NATIONAL QUALITY FORUM

Efficiency and Value in Healthcare: Linking Cost and Quality Measures

A PAPER COMMISSIONED BY THE NATIONAL QUALITY FORUM

Andrew M. Ryan, Ph.D.
Christopher P. Tompkins, Ph.D.

November 14, 2014

This report is funded by the Robert Wood Johnson Foundation under grant #71278.

Acknowledgements

The authors would like to acknowledge the staff at the National Quality Forum, particularly Ashlie Wilbon, Taroon Amin, and Erin O'Rourke, for their contributions and assistance with this project. We would also like to acknowledge the outstanding contributions of the expert panel, particularly Joyce Dubow and Carole Flamm, the panel co-chairs. We would also like to acknowledge the research assistance of Jayme Mendelsohn.

Introduction from the National Quality Forum

While the need to use cost and quality measures together to assess health system efficiency is well-established, there is currently no clear consensus among stakeholders or recognized state of the art on how to do so. To begin exploring the state of measurement science in this area, NQF commissioned this white paper to categorize the various methodologies for and challenges of linking cost and quality measures. To guide the development of this paper, NQF convened a multistakeholder expert panel to provide input through an iterative review process. The goal of this paper is to outline various approaches to measuring efficiency in health care, explore their current and potential uses, and determine possible paths forward.

NQF has built much of its work around cost and resource use measurement on guidance established by prior work, including the Institute of Medicine's definition of efficiency and NQF's Patient Focused Episode of Care Model. This prior work has differentiated efficiency and value into separate concepts. However, the findings of this current project raise important questions around some of our basic assumptions for measurement science. First, while NQF has traditionally defined efficiency and value as separate but related concepts, this work has shown that they may be intrinsically intertwined. Through the environmental scan for this work and stakeholder discussions, it became clear that some approaches to efficiency measurement currently in use inherently require weighting or threshold setting, implying stakeholder preference and value. Perspective, intent, and how efficiency is defined often dictate how measures are selected and weighted for a performance improvement program. Secondly, while NQF's current efforts have focused on endorsing individual measures of quality and cost/resource use, these measures are increasingly being used to together to assess provider performance on efficiency as the system moves towards value-based purchasing to reward high-quality, cost effective care . Finally, there is a strong movement focused on promoting alignment of individual measures to ensure consistent reporting on provider performance. However, we have learned from this work that the approach to combining the measures to understand efficiency is as important as the individual measures themselves. Solely focusing on the alignment of individual measures alone still may not yield the consistency we seek since combining the same measures through various efficiency models may lead to inconsistent, uncorrelated results.

Efficiency measurement is evolving and is increasingly being used to determine provider payments, and the providers that consumers have access to through their health plan's network. However, basing provider payments or even the structure of a health plan network, on efficiency models that produce varying results has been brought to question by this expert panel and the white paper authors. For example, a hospital may receive vastly dissimilar scores from various health plans, each using their own efficiency measurement approach. A hospital may score well in one insurer's program, poorly in another's, and in the middle in a third, even if the same performance measures are used. Such conflicting results make it difficult for the hospital to know how to improve their performance or for a consumer to know where to go to get the best care. This work highlights the need for increased transparency and multistakeholder review of how measures are combined to assess efficiency to better enable providers to engage in focused improvement efforts and for consumers to make informed choices.

NQF strives to be a leader in illuminating measurement issues and seeks to provide guidance on how to make efficiency measurement useful and meaningful to all stakeholders. This project is an important contribution to the field as it raises important principles that should be considered as efficiency measurement is increasingly used in high-stakes applications. First, perspective matters and defining efficiency from any perspective can influence both measure selection and how measures are combined. Second, efficiency measurement should only be done in a transparent and scientifically sound manner. Finally, there is a need for continued evidence generation of the validity and reliability of various efficiency measurement approaches for different applications. By building on this work, the measurement community can help advance the rigor of efficiency measurement and drive system improvements for all stakeholders.

Table of Contents

Purpose of the commissioned paper

The National Quality Forum (NQF) has commissioned a paper to assess alternative approaches to link – or combine – measures of quality and cost for the purpose of measuring efficiency and value in health care. This paper reviews various approaches and considers the implications of alternative methods for profiling and scoring. In addition to assessing the technical issues related to measuring and profiling provider performance, we will consider the implications for using alternative approaches in the context of various programs, such as the creation of tiered insurance networks and value-based payment.

Our goal in writing the commissioned paper is to help build consensus about the key considerations and appropriateness of alternative approaches for combining quality and cost measures into quantifiable measures of efficiency and value. Earlier versions of this paper served as a foundation to inform the deliberations of a multi-stakeholder expert panel, which provided input on the methodological challenges to linking cost and quality measures and the best practices for combining cost and quality measures to assess efficiency or value of care.[1]

A substantial literature has been devoted to understanding and measuring economic efficiency in health care.[2] While questions of efficiency in health care have been of interest for decades,[3,4] this interest has accelerated in recent years.[5] However, as identified by a recent systematic review commissioned by AHRQ, considerations of quality of care have been largely absent from this literature.[5] Instead, researchers have evaluated economic efficiency using a variety of techniques to understand how a given output (e.g. a hospital day) can be optimized for a given set of health care inputs (e.g. physician labor, nurse labor). While the study of economic efficiency in health care is of great importance, it is not the focus of this paper.

In this paper, we are interested in the assessment of efficiency only through the joint consideration of cost and quality. We do not focus on approaches to the measurement of efficiency– such as brand prescribing rates or rates of MRI for patients with back pain – that seek to identify relative resource use and appropriateness.[6] Measuring inappropriate resource use, or "waste", clearly has value but represents an overly narrow interpretation of efficiency.[6,7] There is also a large literature concerned with the relationship between costs and quality,[8-12] and a smaller literature on relationship between economic efficiency and quality.[13] While relevant to the concept of efficiency that we seek to understand, this literature is not primarily concerned with profiling individual providers on the basis of efficiency.

Key Definitions

This paper will reference a number of common terms that may have different connotations for different audiences. Throughout this paper, we will apply a modified version of the definitions from the NQF's Patient-Focused Episodes of Care project:[14]

> **Quality of care:** the degree to which health services for individuals and populations increase the likelihood of desired health and patient experience outcomes and are consistent with professional knowledge[15]
>
> **Cost of care**: measures total health care spending, including total resource use and unit price(s), by payor or consumer, for a health care service or group of health care services associated with a specified patient population, time period, and unit(s) of clinical accountability. Costs of care can be considered from different perspectives, including the patient, the purchaser, the provider, or the societal perceptive. In this paper, we consider costs primarily from the perspective of the payer (either the patient or the third-party purchaser) and consider only financial costs associated with care. Non-financial costs are relevant when considering costs from the perspective of patients (e.g. opportunity costs and travel costs associated with treatment), providers (e.g., administrative costs from interacting with insurers[16]), and society (e.g., the effects of health care costs on the US economy[17]). However, the challenges associated with collecting data on many of these types of costs limits the inclusions of these costs in many applications to measure, profile, and manage health care costs.
>
> The financial costs of care, from the payer perspective, can also be operationalized in several ways. For instance, costs could be defined either as charges for services billed by providers or as "allowed charges," the payment amounts for services that are negotiated between insurers and some providers. The assessment of health care costs may also substitute average or "standardized prices" across the population of health providers in order to remove price variation and allow for costs to be used as a measure of resource use that is due solely to utilization patterns, rather than differential pricing. The merit of these alternative conceptualizations of costs is likely to depend on the application of cost and quality profiling – known as the use case – which we discuss in Section 5 of this report.
>
> **Efficiency of care:** measures the cost of care associated with a specified level of quality of care. "Efficiency of care" is a measure of the relationship of the cost of care associated with a specific level of performance measured with respect to dimensions of quality.
>
> **Value of care:** measures a specified stakeholder's (such as an individual patient's, consumer organization's, payor's, provider's, government's, or society's) preference-weighted assessment of a particular combination of quality and cost of care performance.[i]

---

[i] Quality, cost, efficiency, and value can be measured and assessed for different aspects or segments of care (i.e., episodes of care ranging from management of a condition over time, to specific procedures or other acute events) and across different levels of organizational accountability (e.g. individual physicians, physician organizations,

As used in this paper, the terms efficiency and value correspond to the respective definitions adopted previously by NQF and other stakeholders. Using these definitions, efficiency can be assessed objectively. By profiling providers' quality, cost, and efficiency, and showing the component pieces, it is reasonable to assume that efficiency can be measured and displayed in a way that allows stakeholders to consider "value" as a preference-weighted assessment of the component pieces; i.e., quality, cost, and efficiency. For example, one approach might determine a provider to be "high quality," while also "high cost," based on its performance in relation to averages in both dimensions. An alternate approach is to insert an intermediate step of measuring efficiency. This might conclude that the provider is "high quality," but actually "low cost" when measured only against providers with similarly high quality, and therefore has high efficiency. Stakeholders can make value inferences in either case. The intermediate step serves to clarify the process by making explicit the objective relationships between quality and cost from which general and specific subjectively-weighted inferences are made regarding value.

hospitals, insurance plans, or accountable care organizations). Decisions about the appropriate level of measurement and accountability will depend on the purpose – or "use case" – of combing quality and cost measures. See Section 5.

Section 1. Why it matters to combine quality and cost measures

Improving the efficiency of health care delivery in the United States is critical. Recent attempts at payment reform, such as pay-for-performance and public quality reporting, have failed to reduce cost growth.[18,19] By focusing primarily on quality measures of underuse – such as non-adherence with evidence-based care – these programs have not provided direct incentives for increased efficiency. Previous efforts to reign in cost growth through managed care, such as capitated payment and utilization review, focused primarily on reducing costs rather than improving quality of care.[20]

To address these shortcomings, the Patient Protection and Affordable Care Act created numerous initiatives that are intended to improve the *efficiency* of US health care –not quality or cost alone. These initiatives include the Physician Value-Based Payment Modifier,[21] Hospital Value-Based Purchasing,[22] The Medicare Advantage Quality Bonus Program,[23] Accountable Care Organization programs,[24] and the End-Stage Renal Disease pay-for-performance program. More directly, legislation was introduced in 2009 to replace the standard update to physician payments with a geographically based "value index," which would adjust payments to physicians according to their relative quality and cost.[25]

In the private sector, a number of insurers have developed products with tiered networks that are based ostensibly on measures of efficiency. These products are structured to increase patient cost-sharing for using providers that are designated in a lower-efficiency tier. The first generation of these programs established tiers based almost exclusively on costs.[26] However, insurers have developed a range of increasingly sophisticated approaches to combine indicators of cost and quality to categorize the efficiency of providers. These efforts are related to the rise of high-deductible health plans and consumerism. Patients need both quality and cost information in order to maximize value by making informed choices about the services they need and the providers they should use. In addition, given the price sensitivity to plans currently sold in insurance exchanges created through the ACA,[27] insurers may adopt narrower networks in order to compete on price.[28] This will likely increase insurers' use of tiered networks based on measures of provider efficiency. Other promising private sector efforts, such as reference pricing,[29] will likely need to integrate provider quality measurement explicitly in order to gain greater acceptance and permit inferences about relative value.

These reforms require both quality and cost performance to be measured and assessed together. These ongoing initiatives share a common set of goals: 1) To better identify high- or low-efficiency providers, and 2) To foster incentives for providers to improve efficiency. Broader efforts to better identify the relative value of health care services are related, but rely on a different set of tools and policy measures. While cost-effectiveness and comparative-effectiveness research seeks to understand the relative cost and effectiveness of medical treatments, efficiency profiling seeks to understand the relative efficiency of health care *providers*.

However, the desire to use efficiency measures has outpaced scientific consensus about how best to incorporate these measures into accountability efforts. As shown in section 2 of this paper, there have been several different measures of efficiency across public programs. Also, while many of the private payer efforts to combine quality and cost have similar features, they differ in important ways.

Efforts are moving ahead to measure and profile health care providers' efficiency without a clear sense of the best or optimal approaches. The issues surrounding combining quality and cost measures are certainly challenging: one recent report described the state of efficiency measurement as "woefully inadequate."[30] Two high profile efforts tasked with grappling with these issues failed to recommend a strategy to do so.[31] This paper reflects NQF's intention to develop a consensus framework to identify the trade-offs between alternative approaches to combine quality and cost indicators in order to guide the future development, evaluation, and use of efficiency and value measurement in health care.

