC, Lebrun LA. Health, behavior, and health care disparities: disentangling the effects of income and race in the United States. *Int J Health Serv.* 2012;42(4):607-625.
43. McGrail KM, van DE, Ross NA, et al. Income-related health inequalities in Canada and the United States: a decomposition analysis. *Am J Public Health.* 2009;99(10):1856-1863.
44. Ngo-Metzger Q, Sorkin DH, Billimek J, et al. The effects of financial pressures on adherence and glucose control among racial/ethnically diverse patients with diabetes. *J Gen Intern Med.* 2012;27(4):432-437.
45. Kirst M, Shankardass K, Bomze S, et al. Sociodemographic data collection for health equity measurement: a mixed methods study examining public opinions. *Int J Equity Health.* 2013;12:75.

46. Davern M, Rodin H, Beebe TJ, et al. The effect of income question design in health surveys on family income, poverty and eligibility estimates. *Health Serv Res.* 2005;40(5 Pt 1):1534-1552.
47. Bonito AJ, Bann C, Eicheldinger C, et al. *Creation of New Race-Ethnicity Codes and Socioeconomic Status (Ses) Indicators for Medicare Beneficiaries. Final Report, Sub-Task 2.* Rockville, MD: AHRQ; 2008. Report No: AHRQ Publication Number 08-0029-EF. Available at http://www.ahrq.gov/research/findings/final-reports/medicareindicators/medicareindicators.pdf.
48. Lovasi GS, Weiss JC, Hoskins R, et al. Comparing a single-stage geocoding method to a multi-stage geocoding method: how much and where do they disagree? *Int J Health Geogr.* 2007;6:12.
49. Goldberg DW, Cockburn MG. Improving Geocode Accuracy with Candidate Selection Criteria. *Transactions in GIS.* 2010;14(s1):149-176.
50. Sonderman JS, Mumma MT, Cohen SS, et al. A multi-stage approach to maximizing geocoding success in a large population-based cohort study through automated and interactive processes. *Geospatial Health.* 2012;6:273.
51. Conti G, Heckman J, Urzua S. The education-health gradient. *The American Economic Review.* 2010;100:234-238.
52. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. *J Health Econ.* 2010;29(1):1-28.
53. Kim M, Zaslavsky AM, Cleary PD. Adjusting Pediatric Consumer Assessment of Health Plans Study (CAHPS) Scores to Ensure Fair Comparison of Health Plan Performances. *Med Care.* 2005;43(1):44-52.
54. O'Malley AJ, Zaslavsky AM, Elliott MN, et al. Case-mix adjustment of the CAHPS Hospital Survey. *Health Serv Res.* 2005;40(6 Pt 2):2162-2181.
55. National Committee on Vital and Health Statistics. Core Health Data Elements, Report of the National Committee on Vital and Health Statistics. *National Committee on Vital and Health Statistics, United States Department of health and Human Services.* Available at: http://www.ncvhs.hhs.gov/ncvhsr1.htm. Updated August1996.
56. Jorgensen S, Thorlby R, Weinick RM, et al. Responses of Massachusetts hospitals to a state mandate to collect race, ethnicity and language data from patients: a qualitative study. *BMC Health Serv Res.* 2010;10:352.
57. Escarce JJ, Carreon R, Veselovskiy G, et al. Collection of race and ethnicity data by health plans has grown substantially, but opportunities remain to expand efforts. *Health Aff (Millwood).* 2011;30(10):1984-1991.
58. Roblin DW. Validation of a neighborhood SES index in a managed care organization. *Med Care.* 2013;51(1):e1-e8.
59. Kwok RK, Yankaskas BC. The use of census data for determining race and education as SES indicators: a validation study. *Ann Epidemiol.* 2001;11:171-177.
60. IOM Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records. *Capturing Social and Behavioral Domains in Electronic Health Records: Phase 1,* Washington, DC: Institute of Medicine;2014.
61. Roelfs DJ, Shor E, Davidson KW, et al. Losing life and livelihood: a systematic review and meta-analysis of unemployment and all-cause mortality. 2011;72:840-854.
62. Dupre ME, George LK, Liu G, et al. The cumulative effect of unemployment on risks for acute myocardial infarction. *Arch Intern Med.* 2012;172:1731-1737.
63. Rosenthal L, Carroll-Scott A, Earnshaw VA, et al. The importance of full-time work for urban adults' mental and physical health. *Social Science & Medicine.* 2012;75:1692-1696.
64. Bolton KL, Rodriguez E. Smoking, drinking and body weight after re-employment: does unemployment experience and compensation make a difference? *BMC Public Health.* 2009;9:77.
65. Galobardes B, Shaw M, Lawlor DA, et al. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health.* 2006;60(2):95-101.
66. Wisnivesky JP, Krauskopf K, Wolf MS, et al. The association between language proficiency and outcomes of elderly patients with asthma. *Annals of Allergy, Asthma, & Immunology.* 2012;109:179-184.

67. Wisnivesky JP, Kattan M, Evans D, et al. Assessing the relationship between language proficiency and asthma morbidity among inner-city asthmatics. *Med Care.* 2009;47:243-249.
68. Schenker Y, Wang F, Selig SJ, et al. The impact of language barriers on documentation of informed consent at a hospital with on-site interpreter services. *J Gen Intern Med.* 2007;22 Suppl 2:294-299.
69. Schenker Y, Karter AJ, Schillinger D, et al. The impact of limited English proficiency and physician language concordance on reports of clinical interactions among patients with diabetes: the DISTANCE study. *Patient Education & Counseling.* 2010;81:222-228.
70. McClellan SR, Wu FM, Snowden LR. The impact of threshold language assistance programming on the accessibility of mental health services for persons with limited English proficiency in the Medi-Cal setting. *Med Care.* 2012;50:554-558.
71. de Bocanegra HT, Rostovtseva D, Cetinkaya M, et al. Quality of reproductive health services to limited English proficient (LEP) patients. *Journal of Health Care for the Poor & Underserved.* 2011;22:1167-1178.
72. Scheier DB. Barriers to health care for people with hearing loss: a review of the literature. *J N Y State Nurses Assoc.* 2009;40:4-10.
73. Hadler,MW, Chen,X, Gonzalez,E, and Roby,DH. Limited English proficient HMO enrollees remain vulnerable to communication barriers despite language assistance regulations. Available at: http://healthpolicy.ucla.edu/publications/search/pages/detail.aspx?PubID=1186. Updated 2013.
74. Schyve PM. Language differences as a barrier to quality and safety in health care: the Joint Commission perspective. *J Gen Intern Med.* 2007;22 Suppl 2:360-361.
75. Institute of Medicine, . *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement.* National Academies Press; 2009. Available at http://www.iom.edu/Reports/2009/RaceEthnicityData.aspx.
76. Nerenz DR, Carreon R, Veselovskiy G. Race, ethnicity, and language data collection by health plans: findings from 2010 AHIPF-RWJF survey. *J Health Care Poor Underserved.* 2013;24(4):1769-1783.
77. Bhalla R, Yongue BG, Currie BP. Standardizing race, ethnicity, and preferred language data collection in hospital information systems: results and implications for healthcare delivery and policy. *J Healthc Qual.* 2012;34:44-52.
78. Lawson EH, Carreon R, Veselovskiy G, et al. Collection of language data and services provided by health plans. *Am J Manag Care.* 2011;17(12):e479-e487.
79. DeNavas-Walt C, Proctor B, Smith J. *Income, Poverty and Health Insurance Coverage in the United States: 2011.* Washington, DC: US Government Printing Office; 2012. Report No: US Census Bureau, Current Population Reports, P60-243. Available at http://www.census.gov/prod/2012pubs/p60-243.pdf.
80. McWilliams JM, Zaslavsky AM, Meara E, et al. Impact of Medicare coverage on basic clinical services for previously uninsured adults. *JAMA.* 2003;290:757-764.
81. McWilliams JM, Zaslavsky AM, Meara E, et al. Health Insurance Coverage And Mortality Among The Near-Elderly. *Health Aff (Millwood).* 2004;23:223-233.
82. Abdullah F, Zhang Y, Lardaro T, et al. Analysis of 23 million US hospitalizations: uninsured children have higher all-cause in-hospital mortality. *Journal of Public Health.* 2010;32:236-244.
83. Wilper AP, Woolhandler S, Lasser KE, et al. Hypertension, Diabetes, And Elevated Cholesterol Among Insured And Uninsured U.S. Adults. *Health Aff (Millwood).* 2009;28(6):w1151-w1159.
84. Kogan MD, Newacheck PW, Honberg L, et al. Association between underinsurance and access to care among children with special health care needs in the United States. *Pediatrics.* 2005;116(5):1162-1169.
85. Lavarreda SA, Brown ER, Bolduc CD. Underinsurance in the United States: an interaction of costs to consumers, benefit design, and access to care. *Annu Rev Public Health.* 2011;32:471-482.
86. Voorhees K, Fernald DH, Emsermann C, et al. Underinsurance in primary care: a report from the State Networks of Colorado Ambulatory Practices and Partners (SNOCAP). *J Am Board Fam Med.* 2008;21(4):309-316.
87. Jean-Jacques M, Persell SD, Hasnain-Wynia R, et al. The implications of using adjusted versus unadjusted methods to measure health care disparities at the practice level. *Am J Med Qual.* 2011;26(6):491-501.

88. Lurie N, Fremont A. Looking forward: cross-cutting issues in the collection and use of racial/ethnic data. *Health Serv Res.* 2006;41(4 Pt 1):1519-1533.
89. Shavers VL, Fagan P, Jones D, et al. The State of Research on Racial/Ethnic Discrimination in The Receipt of Health Care. *Am J Public Health.* 2012;102:953-966.
90. Culhane JF, Goldenberg RL. Racial disparities in preterm birth. *Semin Perinatol.* 2011;35(4):234-239.
91. Herman WH, Ma Y, Uwaifo G, et al. Differences in A1C by Race and Ethnicity Among Patients With Impaired Glucose Tolerance in the Diabetes Prevention Program. *Diabetes Care.* 2007;30:2453-2457.
92. Davidson MB, Schriger DL. Effect of age and race/ethnicity on HbA1c levels in people without known diabetes mellitus: Implications for the diagnosis of diabetes. *Diabetes Res Clin Pract.* 2010;87:415-421.
93. Hunt KJ, Gebregziabher M, Lynch CP, et al. Impact of diabetes control on mortality by race in a national cohort of veterans. *Ann Epidemiol.* 2013;23:74-79.
94. Kressin NR, Orner MB, Manze M, et al. Understanding Contributors to Racial Disparities in Blood Pressure Control. *Circulation: Cardiovascular Quality and Outcomes.* 2010;3:173-180.
95. Bosworth HB, Dudley T, Olsen MK, et al. Racial differences in blood pressure control: potential explanatory factors. *Am J Med.* 2006;119:70-15.
96. Weissman JS, Hasnain-Wynia R. Advancing Health Care Equity through Improved Data Collection. *N Engl J Med.* 2011;364:2276-2277.
97. Wynia MK, Ivey SL, Hasnain-Wynia R. Collection of data on patients' race and ethnic group by physician practices. *N Engl J Med.* 2010;362:846-850.
98. Kushel MB, Vittinghoff E, Haas JS. Factors associated with the health care utilization of homeless persons. *JAMA.* 2001;285(2):200-206.
99. Baggett TP, O'Connell JJ, Singer DE, et al. The Unmet Health Care Needs of Homeless Adults: A National Study. *Am J Public Health.* 2010;100:1326-1333.
100. Doran KM, Ragins KT, Iacomacci AL, et al. The revolving hospital door: hospital readmissions among patients who are homeless. *Med Care.* 2013;51(9):767-773.
101. Baggett TP, Hwang SW, O'Connell JJ, et al. Mortality among homeless adults in Boston: shifts in causes of death over a 15-year period. *JAMA Intern Med.* 2013;173(3):189-195.
102. Greysen SR, Allen R, Rosenthal MS, et al. Improving the quality of discharge care for the homeless: a patient-centered approach. *J Health Care Poor Underserved.* 2013;24(2):444-455.
103. U.S. Department of Housing and Urban Development. *Homeless Assistance.* 2009. Available at http://portal.hud.gov/hudportal/HUD?src=/program_offices/comm_planning/homeless. Last accessed March 2014.
104. Stimpson JP, Wilson FA. Cholesterol screening by marital status and sex in the United States. *Preventing Chronic Disease.* 2009;6:A55.
105. Lee S, Cho E, Grodstein F, et al. Effects of marital transitions on changes in dietary and other health behaviours in US women. *Int J Epidemiol.* 2005;34:69-78.
106. Eng PM, Kawachi I, Fitzmaurice G, et al. Effects of marital transitions on changes in dietary and other health behaviours in US male health professionals. *Journal of Epidemiology & Community Health.* 2005;59:56-62.
107. Hughes ME, Waite LJ. Marital Biography and Health at Mid-Life. *J Health Soc Behav.* 2009;50:344-358.
108. Shor E, Roelfs DJ, Bugyi P, et al. Meta-analysis of marital dissolution and mortality: reevaluating the intersection of gender and age. *Soc Sci Med.* 2012;75(1):46-59.
109. Shor E, Roelfs DJ, Curreli M, et al. Widowhood and mortality: a meta-analysis and meta-regression. *Demography.* 2012;49(2):575-606.
110. Kutner M, Greenberg E, Jin Y, et al. *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy.* Washington, DC: Institute of Education Sciences, Natinal Center of Education Statistics, U.S. Department of Education; 2006. Report No: NCES 2006483. Available at https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006483. Last accessed March 2014.

111. Institute of Medicine Committee on Health Literacy. *Health Literacy: A Prescription to End Confusion.* Washington, DC: The National Academies Press; 2004. Available at http://www.iom.edu/Reports/2004/health-literacy-a-prescription-to-end-confusion.aspx .
112. Kaestle CF, Campbell A, Finn JD, et al. *Adult Literacy and Education in America.* Washington, D.C.: National Center for Education Statistics; 2001. Available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2001534.
113. von Wagner C, Steptoe A, Wolf MS, et al. Health literacy and health actions: a review and a framework from health psychology. *Health Education & Behavior.* 2009;36:860-877.
114. DeWalt DA, Callahan LF, Hawk VH, et al. *Health Literacy Universal Precautions Toolkit.* Rockville, MD: Agency for Heatlhcare Quality and Research; 2010. Report No: AHRQ Publication No. 10-0046-EF. Available at http://www.ahrq.gov/qual/literacy/. Last accessed July 2014.
115. Wallston KA, Cawthon C, McNaughton CD, et al. Psychometric properties of the Brief Health Literacy Screen in clinical practice. *J Gen Intern Med.* 2014;29:119-126.
116. Chaix B, Rosvall M, Lynch J, et al. Disentangling contextual effects on cause-specific mortality in a longitudinal 23-year follow-up study: impact of population density or socioeconomic environment? *Int J Epidemiol.* 2006;35:633-643.
117. Ponce,NA. Measuring Neighborhood Effects and the Use of Geo-coded Variables. Available at: http://www.ctsi.ucla.edu/education/files/view/training/docs/Ponce_Neigborhood_Data_Sources_012213.pdf. Updated January2013.
118. Butler DC, Petterson S, Phillips RL, et al. Measures of social deprivation that predict health care access and need within a rational area of primary care service delivery. *Health Serv Res.* 2013;48(2 Pt 1):539-559.
119. Fremont AM, Bierman AS, Wickstrom SL, et al. Use of Indirect Measures of Race/Ethnicity and Socioeconomic Status in Managed Care Settings to Identify Disparities in Cardiovascular and Diabetes Care Quality. *Health Affairs.* 2005;24:526.
120. Lurie N, Fremont A, Somers SA, et al. The National Health Plan Collaborative to Reduce Disparities and Improve Quality. *Joint Commission Journal on Quality & Patient Safety.* 2008;34:256-265.
121. Pardo-Crespo MR, Narla NP, Williams AR, et al. Comparison of individual-level versus area-level socioeconomic measures in assessing health outcomes of children in Olmsted County, Minnesota. *J Epidemiol Community Health.* 2013;67:305-310.
122. Chang CM, Su YC, Lai NS, et al. The combined effect of individual and neighborhood socioeconomic status on cancer survival rates. *PLoS One.* 2012;7(8):e44325.
123. Winkleby M, Cubbin C, Ahn D. Effect of cross-level interaction between individual and neighborhood socioeconomic status on adult mortality rates. *Am J Public Health.* 2006;96(12):2145-2153.
124. Diez-Roux AV, Kiefe CI, Jacobs DR, Jr., et al. Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Ann Epidemiol.* 2001;11:395-405.
125. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Ann Behav Med.* 2010;40:218-227.
126. Mookadam F, Arthur HM. Social support and its relationship to morbidity and mortality after acute myocardial infarction: Systematic overview. *Arch Intern Med.* 2004;164:1514-1518.
127. Hibbard JH, Greene J, Becker ER, et al. Racial/Ethnic Disparities And Consumer Activation In Health. *Health Aff (Millwood).* 2008;27:1442-1453.
128. Pampel FC, Krueger PM, Denney JT. Socioeconomic Disparities in Health Behaviors. *Annu Rev Sociol.* 2010;36:349-370.
129. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health & Social Behavior.* 1997;38:21-37.
130. NQF. *Healthcare Disparities and Cultural Competency Consensus Standards: Disparities-Sensitive Measure Assessment.* Washington, DC: NQF; 2012. Available at http://www.qualityforum.org/Publications/2012/11/Healthcare_Disparities_and_Cultural_Competency_Consensus_Standards__Disparities-Sensitive_Measure_Assessment.aspx.

131. NQF. *Guidance for Evaluating Usability and Use of Performance Measures.* Washington, DC: NQF; 2012. Available at http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=70845.
132. NQF. *Healthcare Disparities and Cultural Competency Consensus Standards: Technical Report.* Washington, DC: NQF; 2012. Available at http://www.qualityforum.org/Publications/2012/09/Healthcare_Disparities_and_Cultural_Competency_Consensus_Standards_Technical_Report.aspx.
133. NQF. *Patient Reported Outcomes (PROs)in Performance Measurement .* Washington, DC: NQF; 2013. Available at http://www.qualityforum.org/Publications/2012/12/Patient-Reported_Outcomes_in_Performance_Measurement.aspx.
134. Braveman PA, Kumanyika S, Fielding J, et al. Health disparities and health equity: the issue is justice. *Am J Public Health.* 2011;101(Suppl 1):S149-S155.
135. Austin PC, Alter DA, Tu JV. The use of fixed- and random-effects models for classifying hospitals as mortality outliers: a Monte Carlo assessment. *Med Decis Making.* 2003;23(6):526-539.
136. Porta, M ed. *A Dictionary of Epidemiology.* 5th ed. New York: Oxford University Press; 2008.

Appendix A: Expert Panel on Risk Adjustment for Sociodemographic Factors and NQF Staff

Expert Panel Roster

Kevin Fiscella, MD, MPH (Co-Chair)
Professor, Family Medicine, Public Health Sciences, Community Health and Oncology
University of Rochester Medical Center
Rochester, NY

David Nerenz, PhD (Co-Chair)
Director, Center for Health Policy & Health Services Research
Director of Outcomes Research for the Neuroscience Institute
Vice-Chair for Research of the Department of Neurosurgery at Henry Ford Hospital
Henry Ford Health System
Detroit, MI

Jean Accius, PhD, PMP
Director, Health and Long-Term Services and Supports
AARP
Washington, DC

Alyce Adams, MPP, PhD
Research Scientist II
Chief of Health Care Delivery and Policy, Division of Research
Kaiser Permanente
Oakland, CA

Mary Barger, PhD, MPH, CNM, FACNM
Associate Professor of Nursing, University of California San Diego
San Diego, CA

Susannah M. Bernheim, MD, MHS
Director of Quality Measurement
Yale New Haven Health System Center for Outcomes Research and Evaluation (CORE)
New Haven, CT

Monica Bharel, MD, MPH
Chief Medical Officer
Boston Health Care for the Homeless Program
Boston, MA

Mary Beth Callahan, ACSW/LCSW
Senior Social Worker
Dallas Transplant Institute
Dallas, TX

Lawrence Casalino, MD, PhD
Livingston Farrand Professor of Public Health
Chief, Division of Outcomes and Effectiveness Research
New York, NY

Alyna Chien, MD, MS
Assistant Professor, Pediatrics, Harvard Medical School
Pediatrician, Children's Hospital Primary Care Center, Boston Children's Hospital
Boston, MA

Marshall Chin, MD, MPH
Richard Parrillo Family Professor of Healthcare Ethics in the Department of Medicine,
University of Chicago
Director of the RWJF Finding Answers: Disparities Research for Change National Program Office
Chicago, IL

Mark Cohen, PhD
Statistical Manager, Continuous Quality Improvement, Division of Research and Optimal Patient Care
American College of Surgeons
Adjunct Associate Professor of Surgery, Feinberg School of Medicine, Northwestern University
Chicago, IL

Nancy Garrett, PhD
Chief Analytics Officer
Hennepin County Medical Center
Minneapolis, MN

Norbert Goldfield, MD
Medical Director, 3M HIS Clinical and Economics Research**,** 3M
Staff Physician, Brightwood Community Health Center
Wallingford, CT

Atul Grover, MD, PhD, FCCP
Chief Public Policy Officer
Association of American Medical Colleges
Washington, DC

David Hopkins, PhD
Senior Advisor**,** Pacific Business Group on Health
Adjunct Associate, Center for Primary Care and Outcomes Research and Center for Health Policy,
Stanford University
San Francisco, CA

Dionne Jimenez, MPP
Research & Policy Coordinator
Service Employees International Union
Los Angeles, CA

Steven Lipstein, MHA
President and CEO
BJC HealthCare
St. Louis, MO

Eugene Nuccio, PhD
Assistant Professor, Division of Health Care Policy and Research, School of Medicine
University of Colorado, Anschutz Medical Campus
Aurora, CO

Sean O'Brien, PhD
Assistant Professor, Biostatistics and Bioinformatics
Duke University Medical Center
Durham, NC

Pam Owens, PhD
Senior Research Scientist and Scientific Director of the AHRQ Quality Indicators™
AHRQ
Rockville, MD

Ninez Ponce, MPP, PhD
Professor, Department of Health Policy and Management
UCLA Fielding School of Public Health
Los Angeles, CA

Thu Quach, PhD, MPH
Research Director
Asian Health Services
Oakland, CA

Tia Goss Sawhney, DrPH, FSA, MAAA
Director of Data, Analytics, and Research
Illinois Department of Healthcare and Family Services
Chicago, IL

Nancy Sugg, MD, MPH
Medical Director
Pioneer Square Clinic & Downtown Homeless Programs, Harborview Medical Center
Associate Professor of Medicine, Division of General Internal Medicine at the University of Washington
Seattle, WA

Rachel Werner, MD, PhD
Associate Professor of Medicine
University of Pennsylvania
Philadelphia, PA

NQF acknowledges technical advice from **Alan Zaslavsky, PhD,** Professor of Health Care Policy, Harvard Medical School, Boston, MA.

Expert Panel Member Biographies

Kevin Fiscella, MD, MPH (Co-Chair)
Professor, Family Medicine, Public Health Sciences, Community Health and Cancer Center, University of Rochester Medical Center
Kevin Fiscella, MD, MPH, is a tenured Professor of Family Medicine, Public Health Sciences, Community Health, and Oncology at the University of Rochester School of Medicine and Dentistry. His scholarly work has focused on both conceptual models and empirical research related to healthcare disparities. His current work addresses practical strategies to mitigate disparities in healthcare quality. He has served on numerous national committees related to health and healthcare disparities. He has published more than 180 papers in peer-reviewed journals and has received major research grants from numerous federal agencies and private foundations.

David Nerenz, PhD (Co-Chair)
Director, Center for Health Policy and Health Services Research, Henry Ford Health System
David R. Nerenz, PhD, is Director of the Center for Health Policy and Health Services Research at Henry Ford Health System in Detroit. He is also Director of Outcomes Research for the Neuroscience Institute and Vice-Chair for Research of the Department of Neurosurgery at Henry Ford Hospital. He was appointed in May of 2012 as a Commissioner on the Medicare Payment Advisory Commission (MedPAC). He recently served as the Chair of the Institute of Medicine Committee on Leading Health Indicators for Healthy People 2020 and Chair of the IOM Subcommittee on Standardized Collection of Race/Ethnicity Data for Healthcare Quality Improvement.

Jean Accius, PhD, PMP
Director, Health and Long-Term Services and Supports, AARP
Jean Accius, PhD, is an expert in health and long-term care policy. His background includes translating research into policy and practice. Currently, Jean is the Director of Health and Long-term Services and Supports (LTSS) at AARP. In this capacity, he leads the policy development process on health and LTSS related issues that guides AARP's legislative, regulatory and litigation activities. He also provides strategic advice and counsel to senior AARP leadership and other departments to ensure policy integration and consistency across the association.

Alyce Adams, MPP, PhD
Research Scientist II and Chief of Health Care Delivery and Policy, Division of Research, Kaiser Permanente
Alyce S. Adams, MPP, PhD, is Research Scientist II and Chief of Health Care Delivery and Policy at the Kaiser Permanente Division of Research in Oakland, California. Her research explores disparities in chronic disease treatment outcomes using longitudinal data methods. Dr. Adams' current studies include a cluster randomized clinical trial to improve diabetic peripheral neuropathy treatment outcomes funded by the Patient Centered Outcomes Research Institute and an evaluation of the impact of Medicare Part D among dual Medicaid and Medicare enrollees funded by the National Institute on Aging. She has a PhD in Health Policy from Harvard University.

Mary Barger, PhD, MPH, CNM, FACNM
Associate Professor of Nursing, University of California San Diego, and American College of Nurse Midwives
Mary Barger PhD, MPH, CNM, FACNM, is a perinatal epidemiologist and has practiced clinical nurse-midwifery for over 25 years. She has taught in a school of public health, a medical school, and two schools of nursing. The focus of her research has been on maternal morbidity and mortality with a focus on cesareans and using administratively collected data, such as hospital discharge data and birth data, to examine processes and outcomes of care. One of her recent studies combined survey data with GIS information to further understand the racial disparities in cesarean rates in California.

Susannah M. Bernheim, MD, MHS
Director, Quality Measurement, Yale New Haven Health System Center for Outcomes Research and Evaluation (CORE)
Susannah M. Bernheim, MD, MHS, is a Family Physician, Health Services Researcher and the Director of Quality Measurement at Yale-New Haven Hospital's Center for Outcomes Research and Evaluation (CORE). She has extensive experience leading teams in measure development, maintenance, NQF endorsement, and implementation. Her research focuses on the intersection of healthcare quality, outcomes and socioeconomic status. She received her MD at the University of California San Francisco, and her Master of Health Sciences at Yale University. Following a research fellowship and prior to joining CORE, she served as Deputy Director of Performance Management for the Yale New Haven Health System.

Monica Bharel, MD, MPH
Chief Medical Officer, Boston Health Care for the Homeless Program
Monica Bharel, MD, is the Chief Medical Officer for the largest nonprofit healthcare organization for homeless individuals in the country. Under her leadership, the organization achieved Level 3 NCQA PCMH recognition and a 3-fold improvement in women's health quality indicators. She was appointed by Governor Patrick to serve on the Behavioral Health Integration Task Force under the Massachusetts payment reform initiative. She has spoken locally and nationally about the needs of vulnerable and homeless individuals. She is currently focused on ensuring that state and national healthcare reform efforts enhance the care for homeless individuals, without inadvertently widening healthcare disparities.

Mary Beth Callahan, ACSW/LCSW
Senior Social Worker, Dallas Transplant Institute
Mary Beth Callahan has worked in nephrology social work since 1984. She is currently Senior Social Worker at Dallas Transplant Institute (DTI) and has had the privilege to work with ESRD patients on hemodialysis, peritoneal dialysis, and transplant. She has served on numerous advisory boards and professional committees. She served as CNSW Chair from 1996-1998 and is co-developer of the CNSW's Outcomes Training Program. Her focus with transplant recipients is to help them prepare to return to work whenever possible and/or to live life to the fullest. One of her ongoing efforts is to encourage other staff members to keep rehabilitation in the forefront of their minds. She hosts Job Club monthly at DTI. Job Club developed from joint research grants from the Society for Transplant Social Workers and the Council of Nephrology Social Workers and provides patients with information on Social Security Work Incentives and connects patients with vocational rehabilitation resources and hope.

Lawrence Casalino, MD, PhD
Livingston Farrand Professor of Public Health; Chief, Division of Outcomes and Effectiveness Research, Weill Cornell Medical College
Lawrence Casalino, MD, PhD, has written some of the seminal articles on unintended consequences of quality measurement and on SES disparities and quality measurement. He has also served on relevant national committees. He has quite a lot of knowledge about the organization of medical practice and hospital care and about the responses of providers to incentives. This knowledge comes from quantitative and qualitative research as well as from the 20 years that he spent as a family physician in full-time practice, during which time he also served as a hospital medical staff president and vice president of a large independent practice association.

Alyna Chien, MD, MS
Assistant Professor, Boston Children's Hospital
Alyna Chien, MD, is a physician health services researcher at Harvard Medical School and Boston Children's Hospital. She is the leading pediatric expert on the use of performance incentives in healthcare and has extensive experience using established risk adjustment methods. Currently, she is examining whether geocoded socioeconomic information can improve pediatric risk adjustment algorithms so that healthcare payments can better reflect pediatric patient complexity. She has used similar geocoding techniques to examine the degree to which socioeconomic factors have affected the ability of very large physician organizations to respond to performance incentives. Her work is funded by AHRQ, NICHD, and RWJF.

Marshall Chin, MD, MPH
Richard Parrillo Family Professor of Healthcare Ethics in the Department of Medicine, University of Chicago
Marshall H. Chin, MD, MPH, FACP, Richard Parrillo Family Professor of Healthcare Ethics in the Department of Medicine at the University of Chicago, is a general internist with extensive experience improving the care of vulnerable patients with chronic disease. Dr. Chin is Director of the RWJF Finding Answers: Disparities Research for Change National Program Office. He was a member of the IOM Committee on Future Directions for the National Healthcare Quality and Disparities Reports. He serves on the NQF MAP Coordinating Committee and was a member of the NQF Healthcare Disparities and Cultural Competency Consensus Standards Steering Committee.

Mark Cohen, PhD
Statistical Manager, Continuous Quality Improvement, Division of Research and Optimal Patient Care, American College of Surgeons
Mark Cohen, PhD, is the Statistical Manager, Continuous Quality Improvement, Division of Research and Optimal Patient Care, American College of Surgeons, and Adjunct Associate Professor of Surgery, Feinberg School of Medicine, Northwestern University. Since 2008, he has managed statistical efforts related to the ACS National Surgical Quality Improvement Program (NSQIP), ACS Universal Surgical Risk Calculator, ACS Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (MBSAQIP), and the ACS NSQIP Pediatric program. He has 125 publications and his current research focuses on optimizing risk-adjustment and reporting methodologies used in these programs. Before joining the ACS, Dr. Cohen was Statistician and, later, Technical Director at the Naval Institute for Dental and Biomedical Research.

Nancy Garrett, PhD
Chief Analytics Officer, Hennepin County Medical Center
Nancy Garrett, PhD, is currently Chief Analytics Officer at Hennepin County Medical Center, where she is developing methods to measure the impact of socioeconomic status on cost and quality measures for HCMC's diverse safety net population. She has an extensive background in applied health services research, and authored a chapter on provider profiling in a managed care textbook. Nancy is on NQF's Cost and Resource Use Steering Committee where she raised issues about adjusting for socioeconomic status that helped lead to the convening of this expert panel. Nancy has a Ph.D. in Demography from the University of Illinois.

Norbert Goldfield, MD
Medical Director, 3M HIS Clinical and Economics Research, 3M
Dr. Goldfield works as a medical director of 3MHIS, developing classification tools linking payment to quality. This work is used throughout the United States and overseas, with public and private payers. Dr. Goldfield is a board certified internist practicing at a community health center. He edits the peer reviewed Journal of Ambulatory Care Management and has published extensively. He is on a number of boards including Health Care for All. He is also the founder and executive director of Healing Across the Divides (www.healingdivides.org), an organization seeking to improve the health of Israelis and Palestinians.

Atul Grover, MD, PhD, FCCP
Chief Public Policy Officer, Association of American Medical Colleges
Atul Grover, MD, PhD, is the Chief Public Policy Officer for the Association of American Medical Colleges (AAMC). In this role, he manages the AAMC's health, educational, and scientific policies. Dr. Grover joined the AAMC in its Center for Workforce Studies, where he managed research activity and directed externally funded workforce studies. Prior to the AAMC, Dr. Grover was a senior consultant in health care finance and applied economics for The Lewin Group, Inc., and also served with the Health Resources and Service Administration. Dr. Grover is a clinical faculty member at the George Washington School of Medicine.

David Hopkins, PhD
Senior Advisor, Pacific Business Group on Health
David S. P. Hopkins, PhD, is Senior Advisor at the Pacific Business Group on Health. Hopkins is also affiliated with the Center for Health Policy and the Clinical Excellence Research Center at Stanford University Medical School. He earned his AB in Biology from Harvard, and his MS in Statistics and PhD in Operations Research from Stanford. Hopkins chaired the California Cooperative Healthcare Reporting Initiative (CCHRI) Executive Committee from 1996-2012 and is the former Chair of the Integrated Healthcare Association Pay-for-Performance Technical Efficiency Committee. He served two terms on NQF's Consensus Standards Approval Committee and currently chairs the NQF Purchaser Council.

Dionne Jimenez, MPP
Research & Policy Coordinator, Service Employees International Union
Dionne Jimenez is a research and policy coordinator for the Service Employees International Union, which represents 2.2 million workers advocating to improve their lives and the services they provide. She performs public policy analysis and develops public policy positions for SEIU on healthcare financing, workforce, quality of care and life, and other key issues related to the healthcare sector. Previous

professional experience includes serving as staff and legislative assistant to Congressman George Miller (D-CA). Dionne is a proud first-generation college graduate. She has a Master of Public Policy degree from the University of California, Los Angeles, School of Public Affairs, and a BA in Political Science from the University of California, Berkeley.

Steven Lipstein, MHA
President and CEO, BJC Healthcare
Steven Lipstein has led BJC Healthcare since 1999. He is highly engaged in ensuring that people everywhere receive high-quality, safe care. BJC is the largest provider of uncompensated care in Missouri. From 2008 to 2010, Mr. Lipstein co-chaired the oversight committee for Missouri Medicaid. Prior to joining BJC, Mr. Lipstein held executive roles at the University of Chicago and The Johns Hopkins Health System. He is vice chair of the Board of Governors for the Patient-Centered Outcomes Research Institute. He graduated from Emory University, has an MHA from Duke University, and completed an administrative fellowship at Massachusetts General Hospital.

Eugene Nuccio, PhD
Assistant Professor, University of Colorado, Anschutz Medical Campus
Eugene Nuccio, PhD, Assistant Professor, holds a doctorate in Education Psychology and has extensive experience with statistical analysis, measurement, and risk adjustment of outcomes for Medicare home care recipients. Since 2004 he has led the development of the last three sets of 40+ prediction models used to risk adjust home health outcomes nationally. He initiated innovations in how to represent OASIS data as well as methodological changes to develop complex multivariate models. Under the direction of MedPAC, Dr. Nuccio linked CMS claims, OASIS, and other data sources to produce experimental quality measures and prediction models. His contributions to the scientific literature on risk adjustment include presentations at AcademyHealth.

Sean O'Brien, PhD
Assistant Professor, Biostatistics and Bioinformatics, Duke University Medical Center
Sean O'Brien, PhD, is an Assistant Professor in the Department of Biostatistics and Bioinformatics at Duke University Medical Center. Since 2005, he has served as statistical director of the Society of Thoracic Surgeons (STS) Data Warehouse and Analysis Center and as co-investigator of several grants and contracts using large registries to study comparative effectiveness and healthcare quality. Dr. O'Brien also works on the development and evaluation of quantitative methods for healthcare provider performance assessment. His research interests include risk adjustment methodology, composite measures, and Bayesian modeling.