Section 2. Options for combining quality and cost measures

Methods for environmental scan

We conducted an environmental scan to identify approaches that combine indicators of quality and cost measures that were used by Medicare, private payers, or other program sponsors. We also identified approaches that link quality and cost indicators which have been developed by researchers but are not currently used by any program sponsor. To be included, an approach must assess cost as an input, and one or more measures of quality as the output.

We searched the PubMed databases for published articles in the English language that appeared in journals between January 1990 and April 2014. Search terms included "quality," "measuring," and "cost." We searched the bibliographies of retrieved articles looking for additional relevant publications. We then searched Google Scholar, the Cochrane Database, and conducted other general internet searches for the same search terms. This provided resources that were not limited to peer-reviewed journals. We also identified applications outside of health care that combine indicators of quality and cost (e.g. Consumer Reports "Best and Worst Cars for the Money" and US News and World Reports "Best Value Schools"). See Appendix A for information on those efforts.

Additionally, we solicited information from the NQF's Expert Panel on Linking Cost and Quality. The materials referred to us by the expert panel frequently led to the discovery of additional approaches. From the panel, we also obtained detailed information on approaches that we knew had been initiated (for instance, in Medicare).

After identifying all of the programs that simultaneously assessed quality and cost, as well as approaches proposed by researchers, we identified and described a set of mutually exclusive approaches that combine quality and cost measures to measure efficiency or value. We then described the basic features of these approaches. Next, we identified the programs that have used quality and cost indicators to profile providers. This includes programs that are currently running as well as those that are now defunct. For these programs, we obtained information on several parameters: the name of the program, the services evaluated (e.g. hospital only, physician only, all services), the level of attribution (e.g. hospital, physician practice, individual physician), the specification of quality, the specification of cost, and the approach used to combine quality and cost indicators.

Approaches used to combine quality and cost measures

We identified seven approaches that are currently in use or have been proposed by researchers to combine quality and cost indicators to measure efficiency.

> **The conditional model**: This approach, described by Timbie and Normand as the "Univariate" approach[32] and by Tompkins et al. as the "Net-Incentive Payment Model"[33] assesses efficiency as the conditional combination of quality and cost. The approach proceeds in four steps: first quality is assessed either by a single indicator or by a composite measure; second cost is assessed, typically by a single measure of total costs; third, either or both of the quality and cost domains are classified into performance groups – frequently as "low", "average", or "high" – using specified criteria; fourth, the quality and cost classifications are combined to assess efficiency. A common approach is to define high efficiency providers as those that are classified as both high quality and

low cost. Alternatively, the Net-Incentive Payment Model assesses the difference in costs between providers within the same quality grouping. The Conditional Model is widely used by private payers to create tiers of providers based on their efficiency.

The Unconditional Model. The unconditional model follows the first two steps of the Conditional Model. Then, the quality and cost domains are assigned weights and combined into a single metric. Thus, in the Unconditional Model, quality and cost are scored independently and then combined. This is the model currently used by Medicare's Hospital Value-Based Purchasing program.

The Quality Hurdle Model and Cost Hurdle Model: A variation on the Conditional Model is the Quality Hurdle Model. This model follows the first three steps of the Conditional Model. Then, providers are subject to a minimum quality standard, the hurdle, before their cost performance is assessed. After meeting this minimum quality standard, providers may be judged on cost performance alone or may be evaluated based on their combination of quality and cost performance. A variation on the Quality Hurdle Model is the Cost Hurdle Model. Here, providers are evaluated on quality performance only after meeting a cost standard, which is often defined as having costs that are below a specified growth rate.[34] Hurdle Models are commonly used for shared savings programs.

The Regression Model: The regression model, proposed by Timbie and Normand,[32] profiles provider quality while conditioning on cost. While it is conceptually similar to the Conditional Model, it has the advantage of using regression analysis to account for the within-provider correlation between quality and cost outcomes. In contrast, the approach taken by the Conditional Model does not account for any correlation between the quality and cost domains. The regression model is not currently used by any program sponsor.

The cost-effectiveness model: The cost-effectiveness model, proposed by Timbie and Normand,[32] differs from the other approaches in that it assigns a dollar value to the patient benefits accrued from the specified quality domain. This can substantially change efficiency profiles. For instance, using the Unconditional or Conditional Model, a hospital with excellent mortality outcomes may be classified as having only moderate efficiency if it also has high costs. However, if the benefit of increased survival is appropriately weighted and the absolute cost differences between this hospital and others are not great, this high cost hospital may in fact have excellent efficiency: it is producing desirable health outputs at a lower cost than other hospitals. A similar approach towards efficiency measurement was developed by Kessler and McClellan to evaluate the cost-effectiveness not of individual providers, but of the characteristics of hospitals.[35]

The Data Envelopment Analysis or Stochastic Frontier Analysis Model: This approach is used to identify the efficient production of quality across all observed levels of cost.[36,37] The efficient frontier is modeled and providers' efficiency can then be evaluated based on their distance from the efficient frontier. One of the key advantages of this approach is that it allows efficiency to be evaluated across continuous measures of cost and quality. It therefore does not require classification of providers into categories based on what may be arbitrary threshold values, a shortcoming of other approaches. This approach has been widely used in academic research to assess economic efficiency in health care, although almost exclusively in cases in which the output

of interest is something other than quality of care.[13] This approach is not currently used by any program sponsors to evaluate provider efficiency.

The Side-by-Side Model: This approach does not combine the quality and cost domains in any way. It follows the first two steps of the Conditional Model, then concludes by displaying the results in summary form. This model typically emphasizes the clear and intuitive display of indicators of quality and cost (e.g. star ratings). However, by leaving the specific combination of cost and quality unspecified when assessing efficiency, this model leads directly to value estimations by stakeholders.

Programs linking cost and quality measures

Exhibit 1 describes identified programs that link indicators of cost and quality to measure efficiency or value. We describe the characteristics of 25 programs for which we were able to obtain sufficiently complete information.

Of these programs, 11 profiled physicians or physician practices, 5 profiled hospitals or surgical centers, 3 profiled both physicians and hospitals, and 6 profiled health systems or health plans. To combine quality and cost indicators, 4 of the identified approaches used the Conditional Model, 6 used the Unconditional Model, 5 used the Side-by-Side Model, and 8 used the Quality Hurdle or Cost Hurdle Model.[ii] The method used to combine quality and cost indicators was unclear for 2 programs.

[ii] While Veterans Affairs hospitals use stochastic frontier analysis to profile the efficiency of hospitals, assessment of efficiency does not consider quality of care as an output.

Section 3. Illustration of models to combine indicators of quality and cost

We illustrated the implementation of several of the models to combine quality and cost measures to provide a clearer idea about their similarities and differences. To do this, we downloaded data on hospital cost and quality from the May 2, 2014 release of Hospital Compare (www.medicare.gov/hospitalcompare). Our measure of cost is Medicare Spending per Beneficiary (MSPB), an NQF endorsed measure (NQF #2158). The measure captures price-adjusted Medicare spending for all services (inpatient, outpatient, home health, hospice, skilled nursing, and durable medical equipment) for acute care hospitals for all admissions in the 3 days prior to admission and 30 days after discharge. We specified cost using the ratio of the national total spending per episode to individual hospitals' total hospital spending per episode. A higher value indicates higher cost performance (i.e., lower cost relative to the national average).

The measure of quality is the Total Performance Score from Hospital Value-Based Purchasing. The Total Performance Score is a composite measure capturing hospital quality performance related to clinical process performance (45%), patient experience (30%), and outcome performance (25%). The measure incorporates both quality attainment and quality improvement. Higher scores indicate higher quality performance.

We merged cost data from 3,260 acute care hospitals with quality data from 2,728 hospitals. Our analytic sample was 2,728 hospitals. Before combining indicators, we standardized the quality and cost indicators by subtracting the mean and dividing by the standard deviation. The distribution of the quality and cost measures are shown in Exhibit 2. We linked quality and cost measures to measure efficiency and value using the following models:

1. **The conditional model**: The conditional model linked quality and cost by assessing cost performance for a given level of quality. We calculated two separate versions of the conditional model that varied with respect to the precision of the quality groupings. The first version classified hospitals into terciles of quality performance and then classified hospitals into cost tritiles: low (bottom 25%), average (middle 50%), and high (top 25%) cost performance. In the second version, hospitals were classified into quality tritiles, and then classified into cost tritiles within each quality tritile. In the second model, we assigned an efficiency score of "9" (the highest score) for the top quality and top cost tritile, decreasing to "1" for the bottom quality and bottom cost tritile.
2. **The unconditional model**: The unconditional model linked quality and cost measures through a weighted combination of measure scores. We calculated two separate versions of the unconditional model, one using 70% quality and 30% cost, the other using 30% quality and 70% cost.
3. **The quality hurdle model**: The quality hurdle model linked quality and cost measures by setting the quality hurdle at the 25^{th} percentile. Below the 25^{th} percentile of quality, hospitals received an efficiency score of 0. Above the 25^{th} percentile of quality, hospitals' efficiency was determined based on their cost performance.

4. **The cost hurdle model**: The cost hurdle was similarly set at the 25th percentile. Below the 25th percentile of cost performance, hospitals received an efficiency score of 0. Above the 25th percentile of cost performance, hospitals' efficiency was determined based on their quality.
5. **The stochastic frontier model**: The stochastic frontier model linked quality and cost measures by estimating quality as a function of cost. Efficiency was then assessed based on hospitals' "technical efficiency", a measure of hospitals' distance from the frontier.

We did not illustrate the linking of cost and quality using the side-by-side model, because this model does not formally combine measures of cost and quality. We also did not link cost and quality measures using the regression model or the cost-effectiveness models because these models require patient-level data.

Exhibit 3 shows the hurdle models, **Exhibit 4** shows the unconditional models, **Exhibit 5** shows the conditional models, and **Exhibit 6** shows the stochastic frontier model. For each of these models, greater cost performance denotes lower cost. Hospitals toward the top right of the scatter plot have higher efficiency while those toward the bottom left have lower efficiency. The precise determination of efficiency depends on the model used to combine quality and cost indicators.

Exhibit 7 shows a correlation matrix between the efficiency scores generated from the alternative models. It indicates a low to high degree of correlation between the efficiency scores generated from the different models. The quality hurdle model has a relatively weak correlation with the other models (with the exception of the unconditional (30% quality) model r=0.78). The cost hurdle model is most strongly correlated with the unconditional (70% quality) model (r=0.81), the conditional (r=0.76), and the frontier model (r=0.87). The unconditional (70% quality) model is also highly correlated with the unconditional (30% quality) (r=0.66) and the frontier model (r=0.95), while the conditional model is strongly correlated with the frontier model (r=0.88). Together, this analysis indicates that the alternative approaches generate meaningfully different efficiency signals. This has important implications for efficiency profiling using these models.

This analysis also gives a sense of some of the pros and cons of different methods for profiling. For instance, as long as quality performance does not affect efficiency profiles after the hurdle is exceeded, the quality hurdle model places much greater emphasis on costs, rather than quality. This can be seen by its correlation with the unconditional model (30% quality). The opposite is true for the cost hurdle model. The analysis also highlights that, while the creation of efficiency tiers is straightforward with the conditional model, deriving value scores from the conditional model requires a separate scoring system that assigns a value to conditional cost and quality performance. Tompkins and colleagues[30] propose one method to do this, but others are possible.

In the analysis of cost and quality data available on Hospital Compare, lower cost is associated with lower quality: a 1% increase in cost performance (lower costs) is associated with a 0.19% decrease in quality performance ($p<.01$). Nonetheless, the analysis indicates that it is possible for hospitals to have both excellent quality performance and excellent cost performance: there are a number of hospitals that are close to two standard deviations higher than the mean for both quality and cost performance.

In other circumstances, there may be greater trade-offs between improving quality and increasing costs. In such cases, program sponsors should accommodate their expectations to the reality of cost and quality trade-offs.

To further illustrate this point, **Exhibit 8** shows the cost per beneficiary and quality scores from a hypothetical sample of hospitals. The vertical axis is spending per beneficiary and the horizontal axis is the hospital's total quality score. The quality scores are expressed here from 0 to 1, with 0 being the lowest quality and 1.00 being the highest. Contrary to the specification of costs to illustrate the alternative models to combine quality and cost, in this example, higher levels of cost indicate worse cost performance. A trend line has been fitted to the data.

As can be seen from **Exhibit 8**, there is a slight positive correlation between cost and quality for these hospitals. This is not to say that cost and quality are slightly positively correlated for all hospitals presently or that this relationship will continue in the future. As the health care system evolves and our ability to measure quality improves, cost and quality may very well become negatively correlated. Moreover, the nature of the relationship between resource requirements and quality may vary across dimensions of quality. For example, improving certain outcomes or adhering to best practices may result in greater resource requirements, suggesting the positive correlation. Meanwhile, quality improvements in patient safety and medical errors may result in lower costs from complications and treatment failures, resulting in a negative correlation between specified levels of quality (patient safety) and total cost of care (including complications and additional services) (Exhibit 9). Similarly, more extensive substitution of hospice and palliative care for higher-cost, marginally futile treatment approaches may have corresponding improvements in patient experience.[12]

After calculating objective efficiency based on principles and empirical calculations, a user could then determine what value to place on that efficiency score based on subjective-preference weighting.

Exhibit 10 provides an illustrative example of how to value hospital performance under a star rating system. The Total Quality Score (horizontal axis) and the efficiency score (vertical axis) are used to assign the value scores (i.e., determine the number of stars). Note that the same efficiency score is valued differently depending on the total quality score: higher total quality results in a greater value (number of stars) for the same efficiency score. Such a star rating system might be suitable for public reporting.

Once the assessment (i.e. number of stars) of the hospital performance has been made, it could be quantified by adjusting a hospital's Total Quality Score (0 to 100 points) upwards or downwards depending on its efficiency rating. An illustrative example is provided in **Exhibit 11**.

Section 4. Summary of findings from environmental scan

Our environmental scan and illustration of alternative models for combining quality and cost indicators highlights a number of key issues related to measuring efficiency or value in health care.

First, there are numerous extant approaches and no clear consensus about best practices. Of the 25 identified programs, we documented five broad approaches to combine quality and cost indicators. There is considerable variation within these approaches. Many of the quality measures included in the quality domains are exclusively measures that are endorsed by the NQF or by professional societies. The cost measures used to assess efficiency, however, have generally not been endorsed by the NQF.

Interestingly, the measure sets used to assess quality for many of the approaches taken by the private payers are more expansive than those used by the public payers. For instance, many of the private efforts profile specialist physicians, who have been largely ignored by public programs. The purpose of efficiency measurement is also different in the public and private efforts: the public efforts seek to use efficiency measurement to adjust provider payments whereas the private efforts use efficiency measurement to create tiered networks or for shared-savings programs.

The alternative approaches used to combine cost and quality measures have a number of pros and cons. The Conditional Model, the Unconditional Model, the Side-by-Side Model, and to a lesser extent the Hurdle Models all have the benefit of being relatively easy to understand. (Many of the program sponsors emphasized the importance of transparency, describing efficiency measurement in simple terms on their website but also publishing detailed methodology reports.) However, these approaches suffer from two separate aggregation problems that may undermine their validity. First, quality is almost always defined using multiple measures, and some kind of weighting scheme is required to summarize the performance of providers on these measures. The opportunity model, in which weights are based on the number of patients that are eligible to receive a given measure, remains a common approach to creating composite measures of quality. Another approach, used by the Alternative Quality Contract, assigns triple the weight to outcome measures relative to process measures. Both of these approaches to weighting measures, however, are largely arbitrary. A recent paper found that among 13 commonly used quality indicators, 7 of them accounted for 93% of the benefits to population health.[38] If weights assigned to individual performance measures do not reflect their importance to the health of patients, weighting schemes will, at a minimum, obscure the signal between observed quality and patient health.[39]

Second, as previously described, profiling has the potential to reach erroneous conclusions about the relative value of providers when the relationship between measured quality and patient health is not well defined. If quality is measured by patient survival, then small improvements have the potential to yield large gains in value, even at large costs. However, if quality is measured by a series of measures that have little relationship with improved patient health, large improvements may not yield gains in value, even at small costs.[32]

Among existing programs, there is a divergence in the practice of price standardization. The public programs (Hospital Value-Based Purchasing, the Physician Value-Based Payment Modifier, and the ACO

programs) standardize payments when measuring efficiency. The private plans vary with respect to price standardization, but tend not to standardize prices.

Variation in the prices of health care services charged by different health care providers, particularly among private payers, is well documented.[40] Variation in prices among private payers is driven largely by the result of negotiations between private payers and individual providers. Measures of health care spending (i.e., cost) that do not first standardize prices will measure costs as the product of price and the quantity of services for individual providers. Measures of spending that standardize prices substitute individual provider prices with average prices across the population of providers.

The decision of a program sponsor to use either unstandardized or standardized prices depends on the needs of the end user. Individual patients are likely to care more about value in terms of quality in relation to out-of-pocket spending. However, given the vast array of insurance products, the information needs are extraordinary to estimate patient cost-sharing associated with a certain procedure or episode of care from a certain provider. For patients with high deductible plans unstandardized prices will likely provide a better guide for out-of-pocket spending. Private insurers that are using tiered benefits designs to encourage patients to get care from lower-priced, higher quality providers are also likely to prefer unstandardized prices. This highlights the importance of not "stripping out" variation that is meaningful for consumers and program sponsors through price standardization. On the other hand, program sponsors with well-established reasons for price variations (e.g. Medicare's index for geographic variation in input prices, and supplemental payments for indirect medical education and disproportionate share for hospitals) may wish to highlight differences in resource use that are affected directly by utilization rates, and therefore use standardized prices.

There also appears to be a general ambivalence on the part of program sponsors with respect to harmonization within the quality and cost domains. This includes harmonization of the quality and cost domains for the same populations of patients (i.e., cost is often assessed for all patients while the quality measures apply to a narrower set of patients), for the same time intervals of measurement (i.e., the quality measures were assessed over much longer time windows than the cost measures), and the methods used to risk-adjust for cost and quality outcomes (e.g. Hospital Value-Based Purchasing uses different approaches for quality and cost).

The early efforts in efficiency profiling focused on hospitals,[26] but many now profile physicians and physician groups. This may have to do with the increase in ambulatory measures and advances in physician attribution methodology but may also reflect the increased bargaining power of hospitals.

Importantly, for the approaches we examined for combining quality and cost measurement, there is virtually no assessment of the reliability and validity of measures linking quality and cost.[5] In almost all cases, a resulting single measure is not defined. Instead, efficiency or value is defined through the joint consideration of quality and cost, with classification typically based on threshold values for both scales. While there is widespread recognition of the small sample-size problem associated with efficiency measurement, the most common solution is to use a sample-size cut-off as an exclusion criterion for providers' data to be profiled. Outside of Hospital Value-Based Purchasing, Bayesian reliability adjustment is not used to increase the reliability of efficiency measurement, although Leapfrog has used reliability adjustment for some surgical mortality measures. [41]

Section 5. Combining indicators of quality and cost for different use cases

Indicators of quality and cost could be combined for a variety of purposes, or "use cases."[iii] Potential use cases include internal efficiency profiling and improvement, public reporting, pay-for-performance, and network design. Key questions relate to how the various aspects related to combining quality and cost indicators may depend on a specific use case. What are the trade-offs that one might consider in combining these measures for a specific purpose? The following are some principles that can be applied to combining quality and cost indicators, including the selection of models, for different use cases:

1. When attempting to measure efficiency or determine value, neither the original cost nor quality signals should be obscured. Importantly, provider-level profiles of efficiency should show side-by-side the indicators of quality, cost, and (where applicable) the derived measures of efficiency or value. This is particularly relevant for public reporting, but is also recommended to ensure transparency for other use cases. Patients can make subjective and idiosyncratic choices about which treatments to receive from which providers, and may wish to formulate their own conclusions about relative value directly from the component measures. Displaying measures of cost and quality can provide inputs to be used for subjective and preference-weighted decisions. When making treatment decisions, patients can supplement these objective measures with past experience, familiarity, convenience, and informal advice from trusted sources. In situations involving terms of regulation or contracting, the disclosure of individual measures in all relevant domains allows stakeholders to understand the individual components, which also should be disclosed for transparency. Few of the current applications that use the side-by-side model to display measures of cost and quality also display a derived measure of efficiency or value. Using one of the models identified to construct a measure of efficiency, and then displaying this efficiency information alongside that of quality and cost, could help patients and other stakeholders interpret the component measures consistently, and still evaluate value based on their own preferences. This display of information is consistent with that of private-sector value ratings (e.g. US News and World Report and Consumer Reports, see Appendix A).
2. Efficiency scores and profiles of different providers should be developed and displayed across the entire relevant range of specific levels of quality. Moreover, levels of quality and corresponding efficiency or value can be measured continuously or discretely. If discrete measures are used (e.g., tiers, bands, or a star system), such categories should reflect meaningful differences across providers rather than arbitrary classifications based on distributions (e.g. centiles). For some applications, such as network design, discrete classifications may be preferred in order to group providers into manageable and meaningfully different network tiers (e.g., with categorically rather than continuously variable copayment amounts). However, discrete classifications add to measurement error by grouping

[iii] The term "use case" borrows from software development where applications are specified to meet the anticipated needs of identified users.

heterogeneous providers in homogenous groups. For example, the actual performance of a given provider close to a categorical threshold may be more similar in performance to another provider just over that boundary than to another provider nominally within the same category but not close to the boundary. Meanwhile, to avoid potentially false precision introduced by the use of continuous scores, variance estimates (such as confidence intervals) should be used whenever possible. Classifications based on rankings (e.g. percentiles) have the potential to magnify the importance of small differences in efficiency if scores are clustered close to threshold values. In some cases, this problem can be addressed through measure selection, i.e., by excluding quality measures that are "topped out;" (e.g., average scores close to the theoretical maximum performance level). Generally, it may be fine to display such measures along with others side-by-side in order to inform audiences about degrees of "mastery" on which most or all providers score well. However, distinguishing performance requires "grading on the curve" and focusing attention on measures for which meaningful variance can be displayed and used to determine relative efficiency and value.

3. It is important to anticipate the perspective of the decision-maker within each use case when selecting measures of cost. Third-party payers are concerned with payments they make for covered services related to the particular focus of measurement, which may include broad classes of care such as ambulatory surgeries, inpatient admissions, or primary care management of various acute and chronic illnesses. For example, a health plan would evaluate the efficiency or value of surgical procedures based on formulaic or negotiated payment rates for facility and professional services (separately or bundled). A consumer perspective would focus on out-of-pocket payments for deductibles, coinsurance, and copayments.

4. Generally, providers are not identical or necessarily similar in their relative quality, cost, efficiency, or value across lines of service; hence, the NQF framework for measuring resource use differentially for specific patient-focused episodes of care. It is often necessary to consider a full episode of care in order both to describe cost and quality dimensions in meaningful ways, and to render inferences about performance in terms of efficiency and value. Patients do not shop for "line-item" medical services such as pre-op services separately from anesthesia, surgical assistance, recovery and rehab; instead, they need pricing and quality information about the full anticipated treatment experience. Cost measures for episodes need to consider the full cost including potential complication costs, and the relevant dimensions of quality such as complication rates, risks, symptom-reduction, and functional status. These patient-focused episodes can provide a comprehensive and coherent framework for evaluating clinically meaningful performance in quality, efficiency, and value. By calculating efficiency and value separately across episodes, the resulting information can facilitate relevant decision-making, identify differential performance, and provide actionable information for care improvement. Meanwhile, the framework provides a standard structure for understanding the alignment of quality and cost measures, such as interpreting the value of services and costs occurring at certain times in light of benefits to patients, which may accrue immediately, eventually, or cumulatively over longer time periods.

5. Models of value that combine indicators of quality, cost, and efficiency differ with respect to the relative weight or importance that they place on quality and cost measures. For instance, the quality hurdle model attributes no merit to providers below the minimum threshold, and then in turn, places greater emphasis on differences in cost performance for providers above the threshold. Similarly, the cost hurdle places greater emphasis on quality performance among providers who perform better than minimal requirements (see **Exhibit 3** and **Exhibit 7**). The choice of model used to combine quality and cost measures could have a significant impact on the relative importance of incentives to reduce costs or improve quality. Generally, failure to distinguish differences in performance in all cases above or below a hurdle or threshold correspondingly reduces incentives for achieving better performance within such wide ranges. Therefore, policymakers and stakeholders should carefully consider how the choice of model to combine quality and cost measures best meets the goals of the use case.

6. When combining measures of quality and cost to assess provider efficiency, it is essential that risk-adjustment procedures are appropriately implemented to distinguish between sources of variation according to their appropriateness in evaluating efficiency and value. One dimension of this is time: it may be essential to measure and control for differences in initial patient severity in order to compare providers fairly in terms of subsequent performance. However, a key term is "initial" severity because in many instances patient severity has resulted from prior differences in provider performance, such as rates of complications or disease progression. Thus, a patient-focused framework may continuously update severity in order to reset "initial" conditions for evaluating efficiency and value, whereas more aggregate performance evaluations such as health plans, ACOs, etc. may hold initial severity constant much longer in order to distinguish longer term differences in performance that could involve managing severity levels.