Pam Owens, PhD
Senior Research Scientist, AHRQ
Pamela Owens, PhD, is a senior research scientist at the Agency for Healthcare Research and Quality (AHRQ). Dr. Owens is the Scientific Director of the AHRQ Quality Indicators™ (QIs) and co-leads Healthcare Cost and Utilization Project (HCUP) outpatient data development. Dr. Owens' research experience includes the quality and access to care for various populations, conditions and settings, including children, low income, mental health, asthma, readmissions, ambulatory surgery, emergency department and inpatient settings. Her work has appeared in journals such as the *JAMA, Medical Care, Health Services Research, Annals of Internal Medicine, Pediatrics, Academic Emergency Medicine, Psychiatric Services,* and *Journal of Preventive Medicine*. Dr. Owens received a PhD in epidemiology and

health policy from Yale University and completed a postdoctoral fellowship at Johns Hopkins. She also has six years of clinical experience as an occupational therapist.

Ninez Ponce, MPP, PhD
Professor, Department of Health Policy and Management – UCLA Fielding School of Public Health
Ninez Ponce, MPP, PhD, is a professor in the Department of Health Policy and Management at the UCLA Fielding School of Public Health. In 14 years at UCLA, she has taught courses on health insurance, health economics, health policy, and research methods, with a research focus on racial/ethnic disparities in cancer prevention and control. She also conducted program evaluation, research, and public policy for a W.K. Kellogg Foundation national initiative to improve healthcare for the underserved. She has served on expert advisory groups for the Institute of Medicine, the Office of the Patient Advocate, and the UCLA Department of Health Services.

Thu Quach, PhD, MPH
Research Director, Asian Health Services
Thu Quach, PhD, MPH, is an epidemiologist and primary research interest has focused on the influence of environmental and sociocultural factors on immigrant population health. As a research scientist at the Cancer Prevention Institute of California, a nonprofit research organization, she leads research studies focusing on the booming nail salon workforce, comprised mainly of Vietnamese immigrants. In 2011, after years of research collaboration, she was recruited by Asian Health Services to become the inaugural research director at this community health center (CHC) serving low-income Asian American patients. She spearheads efforts across several CHCs to incorporate social determinants of health factors in risk adjustment.

Tia Goss Sawhney, DrPH, FSA, MAAA
Director of Data, Analytics, and Research, Illinois Department of Healthcare and Family Services
Tia Goss Sawhney, DrPH, FSA, MAAA, is the Director of Data, Research, and Analytics with the Illinois Medicaid plan. She is the author of the 2010 paper "Health Insurance Risk Adjustment: The Income Effect". The paper is included in her 2012 dissertation "Controlling Indirect Selection under Healthcare Reform" available at *www.soa.org/files/sections/health-dissertation-**sawhney**.pdf*. She is Fellow of the Society of Actuaries and a Member of the American Academy of Actuaries and active in each organization.

Nancy Sugg, MD, MPH
Medical Director Pioneer Square Clinic & Downtown Homeless Programs, Harborview Medical Center
Nancy Sugg, MD, MPH, is Associate Professor of Medicine in the Division of General Internal Medicine at the University of Washington and Medical Director of Harborview Medical Center's Pioneer Square Clinic and Downtown Homeless Programs. She is the Chair of the Care Management Committee at Harborview Medical Center, focusing on decreasing inpatient lengths of stay and readmissions and improving transitions of care. She works closely with Seattle-King County Public Health's Healthcare for the Homeless Network developing integrated medical services for homeless and outcomes measures for clinics caring for underserved populations. Dr. Sugg directs research projects and mentors future primary care providers and policymakers for underserved populations.

Rachel Werner, MD, PhD

Associate Professor of Medicine, University of Pennsylvania

Rachel Werner, MD, PhD, is an Associate Professor of Medicine at the University of Pennsylvania. She received her medical degree from the University of Pennsylvania School of Medicine, where she also did her residency in Internal Medicine. While completing a clinical fellowship in general internal medicine, she also received a PhD in health economics from the Wharton School at the University of Pennsylvania. Dr. Werner's research seeks to understand the effect of healthcare policies and delivery systems on quality of care. In particular, she has examined the role of provider payment and financial incentives on provider behavior, the organization of healthcare, racial disparities, and overall healthcare quality. Her work has empirically investigated numerous unintended consequences to quality improvement incentives and was among the first to recognize that public reporting of quality information may worsen racial disparities. She is currently principal investigator of an R01 from the Agency of Healthcare Research and Quality (examining how pay-for-performance in hospitals changed the value of healthcare) and an R01 from the National Institute of Aging (examining the effect of Medicaid pay-for-performance for nursing homes on delivery of nursing home care). She also directs one of five national centers to evaluate the effectiveness of the medical home by the Veterans Health Administration. She has received numerous awards including the Dissertation Award and the Alice Hersh New Investigator Award from Academy Health and the Presidential Early Career Award for Scientists and Engineers. Her research has been published in high-impact, peer-reviewed journals, including *JAMA*, *Journal of Health Economics, Health Services Research,* and *Health Affairs*. In addition to her research, Dr. Werner is a practicing primary care internist at the Philadelphia VA Medical Center and regularly attends the hospital's internal medicine service. She supervises healthcare provided by Hospital of the University of Pennsylvania.

NQF Staff

Helen Burstin, MD, MPH
Chief Scientific Officer

Karen Beckman Pace, PhD, MSN
Senior Director
Performance Measurement

Taroon Amin, MA, MPH
Senior Director
Performance Measurement

Karen Johnson, MS
Senior Director
Performance Measurement

Suzanne Theberge, MPH
Senior Project Manager
Performance Measurement

Erin O'Rourke
Senior Project Manager
Performance Measurement

Appendix B: Glossary

Accountability Applications – Use of performance results about identifiable, accountable entities to make judgments and decisions as a consequence of performance, such as reward, recognition, punishment, payment, or selection (e.g., public reporting, accreditation, licensure, professional certification, health information technology incentives, performance-based payment, network inclusion/exclusion).[131]

Confounding – The distortion in the degree of association between an exposure (independent variable) and an outcome (dependent variable) due to a mixing of effects between the exposure and an incidental (confounding) factor. Confounding represents systematic error and threatens the internal validity of an epidemiologic study since it can lead to false conclusions regarding the true relationship between an exposure and outcome.

Health Disparity – Healthy People 2020 defines a health disparity as "a particular type of health difference that is closely linked with social, economic, and/or environmental disadvantage. Health disparities adversely affect groups of people who have systematically experienced greater obstacles to health based on their racial or ethnic group; religion; socioeconomic status; gender; age; mental health; cognitive, sensory, or physical disability; sexual orientation or gender identity; geographic location; or other characteristics historically linked to discrimination or exclusion."

Healthcare Disparity – Differences in healthcare quality, access, and outcomes adversely affecting members of racial and ethnic minority groups and socially disadvantaged populations. [132]

Outcome – The result of providing healthcare. The term outcome will be used to broadly include the following types of outcomes relevant to performance measurement:

- Quality outcomes include:
  - Health outcome is the health status of a patient (or change in health status) resulting from healthcare—desirable or adverse.
  - In some situations, resource use may be considered a proxy for a health state (e.g., hospitalization may represent deterioration in health status).
  - Intermediate clinical outcome is a change in physiologic state that leads to a longer term health outcome (e.g., hemoglobin, blood pressure).
  - Patient-reported outcome is any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else. The domains of PROs include health-related quality of life/functional status, symptom/symptom burden, experience with care (including engagement, activation), and health-related behaviors.[133]
- Economic outcomes include the cost and resource use associated with providing healthcare services. (Although efficiency is considered one aspect of quality, cost and resource use alone without consideration of quality is not considered a quality performance measure.)

Peer groups for comparison – Creation of peer groups of providers caring for a similar mix of patients, within which to examine performance scores.

Performance measure – Numeric quantification of healthcare quality for a designated accountable entity such as hospital, health plan, nursing home, clinician, etc. (NQF measure testing report).

Risk Adjustment (also known as case-mix adjustment) – Statistical methods to control or account for patient-related factors when computing performance measure scores; methods include multivariable modeling, indirect standardization, or direct standardization. These methods can be used to produce a ratio of observed-to-expected, a risk-adjusted rate, or other estimate of performance.

Social Determinants of Health – Healthy People 2020 defines social determinants of health as conditions in the environments in which people are born, live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks. Conditions (e.g., social, economic, and physical) in these various environments and settings (e.g., school, church, workplace, and neighborhood) have been referred to as "place." In addition to the more material attributes of "place," the patterns of social engagement and sense of security and well-being are also affected by where people live. Resources that enhance quality of life can have a significant influence on population health outcomes. Examples of these resources include safe and affordable housing, access to education, public safety, availability of healthy foods, local emergency/health services, and environments free of life-threatening toxins.

Social disadvantage – Braveman et al. define social disadvantage as "Unfavorable social, economic, or political conditions that some groups of people systematically experience based on their relative position in social hierarchies.[134] Social disadvantage indicates restricted ability to participate fully in society and enjoy the benefits of progress. Social disadvantage is reflected, for example, by low levels of wealth, income, education, or occupational rank, or by less representation at high levels of political office.

Sociodemographic – Broad term referring to a variety of socioeconomic (e.g., income, education, occupation) and demographic factors (age, race, ethnicity, primary language).

Socioeconomic Status – Broadly conceptualized as one's relative position within society. Socioeconomic status has traditionally been defined and measured by education, income, and occupation.[9]

Stratification – Computing performance scores separately for different strata or groupings of patients based on some characteristic(s) — i.e., each healthcare unit has multiple performance scores (one for each stratum) rather than one overall performance score.

Appendix C: Outcome Performance Measures and Risk Adjustment – the Basics

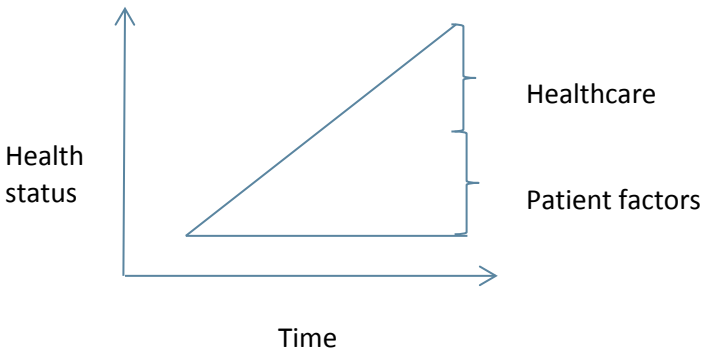
Outcome performance measures aggregate the data on individual patient outcomes for an accountable entity (e.g., hospital, clinician, nursing home). Outcomes generally are a function of several inputs including patient factors, treatment effectiveness, quality of care, and random events. This can be represented as an equation:

> **Outcomes = *f*** (intrinsic patient factors, treatment effectiveness, quality of care, random chance) [6, p. 5]

This equation is a simplified description because outcomes also may be a function of complex interaction among these factors.

Outcomes often represent a change in some health status indicator (e.g., function, pain) over time; that change can be due to both healthcare and patient factors as represented in Figure C1. Some outcomes, such as hospital readmission, are considered a proxy for a change in health status.

Figure C1. Outcome as a Change over Time



Risk Factors

Iezzoni[6, p. 31] identified the major categories for the potential patient factors that may influence outcomes to include the following. This is not a comprehensive list and concepts may overlap. Additionally, not all factors may affect every outcome.

- Genetics (e.g., predisposition to conditions or health-related behaviors)
- Demographic characteristics (e.g., age, sex, race, ethnicity, primary language)
- Clinical factors (diagnoses, conditions and severity; physiologic stability; physical, mental, cognitive function)
- Psychosocial factors, socioeconomic, and environmental factors (e.g., family, education, occupation, economic resources, health insurance, neighborhood)
- Health-related behaviors and activities (tobacco, diet, physical activity)
- Quality of life, attitudes, and perceptions (health-related quality of life and overall health status; preferences; cultural, religious beliefs, and behavior)

The final selection of risk factors involves an iterative process using the guidelines identified in Table 5 including:

- Clinical/conceptual relationship with the outcome of interest
- Empirical association with the outcome of interest
- Variation in prevalence of the factor across the measured entities
- Present at the start of care
- Is not an indicator or characteristic of the care provided (e.g., treatments, expertise of staff)
- Resistant to manipulation or gaming
- Accurate data that can be reliably and feasibly captured
- Contribution of unique variation in the outcome (not redundant)
- Potentially, improvement in risk model metrics of discrimination and/or calibration
- Potentially, face validity and acceptability

Risk Adjustment in Outcome Performance Measurement

The ultimate goal of performance measurement is to facilitate improvement in healthcare and health. Measurement is used to identify differences in quality of healthcare and identify opportunities for improvement. Unlike many process performance measures, which are focused on care practices that should be delivered to all patients in a specified target population, the goal for outcome performance may not be 100% (or 0%). Due to the limits of science, not all patients will achieve the outcome (e.g., survive), and the "right" rate may not be known. Consequently, it is through comparison across providers that opportunities for improvement are identified. Providers with superior risk-adjusted outcomes set the goal for what is possible to achieve. In order for performance results to be meaningful and valid for identifying differences in performance across providers, outcome performance measures must be adjusted for different levels of risk in the patients served.

Outcome performance measurement is intended to identify the effect of care on the outcome of interest in order to make a conclusion about quality and direct efforts for quality improvement. As indicated in the equation and Figure 1, the relationship between healthcare and the outcome may be confounded by various patient factors. That is, patient factors (e.g., severity or complexity) are also correlated with the outcome and provide an alternative explanation for the outcome. Confounding factors need to be controlled or adjusted in order to make conclusions about the quality of care based on performance on the outcome measure.

Risk adjustment (also known as case-mix adjustment) refers to statistical methods to control or account for patient-related factors when computing performance measure scores; methods include multivariable modeling, indirect standardization, or direct standardization. These methods can be used to produce a ratio of observed-to-expected, a risk-adjusted rate, or other estimate of performance. Risk adjustment refers to the operations performed during the calculation of the performance score. Methods include:

- Comparison of observed-to-expected outcomes for an accountable entity
    - Indirect standardization where the expected number of outcomes is determined by applying stratum-specific rates determined from all patients to the number of cases in

each stratum for each provider — i.e., what is expected if the hypothetical average provider cared for the specific mix of patients
    - Extension of indirect standardization to multivariable statistical models [6]

- Direct standardization where provider-specific rates are calculated in each stratum and applied to the standard population case mix, producing an estimate of what would be expected if the provider were to treat the standard case mix. [6] This approach is not commonly used to profile performance.

Risk Model Evaluation

Statistical risk models are often evaluated on model discrimination (extent to which the model predicts higher probabilities of the outcome for patients who experienced the outcome than for those who did not) and calibration (the match between predicted and actual outcome rates within subgroups of the data such as risk deciles). It is important to recognize when assessing risk models used for outcome performance measures, the metrics of model discrimination such as C-statistic or R-squared are not necessarily expected to achieve comparable values as models that include and are intended to explain the contribution of all variables that influence the outcome. In risk models, the independent variables are purposely limited to patient risk factors; variables related to care processes or structures are not included so that differences in risk-adjusted outcome rates can be attributed to differences in the care provided, i.e., differences in quality.

Approaches to Statistical Modeling

Statistical modeling to estimate the provider score on the outcome involves choosing from among a variety of options including:

- Random effects with shrinkage estimators vs. fixed effects
- Shrinking toward the overall average or some other benchmark (e.g., average of “like” providers)
- Hierarchical models
- Bayesian analysis

The various methods may have different trade-offs and policy implications. For example, fixed effects models identify more outliers, some of which will be false positives; whereas, random effects models identify fewer outliers, some of which will be false negatives.[135]

Appendix D: Confounding – the Basics

Confounding is an epidemiological term that refers to the distortion in the degree of association between an exposure (independent variable) and an outcome (dependent variable) due to a mixing of effects between the exposure and an incidental (confounding) factor. [136]Confounding represents systematic error and threatens the internal validity of an epidemiologic study because it can lead to false conclusions regarding the true relationship between an exposure and outcome.

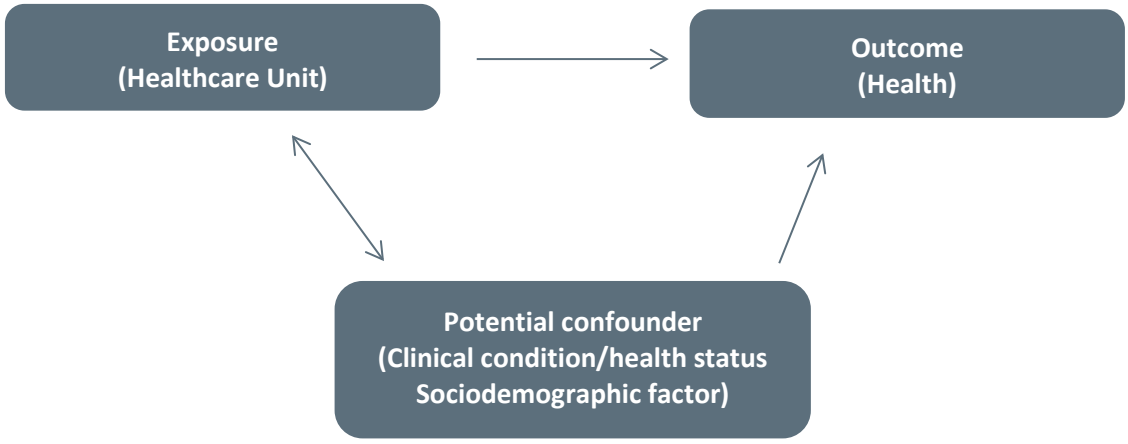
In the field of epidemiology, researchers often are interested in determining whether, how, and to what extent—an "exposure" to a particular entity (e.g., a microbe, a medication, or a procedure) is related to a particular outcome (e.g., a sickness, a recovery, or an improvement). The direction and magnitude of that relationship between the exposure of interest and the outcome is known as the "effect size"; it can be positive or negative, large or small, and statistically significant or not. In the case of outcome performance measurement, the "exposure" of interest is the healthcare unit's structures and processes of care that influence some particular outcome (e.g., mortality).

Usually, however, there are other factors—in addition to the exposure of interest—that are associated with that particular outcome. If such factors are related to the exposure of interest and are causally related to the outcome of interest, they can distort the effect size. This distortion is known as *confounding* and those other factors are known as potential c*onfounders*. The three characteristics of potential confounders are as follows:

- they are a risk factor for the outcome of interest,
- they are associated with the exposure of interest, and
- they are not affected by either the exposure or the outcome.[34]

Importantly, the third characteristic indicates that potential confounders do not represent an intermediate step in the causal pathway between the exposure of interest and the outcome; also, it can be satisfied by factors that precede both the exposure of interest and the outcome. The relationship between the exposure of interest, the outcome of interest, and potential confounders is shown in Figure D1.