   Furthermore, it is important to consider the composition of comparison patient cohorts for any given provider's attributed patient panel. Standard risk-adjustment procedures use "indirect standardization" in which regression analysis applied to the entire sample of patients is used to assign severity weights to individual comorbidities and risk factors. These weights are used to calculate "observed" outcomes, [42] as well as "expected" outcomes,[43] and combine these to evaluate providers' outcome performance while holding selected risk factors constant. However, this approach may not be optimal for two reasons: 1) if samples do not overlap on risk factors; and 2) if the functional form of the regression model does not fully account for differences in risk. In either case, provider outcome profiles may be confounded by specific characteristics that are unrepresented in other providers. Under these circumstances, matching using "direct standardization" may account for differences in severity across providers more effectively because comparison groups are formed separately and intentionally to match the characteristics and circumstances affecting the provider being evaluated.[44] In addition, when presenting quality and cost measures, quality performance should be displayed in its original form, and not adjusted for cost (and vice versa). Such adjustment would not allow quality and cost to be evaluated as separate constructs, which is critical for side-by-side comparisons.

These principles can be applied to the different use cases to identify some guidelines for efficiency measurement. To guide this effort, we performed key informant interviews with several members of the expert panel, each focusing on one of the identified use cases (internal efficiency improvement, public reporting, pay-for-performance, and network design). We also performed a literature search related to linking cost and quality measures for the cases, analyzing the key considerations for combining the quality and cost measures for the different use cases. Table 1 identifies a number of key issues related to the linking of cost and quality measures to assess efficiency or value for alternative use cases.

Internal efficiency improvement: Internal efficiency improvement can take many forms, from homegrown efforts within quality improvement departments to national profiling programs, such as efforts through the Cystic Fibrosis Patient Registry,[45] the Virginia Cardiac Surgery Quality Initiative[46], and the CMS Quality and Resource Use Reports for physician practices.[47] Other organizations, such as Premier Inc., provide reports and analytics related to cost and quality to their members or customers. While the scope and intentions of these efforts vary, they share the same intention, namely, to supply providers with actionable information on quality and cost, often confidentially, to facilitate efficiency improvement.

The primary audience for efficiency improvement efforts are providers themselves. This sets apart this use case from the others. One implication of this is that the only mechanism for improvement is for providers to use this information to improve their processes of care. Patients will not be steered towards higher efficiency providers through internal or confidential profiling. While providers may be benchmarked against similar providers or national rates, the standard basis of comparison for internal efficiency efforts is within-provider over time: is the provider improving on a given metric? Because the focus of these efforts is internal, providers can benefit most from granular, actionable information. This includes details on processes of care. Outcome measures are important metrics for providers to understand their comparison relative to others; however, by itself, this information is not sufficient to drive improvement without further investigation into causal differences related to structures or processes that are amenable to adaptive change.[48]

The cost measurement that is most relevant for internal efficiency relates to “waste,” “appropriateness,” and production costs. Some efficiency fixes, such as to order lower-cost supplies, are a one-time fix for providers. For example, switching vendors or negotiating lower prices can establish permanently (or at least indefinitely) lower costs. Other improvements may address ongoing processes and require consistent effort and attention. For instance, identifying the inappropriate use of expensive antibiotics can be a mechanism to reduce costs, but may require the continuous monitoring of prescribing patterns over time. Similarly, altering referral patterns or avoiding inappropriate procedures can also increase efficiency but must be monitored or revisited over time. Increasingly, sophisticated software can help hospitals and other providers to identify these cost drivers.

Ideally, metrics on cost and quality are aggregated at the level of a clinical unit or service line that is in position to improve efficiency. In some cases, increasing efficiency can be reasonably expected to also improve quality. For instance, once patients are taking more than a certain number of medications (e.g.

five), adding more medications will almost invariably lead to interactions. Careful monitoring or even limiting additional prescribing at this point may improve both efficiency and quality.

Standardized reports like those produced by CMS may be a good place to start. However, more detailed drill-downs may be necessary for providers to identify waste.[49] Ideally, internal quality reports would be aligned with the metrics used by other sponsors of efficiency improvement programs (e.g. payers), although this is not always possible. A problem with standardized reports is that they come from entities (e.g. payers) without full information about providers' processes. As a result, they may not be able to identify key sources of inefficiency. For example, claims data may not be sufficiently detailed to determine whether a procedure was appropriate.

Public reporting: For public reporting, the audiences include consumers, as well as health care providers and other stakeholders, such as policymakers. Public reporting could drive value improvement if patients migrate toward higher efficiency providers, and away from lower efficiency providers. The potential for volume loss, and the heightened salience of efficiency from the public reports, may in turn lead providers themselves to attempt to improve in response to the reports.

Patients' responsiveness to public quality reporting has been the subject of intense research over the last several decades.[50, 51] There is now some consensus about how these programs should be designed with respect to the measures displayed (those that are most relevant to patients (e.g. patient experience and patient outcomes) and how the information is displayed (e.g. summary measures).[52] Nonetheless, the measures used for quality reports remain a fundamental limitation. While patients are interested primarily in outcome performance, outcome measures (such as mortality or complication rates) remain rare for many diagnoses while others (such as readmission rates) are more difficult to interpret.[53] Expanding the collection of patient-reported outcomes on functional status could enhance both the validity of public reports and patients' response to the published information.[54] Some measures of patient experience also can be important, relevant, and understandable to patients who are making decisions about which provider they should choose. Nonetheless, measuring the key dimensions of this experience remain challenging. Attempting to improve patient experience scores may also, in some cases, decrease other dimensions of quality.[55] For instance, a hospital that served meals that are relatively rich and more familiar to patients, although not healthy, may receive higher patient experience scores than one that limited patients' food options to "healthier" or more familiar choices following heart surgery. In contrast, educating patients about nutritional requirements and providing meals that are both healthy and appreciated may help as much or more to improve reported patient experience and clinical outcomes.

In contrast to quality, public reporting of cost is relatively new, and has received considerably less attention in the literature.[49] There are a number of unique features related to the public reporting of cost. First, patients are concerned primarily with the out-of-pocket cost that they would be expected to pay for a given episode, rather than or separate from the cost borne by a third-party. Consumers want to assess value by evaluating choices based on what they can expect in terms of outcomes and the price they would pay. However, it is not possible for program sponsors to display a reasonably accurate expected cost for a given patient. This is because, even within a provider or facility, cost will vary by a

patient's diagnosis, severity, comorbidity burden, risk of complications, and the specific care that they would receive. In addition, even the expected cost adjusted to reflect patient circumstances would also vary across insurers and specific plans, given differences in benefits and coverage, and could vary within a given plan (e.g., whether a patient has not yet reached her deductible or has exceeded the out-of-pocket limit). The price may also vary according to whether a provider is in the network of a given insurance product. Compared to buying a car, where the price will be well known, publicly reported cost for health care providers will likely be a weak signal for an individual patient.

Instead of providing accurate information specific to patients about the expected costs of their care, public reporting of costs could provide a basic benchmark for the relative prices and intensity of services across providers. This could provide patients with a general sense of their likely expense from seeing care from different providers based on their tendencies to be relatively efficient or "high value" for similar patients and lines of service. More specific information about pricing could be obtained from their insurance carrier.

When paired with quality information, publicly reported costs (such as episode costs) and general indicators of efficiency and value can provide incrementally more information than a patient would otherwise have when choosing a health care provider. An appropriate model could combine the quality and cost domains and then rank providers in a given market area based on their efficiency. An optimal reporting system would simultaneously display cost, quality, and efficiency domains in aggregated form (e.g. stars). While a default presentation may rank providers in a specified market area by their efficiency performance, a flexible reporting system would allow patients to reweight the cost, quality, and efficiency domains and re-order rankings based on their preferences in order to reach their own conclusions about value.

The importance of displaying information in a valid, understandable and actionable format is paramount. The goals of public reporting – to change patient and physician behavior – are not the same as those of transparency – which is simply to allow access to relevant information. A maximally flexible system would allow patients and other stakeholders to drill down to performance for different episodes, whether for conditions or anticipated types of treatment, and for individual performance measures. This would allow consumers to customize reports to their needs and allow them to understand the factors that contributed to provider performance.

Because public reporting is intended to steer patients toward high quality and high efficiency providers, and away from low quality and low efficiency providers, the model used to combine cost and quality measures should be focused on identifying these extremes. As a result, the considerations for doing so may be similar to those used for network design (see below).

Pay-for-performance/value-based purchasing: The first generation of pay-for-performance programs were aimed primarily at measures of underuse (e.g. diabetes care, cancer screening).[56] The hypothesis that greater adherence with evidence-based care would lead to improved health and to cost reductions was not supported by most program results.[19] The second generation of pay-for-performance, commonly referred to as value-based purchasing, is more explicitly focused on cost reductions through

several pathways. First, some programs link incentives to measures of appropriateness (such as rates of imaging) rather than measures of underuse. Other programs attempt to lower cost or resource use by including cost or resource use as a performance measure (e.g. Hospital Value-Based Purchasing). Finally, other programs use a shared savings approach in which bonuses for high quality care are financed by the savings to payers generated by providers, typically by having spending that is less than a projected trend.[34] At the same time, the second generation pay-for-performance programs are moving towards incorporating outcome quality measures, and away from process measures.

By providing incentives directly to providers to improve efficiency, provider behavior change is the mechanism for value improvement in pay-for-performance programs. Consumer choice is a less important mechanism for facilitating value improvement in these programs. As a result, the programs need to focus on those quality and cost measures that are central to the payers themselves.

In general, pay-for-performance programs are designed to facilitate incremental improvement among the population of providers. Programs can be designed to reward the best providers, or punish the worst performing providers, but these are not as common. If the intention is to reward or punish, hurdle models may be more appropriate because they isolate and deny merit to providers who fail to meet threshold requirements.[57] To facilitate incremental improvement, a model that generates continuous efficiency or value scores across the range of the quality distribution (e.g. the cost-effectiveness model) rather than models that assign providers into discrete groups (e.g. the conditional model) may be preferred. Similarly, rewarding improvement over time may facilitate generalized improvement in addition to rewards for cross-sectional comparisons within performance periods.

There is also the question of how a program should rank the relative efficiency or value of providers that have high cost and high quality compared to those that have low cost and low quality. For instance, the Physician Value-Based Payment Modifier treats these physician practices as providing comparable efficiency (giving them no payment adjustment). In contrast cost-effectiveness models assess the cost/QALYs for each provider, and providers are ranked based on this formula-based measure. This can help to distinguish whether the cost/QALY is lower (i.e. higher value) among the high quality/high cost provider or the low quality/low cost provider.[32, 38]

Network design: Commercial insurers are increasingly using measures of quality and cost to determine which providers to include in their networks. They are also using these metrics to assign providers to different "tiers" for which patients face different levels of cost sharing. These efforts are being focused on high cost specialty care (see Exhibit 1). When creating these tiers, insurers must balance a number of objectives, including: 1) Getting the efficiency scores right in order to make the correct determinations about which providers are included in a network, and their tiers correct from the perspective of the payer; 2) Getting buy-in from providers in the form of acceptance of the contractual terms; 3) Getting buy-in from patients in the form of responsiveness and behavioral change; and 4) Managing the regulatory environment by conforming to applicable standards and keeping pace with expectations. For instance, the New York State Attorney General investigated several insurers that created tiered networks on the basis of cost alone, not quality.[58]

Profiling quality and cost for network design is intended primarily to steer patients towards those providers that are deemed higher value. As with the other use cases, measuring quality appropriately remains a challenge. The quality measures that are collected (e.g. preventive screening) are often unrelated to major cost drivers in the near-term. Better measurement around patient-reported outcomes could facilitate much better efficiency profiling and tiering. In addition, consumers may not understand how the measures are constructed, and may face difficulty determining the credibility of measures that are not intuitive. This may be why tiers based on cost and quality have not driven major changes in patient behavior.[59] Conceptually, insurance tiers are a weaker version of reference pricing, in which patients bear the entire costs of care above an established reference price. A reference pricing experiment in California appears to have led to larger reductions in prices, apparently because providers with higher prices feared volume loss.[29] These concepts could be combined if a reference price was established for care based on the price of a high efficiency provider.

The clinical areas that have been targeted by insurance tiering tend to be high cost specialty care that have high price variation and presumably the most potential for significant cost savings (e.g. spine surgery, knee and hip replacement, cardiac care, transplants, bariatric surgery, and complex and rare cancers). Price variation is a leading driver of cost variation for this care, and insurers have an incentive to use unstandardized prices to profile providers for network design.

The choice of model for combining cost and quality measures to profile efficiency for network design is likely driven by insurers' interest in reducing costs. While it is important to have a quality standard – particularly to meet regulatory requirements – a basic standard of quality may be sufficient. For this reason, the quality hurdle is aligned with interest of commercial insurers.

Section 6. Implications for the National Quality Forum

The NQF has a number of options for how it could advance evidence-based methods to link quality and cost measures to assess efficiency in health care. First, NQF could use its existing endorsement process in several ways:

1. **Request that developers of cost or resource use measures specify a link with quality measures**. The submission form for cost and resource use measures could include a section asking developers to include a variety of additional information related to how a cost or resource use measure would be linked to quality performance. Developers would not be required to submit this information in order for a cost or resource use measure to be endorsed, but this information could help NQF committee members and eventually stakeholders understand the intended use of the measures in practice. The following could be requested of developers of cost or resources use measures in the endorsement process:
    a. **Identify quality measures that are relevant for the proposed cost or resource use measure.** The selected quality measures should be reliable, valid, useable, feasible to collect, and related to the proposed cost or resource use measure (e.g., both measures assesse performance for patients with the same diagnosis or patients receiving care for the same procedure). The selected quality measures should be endorsed by NQF, unless other measures are identifiable that are more aligned with the cost measure, or more appropriate for linkage. The type of quality measure (structure, process, outcome, or patient experience) should depend on the use case (see below). There is a place for process measures, which often are useable and actionable, but they should be proximal to an outcome. Whenever possible, it is preferable for the specifications of the cost and quality indicators to be harmonized. This includes measuring cost and quality for comparable populations of patients, for the same time intervals of measurement, and the methods used to risk-adjust for cost and quality outcomes. Optimally, this would be done around common episodes. However, it may often not be possible or reasonable to harmonize cost and quality measures given prevalent limitations in current measures. One key reason for that is composite measures are often used to measure quality performance, and the individual measures contained in composite measures often have different data capture periods and apply to different populations. Nonetheless, this is a principle to strive for in future measure development.

    b. **Determine whether and what type of composite measures will be used to measure quality**. Composite measures have important uses. For instance, consumers may prefer a single score that is easy to interpret, and program sponsors may need a single score to evaluate providers (e.g. for pay-for-performance). There are a number of approaches to create composite measures. These include all-or-none composites (requiring that a patient receive all recommended care for the composite to be met), composites based on opportunities (equal to the sum of successfully achieved processes of care divided by the opportunities to provide

recommended care), and composites that assign different weights to different types of measures (e.g., weighting outcome measures more heavily).[60] There are also a number of NQF-endorsed composite measures. While it is preferable for programs to use NQF-endorsed composite measures, the dearth of these measures makes this unreasonable currently in most cases. Instead, programs should have a reasonable justification for the weighting of individual measures, including the known correlation between measures and patient outcomes.[38]

c. **Combine the quality and cost indicators in a manner that is most appropriate for a likely use case**. Issues related to combining cost and quality measures for different use cases are described in the previous section. In some instances, the same measures and models to combine indicators may be used for multiple use cases (e.g., pay-for-performance and public quality reporting). Side-by-side displays of measures, for example for public reporting, can include measures that are scored on mastery, rather than relative performance or rankings. For example, if a large majority of providers have similar or even identical scores on a measure, it may be informative and reassuring for consumers to be aware that their options are similar, or possibly all excellent, on that measure. Distinguishing relative performance, on the other hand, requires "grading on the curve," with meaningful underlying differences that are measured reliably.

The benefit of simply requesting that measure developers provide this additional information is that this detail would likely be useful for NQF committees making endorsement decisions. The optional nature of this information would also minimize the burden for developers. The downside of this approach is that – because this information would not be required for endorsement – it may not be provided by developers.

2. **Require that developers of performance measures specify a link with corresponding quality or cost measures**. Instead of requesting information from developers about how cost or resource use measures could be linked to quality, NQF could require developers to provide this information. This could be done by modifying the current "usability and use" criterion. This criterion is defined as the "Extent to which potential audiences (e.g., consumers, purchasers, providers, policymakers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations." NQF could require developers to satisfactorily provide the information described above in order to meet this criterion.
   The benefit of requiring developers to provide this information is that it would ensure that there was a specified plan through which cost or use measures would be linked to quality measures in a reasonable and valid manner. The main drawback of this approach is that it would increase the burden of developers when submitting measures for endorsement. This may decrease the incentives of developers to submit these measures to NQF.
3. **Create a pathway to endorse efficiency measures that link cost and quality measures.** To date, few stand-alone measures are being used to assess efficiency. Instead, efficiency is largely being assessed as the output of alternative models that link quality and cost

measures (see Section 2). However, a number of measure-developers have established sufficiently detailed processes to measure efficiency that they could submit for NQF endorsement. If the NQF decided to endorse approaches to efficiency measurement it could consider a number of guidelines. First, the NQF could stipulate that the quality and cost measures used to evaluate efficiency should have been previously endorsed. If not, the developer would have to provide a sufficient reason. Second, the NQF could provide guidance with respect to whether specifications of quality and cost measures should be harmonized. This may result in the modification of the specifications of measures that have previously been endorsed by the NQF. Third, the output of the efficiency measures should meet the standards of scientific acceptability established by the NQF. Specifically, efficiency classifications should be reliable and valid, and statistical testing should be able to demonstrate this. If efficiency measures were endorsed, NQF could provide guidance about how these measures should be used in accountability programs (e.g., that they be displayed side-by-side with cost and quality measures for any reporting application).

If NQF chose to endorse standalone efficiency measures, the quality of these measures would likely improve, as the measures would need to pass the rigorous scientific criteria that are required for endorsement. The endorsement process may also elevate the profile of efficiency measures, encouraging their use. The downside of NQF choosing to endorse efficiency measures is that, if this strategy was pursued in lieu of recommendations #1 and #2, cost and resource use measures could continue to be endorsed without an explicit link to quality measures. NQF and others would then have to wait for developers to submit an efficiency measure that linked a newly endorsed cost or resource use measure, which may never occur.

4. **Use the Measures Application Partnership to advance the linking of cost and quality programs.** Apart from using the endorsement process, NQF could use its Measures Application Partnership to promulgate evidence-based efficiency assessment. The Measures Application Partnership is a multi-stakeholder partnership organized by NQF to provide guidance to the Department of Health and Human Services about the use of performance measures in public accountability programs. The scope of the Measures Application Partnership could be expanded to offer recommendations about linking cost and quality measures to assess efficiency in health care. The Measures Application Partnership could use this report as a foundation to provide this kind of guidance to influence evidence-based policymaking.

These recommendations could be pursued either alone, or in combination. For instance, measure developers could be asked to require additional information about how quality would be linked to cost or resource use measures (recommendation #1) and NQF could allow standalone resource use measures to be separately endorsed (recommendation #3). Also, the use of the Measures Application Process to encourage evidence-based efficiency assessment could be pursued independent of recommendations concerning the endorsement process.

There are a number of outstanding questions about linking quality and cost to measure efficiency in health care. What is the reliability and validity of the alternative models of linking quality and cost?

Would certain models systematically favor certain types of providers? How do consumers understand alternative models and displays of quality, cost, and efficiency information? How might alternative models create different incentives for provider behavior change? Future research should address these questions.

List of Exhibits

Exhibit 1. Summary of programs that combine quality and cost indicators to measure efficiency[iv]

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
1.	Aetna Aexcel [61]	12 categories of specialty services[v]	Specialist and physician practice level	Volume (at least 20 episodes in the last year) clinical performance structure measures (use of technology, certification) completion of performance-based improvement module claims based measures (HEDIS, readmissions, in-hospital complications)	All costs attributed to specialists for specific episodes of care	Variation on Quality Hurdle model. Quality and volume are assessed first. If costs are lower than threshold based on peer performance, providers are designated for Aexcel network.
2.	Anthem Blue Precision	5 categories of specialty services[vi]	Specialist and physician practice level	Receipt of designation from National Committee for Quality Assurance (NCQA) or Bridges to Excellence, or performance on clinical process measures.	Combination of all attributed costs, diagnostic testing, prescription use, procedures and follow-up care, and hospital care.	Quality hurdle model. Physicians must first be recognized for quality. Physicians are then designated for recognition if their cost performance is not significantly higher ($p <$ .10) of 110% the geographic average.

[iv] In addition to the programs identified in this table, we are aware of a number of other programs that appeared to combine indicators of quality and cost to measure efficiency. These include programs initiated by Castlight Health, the Minnesota Smart Buy Alliance, PacifiCare, the Puget Sound Health Alliance, Blue Shield of Oregon, Tufts Health Plan, and the Wisconsin Department of Employee Trust Funds. However, we were unable to obtain detailed information about how the programs were specified.

[v] Cardiology, Cardiothoracic surgery, Gastroenterology, General surgery, Neurology, Neurosurgery, Obstetrics and gynecology, Orthopedics, Otolaryngology/ENT, Plastic surgery, Urology, and Vascular surgery

[vi] Rheumatology, cardiology, obstetrics/gynecology, endocrinology, and pulmonary medicine.

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
3.	Blue Cross and Blue Shield Blue Distinction Centers® for Specialty Care[62-64]	6 categories of specialty services[vii]	Hospital	Nationally consistent measures based on structure, process, outcomes, and patient experience. Hospitals must meet quality thresholds for each domain. Measures were developed with input from the medical community.	All costs for specific episodes of care (including facility, professional, other). Each provider's cost of care is calculated on an episode basis, using allowed amounts based on Blue Plans' claims data. The cost of care criteria takes into account outliers, patient level risk factors, and geographic variation, before each facility is assessed against a consistent national benchmark.	Quality Hurdle Model

[vii] Six specialty care areas are included: Spine Surgery, Knee and Hip Replacement, Cardiac Care, Transplants, Bariatric Surgery and Complex and Rare Cancers. The three specialty care areas with asterisks have Blue Distinction Centers; Blue Distinction Center+ designations will continue to roll out in other areas.

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
4.	Blue Cross Blue Shield of Illinois and Advocate health care[65]	All covered services for Advocate health care, a not-for-profit integrate system	System level	Performance for 12 measures	Global budget target	Combination of Quality Hurdle and Cost Hurdle Models.
5.	Blue Cross Blue Shield of Michigan Hospital P4P program[66]	Hospitalized patients with specific index admissions	Hospital	Composite index of collaborative Quality Initiatives, population based, performance, all-cause readmissions	Diagnosis standardized cost-per-case	Unconditional Model. Payments are based on the weighted sum of quality and cost domains
6.	Blue Cross Massachusetts Alternative Quality Contract [67]	All covered services	Alternative Quality Contract provider organizations	32 ambulatory measures, 32 hospital measures. 5 Quality "gates" for each measure, resulting in different bonus payments. Outcome measures are triple weighted. Non-linear function between quality score and payout. [68,69]	Global budget target	Unconditional Model. High quality is rewarded as a bonus, can equal up to 10% of global budget.[viii]
7.	Buyers Health Care Action Group Purchasing Initiative[70,71]	All services	Care systems in Minneapolis/St. Paul	Patient experience and participation in quality improvement initiatives.	Total costs	Side-by-Side Model

[viii] The AQC can be conceptualized two different programs that are not directly connected: a shared savings program and a quality bonus program.

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
8.	Cigna Care Designation[72]	22 categories of specialist services[ix]	Physicians and physician groups	5 domains related to National Committee for Quality Assurance (NCQA) Physician Recognition; Group Board Certification; Composite quality index on adherence to 101 Evidence-Based Medicine (EBM) Rules; American Board of Internal Medicine Process Improvement Module Completion; Certified Bariatric Center Affiliated Surgeons.	Costs related to Episode Treatment Group (ETG) methodology	Conditional Model. Providers are compared by specialty within markets.
9.	Cigna Collaborative Accountable Care[73]	All covered services	Large primary care or multispecialty practices, integrated delivery system, of physician-hospital organization	Composite measure assessing adherence to evidence based medicine for preventive care, chronic care, and acute care.	Unclear	Quality Hurdle Model

[ix] Allergy and Immunology, Cardiology, Cardio-Thoracic Surgery, Colon and Rectal Surgery, Dermatology, Ear, Nose and Throat, Endocrinology, Family Practice, Gastroenterology, General Surgery, Hematology and Oncology, Internal Medicine, Nephrology, Neurology, Neurosurgery, Obstetrics and Gynecology, Ophthalmology, Orthopedics and Surgery, Pediatrics, Pulmonary, Rheumatology, and Urology

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
10.	Health Partners Relative Resource Use[74]	Primary care, specialty care, and hospitals	Physicians, physician practices, and hospitals	Separate composite measures for primary care, specialty care, and hospitals. Components of composite differ for different types of services.	Uses NQF endorsed total cost of care measure. Encompasses all services with/without price standardization.	Side-by-Side Model
11.	Hospital Value-Based Purchasing	Part A and Part B Medicare services	Hospital	Sum of performance score (incorporating attainment and improvement) for individual measures in various domains (outcomes, clinical process, and patient experience)	Episode covering standardized payments from 3 days prior and 30 days following hospitalization.	Unconditional Model
12.	Integrated Healthcare Association Value Based pay-for-performance program	Patient care among seven health plans in California.	Physician organizations	31 clinical quality measures, 15 meaningful use measures, 6 patient experience measures, 12 appropriate resource use measures.	Per member per months total cost of care, including physician, hospital, pharmacy and ancillary payments.	Quality Hurdle and Cost Hurdle Models are used together. Shared savings model then adjusts savings by quality performance.