Figure D1. **Relationship between exposure, outcome, and potential confounders**



Depending on the strength of the relationships between potential confounders and the exposure and outcome, the type and degree of distortion in the effect size can vary. For example, confounding can make an effect size appear to be statistically significant when it is not (that is, there may appear to be an actual relationship between an exposure of interest and a particular outcome, even when there is not one) or vice-versa. Confounding also can change the direction or magnitude of the effect size (that is, the relationship may appear to be a positive one when in fact it is negative, or it may appear larger [or smaller] than it actually is). Because confounders obscure the relationship between the exposure of interest and the outcome, researchers try to eliminate (or at least minimize) the distortion by "adjusting" for confounding factors in some way (often using statistical techniques).

The discussion of confounders can be extended conceptually to outcome performance measurement. The purpose of outcome performance measurement is to identify the effect of care on health-related outcomes, in order to make a conclusion about quality. In this case, the "exposure" of interest is to the health care unit and its various structures and processes of care that influence some particular outcome (e.g., mortality). As noted above, if other factors are associated with—but not the result of—actual structures or processes used in the provision of care and also influence the outcome of interest, the true "contributions" of the care structures/processes to the outcome may be obscured, because they are "mixed with" or distorted by the contributions of those other factors. In order to make correct conclusions about quality, adjustment for potential confounders is needed. In outcome performance measurement, potential confounders include patient-level characteristics that are risk factors for the outcome of interest that are present prior to the provision of care. Adjustment for such factors is known as risk adjustment or case mix adjustment.

Clinical factors present at the start of care (e.g., severity of illness) and/or other health status factors (e.g., self-reported health) typically are considered potential confounders in outcome performance measurement and therefore are included in risk adjustment strategies. However, risk factors such as genetic characteristics, sociodemographic factors, health-related behaviors, and less commonly available patient-level factors such as beliefs, attitudes, and perceptions may also be potential confounders and if so, should be included in the risk-adjustment strategy, as failure to do so may conceal the true relationship between the structure/processes of care and lead to incorrect conclusions about the quality of that care.

A complication that develops during the course of care can affect the outcome, but should not be considered a confounder because it is in the causal pathway between the exposure to the healthcare unit and the outcome.

Appendix E: Example of Checking for Between-Unit Effect

An example to analyze within- versus between-unit covariate effects when a patient-level variable "low income" is significant is to add to the model a unit-level variable "percent of patients of low income." The regression coefficient for the patient-level covariate is summarizing outcome differences of low income vs. non-low income patients at providers who are matched on the percent of low-income patients. The regression coefficient for the provider-level covariate "percent of patients of low income" is summarizing outcome differences for patients of the same income category who are treated by providers that differ with respect to their percent of low-income patients. If the patient-level covariate is negligible and the provider-level covariate is large, this is consistent with the interpretation that the association between income and outcome is related to systematic differences in qualify of providers who tend to treat more vs. fewer low-income patients (and not due to differences in outcomes for low versus non-low income patients within the same provider). In that case, care is needed because certain adjustment methods which fail to distinguish within- versus between-provider income differences may produce biased comparisons of providers. If the patient-level covariate is large and provider-level covariate is negligible, this suggests it may be important to adjust for income (to the extent that the mix of low-income patients varies across providers) and that failure to distinguish within- versus between-provider income differences may have negligible impact. If both patient-level and provider-level are large, then adjustment *methods should be used that remove the effects of within-unit differences (as they interact with varying unit proportions in the disadvantaged groups) but do not mask the quality differences among units.*

Examples in Literature (PubMed citation)

Feaster D, Brincks A, Robbins M, et al. Multilevel models to identify contextual effects on individual group member outcomes: a family example. *Fam Process,* 2011;50(2):167-183.
Abstract: This manuscript illustrates methods for utilizing measurements of individuals to identify group contextual effects on individual outcomes. Contextual effects can be identified by 1 of 3 methods: (1) divergence of the simple within- and between-group regression coefficients, (2) the presence of a cross-level interaction of the within- and between-group predictor variable, or (3) the effect of discrepancies within the group. These methods can be used to incorporate group context into an individual model and can be utilized for any individual process variable that might be affected by a group context. Example data include measures of hassles and coping adequacy of inner city, poor, African American new mothers, and their family members.

Reames BN, Birkmeyer NJ, Dimick JB, et al. Socioeconomic disparities in mortality after cancer surgery: failure to rescue. *JAMA Surg,* 2014;149(5):475-481.
Abstract: IMPORTANCE. Disparities in operative mortality due to socioeconomic status (SES) have been consistently demonstrated, but the mechanisms underlying this disparity are not well understood. OBJECTIVE. To determine whether variations in failure to rescue (FTR) contribute to socioeconomic disparities in mortality after major cancer surgery. DESIGN, SETTING, AND PARTICIPANTS. We performed a retrospective cohort study using the Medicare Provider Analysis and Review File and the Medicare Denominator File. A summary measure of SES was created for each ZIP Code using 2000 U.S. Census data linked to residence. Multivariable logistic regression was used to examine the influence of SES on rates of FTR, and fixed-effects hierarchical regression was used to evaluate the extent to which disparities could be attributed to differences among hospitals. A total of 596,222 patients undergoing esophagectomy, pancreatectomy, partial or total gastrectomy, colectomy, lung resection, and

cystectomy for cancer from 2003 through 2007 were studied. MAIN OUTCOMES AND MEASURES. Operative mortality, postoperative complications, and FTR (case fatality after >/=1 major complication). RESULTS. Patients in the lowest quintile of SES had mildly increased rates of complications (25.6% in the lowest quintile vs 23.8% in the highest quintile, P = .003), a larger increase in mortality (10.2% vs 7.7%, P = .0009), and the greatest increase in rates of FTR (26.7% vs 23.2%, P = .007). Analysis of hospitals revealed a higher FTR rate for all patients (regardless of SES) at centers treating the largest proportion of patients with low SES. The adjusted odds ratios (95% CIs) of FTR according to SES ranged from 1.04 (0.95-1.14) for gastrectomy to 1.45 (1.21-1.73) for pancreatectomy. Additional adjustment for hospital effect nearly eliminated the disparity observed in FTR across levels of SES. CONCLUSIONS AND RELEVANCE. Patients in the lowest quintile of SES have significantly increased rates of FTR. This finding appears to be in part a function of the hospital where patients with low SES are treated. Future efforts to improve socioeconomic disparities should concentrate on hospital processes and characteristics that contribute to successful rescue.

Appendix F: Illustration of Adjustment using Direct Standardization

With direct standardization, unit-specific rates are computed in each stratum and applied to a standard population case mix, producing an estimate of what might be expected if the provider were to treat the standard patient mix. You do not use a population average for the strata as in indirect standardization. This method sometimes becomes problematic if cell sizes are very small.

Table F1. Illustration of Risk Adjustment using Direct Standardization

	All Patients in National Population		**Unit A**		**Unit B**		**Unit C**	
SDS Strata	Patient Mix N/Percent	Clinically-Adjusted Deaths N/Percent	Patient Mix n/Percent	Clinically-Adjusted Deaths n/Percent	Patient Mix n/Percent	Clinically-Adjusted Deaths n/Percent	Patient Mix n/Percent	Clinically-Adjusted Deaths n/Percent
All Patients	1,000,000 100%	22,000 **2.2%**	1000	22 **2.2%**	1000	26 **2.6%**	1000	29 **2.9%**
Average-High Income	800,000 80%	16,000 2.0%	800	16 2.0%	400	8 2.0%	400	8 2.0%
Low Income	200,000 20%	6,000 3.0%	200	6 3.0%	600	18 3.0%	600	21 3.5%
Provider rate for average-high income stratum applied to national proportion of average-high income				2% * 80%= 1.6%		2% * 80%= 1.6%		2% * 80% = 1.6%
Provider rate for low-income stratum applied to national proportion of low income				3% * 20%= 0.6%		3% * 20%= 0.6%		3.5% * 20% = 0.7%
Risk adjusted death rate accounting for sociodemographic risk is what might be expected if provider were to treat the national standard patient mix equals the SUM of provider stratum rate multiplied by the national proportion for the stratum				**2.2%**		**2.2%**		**2.3%**

Appendix G: Responses to Comments

Following are the Expert Panel's responses to the major themes identified in the comments received during the public comment period.

1. Masking disparities, masking quality problems, different standards

Commenters agreed with the recommendation that stratification was the appropriate method to identify disparities. However, some commenters objected to sociodemographic adjustment for purposes of public reporting and pay-for-performance and urged continuation of NQF's existing criteria and guidance. They expressed concerns that adjusting for sociodemographic factors masks disparities in outcomes, masks quality problems, creates different standards, and reduces the incentive to improve and reduce disparities. Other commenters noted that the analyses that are needed to include adjustment for sociodemographic factors would highlight where there are disparities (i.e., significant coefficient in a risk model). Some commenters suggested that both SDS-adjusted and stratified data be publicly reported.

Response

The term "masking disparities" is a misnomer because disparities are not visible using current clinically-adjusted measures. Masking disparities in outcomes (or processes), masking disparities in quality, and setting different standards, while related, represent distinct concerns. The Expert Panel provides two responses — one methodological and one to provide for greater transparency about disparities. Both of these are discussed in detail in the final report.

- The Expert Panel developed an in-depth discussion of the methodological basis for SDS adjustment, which is provided in section 4.
- The Expert Panel recommended that if a measure was SDS adjusted, then specifications also include instructions for stratification.

2. Evidence of harm

Some of the objections to sociodemographic adjustment were based on the perception that the primary reason for the recommendations was potential harms to disadvantaged patients related to not adjusting for sociodemographic factors and that there was insufficient evidence of such harms. Therefore, they concluded that a change in the criteria related to adjusting for sociodemographic factors is not warranted.

Response

a. Whether to adjust for sociodemographic factors or not, and how, is first and foremost based on sound methods for quality measurement. That is, the Panel first asked the question: "Will consideration of sociodemographic adjustment improve comparability of performance between providers?" Sound measurement represents the central tenet of performance assessment and enables optimal decision making among patients, purchasers, and payers to make informed comparisons between providers and inferences about their relative quality. Sound measurement also improves perceptions of fairness among those being assessed. The majority of the Panel thinks that sociodemographic adjustment, under the conditions identified in the report and in the detailed discussion of methods (see section 4) will produce performance measures that will provide more

valid, meaningful, and fair comparisons among plans and providers on key dimensions of quality of care. This focus on best possible comparative measurement of quality is consistent with NQF's focus on quality measurement per se, rather than on the actual consequences of uses of measures (which it does not directly control).

b. The primary evidence that is relevant to the question of whether or not to adjust for sociodemographic factors is the substantial body of evidence that demonstrates the relationship between a variety of sociodemographic factors and a variety of health outcomes (and some processes). However, it is important to note that the recommendations do not suggest sociodemographic adjustment of all performance measures, or even all outcome performance measures. The decision on whether to include sociodemographic factors needs to be made for each individual performance measure based on the conceptual and empirical relationships that exist between the factors and the outcome or process being measured as well as working through the guidelines for selecting risk factors. Therefore, a body of evidence about the relationship between sociodemographic factors and outcomes (or processes) provides only a starting point for considering sociodemographic factors as confounders and potential risk adjustment.
c. The potential harms from not adjusting for sociodemographic factors identified in the report are potential consequences of not following accepted and sound methods to control for confounding (see the methods discussion in section 4). The Panel reviewed a number of published studies documenting harm to safety net providers, primarily through financial penalties. Fewer studies addressed potential reputational harm to providers. No studies directly assessed harm to patients under the current policies. The Panel recognized that it is a plausible, but unproven assumption that reducing revenue to financially strapped safety net organizations could eventually result in fewer resources devoted to care for disadvantaged patients resulting in worse outcomes. A few additional references related to potential harms have been identified, but that is not the primary evidence question.
d. The Panel notes that the current policy prohibiting sociodemographic adjustment was not based on empirical evidence of benefit or harm to patients. It also notes that the National Healthcare Disparities Report produced by AHRQ shows little consistent progress in reducing healthcare disparities during the time of the current policy of prohibiting adjustment for sociodemographic factors. There also is not a body of evidence on potential harms to patients related to allowing sociodemographic adjustment (e.g., setting different standards and reducing incentives to improve).
e. Therefore, the recommendations are based on sound principles of measurement science, and the decision of whether to adjust for sociodemographic factors needs to be made for each individual measure based on the conditions laid out in the recommendations.
f. We have corrected the cited reference (#24 in the draft report) – the text was correct, but the correct citation: is Joynt, KE, Jha, AK (2013). Characteristics of hospitals receiving penalties under the Hospital Readmissions Reduction Program. *JAMA*, 309(4), 342-343).

3. Definition of quality, healthcare responsibility, reduce incentive to improve, impede progress on outcomes such as readmission

Some commenters thought that the discussion about what healthcare plans or providers can control or influence reflected a narrow view of healthcare quality and provider responsibility to adjust care based on sociodemographic factors. Some expressed concern that sociodemographic adjustment would impede progress that is being made on hospital readmissions and that hospitals would abandon efforts to reduce readmissions (or potentially other important outcomes) as a result of sociodemographic adjustment.

Response

a. The Expert Panel agrees that healthcare should be based on the characteristics of the patients served; should not lower goals or standards when providing care to disadvantaged patients; and the need to identify and reduce disparities.

b. That said, the vast majority of comments received during the public comment period made some mention of factors outside of providers' or health plans' control that influence measured outcomes. Most outcomes are clearly a function not only of what plans and providers do, but of other factors operating at the individual, household, community, and broad societal levels. There is no widely-accepted definition of quality of care that holds doctors, hospitals, health plans, and other sorts of "providers" responsible for ALL factors leading to many measured outcomes.

c. Sociodemographic risk adjustment does not contradict broad definitions of healthcare quality reflected in the IOM definition of healthcare quality; or others such as AHRQ's: "Doing the right thing for the right patient, at the right time, in the right way to achieve the best possible results"; or CMS definition from its QI Roadmap: "Right care for every person every time." All of these definitions focus on what healthcare entities do, not about what society does or does not do.

d. Risk adjustment for certain factors does not absolve providers/plans from the responsibility to use interventions appropriate for those factors when present in the patients served whether clinical factors (e.g., recognizing and addressing comorbidities) or sociodemographic factors (e.g., recognizing and addressing non-English speaking persons). This holds whether clinical factors or sociodemographic related factors are being considered for adjustment.

e. Adjustment for sociodemographic factors when indicated improves comparability among providers/plans. It does not place a limit on the scope of interventions that could be used to mitigate the effects of sociodemographic factors such as the number of language translations or interpreters available or "discharge clinics" for patients without primary care providers. Risk adjustment creates a "level playing field" so that differences across providers/plans in addressing or not addressing the sociodemographic factors will be reflected in the adjusted performance measure scores.

f. Risk adjustment does change the estimate of the provider's performance (either up or down) depending on the proportion of patients in the SDS categories. This is appropriate in the context of the question: ***how would the outcomes of various units compare if hypothetically they had the same mix of patients*****?** (See section 4.) However, if the question is: ***how do the outcomes of patients with different characteristics compare (either within an individual unit or at the population level)*****?** then a different analysis is indicated. As recommended by the Expert Panel and in prior NQF projects, identifying disparities in either outcomes or processes requires additional information and analysis (e.g., stratification by relevant sociodemographic characteristic).

g. Adjustment for sociodemographic factors when indicated does not necessarily remove the focus of improvement or the need to work collaboratively with other settings, depending on the performance measure. By measuring and comparing performance on risk-adjusted rates, providers/plans, and others can identify when performance is lagging and providers/plans that are achieving excellent performance. For improvement, providers/plans always need to examine their own data stratified by relevant clinical and/or sociodemographic characteristics to identify patients who are and are not achieving desired outcomes and potential strategies to improve. Additionally, risk adjustment procedures should be updated on a periodic basis so that improvements are reflected in updated model coefficients.

4. Methods

Some comments were about methods or description of methods in the report:

- ***"Not primarily mediated by quality" should not be a requirement for selecting risk factors***
  Comments by a statistician and an epidemiologist caution against focusing on causal pathways. The statisticians on the Panel also recommended that this language is not needed. It is difficult to define in order to operationalize and therefore, could potentially add burden to the measure development process.

Response

a. Based on epidemiologic principles related to confounding and statistical theory of causal inference, the language "not primarily mediated by quality" is not needed (see methods discussion in section 4) and has been omitted from the revised recommendations. The decision on whether to include sociodemographic factors needs to be made for each individual performance measure based on the conceptual and empirical relationships that exist between the factors and the outcome as well as working through the guidelines for selecting risk factors. However, an assessment of a conceptual relationship between an SDS factor and outcome of interest includes a consideration of whether the effect of the SDS is primarily mediated by the quality of care delivered. This is discussed in section 6.

- ***Disagree with characterization of sociodemographic adjustment making more "accurate" or "correct" conclusions and suggest language that risk adjustment improves comparability***

Response

b. One of the core principles used the language "avoid making incorrect inferences about performance" and is an appropriate statement related to risk adjustment. Making correct conclusions is a logical statement of the same concept from a positive perspective. Language used to describe validity (to which risk adjustment relates) often refers to "accurate" and "correct" but varies by disciplines and preference. The term "accurate" is also used sometimes to indicate precision and could be confusing. The references to "accurate" have been replaced with the terms: avoid incorrect inferences, improve comparability, and make unbiased estimates (statistical term used in the methods report) depending on the context.

5. Implementation is the issue, not measurement

Some of the objections to the recommendations were based on the perspective that the issue (harm to providers or patients through lack of adjustment) was really about how the measures were used in pay-for-performance programs and not about measurement per se. Some suggested alternative ways to structure incentive programs. Some advocated for peer group comparisons as recommended by MedPAC for the hospital readmission measure. However, some other commenters suggested that the alternative of peer groups for comparison explicitly accepts or creates different standards for plans or providers grouped by a sociodemographic variable. One commenter noted specific mechanisms for adjusting payment for services based on higher needs related to sociodemographic factors and therefore, adjustment for performance measures could result in overpayment.

Response

a. The panel focused primarily on the question of whether consideration of sociodemographic adjustment improves the performance measure for comparisons and avoids incorrect inferences about quality. Although concerns about the impact of payment incentive programs might have been the impetus to re-examine NQF's policy on adjusting for sociodemographic factors, the primary basis for the recommendations is that they are consistent with accepted practices and guidelines for selecting risk factors for performance measurement and epidemiologic and statistical approaches to handle confounding in order to enable comparisons and avoid incorrect inferences about quality regardless of the specific accountability application.
b. The concerns of the Panel have not just been limited to issues of payment incentive programs. Rather, the concerns of the Panel are also set in the context of public reporting and the validity of inferences or comparisons made with performance measures that are not adjusted for sociodemographic factors when appropriate. Alternatives to adjustment that may be useful in pay-for-performance contexts do not address a deeper concern that failing to consider sociodemographic adjustment can yield performance measures that may be fundamentally misleading to patients, consumers, purchasers, payers, and regulators who are engaged in making comparisons among plans or providers.
c. Appropriate adjustment for sociodemographic factors may not be sufficient to address the financial issues of safety net providers/plans; however, the performance measures used in such programs should provide an unbiased estimate (i.e., without systematic deviation from the true value) of performance on the quality measure for the entity being measured and compared.
d. Peer groups for comparison: As noted by some commenters, unlike model-based adjustment, this approach does have the potential to mask quality differences. One commenter elaborated: "The two approaches are fundamentally different in that risk adjustment adjusts for the distribution of patient characteristics (such as poverty), while peer group comparison adjusts for unit characteristics. For example, if comparisons are made within a peer group of hospitals that have trouble providing high quality care because they are under resourced and poorly reimbursed, we might say a hospital is superior to its peer group even though the same patients would have received superior care at another hospital outside the group. Conversely if a hospital is superior in risk adjusted scores, it suggests that the same group of patients would do better there than at another hospital. Peer group comparison may have a place as a tool of the incentive system rather than as part of the construction of the measure itself."
e. Stratification: The Expert Panel discussed the statistical limitations, mainly in the form of small sample sizes for computing performance scores for each stratum for an individual physician, or small physician group, or small hospital. One of the commenters elaborated: "Sample sizes for some measures adequate for estimation (with adequate reliability) of a single measure for a unit, but not for separate estimation of measures for strata (subgroups), especially when some strata have only sparse representation in some units. This is unlikely to be a problem, however, for model-based statistical adjustment, since model parameters may be estimated from the combined data from a multitude of units. Furthermore, these model parameters give a summary measure of within-unit disparities that typically is more sensitive than what can be discerned from perusing a set of stratified results."

6. Burden to developers, guidance to developers

Some of the objections were based on burden to measure developers and concern that developers would not develop performance measures that required sociodemographic risk adjustment. Other

commenters cautioned about potential developer burden and suggested more guidance for developers would be needed.

Response

a. Risk adjustment is a complex and nuanced area of methodology and requires expertise that may not be present in all measure developers. It is difficult to anticipate all possible scenarios to create more prescriptive directives and rules. Measure developers need the flexibility to use the methods that are indicated in a particular situation.
b. Although plans and providers may not directly pay measure developers for their work, the support for, and potentially greater acceptance of, sociodemographically adjusted measures by plans and providers will give some measure developers an opportunity rather than a burden.
c. The measure submission questions should guide what measure developers are expected to present for review and evaluation.
d. Nothing in the Panel's recommendations asks or demands that measure developers collect or analyze primary data. The obligation on measure developers is presumably exactly like the obligation that already exists with regard to clinical variables used for adjustment. Measure developers will be obliged to recognize, and incorporate when possible, existing valid empirical data on the association between sociodemographic factors and "outcomes" (or some processes). When such data exist, developers may have to do more work than under the current policy prohibiting sociodemographic risk adjustment; this additional work should not be a barrier to the creation of measures that, in use, will provide more valid and informative comparisons among plans and providers.
e. Initially, data limitations may constrain what is feasible, and NQF Committees will need to recognize that. If the recommendations are implemented, SDS data will improve over time.

7. Data burden, feasibility

Some commenters saw sociodemographic data limitations as a reason to delay implementation. Other commenters cautioned about the potential of making data collection too burdensome. Some commenters noted that potential adjustment for sociodemographic factors would provide incentive to collect the necessary data. Some commenters noted other efforts related to data on sociodemographic factors, specifically recent IOM work *Capturing Social and Behavioral Domains in Electronic Health Records: Phase 1* .

Response

a. Initially, data limitations may constrain what is feasible either in the sense of development and testing of adjustment models or in the sense of using an SDS-adjusted measure in public reporting or pay-for-performance. NQF Committees will need to recognize that. The collection and availability of sociodemographic data are likely to advance as follows:
    - Initially, developers will primarily need to use variables readily available in existing data sets (e.g., Medicaid status); then
    - patient or member address for geocoding to census tract data; then
    - standard definitions and data collection processes as defined and supported by groups such as AHRQ, IOM, and CMS.

8. Additional sociodemographic factors, community factors

Some commenters suggested additional factors that should be considered or that more attention should have been given to community-level factors.

Response

a. Potential sociodemographic factors were identified in the report, but currently there is no basis for being more prescriptive about specific risk factors, especially when decisions about risk adjustment need to be determined for each individual measure.
b. Sociodemographic factors could be obtained from three sources:
    - socicodemographic data collected from each individual (e.g., race/ethnicity, literacy, homelessness, English proficiency, marital status, etc.);
    - census variables obtained through address geocoding usually at the census tract level, but could be identified for other levels like ZIP Code (e.g., percent below poverty level, percent employed, average education level); and
    - community resource variables that come from sources other than census data (e.g., strength of primary care network in a community, availability of visiting home nurses, Meals on Wheels, public transportation, community health centers, etc.).
c. The Panel agrees that community factors such as availability of public transportation, size and strength of community health center network, availability of primary care, availability of support services like Meals on Wheels, etc. can have a profound effect on patient outcomes. Risk adjustment addresses patient characteristics (see section 4) so community characteristics need to be specific to each patient. Community characteristics could be assigned to each individual patient (e.g., percent poverty or public transportation in community ***where the patient resides***). Generally, characteristics of the healthcare unit are not considered patient risk factors (e.g., percent poverty or availability of public transportation in the community ***where the unit is located*** or ***for the patient population served***). However, some community factors such as public funding of safety net providers have implications for the capacity of healthcare units to deliver quality service and policy implications for the response to performance assessment.

9. Implementing the recommendations and monitoring impact

Some commenters suggested more research, incremental approaches to implementation, and monitoring impact. Other commenters suggested immediate implementation and review of endorsed measures to identify those that might require an ad hoc review.

a. Adoption of the Panel's recommendations about sociodemographic adjustment and stratification will inevitably be "incremental." That is, measures currently in use will not have to be considered for sociodemographic adjustment until the next review cycles for those measures come up. Some measures for which a strong conceptual argument for adjustment exists will not be able to be implemented with sociodemographic adjustment because data constraints prevent development and validation of an adjustment model. For other measures, the data may be available to develop and validate a model, but not be available to routine use in a large population of plans or providers.
b. The limited evidence available to date about the effects of sociodemographic adjustment suggests that the effects will not be profound. That is, providers or plans may move to some extent up or down in relative rankings, but "good" will not instantly become "bad" and vice-versa. Effects of adjustment will likely be modest, based on analyses that have been done and reported to date, but could be substantial for some healthcare units.

c. To address these concerns, the Expert Panel recommended a transition period to monitor implementation (Recommendation #2) and that NQF appoint a new standing committee focused on disparities to review implementation and monitor o unintended consequences (Recommendation #3).

10. Clarifications

Some comments requested specific clarifications or indicated the need for clarification. Following are some specific clarifications.

Are health plans included?

Are cost and resource use measures included?

Some comments seemed to imply that all performance measures would be adjusted for sociodemographic factors.

Response

a. The recommendations apply to performance measurement for any setting or unit of analysis, including health plans.
b. The recommendations apply to outcome performance measures (including cost and resource use and PRO-based performance measures) and some process measures depending on the specific circumstances. The recommendations are purposely not prescriptive in terms of factors and methods — that needs to be determined for each individual measure.
c. The recommendations do not mean that all performance measures should be adjusted for sociodemographic factors — that has to be determined for each individual performance measure. The Panel's recommendations and supporting text are clear that the recommendation about sociodemographic adjustment applies only in specific circumstances. Examples of measures that would generally not be adjusted are provided in the report.

11. Opposed to NQF having a role in guidance on implementation of endorsed performance measures (Recommendation 7)

Five commenters who were in support of most of the recommendations did not agree that NQF should have a role in providing guidance on implementation and use of endorsed performance measures. The commenters think that this is outside NQF's role for endorsing performance measures and overlaps with the role of the Measure Applications Partnership (MAP).

Response

How a measure is implemented involves multiple decisions that could affect the validity of conclusions (inferences) made about quality of care and potential unintended consequences. The recommendation is for NQF to consider expanding its role to include guidance on implementation of performance measures**.** This will require NQF's decisionmaking bodies (CSAC and Board) to explore the pros and cons and implications for endorsement and measure selection for specific program uses. This fits with work already underway at NQF to explore ways to make the measure endorsement and measure selection processes more coherent and efficient.

National Quality Forum
1030 15th St NW, Suite 800
Washington, DC 20005
http://www.qualityforum.org

ISBN 978-1-933875-66-8

Attachment 4:

"Safeguarding Access to Medically Complex Care for Children by Requiring Health Plans and Exchanges to Develop Adequate Provider Networks"

CHILDREN'S HOSPITAL ASSOCIATION

600 13TH STREET, NW, SUITE 500
WASHINGTON, D.C. 20005
p | 202-753-5500
f | 202-347-5147

6803 WEST 64TH STREET
OVERLAND PARK, KS 66202
p | 913-262-1436
f | 913-262-1575

WWW.CHILDRENSHOSPITALS.ORG

Children's Hospital Association Consensus Statement

Safeguarding Access to Medically Complex Care for Children
by Requiring Health Plans and Exchanges to Develop Adequate Provider Networks
June 2014

Safeguarding Children's Access to Care

- Children are a unique population. To appropriately address the health care needs of all children, regardless of age[1], network adequacy standards must be specifically evaluated with respect to any insurance product that is offered in the commercial marketplace or by state or federal governments, and which purports to provide services to children. This includes traditional insurance products, managed care products, coverage offered through Medicaid and/or the Children's Health Insurance Program (CHIP), and products sold on the new health care Exchanges.

- Network adequacy should be evaluated for two specific subpopulations of children:

    - First, plans should demonstrate that adequate primary care access is available and affordable for all children that are generally well. The NCQA HEDIS measures and others are a good starting place for evaluating access and performance of well-child networks. It should be noted that there have historically been significant problems with access to dental and mental health services for generally well children. Health plans should demonstrate that they can provide adequate and affordable access to pediatric-appropriate specific services, using the Medicaid Early Periodic Screening Diagnostic and Treatment benefit as a model.

    - Second, network adequacy must be specifically evaluated for children with more complex or chronic health care needs (children with special health care needs). For example, a recent Health Affairs[2] article found that 10 percent of children enrolled in Medicaid and/or CHIP account for approximately 70 percent of the resource utilization in pediatrics. As such, payers should demonstrate that their networks are adequate and are affordable for the subpopulation of children that represents the majority of their anticipated pediatric spending in any given year.

This consensus statement focuses on the subpopulation of children with special health care needs, which is defined in simple terms as the top 10 percent of utilizers in the most recent rate year.

---

[1] This consensus statement addresses the provision of health care services to children of all ages, including adolescents.
[2] G. Kenney, J. Ruhter, and T. Selden, "Containing Costs and Improving Care for Children in Medicaid and CHIP," Health Affairs 28:6 (2007): w1012- w1036 web exclusive.

Essential pediatric community providers[3]

In order to serve children adequately, networks must include one or more pediatric hospital providers that are in the geographic area and maintain comprehensive pediatric specialty services. At a minimum, these hospitals should have the capacity to provide neonatal services, critical and intensive care, surgical, emergency/trauma services, and the relevant range of ancillary supports, such as occupational, physical, and speech therapy services; specialized pharmacies; anesthesia; durable medical equipment; and linguistically and culturally responsive providers/services.

- Public and private insurance products should provide access to the full range of pediatric subspecialty services typically required to care for sick children of all ages. This should include, but not be limited to, access to pediatric cardiology, neurology, nephrology, developmental medicine, psychiatry, gastroenterology, orthopedics, and radiology. In general, public and private insurance plans should demonstrate that the specialists in question have been trained and credentialed or have substantial experience treating children.

- Network adequacy should be evaluated in consideration of providing access to an appropriate pediatric facility—which may be in another state. It may be necessary for children to travel significant distances and to other states to receive treatment. Therefore, if an appropriate pediatric specialty hospital is not available in the immediate geographic area, arrangements should be made to refer the patient to an appropriate facility in the same or different state, at no greater out-of-pocket expense.

    - Health plans should specifically contract with these providers "in network"; it is unacceptable to allow payers to contract for the majority of services required on an out-of-network or single-case agreement basis.

- Plan networks should be designed to provide services for all levels of complexity, including for rare conditions, without administrative or cost barriers for consumers. In plans with tiered provider networks, pediatric specialty care providers should be in no higher than the second least expensive tier. When pediatric specialty care is predominantly provided in one or two provider settings, those settings should be in the lowest cost tier. Consumers must be informed of cost sharing requirements associated with the tiers.

- Pediatric provider networks should be stable, and public and private insurers should be required to inform consumers well in advance of any significant changes to their networks. Transitional coverage should be mandated for care that is "in process" (active treatment) at the time that network changes are being made.

---

[3] The following requirements could be readily supplemented or refined based on the historic experience of existing insurance products for the population of children and adolescents with special health care needs.

- Children with special health care needs require a network that can support an appropriate transition to adult care providers. The network must include adult providers who are able to provide for their unique needs or pediatric and/or adolescent specialists who are able to provide specialized care through young adulthood.

Out-of-network requirements

Children should not be penalized by the health plan if their care is provided by an out-of-network provider because there is no provider available within network that is capable of providing a covered benefit.

- When out-of-network care is received, cost sharing and other requirements for the consumer should be the same as if the plan was contracted and in-network.

- Plans should demonstrate that they maintain an adequate and timely approval process for out-of-network services, utilize appropriate pediatric clinical standards in evaluating requests, and have an appeals process for denied services. Regulators should assure that the insurance company does not impose additional barriers to access, including onerous prior authorization processes.

- Out-of-network arrangements should only be used as an exception for extremely rare services, such as transplantation.

- Single case agreements are not an acceptable alternative to plans having a generally adequate pediatric network.

- Payers should be required to reimburse hospitals the reasonable and customary value for services that are provided on an out-of-network basis.

Monitoring requirements

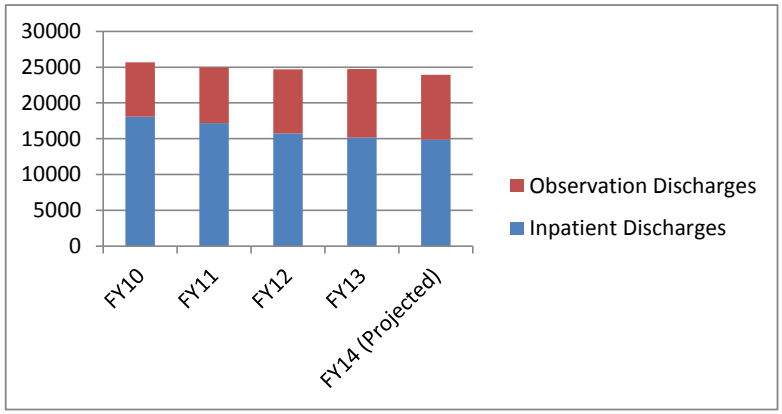
The insurer should proactively address access to pediatric specialty services when filing provider network information.

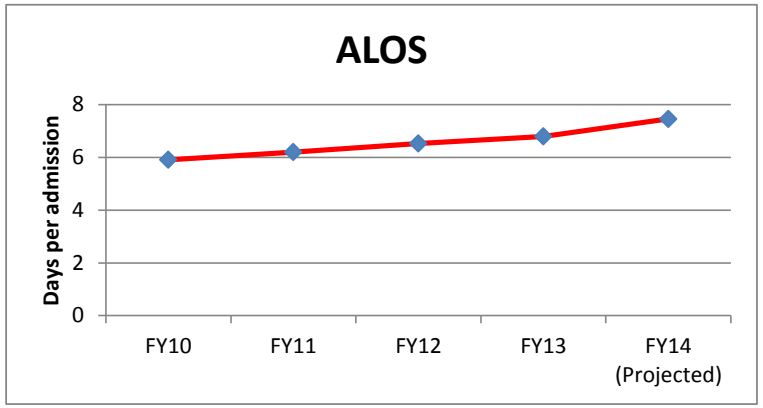
- Monitoring agencies, such as state insurance departments, should establish reporting requirements that specifically address access to care for children with special health care needs. Plans should expressly report on the utilization of services by this population, including the nature of the services, their location, and the extent to which these services were provided outside the contracted networks.

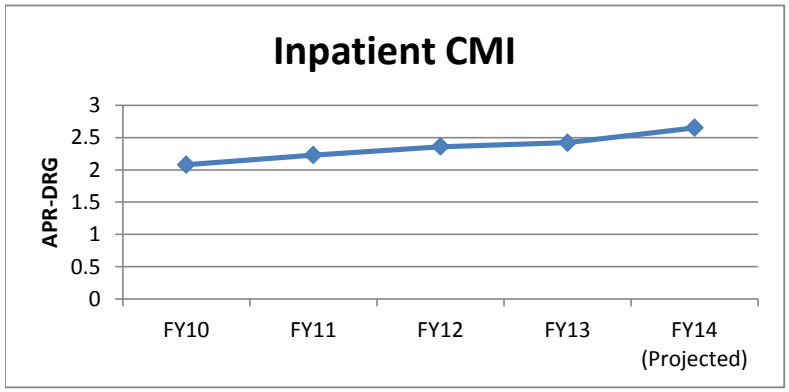
- Plans should report on all existing externally benchmarked, risk-adjusted pediatric quality and outcomes standards utilized in their contracts with hospitals and subspecialty networks. In addition, plans should identify which of these standards were used to determine the relative quality of the limited or tiered network plans they are intending to market in the event those types of networks are established.