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
13.	Kroger Centers of Excellence Program	Hip and knee replacement and spine surgery	Hospitals and surgical centers	Used Blue Cross and Blue Shield Distinction criteria (see above) supplemented with information on facility rankings from published reports (e.g. US News and World Report's Top 50 Orthopedic Facilities)	Unstandardized reimbursement for episode cost.	Unconditional model. Quality and cost performance was combined to form a single composite. Facilities were classified into four tiers based on composite performance.
14.	Leapfrog Hospital Rewards Program[75]	Patients hospitalized with AMI, pneumonia, or child birth, or receiving CABG or PCI.	Hospital	Composite score of multiple measures. Uses a two-level weighting approach based on potential of indicator to reduce mortality and the importance of the indicator to the employer.	Inpatient costs	Conditional Model
15.	Maine Health Management Coalition	Adult care, pediatric care, and hospital care	Physicians, physician practices, and hospitals	Composite measure based on Bridges to Excellence / Hospital Compare measures categorized into "low", "good", "better", and "best"	Whether practice is "working to control cost"	Side-by-Side Model
16.	Maryland multi-payer patient-centered medical home program[65]	All covered services	Primary care practices	21 quality measures; and reductions in use of high-cost services.	Total costs for assigned patients.	Cost Hurdle Model

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
17.	Massachusetts Group Insurance Commission value-tiering program[76,77]	All covered services	Physician-level. Physician profiles various participating plans	Composite based on 79 quality measures relevant to particular providers	Price standardized episode costs based on Episode Treatment Group methodology[59]	Unclear
18.	Medica and Fairview health services[65]	All covered services for Fairview Health Services, a non-profit health system	System level	Minimum quality gate, then confidential algorithm	Global budget target	Unclear
19.	Medicare Physician Group Practice Demonstration	Part A and Part B Medicare services	Integrated delivery systems	Performance for 32 ambulatory care performance measures.	Total costs per capita for aligned beneficiaries	Unconditional Model[x]

[x] While it appeared that the Unconditional Model was used in the Physician Group Practice Demonstration, there was some uncertainty about this classification

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
20.	Medicare Shared Savings and Pioneer Accountable Care Organization programs[78]	Part A and Part B Medicare services	Accountable Care Organization	Composite measure of patient/caregiver experience (7 measures); Care coordination/patient safety (6 measures); Preventive health (8 measures); At-risk population: Diabetes (1 measure and 1 composite consisting of five measures); Hypertension (1 measure) Ischemic Vascular Disease (2 measures); Heart Failure (1 measure); Coronary Artery Disease (1 composite consisting of 2 measures).	Payment standardized total costs per capita for aligned beneficiaries	Quality Hurdle Model
21.	NCQA relative resource use[79]	Condition-specific costs for people with specified chronic diseases.[xi]	Health plan level by product (e.g. HMO, PPO)	Composite measure based on HEDIS indicators relevant to disease area	Annual condition-specific costs for all relevant services	Side-by-Side Model

[xi] Asthma, cardiovascular conditions, COPD, diabetes, and hypertension

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
22.	Physician Value-Based payment modifier	Part A and Part B Medicare services	Physician practice	Composite measure of clinical care, patient experience, population/community health, patient safety, care coordination, and efficiency.	Composite measure of total costs per capita for attributed beneficiaries, and for beneficiaries with specific chronic disease	Conditional Model
23.	Tufts Health Plan primary care "Blue Ribbon" program[80]	Primary care	Physician practice	7 HEDIS process of care measures and 7 patient experience measures. Calculate adjusted composite process scores (z-scores), and composite scores for patients experience (z-scores). Scores were then summed and renormalized.	Primary care Episode Treatment Groups	The Conditional Model. The quality and cost domains are standardized and combined with equal weighting. To be designated with the "Blue Ribbon", providers must be above the median on both the quality and cost domains, as well as the combined domain.

	Name of program	Services evaluated	Level of attribution	Specification of quality	Specification of cost	Approach to combining quality and cost
24.	UnitedHealth Premium[81-83]	25 categories of specialist services.[xii]	Physician, physician practices	Composite score based on evidence based measures related to preventive care, appropriate care, chronic disease care, patient safety, sequencing of care, and care outcomes.	Risk adjusted total cost of care (population cost), and episode cost measurement.	The Unconditional Model. Provider designations are made separately for cost and quality based on statistical criteria. It's unclear how the different designations translate into payment or cost sharing differences.
25.	Virginia Cardiac Surgery Quality Initiative[46]	All cardiac surgical patients	Surgeon and hospital	Extensive structure (volume), process, and outcome (mortality and complication) measures.	Normalized hospital and surgeon charges[84]	Side-by-side Model. Comparisons are made for anonymized hospitals and are primarily on quality measures.

[xii] Allergy, Cardiology, Cardiology – Electrophysiology, Cardiology – Interventional, Endocrinology, Family Medicine, General Surgery, General Surgery - Colon/Rectal, Internal Medicine, Nephrology, Neurology, Neurosurgery – Spine, Ophthalmology, Obstetrics and Gynecology, Orthopedics - Foot/Ankle, Orthopedics – General, Orthopedics – Hand, Orthopedics - Hip/Knee, Orthopedics - Shoulder/Elbow, Orthopedics – Spine, Orthopedics – Sports Medicine, Pediatrics, Pulmonology, Rheumatology, and Urology

Exhibit 2. Distribution of quality and cost measures used in models

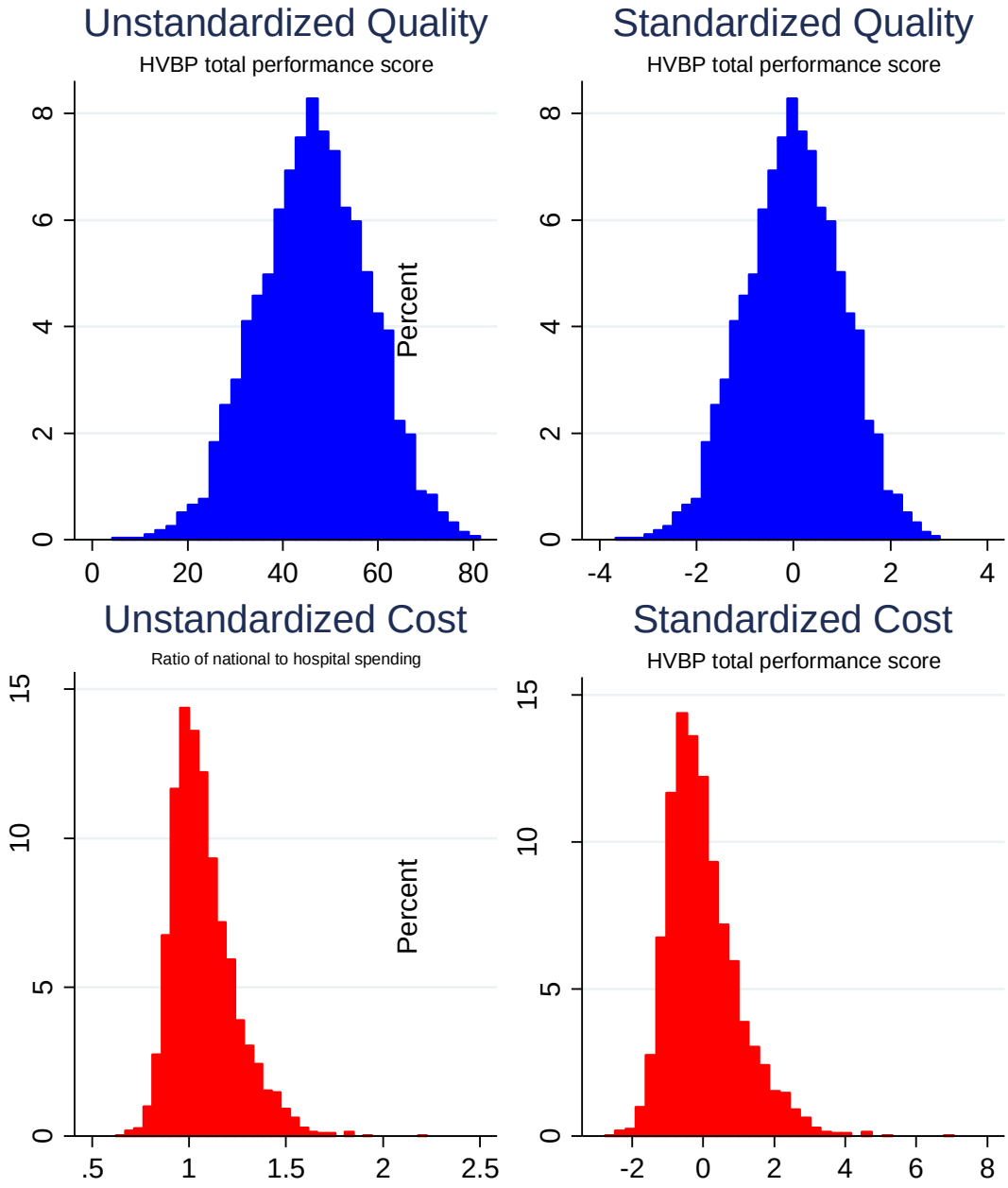


Exhibit 3. Illustration of quality hurdle and cost hurdle models

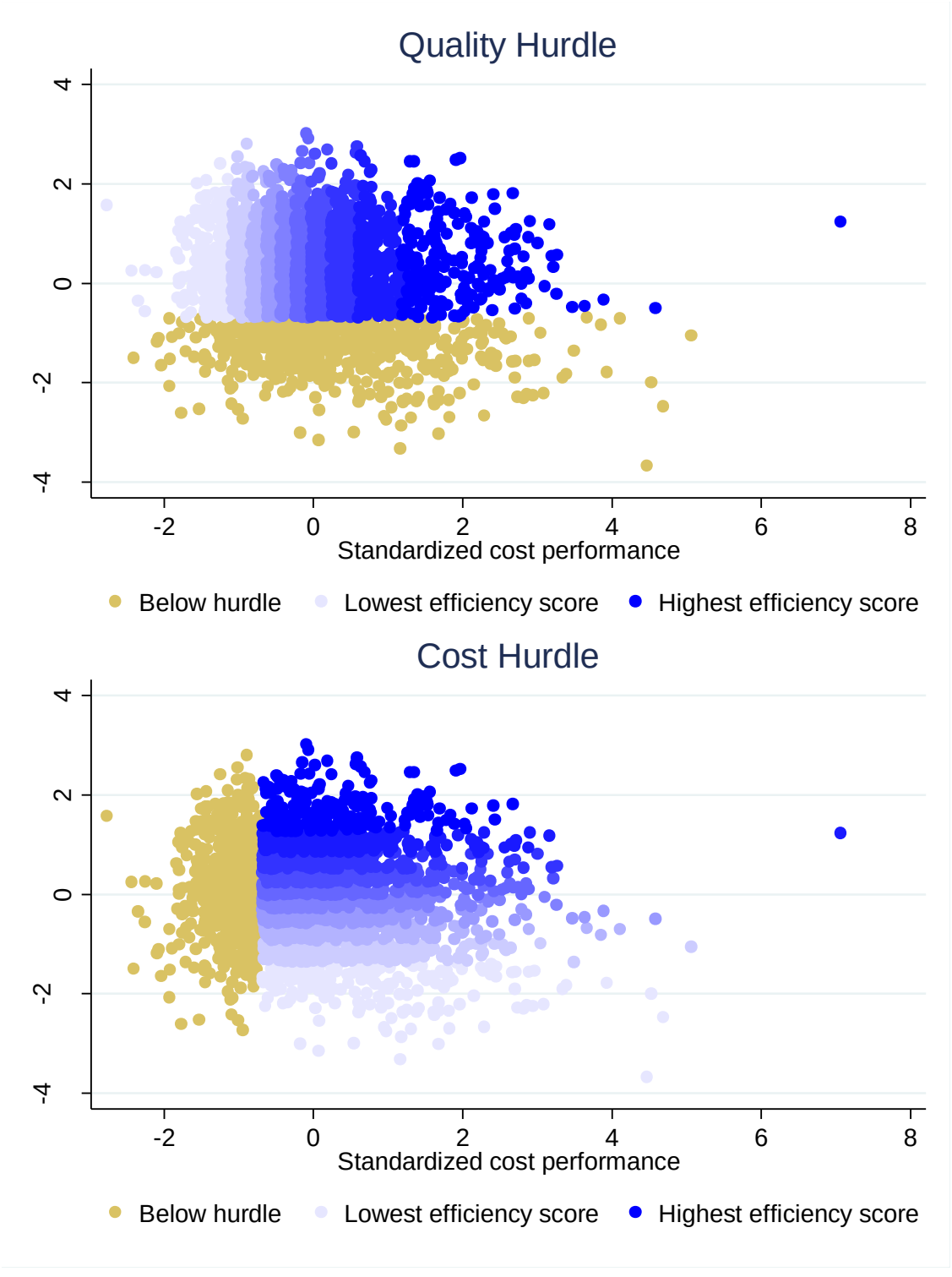


Exhibit 4. Illustration of unconditional model

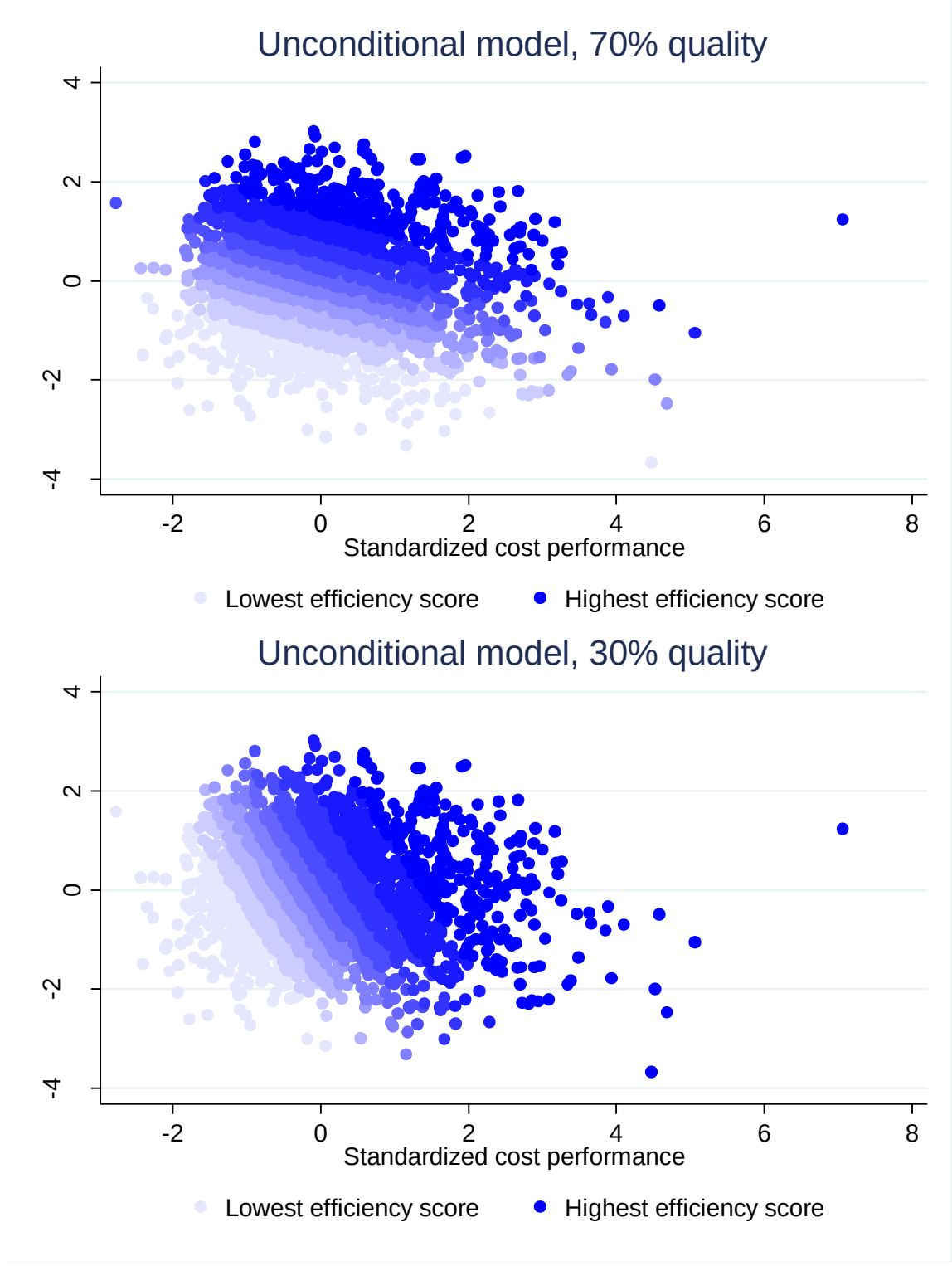
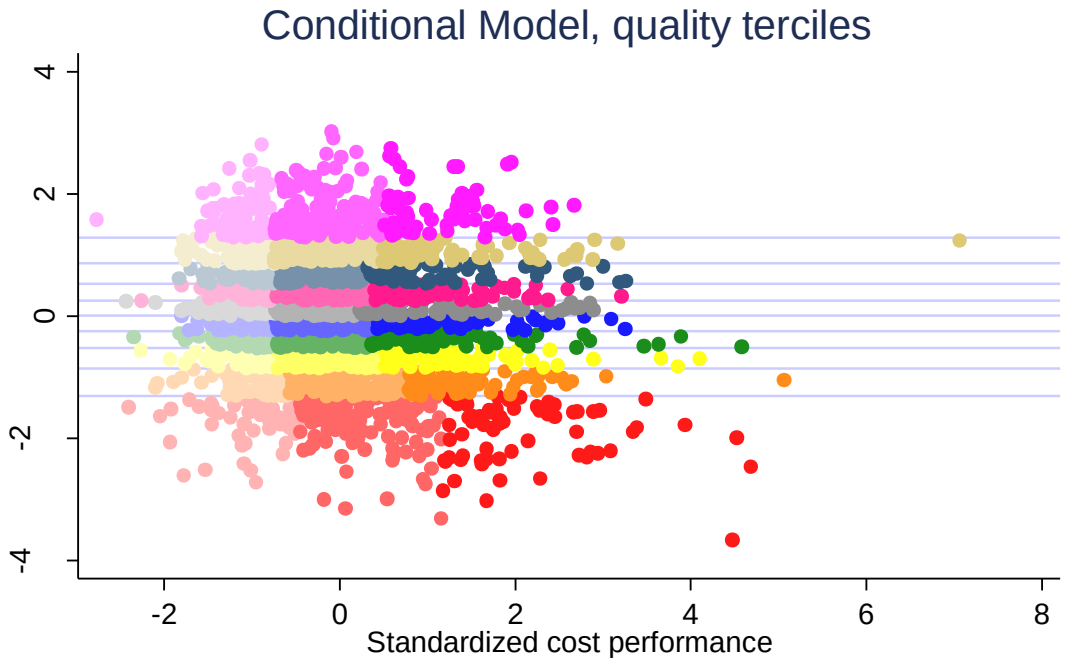
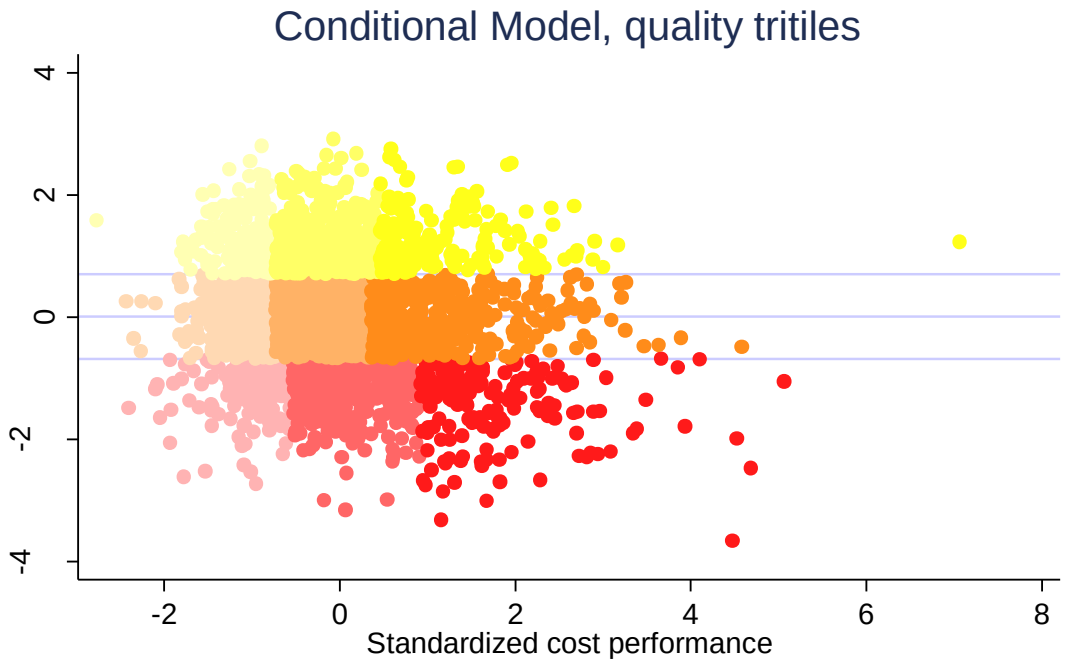


Exhibit 5. Illustration of conditional model



Note 1: Darker shades denote higher Cost performance
Note 2: Horizontal lines denote quality terciles



Note 1: Darker shades denote higher Cost performance
Note 2: Horizontal lines denote quality terciles

Exhibit 6. Illustration of stochastic frontier model

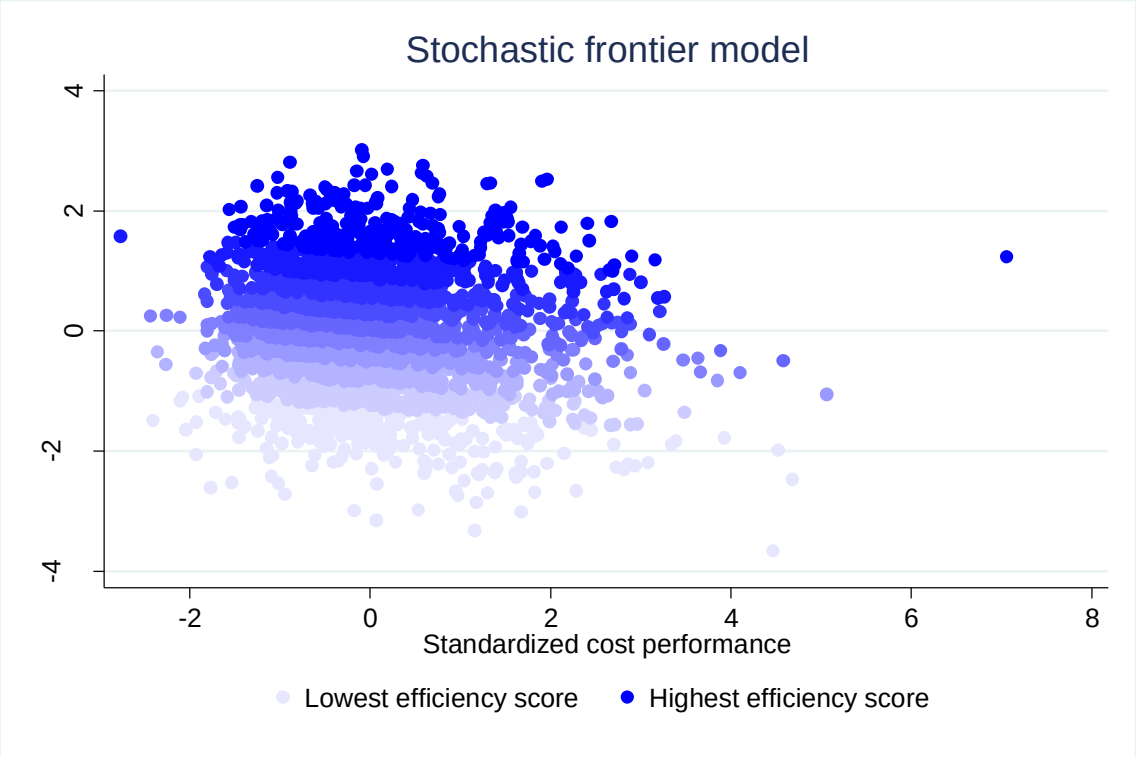


Exhibit 7. Correlation between efficiency scores generated from alternative models linking cost and quality indicators