Chart 1:

Recent Hospital Trends


30000
25000
20000
15000
10000
5000
0
FY10
FY11
FY12
FY13
FY14 (Projected)
Observation Discharges
Inpatient Discharges


ALOS
Days per admission
8
6
4
2
0
FY10
FY11
FY12
FY13
FY14
(Projected)


Inpatient CMI
APR-DRG
3
2.5
2
1.5
1
0.5
0
FY10
FY11
FY12
FY13
FY14
(Projected)

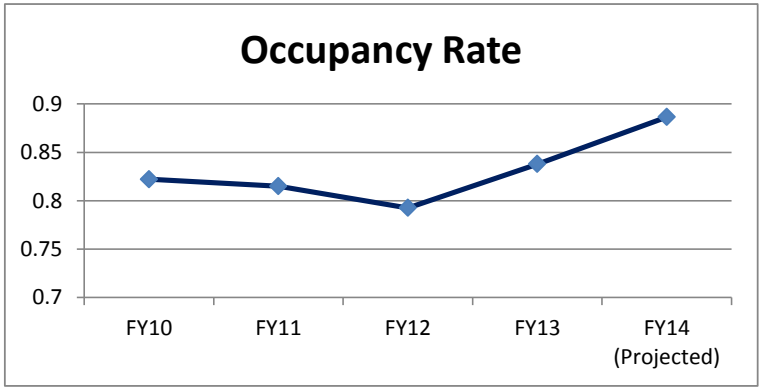

Occupancy Rate
0.9
0.85
0.8
0.75
0.7
FY10
FY11
FY12
FY13
FY14
(Projected)

Chart 2:

Cumulative Savings and Cost Containment

Cumulative Savings and Cost Containment

Boston Children's Hospital has been working on efforts to reduce costs from every angle, including unit price, efficiency and utilization. The hospital has decreased our overall per unit cost, volume adjusted, each year for the last five years. We have taken over $125M of expenses out of our system. In FY2013 we implemented $76M in clinical cost savings and early indications suggest an additional $24M in cost savings for FY2014. If successful, we will surpass reductions in costs of over $200M over the last several years. This chart provides a depiction in our unit costs relative to CPI and CPI-M benchmarks since 2009. Early end-of-fiscal year projections suggest that the trend will be better than budgeted for FY14

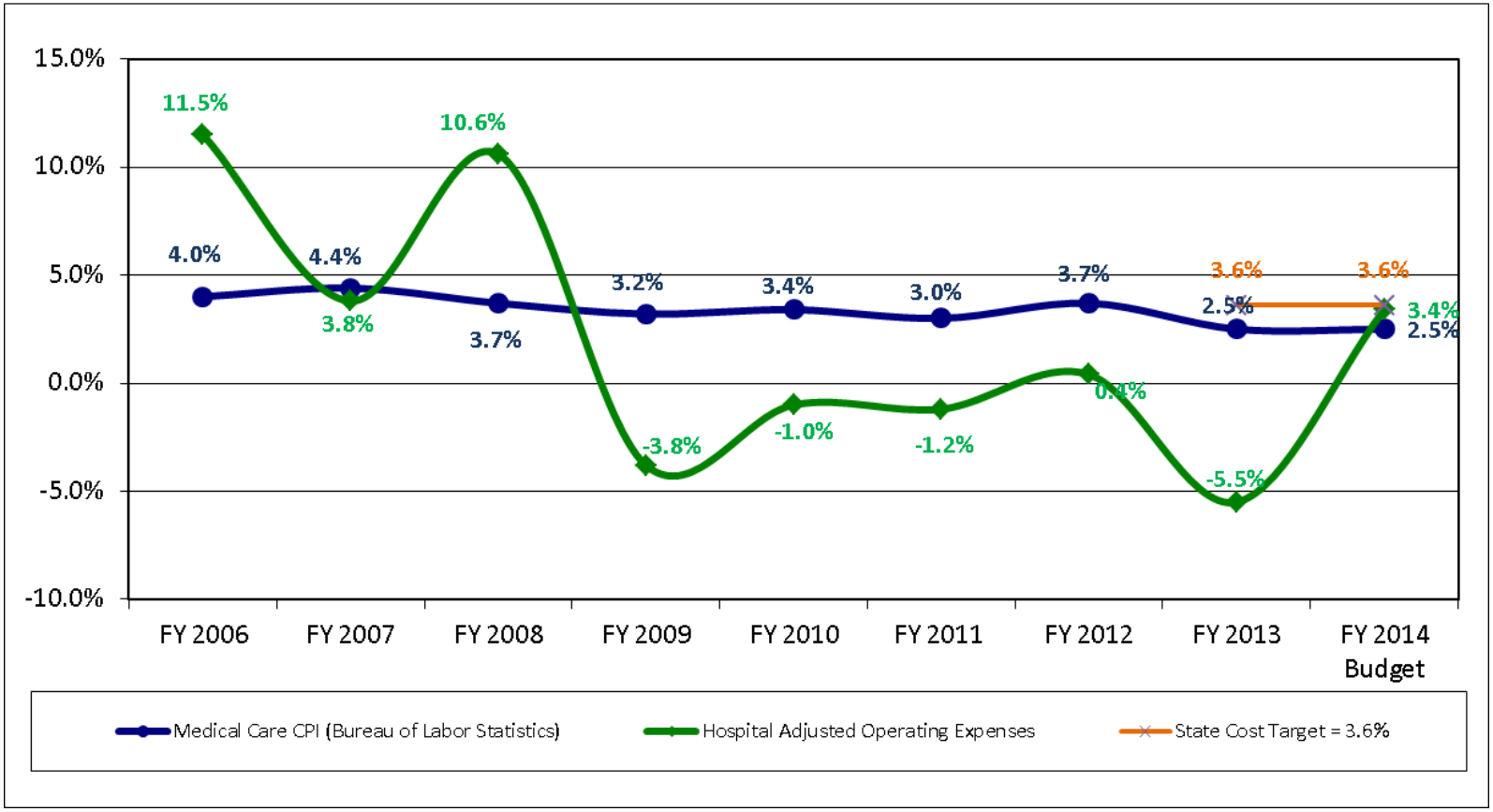


Chart 3:

Medicaid Losses from FY05 – FY14

Medicaid Losses FY05-FY14

Includes MA FFS, Medicaid MCO's, Supplemental & P4P payments, HST

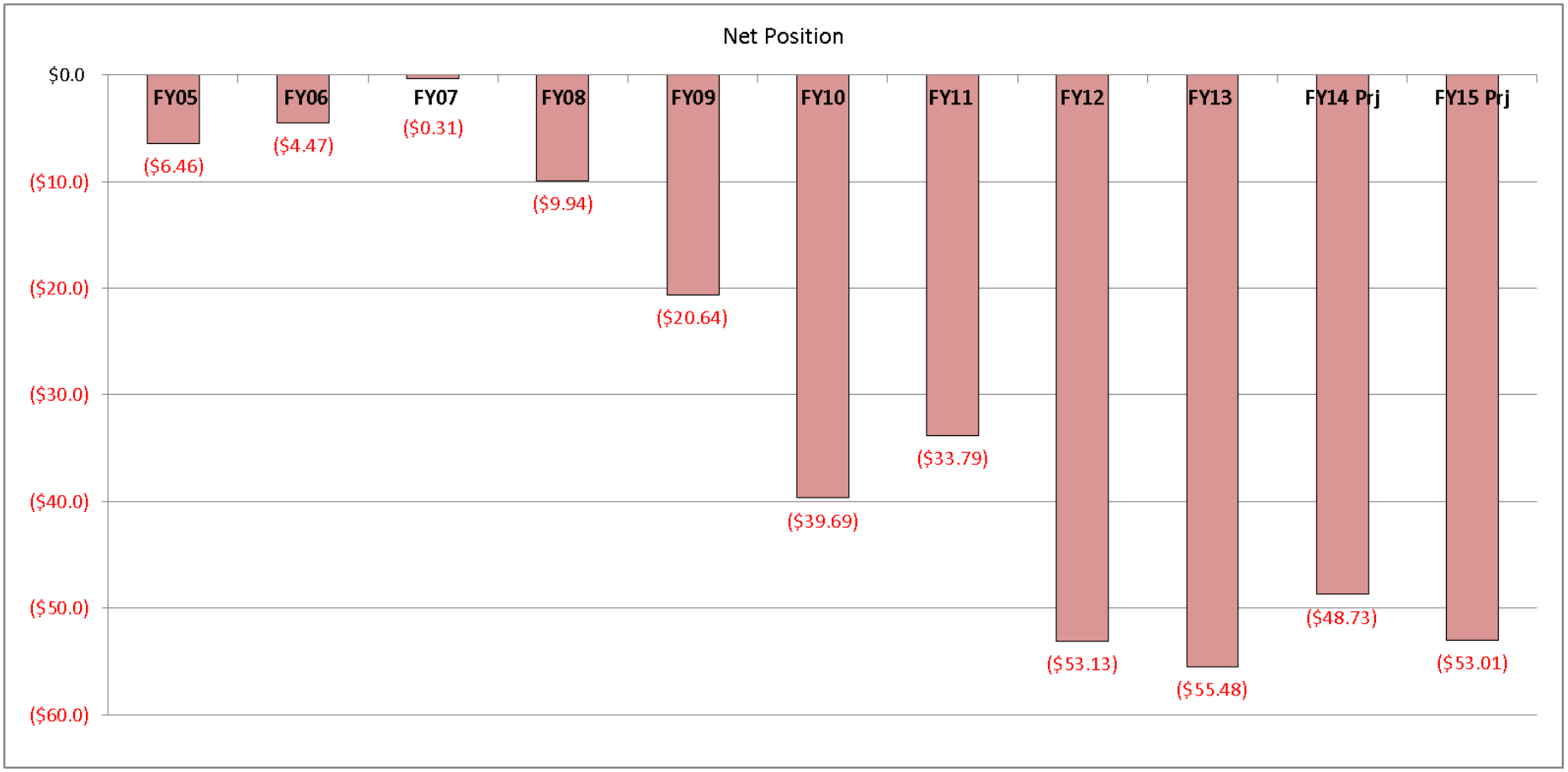


Chart 4:

Integrated Care Organization Budget

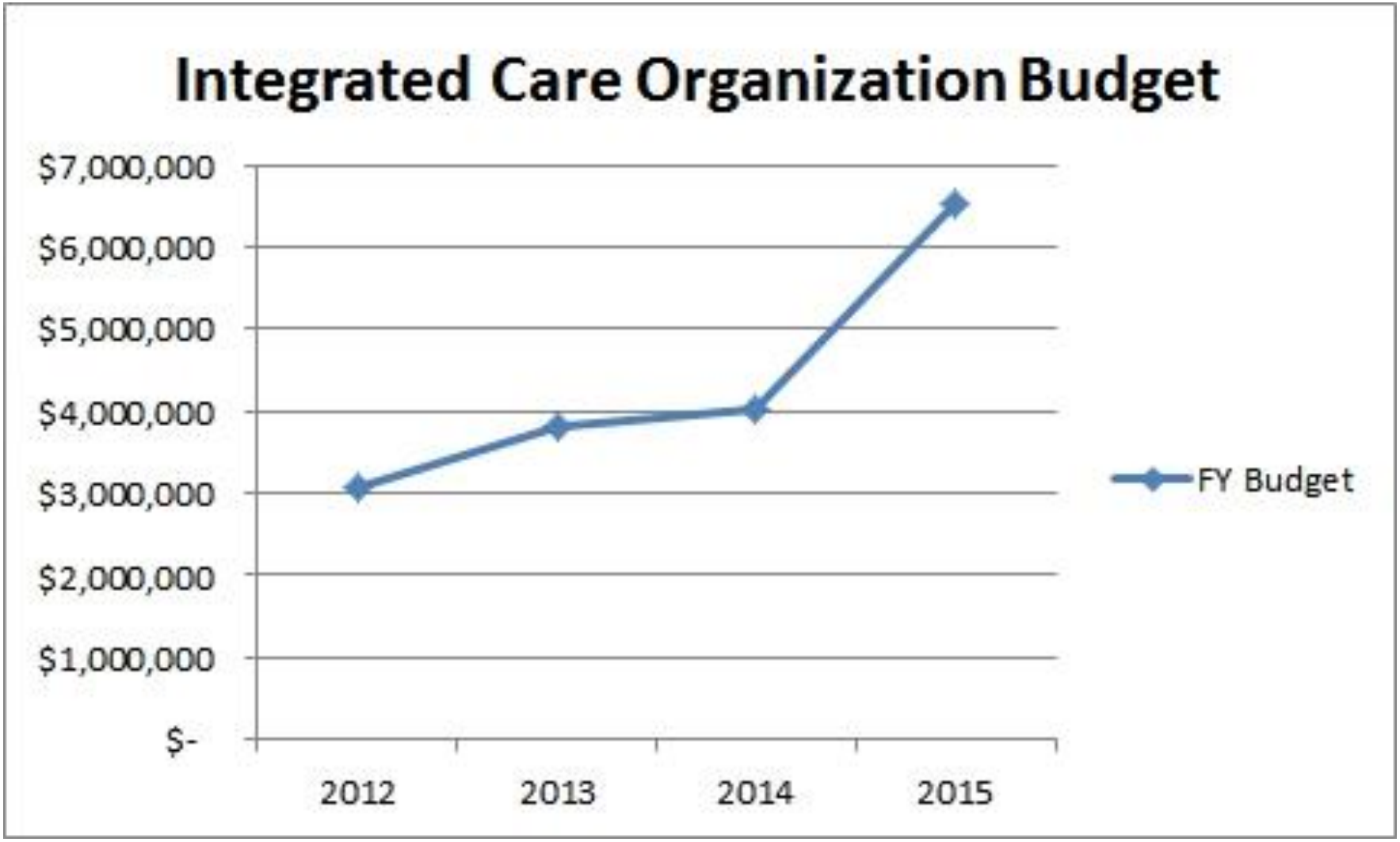

Integrated Care Organization Budget
$7,000,000
$6,000,000
$5,000,000
$4,000,000
$3,000,000
$2,000,000
$1,000,000
$-
2012
2013
2014
2015
FY Budget

Exhibit 1:

Revenue Summary (In-State and Out-of-State)

Exhibit 1 AGO Questions to Hospitals

NOTES:

1. Data entered in worksheets is **hypothetical** and solely for illustrative purposes, provided as a guide to completing this spreadsheet. Respondent may provide explanatory notes and additional information at its discretion.
2. For hospitals, please include professional and technical/facility revenue components.
3. Please include POS payments under HMO.
4. Please include Indemnity payments under PPO.
5. **P4P Contracts** are pay for performance arrangements with a public or commercial payer that reimburse providers for achieving certain quality or efficiency benchmarks. For purposes of this excel, P4P Contracts do not include Risk Contracts.
6. **Risk Contracts** are contracts with a public or commercial payer for payment for health care services that incorporate a per member per month budget against which claims costs are settled for purposes of determining the withhold returned, surplus paid, and/or deficit charged to you, including contracts that subject you to very limited or minimal "downside" risk.
7. **FFS Arrangements** are those where a payer pays a provider for each service rendered, based on an agreed upon price for each service. For purposes of this excel, FFS Arrangements do not include payments under P4P Contracts or Risk Contracts.
8. **Other Revenue** is revenue under P4P Contracts, Risk Contracts, or FFS Arrangements other than those categories already identified, such as management fees and supplemental fees (and other non-claims based, non-incentive, non-surplus/deficit, non-quality bonus revenue).
9. **Claims-Based Revenue** is the total revenue that a provider received from a public or commercial payer under a P4P Contract or a Risk Contract for each service rendered, based on an agreed upon price for each service before any retraction for risk settlement is made.
10. **Incentive-Based Revenue** is the total revenue a provider received under a P4P Contract that is related to quality or efficiency targets or benchmarks established by a public or commercial payer.
11. **Budget Surplus/(Deficit) Revenue** is the total revenue a provider received or was retracted upon settlement of the efficiency-related budgets or benchmarks established in a Risk Contract.
12. **Quality Incentive Revenue** is the total revenue that a provider received from a public or commercial payer under a Risk Contract for quality-related targets or benchmarks established by a public or commercial payer.

2010

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					4,739,329											244,224,041
Tufts Health Plan					607,174											63,730,721
Harvard Pilgrim Health Care																91,243,476
Fallon Community Health Plan																8,075,431
CIGNA																11,665,236
United Healthcare																18,261,465
Aetna																27,344,784
Other Commercial																9,988,609
Total Commercial																**474,533,762**
Network Health																18,111,414
Neighborhood Health Plan																59,648,473
BMC HealthNet, Inc.																27,908,182
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																3,333,705
Total Managed Medicaid																**109,001,775**
Children's Medical Security Plan																**593,456**
Medicaid Out of State																**27,486,500**
MassHealth			570,903													**69,582,646**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**8,771,837**
Out of State Commercial																**182,986,394**
Other																**84,155,206**
GRAND TOTAL																**957,111,576**

2011

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					4,229,376											225,154,412
Tufts Health Plan																61,981,078
Harvard Pilgrim Health Care																93,024,232
Fallon Community Health Plan																7,207,027
CIGNA																10,721,844
United Healthcare																16,831,800
Aetna																21,472,859
Other Commercial																8,344,931
Total Commercial																**444,738,183**
Network Health																16,681,553
Neighborhood Health Plan																48,310,153
BMC HealthNet, Inc.																26,702,794
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																4,197,535
Total Managed Medicaid																**95,892,035**
Children's Medical Security Plan																**914,017**
Medicaid Out of State																**34,856,835**
MassHealth			1,220,409													**70,119,551**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**9,932,739**
Out of State Commercial																**200,872,467**
Other																**87,501,353**
GRAND TOTAL																**944,827,180**

2012

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					1,724,960	17,899,812		(9,755)		1,332,351				96,990		184,570,047
Tufts Health Plan																62,479,716
Harvard Pilgrim Health Care																94,966,355
Fallon Community Health Plan																7,626,423
CIGNA																10,265,785
United Healthcare																19,842,059
Aetna																19,759,986
Other Commercial																9,518,299
Total Commercial																**409,028,671**
Network Health																13,537,172
Neighborhood Health Plan																42,708,188
BMC HealthNet, Inc.																21,334,393
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																3,263,460
Total Managed Medicaid																**80,843,213**
Children's Medical Security Plan																**882,646**
Medicaid Out of State																**30,344,579**
MassHealth			1,384,901													**70,802,813**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**7,893,502**
Out of State Commercial																**196,424,960**
Other																**79,429,607**
GRAND TOTAL																**875,649,991**

Notes

Excluding claims based revenue from totals

2013

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					1,778,272	21,038,136		164,206		2,214,770				93,741		180,987,010
Tufts Health Plan																60,711,336
Harvard Pilgrim Health Care					509,983											103,841,644
Fallon Community Health Plan																8,815,255
CIGNA																13,221,785
United Healthcare																19,941,846
Aetna																17,068,505
Other Commercial																8,151,543
Total Commercial																**412,738,924**
Network Health																19,515,248
Neighborhood Health Plan																42,538,481
BMC HealthNet, Inc.																24,364,242
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																3,420,352
Total Managed Medicaid																**89,838,323**
Children's Medical Security Plan																**622,850**
Medicaid Out of State																**30,384,555**
MassHealth			424,370													**66,398,818**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**9,873,075**
Out of State Commercial																**197,087,138**
Other																**83,242,899**
GRAND TOTAL																**890,186,582**

Notes

Excluding claims based revenue from totals

Plan		2010	2011	2012	2013
BCBS	Revenue	$244,224,041	$225,154,412	$184,570,047	$180,987,010
	Overall Rate Increase	3.0%	6.5%	2.0%	2.0%
	% of Base @ Risk	2.0%	2.0%		
	Rate Increase @ Risk			1.0%	1.0%
	Revenue @ Risk	$4,739,329	$4,229,376	$1,724,960	$1,778,272
HPHC	Revenue				$103,841,644
	Overall Rate Increase				2.0%
	Rate Increase @ Risk				0.5%
	Revenue @ Risk				$509,983
THP	Revenue	$63,730,721			
	Overall Rate Increase	5.0%			
	Rate Increase @ Risk	1.0%			
	Revenue @ Risk	$607,174			

Table 1: Rate Increase Blend

	THP 2010 Rate Increase	FY 2011 THP Revenue IP/OP % Split	% Contribution
Inpatient	5.50%	45%	2.5%
Outpatient	4.62%	55%	2.5%
		Blended Rate Increase	**5.0%**

Exhibit 2:

Summary of Hospital Costs

BOSTON CHILDREN'S HOSPITAL
DHCFP-403 - Schedule VI
GPSR Comparison
Fiscal 2013

Line No.	Cost Center Description	403 C2 FY 2013 Total Revenue (Col. 3+4)	403 C3 FY 2013 Routine GPSR	403 C4 FY 2013 Ancillary GPSR (Col. 5-38)	403 C5 FY 2013 Surgery	403 C6 FY 2013 Labor & Delivery	403 C7 FY 2013 Recovery Room	403 C8 FY 2013 Anesthesiology	403 C9 FY 2013 IV Therapy	403 C10 FY 2013 Med Sup Special	403 C11 FY 2013 Drugs Special	403 C12 FY 2013 Laboratory	403 C13 FY 2013 Blood	403 C14 FY 2013 Blood Processing & Storages	403 C15 FY 2013 EKG	403 C16 FY 2013 Cardiac Cath Lab	403 C17 FY 2013 Diagnostic Radiology
	Routine Inpatient Care Services																
1	Medical & Surgical Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	Pediatric Acute	340,313,146	149,921,255	190,391,891	57,611,924	-	5,290,336	6,142,392	-	-	51,090,820	16,191,449	3,934,896	-	3,867,276	8,540,532	12,288,184
3	Obstetric Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	Psychiatric Acute	11,697,224	10,436,451	1,260,773	23,971	-	6,455	6,568	-	-	967,762	116,605	136	-	11,548	-	12,521
5	Ventilator Unit	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	Skilled Nursing Facilities	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	Clinical Research	10,993,188	5,056,725	5,936,463	486,030	-	75,329	117,876	-	-	2,575,857	769,195	594,858	-	108,624	1,348	729,751
8	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	**Subtotal (Lines 1-9)**	363,003,558	165,414,431	197,589,127	58,121,925	-	5,372,120	6,266,836	-	-	54,634,439	17,077,249	4,529,890	-	3,987,448	8,541,880	13,030,456
11	Med / Surg Intensive Care	310,584,761	107,507,095	203,077,666	69,376,337	-	107,766	4,176,448	-	-	36,416,430	18,504,639	12,866,491	-	4,705,086	12,354,643	8,531,859
12	Coronary Intensive Care	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13	Neonatal Intensive Care	60,681,562	38,097,756	22,583,806	3,376,511	-	1,638	376,604	-	-	5,683,715	2,040,031	932,624	-	505,892	72,634	1,692,098
14	Other ICU (Bone Marrow Unit)	32,181,624	15,543,723	16,637,901	615,489	-	98,254	136,788	-	-	8,949,978	1,673,474	2,411,090	-	75,627	1,107	2,009,321
19	**Subtotal (Lines 11-18)**	403,447,947	161,148,574	242,299,373	73,368,337	-	207,658	4,689,840	-	-	51,050,123	22,218,144	16,210,205	-	5,286,605	12,428,384	12,233,278
20	Newborn Nursery	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21	Chronic / Rehabilitation	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22	**Subtotal (Lines 10+19+20+21)**	766,451,505	326,563,005	439,888,500	131,490,262	-	5,579,778	10,956,676	-	-	105,684,562	39,295,393	20,740,095	-	9,274,053	20,970,264	25,263,734
	Routine Ambulatory Care Services						-										
23	Emergency Services	81,580,216	69,913,493	11,666,723	-	-	-	-	-	-	1,681,627	1,834,976	141,874	-	53,355	5,671	5,710,447
24	Clinic / Ambulatory Care Services	82,212,791	48,326,375	33,886,416	-	-	-	-	-	-	24,566,506	4,671,922	2,254,354	-	328,842	129,029	1,548,289
25	Satellite Clinic Services	73,568,026	9,482,629	64,085,397	45,601,173	-	-	-	-	-	14,521,239	2,716,434	217,178	-	639,136	-	339,074
26	Ambulatory Surgical Services	131,240,591	82,314,730	48,925,861	-	-	22,384,239	17,894,831	-	-	4,077,204	2,222,980	149,630	-	121,102	119,686	1,115,198
27	Ambulatory Renal Dialysis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Home Dialysis Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
29	Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	Home Health Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	Observation Beds	57,408,369	17,043,367	40,365,002	-	-	-	-	-	-	3,118,772	2,162,357	419,262	-	314,195	29,487,259	2,005,961
32	Private Referrals	249,735,496	-	249,735,496	-	-	-	-	-	-	3,202,294	54,771,792	6,353,871	-	26,963,198	15,047,115	44,445,672
33	Other Ambulatory -MEHC	14,267,238	12,967,870	1,299,368	-	-	-	-	-	-	34,899	1,256,937	3,120	-	815	-	2,265
36	**Subtotal Routine Ambulatory**	690,012,727	240,048,464	449,964,263	45,601,173	-	22,384,239	17,894,831	-	-	51,202,541	69,637,398	9,539,289	-	28,420,643	44,788,760	55,166,906
	Services (Lines 23-35)																
37	**Total Patient Care**	1,456,464,232	566,611,469	889,852,763	177,091,435	-	27,964,017	28,851,507	-	-	156,887,103	108,932,791	30,279,384	-	37,694,696	65,759,024	80,430,640

BOSTON CHILDREN'S HOSPIT/
DHCFP-403 - Schedule VI
GPSR Comparison
Fiscal 2013

Line No.	Cost Center Description	403 C18 FY 2013 Therapeutic Radiology	403 C19 FY 2013 CT Scanner	403 C20 FY 2013 Nuclear Medicine	403 C21 FY 2013 Respir Therapy	403 C22 FY 2013 Pulmonary Function	403 C23 FY 2013 EEG	403 C24 FY 2013 Electro-myography	403 C25 FY 2013 Physical Therapy	403 C26 FY 2013 Occupatnl Therapy	403 C27 FY 2013 Speech Therapy	403 C28 FY 2013 Recreatnl Therapy	403 C29 FY 2013 Audiology Therapy	403 C30 FY 2013 Psychology Psychiatry	403 C31 FY 2013 Renal Dialysis	403 C32 FY 2013 Organ Acquistn	403 C33 FY 2013 Ambulance
	Routine Inpatient Care Services																
1	Medical & Surgical Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	Pediatric Acute	-	2,149,614	1,541,993	2,020,639	179,110	7,464,292	-	3,013,646	-	123,090	-	-	-	1,019,228	2,143,658	-
3	Obstetric Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	Psychiatric Acute	-	6,734	1,497	1,701	323	6,864	-	16,114	-	902	-	-	-	67,547	-	-
5	Ventilator Unit	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	Skilled Nursing Facilities	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	Clinical Research	-	92,504	98,889	9,157	8,652	14,010	-	35,778	-	6,496	-	-	-	-	-	-
8	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	**Subtotal (Lines 1-9)**	-	2,248,852	1,642,379	2,031,497	188,085	7,485,166	-	3,065,538	-	130,488	-	-	-	1,086,775	2,143,658	-
11	Med / Surg Intensive Care	-	1,065,857	128,177	25,354,331	11,716	2,707,168	-	738,740	-	21,126	-	-	-	532,811	3,402,966	-
12	Coronary Intensive Care	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13	Neonatal Intensive Care	-	52,747	15,331	6,493,859	-	460,596	-	68,461	-	31,683	-	-	-	297,089	-	-
14	Other ICU (Bone Marrow Unit)	-	153,089	147,765	83,495	6,143	1,930	-	115,015	-	4,246	-	-	-	2,261	-	-
19	**Subtotal (Lines 11-18)**	-	1,271,693	291,273	31,931,685	17,859	3,169,694	-	922,216	-	57,055	-	-	-	832,161	3,402,966	-
20	Newborn Nursery	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21	Chronic / Rehabilitation	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22	**Subtotal (Lines 10+19+20+21)**	-	3,520,545	1,933,652	33,963,182	205,944	10,654,860	-	3,987,754	-	187,543	-	-	-	1,918,936	5,546,624	-
	Routine Ambulatory Care Services																
23	Emergency Services	-	905,631	1,937	29,762	1,597	8,969	-	341	-	939	-	-	-	1,191	-	-
24	Clinic / Ambulatory Care Services	-	21,786	22,127	-	18,304	965	-	2,421	-	7,232	-	-	-	(690)	-	-
25	Satellite Clinic Services	-	1,007	772	-	5,603	-	-	29,339	-	2,201	-	-	-	-	-	-
26	Ambulatory Surgical Services	-	59,402	15,991	142,091	1,109	399,827	-	121,921	-	9,056	-	-	-	4,127	-	-
27	Ambulatory Renal Dialysis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Home Dialysis Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
29	Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	Home Health Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	Observation Beds	-	580,632	76,028	226,024	8,059	998,931	-	96,826	-	4,104	-	-	-	14,407	-	-
32	Private Referrals	-	5,347,794	6,171,840	86,056	3,152,824	13,593,346	-	5,482,597	-	13,583,209	-	-	-	5,467,011	-	-
33	Other Ambulatory -MEHC	-	1,045	-	-	-	-	-	287	-	-	-	-	-	-	-	-
36	**Subtotal Routine Ambulatory**	-	6,917,297	6,288,695	483,933	3,187,496	15,002,038	-	5,733,732	-	13,606,741	-	-	-	5,486,046	-	-
	Services (Lines 23-35)																
37	**Total Patient Care**	-	10,437,842	8,222,347	34,447,115	3,393,440	25,656,898	-	9,721,486	-	13,794,284	-	-	-	7,404,982	5,546,624	-

BOSTON CHILDREN'S HOSPIT/
DHCFP-403 - Schedule VI
GPSR Comparison
Fiscal 2013

Line No.	Cost Center Description	403 C34 FY 2013 MRI	403 C35 FY 2013 Plaster Rooms	403 C36 FY 2013 Orthotic	403 C37 FY 2013 Other Ancillary
	Routine Inpatient Care Services				
1	Medical & Surgical Acute	-	-	-	-
2	Pediatric Acute	5,670,148	4,628	104,036	-
3	Obstetric Acute	-	-	-	-
4	Psychiatric Acute	13,525	-	-	-
5	Ventilator Unit	-	-	-	-
6	Skilled Nursing Facilities	-	-	-	-
7	Clinical Research	211,697	412	-	-
8	Other Acute (Specify)	-	-	-	-
9	Other Acute (Specify)	-	-	-	-
10	**Subtotal (Lines 1-9)**	5,895,370	5,040	104,036	-
11	Med / Surg Intensive Care	2,042,649	-	32,426	-
12	Coronary Intensive Care	-	-	-	-
13	Neonatal Intensive Care	476,585	-	5,708	-
14	Other ICU (Bone Marrow Unit)	147,523	-	5,306	-
19	**Subtotal (Lines 11-18)**	2,666,757	-	43,440	-
20	Newborn Nursery	-	-	-	-
21	Chronic / Rehabilitation	-	-	-	-
22	**Subtotal (Lines 10+19+20+21)**	8,562,127	5,040	147,476	-
	Routine Ambulatory Care Services				
23	Emergency Services	1,289,254	(848)	-	-
24	Clinic / Ambulatory Care Services	144,535	170,661	133	-
25	Satellite Clinic Services	11,530	711	-	-
26	Ambulatory Surgical Services	86,094	2,565	(1,192)	-
27	Ambulatory Renal Dialysis	-	-	-	-
28	Home Dialysis Services	-	-	-	-
29	Psychiatry	-	-	-	-
30	Home Health Services	-	-	-	-
31	Observation Beds	849,164	229	2,792	-
32	Private Referrals	42,544,606	3,522,045	226	-
33	Other Ambulatory -MEHC	-	-	-	-
36	**Subtotal Routine Ambulatory**	44,925,183	3,695,363	1,959	-
	Services (Lines 23-35)				
37	**Total Patient Care**	53,487,310	3,700,403	149,435	-

BOSTON CHILDREN'S HOSPITAL
GSPR by Service Category (Based on 403 Schedule VI)
FY 2010 - 2013

Service Category	FY 2010	FY 2011	FY 2012	FY 2013
Routine	$552,212,881	$546,765,020	$538,011,148	$566,611,469
Surgery	$169,867,391	$164,502,494	$164,178,926	$177,091,435
Recovery Room	$31,661,953	$30,060,605	$29,525,574	$27,964,017
Anesthesiology	$40,875,030	$45,630,952	$38,790,849	$28,851,507
Drugs	$143,347,384	$144,909,084	$150,902,284	$156,887,103
Laboratory	$145,186,310	$127,955,902	$115,705,451	$108,932,791
Blood	$39,721,712	$40,499,544	$35,824,296	$30,279,384
EKG	$10,624,997	$39,807,432	$39,153,826	$37,694,696
Cardiac Cath Lab	$95,488,626	$63,463,997	$64,077,095	$65,759,024
Diagnostic Radiology	$132,794,978	$133,575,881	$115,308,447	$80,430,640
CT Scanner	$26,707,611	$17,382,532	$14,169,780	$10,437,842
Nuclear Medicine	$15,703,360	$11,724,032	$9,570,107	$8,222,347
Respiratory Therapy	$31,522,167	$34,660,250	$34,855,729	$34,447,115
Pulmonary Function	$4,249,892	$4,267,432	$4,044,012	$3,393,440
EEG	$22,066,758	$28,043,259	$26,454,429	$25,656,898
Physical Therapy	$11,671,039	$11,642,075	$10,206,052	$9,721,486
Speech Therapy	$14,539,896	$14,745,893	$14,145,072	$13,794,284
Renal Dialysis	$5,200,396	$6,189,648	$6,788,315	$7,404,982
Organ Acquisition	$4,228,431	$2,955,094	$4,089,093	$5,546,624
MRI	$89,320,004	$72,680,005	$69,375,985	$53,487,310
Plaster Rooms	$4,338,978	$4,629,048	$4,359,377	$3,700,403
Orthotic	$270,563	$300,464	$213,393	$149,435
Totals	**$1,591,600,357**	**$1,546,390,643**	**$1,489,749,240**	**$1,456,464,232**

Exhibit 1 AGO Questions to Hospitals

NOTES:

1. Data entered in worksheets is **hypothetical** and solely for illustrative purposes, provided as a guide to completing this spreadsheet. Respondent may provide explanatory notes and additional information at its discretion.
2. For hospitals, please include professional and technical/facility revenue components.
3. Please include POS payments under HMO.
4. Please include Indemnity payments under PPO.
5. **P4P Contracts** are pay for performance arrangements with a public or commercial payer that reimburse providers for achieving certain quality or efficiency benchmarks. For purposes of this excel, P4P Contracts do not include Risk Contracts.
6. **Risk Contracts** are contracts with a public or commercial payer for payment for health care services that incorporate a per member per month budget against which claims costs are settled for purposes of determining the withhold returned, surplus paid, and/or deficit charged to you, including contracts that subject you to very limited or minimal "downside" risk.
7. **FFS Arrangements** are those where a payer pays a provider for each service rendered, based on an agreed upon price for each service. For purposes of this excel, FFS Arrangements do not include payments under P4P Contracts or Risk Contracts.

8. **Other Revenue** is revenue under P4P Contracts, Risk Contracts, or FFS Arrangements other than those categories already identified, such as management fees and supplemental fees (and other non-claims based, non-incentive, non-surplus/deficit, non-quality bonus revenue).
9. **Claims-Based Revenue** is the total revenue that a provider received from a public or commercial payer under a P4P Contract or a Risk Contract for each service rendered, based on an agreed upon price for each service before any retraction for risk settlement is made.

10. **Incentive-Based Revenue** is the total revenue a provider received under a P4P Contract that is related to quality or efficiency targets or benchmarks established by a public or commercial payer.
11. **Budget Surplus/(Deficit) Revenue** is the total revenue a provider received or was retracted upon settlement of the efficiency-related budgets or benchmarks established in a Risk Contract.
12. **Quality Incentive Revenue** is the total revenue that a provider received from a public or commercial payer under a Risk Contract for quality-related targets or benchmarks established by a public or commercial payer.

2010

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					4,739,329											239,484,712
Tufts Health Plan					607,174											63,123,547
Harvard Pilgrim Health Care																91,243,476
Fallon Community Health Plan																8,075,431
CIGNA																11,665,236
United Healthcare																18,261,465
Aetna																27,344,784
Other Commercial																9,988,609
Total Commercial																**469,187,259**
Network Health																18,111,414
Neighborhood Health Plan																59,648,473
BMC HealthNet, Inc.																27,908,182
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																3,333,705
Total Managed Medicaid																**109,001,775**
Children's Medical Security Plan																**593,456**
Medicaid Out of State																**27,486,500**
MassHealth			570,903													**69,582,646**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**29,891,612**
Out of State Commercial																**182,986,394**
Other																**87,502,347**
GRAND TOTAL																**976,231,989**

Notes
Excluding claims based revenue from totals
Excluding claims based revenue from totals

2011

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					4,229,376											220,925,036
Tufts Health Plan																61,981,078
Harvard Pilgrim Health Care																93,024,232
Fallon Community Health Plan																7,207,027
CIGNA																10,721,844
United Healthcare																16,831,800
Aetna																21,472,859
Other Commercial																8,344,931
Total Commercial																**440,508,807**
Network Health																16,681,553
Neighborhood Health Plan																48,310,153
BMC HealthNet, Inc.																26,702,794
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																4,197,535
Total Managed Medicaid																**95,892,035**
Children's Medical Security Plan																**914,017**
Medicaid Out of State																**34,856,835**
MassHealth			1,220,409													**70,119,551**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**27,639,225**
Out of State Commercial																**200,872,467**
Other																**91,373,187**
GRAND TOTAL																**962,176,124**

Notes
Excluding claims based revenue from totals

2012

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					1,724,960	16,174,852		(9,755)		1,332,351				96,990		184,570,047
Tufts Health Plan																62,479,716
Harvard Pilgrim Health Care																94,966,355
Fallon Community Health Plan																7,626,423
CIGNA																10,265,785
United Healthcare																19,842,059
Aetna																19,759,986
Other Commercial																9,518,299
Total Commercial																**409,028,671**
Network Health																13,537,172
Neighborhood Health Plan																42,708,188
BMC HealthNet, Inc.																21,334,393
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																3,263,460
Total Managed Medicaid																**80,843,213**
Children's Medical Security Plan																**882,646**
Medicaid Out of State																**30,344,579**
MassHealth			1,384,901													**70,802,813**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**25,912,658**
Out of State Commercial																**196,424,960**
Other																**83,279,603**
GRAND TOTAL																**897,519,143**

Notes

Excluding claims based revenue from totals

2013

	P4P Contracts					Risk Contracts						FFS Arrangements		Other Revenue		
	Claims-Based Revenue		Incentive-Based Revenue			Claims-Based Revenue		Budget Surplus/ (Deficit) Revenue		Quality Incentive Revenue						
	HMO	PPO	HMO	PPO	TOTAL	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	HMO	PPO	Both
Blue Cross Blue Shield					1,778,272	19,259,864		164,206		2,214,770				93,741		180,987,010
Tufts Health Plan																60,711,336
Harvard Pilgrim Health Care					509,983											103,331,661
Fallon Community Health Plan																8,815,255
CIGNA																13,221,785
United Healthcare																19,941,846
Aetna																17,068,505
Other Commercial																8,151,543
Total Commercial																**412,228,941**
Network Health																19,515,248
Neighborhood Health Plan																42,538,481
BMC HealthNet, Inc.																24,364,242
Health New England																-
Fallon Community Health Plan																-
Other Managed Medicaid																3,420,352
Total Managed Medicaid																**89,838,323**
Children's Medical Security Plan																**622,850**
Medicaid Out of State																**30,384,555**
MassHealth			424,370													**66,398,818**
Tufts Medicare Preferred																-
Blue Cross Senior Options																-
Other Comm Medicare																-
Commercial Medicare Subtotal																**-**
Medicare																**27,887,440**
Out of State Commercial																**197,087,138**
Other																**87,092,895**
GRAND TOTAL																**911,540,960**

Notes

Excluding claims based revenue from totals

Plan		2010	2011	2012	2013
BCBS	Revenue	$244,224,041	$225,154,412	$184,570,047	$180,987,010
	Overall Rate Increase	3.0%	6.5%	2.0%	2.0%
	% of Base @ Risk	2.0%	2.0%		
	Rate Increase @ Risk			1.0%	1.0%
	Revenue @ Risk	$4,739,329	$4,229,376	$1,724,960	$1,778,272
HPHC	Revenue				$103,841,644
	Overall Rate Increase				2.0%
	Rate Increase @ Risk				0.5%
	Revenue @ Risk				$509,983
THP	Revenue	$63,730,721			
	Overall Rate Increase	5.0%			
	Rate Increase @ Risk	1.0%			
	Revenue @ Risk	$607,174			

Table 1: Rate Increase Blend

	THP 2010 Rate Increase	FY 2011 THP Revenue IP/OP % Split	% Contribution
Inpatient	5.50%	45%	2.5%
Outpatient	4.62%	55%	2.5%
		Blended Rate Increase	**5.0%**

BOSTON CHILDREN'S HOSPITAL
DHCFP-403 - Schedule VI
GPSR Comparison
Fiscal 2013

Line No.	Cost Center Description	403 C2 FY 2013 Total Revenue (Col. 3+4)	403 C3 FY 2013 Routine GPSR	403 C4 FY 2013 Ancillary GPSR (Col. 5-38)	403 C5 FY 2013 Surgery	403 C6 FY 2013 Labor & Delivery	403 C7 FY 2013 Recovery Room	403 C8 FY 2013 Anethes-iology	403 C9 FY 2013 IV Therapy	403 C10 FY 2013 Med Sup Special	403 C11 FY 2013 Drugs Special	403 C12 FY 2013 Laboratory	403 C13 FY 2013 Blood	403 C14 FY 2013 Blood Processing & Storages	403 C15 FY 2013 EKG	403 C16 FY 2013 Cardiac Cath Lab	403 C17 FY 2013 Diagnostic Radiology
	Routine Inpatient Care Services																
1	Medical & Surgical Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	Pediatric Acute	340,313,146	149,921,255	190,391,891	57,611,924	-	5,290,336	6,142,392	-	-	51,090,820	16,191,449	3,934,896	-	3,867,276	8,540,532	12,288,184
3	Obstetric Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	Psychiatric Acute	11,697,224	10,436,451	1,260,773	23,971	-	6,455	6,568	-	-	967,762	116,605	136	-	11,548	-	12,521
5	Ventilator Unit	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	Skilled Nursing Facilities	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	Clinical Research	10,993,188	5,056,725	5,936,463	486,030	-	75,329	117,876	-	-	2,575,857	769,195	594,858	-	108,624	1,348	729,751
8	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	**Subtotal (Lines 1-9)**	363,003,558	165,414,431	197,589,127	58,121,925	-	5,372,120	6,266,836	-	-	54,634,439	17,077,249	4,529,890	-	3,987,448	8,541,880	13,030,456
11	Med / Surg Intensive Care	310,584,761	107,507,095	203,077,666	69,376,337	-	107,766	4,176,448	-	-	36,416,430	18,504,639	12,866,491	-	4,705,086	12,354,643	8,531,859
12	Coronary Intensive Care	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13	Neonatal Intensive Care	60,681,562	38,097,756	22,583,806	3,376,511	-	1,638	376,604	-	-	5,683,715	2,040,031	932,624	-	505,892	72,634	1,692,098
14	Other ICU (Bone Marrow Unit)	32,181,624	15,543,723	16,637,901	615,489	-	98,254	136,788	-	-	8,949,978	1,673,474	2,411,090	-	75,627	1,107	2,009,321
19	**Subtotal (Lines 11-18)**	403,447,947	161,148,574	242,299,373	73,368,337	-	207,658	4,689,840	-	-	51,050,123	22,218,144	16,210,205	-	5,286,605	12,428,384	12,233,278
20	Newborn Nursery	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21	Chronic / Rehabilitation	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22	**Subtotal (Lines 10+19+20+21)**	766,451,505	326,563,005	439,888,500	131,490,262	-	5,579,778	10,956,676	-	-	105,684,562	39,295,393	20,740,095	-	9,274,053	20,970,264	25,263,734
	Routine Ambulatory Care Services						-										
23	Emergency Services	81,580,216	69,913,493	11,666,723	-	-	-	-	-	-	1,681,627	1,834,976	141,874	-	53,355	5,671	5,710,447
24	Clinic / Ambulatory Care Services	82,212,791	48,326,375	33,886,416	-	-	-	-	-	-	24,566,506	4,671,922	2,254,354	-	328,842	129,029	1,548,289
25	Satellite Clinic Services	73,568,026	9,482,629	64,085,397	45,601,173	-	-	-	-	-	14,521,239	2,716,434	217,178	-	639,136	-	339,074
26	Ambulatory Surgical Services	131,240,591	82,314,730	48,925,861	-	-	22,384,239	17,894,831	-	-	4,077,204	2,222,980	149,630	-	121,102	119,686	1,115,198
27	Ambulatory Renal Dialysis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Home Dialysis Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
29	Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	Home Health Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	Observation Beds	57,408,369	17,043,367	40,365,002	-	-	-	-	-	-	3,118,772	2,162,357	419,262	-	314,195	29,487,259	2,005,961
32	Private Referrals	249,735,496	-	249,735,496	-	-	-	-	-	-	3,202,294	54,771,792	6,353,871	-	26,963,198	15,047,115	44,445,672
33	Other Ambulatory -MEHC	14,267,238	12,967,870	1,299,368	-	-	-	-	-	-	34,899	1,256,937	3,120	-	815	-	2,265
36	**Subtotal Routine Ambulatory**	690,012,727	240,048,464	449,964,263	45,601,173	-	22,384,239	17,894,831	-	-	51,202,541	69,637,398	9,539,289	-	28,420,643	44,788,760	55,166,906
	Services (Lines 23-35)																
37	**Total Patient Care**	1,456,464,232	566,611,469	889,852,763	177,091,435	-	27,964,017	28,851,507	-	-	156,887,103	108,932,791	30,279,384	-	37,694,696	65,759,024	80,430,640

BOSTON CHILDREN'S HOSPIT/
DHCFP-403 - Schedule VI
GPSR Comparison
Fiscal 2013

Line No.	Cost Center Description	403 C18 FY 2013 Therapeutic Radiology	403 C19 FY 2013 CT Scanner	403 C20 FY 2013 Nuclear Medicine	403 C21 FY 2013 Respir Therapy	403 C22 FY 2013 Pulmonary Function	403 C23 FY 2013 EEG	403 C24 FY 2013 Electro-myography	403 C25 FY 2013 Physical Therapy	403 C26 FY 2013 Occupatnl Therapy	403 C27 FY 2013 Speech Therapy	403 C28 FY 2013 Recreatnl Therapy	403 C29 FY 2013 Audiology Therapy	403 C30 FY 2013 Psychology Psychiatry	403 C31 FY 2013 Renal Dialysis	403 C32 FY 2013 Organ Acquistn	403 C33 FY 2013 Ambulance	403 C34 FY 2013 MRI
	Routine Inpatient Care Services																	
1	Medical & Surgical Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	Pediatric Acute	-	2,149,614	1,541,993	2,020,639	179,110	7,464,292	-	3,013,646	-	123,090	-	-	-	1,019,228	2,143,658	-	5,670,148
3	Obstetric Acute	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	Psychiatric Acute	-	6,734	1,497	1,701	323	6,864	-	16,114	-	902	-	-	-	67,547	-	-	13,525
5	Ventilator Unit	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	Skilled Nursing Facilities	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	Clinical Research	-	92,504	98,889	9,157	8,652	14,010	-	35,778	-	6,496	-	-	-	-	-	-	211,697
8	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	Other Acute (Specify)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	**Subtotal (Lines 1-9)**	-	2,248,852	1,642,379	2,031,497	188,085	7,485,166	-	3,065,538	-	130,488	-	-	-	1,086,775	2,143,658	-	5,895,370
11	Med / Surg Intensive Care	-	1,065,857	128,177	25,354,331	11,716	2,707,168	-	738,740	-	21,126	-	-	-	532,811	3,402,966	-	2,042,649
12	Coronary Intensive Care	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13	Neonatal Intensive Care	-	52,747	15,331	6,493,859	-	460,596	-	68,461	-	31,683	-	-	-	297,089	-	-	476,585
14	Other ICU (Bone Marrow Unit)	-	153,089	147,765	83,495	6,143	1,930	-	115,015	-	4,246	-	-	-	2,261	-	-	147,523
19	**Subtotal (Lines 11-18)**	-	1,271,693	291,273	31,931,685	17,859	3,169,694	-	922,216	-	57,055	-	-	-	832,161	3,402,966	-	2,666,757
20	Newborn Nursery	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21	Chronic / Rehabilitation	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22	**Subtotal (Lines 10+19+20+21)**	-	3,520,545	1,933,652	33,963,182	205,944	10,654,860	-	3,987,754	-	187,543	-	-	-	1,918,936	5,546,624	-	8,562,127
	Routine Ambulatory Care Services																	
23	Emergency Services	-	905,631	1,937	29,762	1,597	8,969	-	341	-	939	-	-	-	1,191	-	-	1,289,254
24	Clinic / Ambulatory Care Services	-	21,786	22,127	-	18,304	965	-	2,421	-	7,232	-	-	-	(690)	-	-	144,535
25	Satellite Clinic Services	-	1,007	772	-	5,603	-	-	29,339	-	2,201	-	-	-	-	-	-	11,530
26	Ambulatory Surgical Services	-	59,402	15,991	142,091	1,109	399,827	-	121,921	-	9,056	-	-	-	4,127	-	-	86,094
27	Ambulatory Renal Dialysis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Home Dialysis Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
29	Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	Home Health Services	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	Observation Beds	-	580,632	76,028	226,024	8,059	998,931	-	96,826	-	4,104	-	-	-	14,407	-	-	849,164
32	Private Referrals	-	5,347,794	6,171,840	86,056	3,152,824	13,593,346	-	5,482,597	-	13,583,209	-	-	-	5,467,011	-	-	42,544,606
33	Other Ambulatory -MEHC	-	1,045	-	-	-	-	-	287	-	-	-	-	-	-	-	-	-
36	**Subtotal Routine Ambulatory**	-	6,917,297	6,288,695	483,933	3,187,496	15,002,038	-	5,733,732	-	13,606,741	-	-	-	5,486,046	-	-	44,925,183
	Services (Lines 23-35)																	
37	**Total Patient Care**	-	10,437,842	8,222,347	34,447,115	3,393,440	25,656,898	-	9,721,486	-	13,794,284	-	-	-	7,404,982	5,546,624	-	53,487,310

BOSTON CHILDREN'S HOSPIT/
DHCFP-403 - Schedule VI
GPSR Comparison
Fiscal 2013

Line No.	Cost Center Description	403 C35 FY 2013 Plaster Rooms	403 C36 FY 2013 Orthotic	403 C37 FY 2013 Other Ancillary
	Routine Inpatient Care Services			
1	Medical & Surgical Acute	-	-	-
2	Pediatric Acute	4,628	104,036	-
3	Obstetric Acute	-	-	-
4	Psychiatric Acute	-	-	-
5	Ventilator Unit	-	-	-
6	Skilled Nursing Facilities	-	-	-
7	Clinical Research	412	-	-
8	Other Acute (Specify)	-	-	-
9	Other Acute (Specify)	-	-	-
10	**Subtotal (Lines 1-9)**	5,040	104,036	-
11	Med / Surg Intensive Care	-	32,426	-
12	Coronary Intensive Care	-	-	-
13	Neonatal Intensive Care	-	5,708	-
14	Other ICU (Bone Marrow Unit)	-	5,306	-
19	**Subtotal (Lines 11-18)**	-	43,440	-
20	Newborn Nursery	-	-	-
21	Chronic / Rehabilitation	-	-	-
22	**Subtotal (Lines 10+19+20+21)**	5,040	147,476	-
	Routine Ambulatory Care Services			
23	Emergency Services	(848)	-	-
24	Clinic / Ambulatory Care Services	170,661	133	-
25	Satellite Clinic Services	711	-	-
26	Ambulatory Surgical Services	2,565	(1,192)	-
27	Ambulatory Renal Dialysis	-	-	-
28	Home Dialysis Services	-	-	-
29	Psychiatry	-	-	-
30	Home Health Services	-	-	-
31	Observation Beds	229	2,792	-
32	Private Referrals	3,522,045	226	-
33	Other Ambulatory -MEHC	-	-	-
36	**Subtotal Routine Ambulatory**	3,695,363	1,959	-
	Services (Lines 23-35)			
37	**Total Patient Care**	3,700,403	149,435	-

BOSTON CHILDREN'S HOSPITAL
GSPR by Service Category (Based on 403 Schedule VI)
FY 2010 - 2013

Service Category	FY 2010	FY 2011	FY 2012	FY 2013
Routine	$552,212,881	$546,765,020	$538,011,148	$566,611,469
Surgery	$169,867,391	$164,502,494	$164,178,926	$177,091,435
Recovery Room	$31,661,953	$30,060,605	$29,525,574	$27,964,017
Anesthesiology	$40,875,030	$45,630,952	$38,790,849	$28,851,507
Drugs	$143,347,384	$144,909,084	$150,902,284	$156,887,103
Laboratory	$145,186,310	$127,955,902	$115,705,451	$108,932,791
Blood	$39,721,712	$40,499,544	$35,824,296	$30,279,384
EKG	$10,624,997	$39,807,432	$39,153,826	$37,694,696
Cardiac Cath Lab	$95,488,626	$63,463,997	$64,077,095	$65,759,024
Diagnostic Radiology	$132,794,978	$133,575,881	$115,308,447	$80,430,640
CT Scanner	$26,707,611	$17,382,532	$14,169,780	$10,437,842
Nuclear Medicine	$15,703,360	$11,724,032	$9,570,107	$8,222,347
Respiratory Therapy	$31,522,167	$34,660,250	$34,855,729	$34,447,115
Pulmonary Function	$4,249,892	$4,267,432	$4,044,012	$3,393,440
EEG	$22,066,758	$28,043,259	$26,454,429	$25,656,898
Physical Therapy	$11,671,039	$11,642,075	$10,206,052	$9,721,486
Speech Therapy	$14,539,896	$14,745,893	$14,145,072	$13,794,284
Renal Dialysis	$5,200,396	$6,189,648	$6,788,315	$7,404,982
Organ Acquisition	$4,228,431	$2,955,094	$4,089,093	$5,546,624
MRI	$89,320,004	$72,680,005	$69,375,985	$53,487,310
Plaster Rooms	$4,338,978	$4,629,048	$4,359,377	$3,700,403
Orthotic	$270,563	$300,464	$213,393	$149,435
Totals	**$1,591,600,357**	**$1,546,390,643**	**$1,489,749,240**	**$1,456,464,232**