Model	Quality hurdle	Cost hurdle	Unconditional (70% quality)	Unconditional (30% quality)	Conditional
Quality hurdle	-	-	-	-	-
Cost hurdle	0.1003	-	-	-	-
Unconditional (70% quality)	0.3196	0.8055	-	-	-
Unconditional (30% quality)	0.7802	0.2590	0.6610	-	-
Conditional	0.2122	0.7591	0.8906	0.5118	-
Frontier	0.0718	0.8745	0.9492	0.3992	0.8753

Exhibit 8. Hypothetical example of cost per beneficiary and total quality scores for sample hospitals (positive correlation)

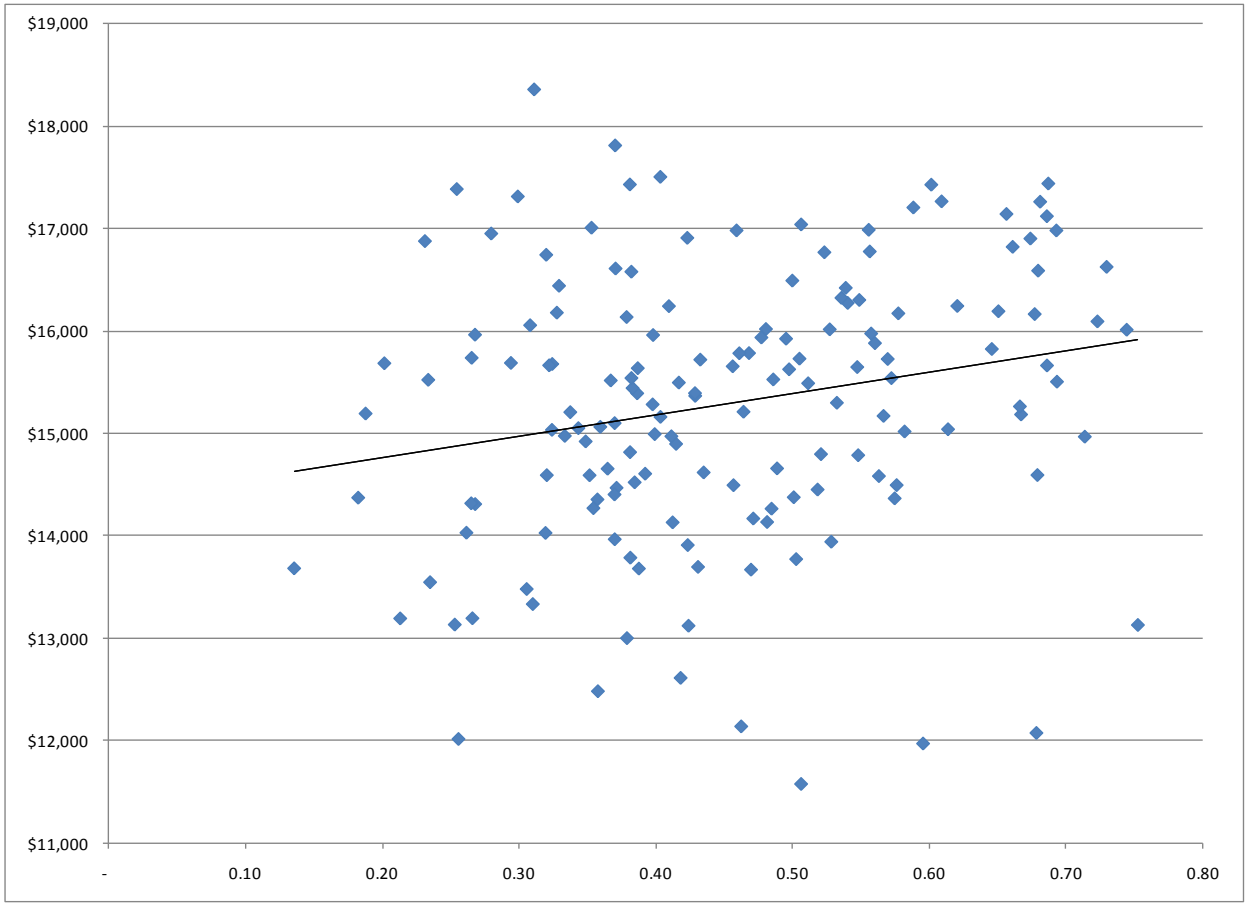


Exhibit 9. Hypothetical example of cost per Beneficiary and quality scores for modified sample (negative correlation)

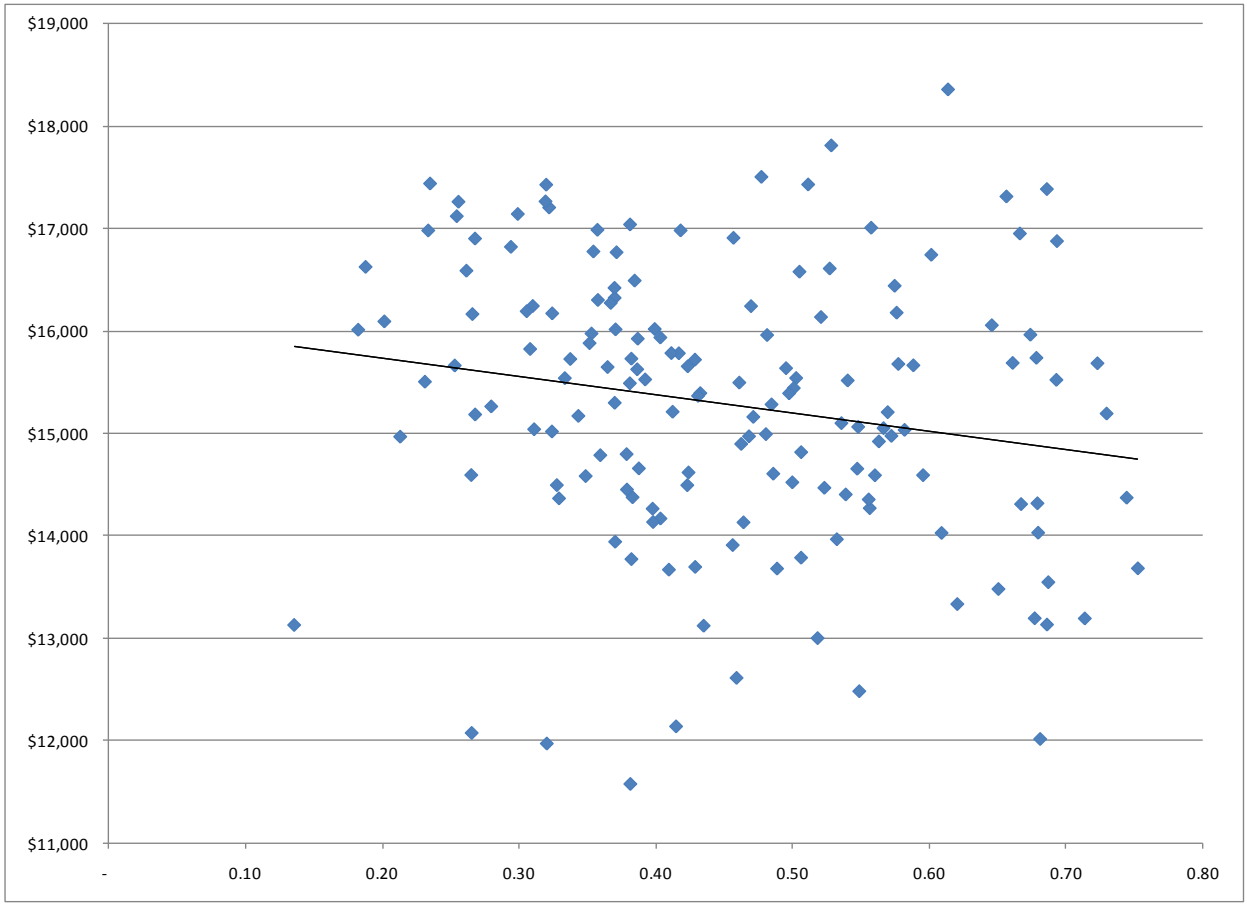


Exhibit 10. Illustrative efficiency value system

	Higher Eff	↑	10	**	**	***	***	****	****	*****	*****	*****	*****
			9	*	**	***	***	****	****	****	*****	*****	*****
			8	*	**	**	***	***	****	****	*****	*****	*****
			7	*	**	**	**	***	***	****	****	*****	*****
Efficiency			6	*	*	**	**	***	***	****	****	*****	*****
Score			5	*	*	**	**	***	***	****	****	*****	*****
			4	*	*	**	**	***	***	****	****	****	*****
			3	*	*	*	**	**	***	***	****	****	*****
			2	*	*	*	**	**	**	***	***	****	*****
			1	*	*	*	*	**	**	***	***	****	****
				1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
				→									
				Higher Quality									
				Total Quality Score									

Legend

***** = **1 Star Efficiency Rating (Lowest Rating)**

****** = **2 Star Efficiency Rating**

******* = **3 Star Efficiency Rating**

******** = **4 Star Efficiency Rating**

********* = **5 Star Efficiency Rating (Highest Rating)**

Exhibit 11. Adjusting the total quality score for efficiency [xiii]

	10	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+2 pts	+2 pts	+2 pts	+2 pts
	9	-2 pts	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+ 1 pt	+2 pts	+2 pts	+2 pts
	8	-2 pts	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+2 pts	+2 pts	+2 pts
	7	-2 pts	-1 pt	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+2 pts	+2 pts
Efficiency	**6**	-2 pts	-2 pts	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+2 pts	+2 pts
Score	**5**	-2 pts	-2 pts	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+2 pts	+2 pts
	4	-2 pts	-2 pts	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+ 1 pt	+2 pts
	3	-2 pts	-2 pts	-2 pts	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt	+2 pts
	2	-2 pts	-2 pts	-2 pts	-1 pt	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+2 pts
	1	-2 pts	-2 pts	-2 pts	-2 pts	-1 pt	-1 pt	0 pts	0 pts	+ 1 pt	+ 1 pt
		1-10	**11-20**	**21-30**	**31-40**	**41-50**	**51-60**	**61-70**	**71-80**	**81-90**	**91-100**
		Total Quality Score									

Summary Table

*** = 1 Star Efficiency Rating (Lowest Rating) = Total Quality Score – 2 pts**

**** = 2 Star Efficiency Rating = Quality Points – 1 pt**

***** = 3 Star Efficiency Rating = Quality Points + 0 pts**

****** = 4 Star Efficiency Rating = Quality Points + 1 pt**

******* = 5 Star Efficiency Rating (Highest Rating) = Total Quality Score + 2 pts**

[xiii] It is also possible to construct a function that approximates the point assignments in Exhibit 11 (and by extension the star assignment in Exhibit 10) rather than make individual determinations of the point values that should be assigned to each cell. For example, the grid in Exhibit 10 represents a function where quality is weighted twice as much as efficiency in determining the point totals. The function is of the form: a*quality score + b*efficiency score + k ,with a max/min of +/- 2. If the max/min of +/-2 is removed, the highest point adjustments would be +/-3 rather than +/- 2.

Exhibit 12. Distinguishing use cases for linking quality and cost measures

Aspect of combining quality and cost measures	Internal efficiency improvement	Public Reporting	Pay-for-performance (value-based purchasing)	Network design
Primary audience	• Providers	• Patients • Providers • Policymakers	• Providers	• Patients
Mechanism for value improvement	• Improve care within the provider organization	• Patient volume shifts to higher value providers • Providers are motivated to improve performance • Policymakers monitor performance gaps and improvement	• Financial incentives for providers to improve efficiency and value	• Patients select providers according to their out-of-pocket costs • Provider's network status affects patients' out-of-pocket costs • Provider performance affects network status
Primary comparisons	• Same provider over time • May be benchmarked against similar providers	• For patients: between accessible providers • For providers: similar providers in similar markets.	• Between providers that offer similar services; within-network or in similar markets	• Between peer providers in a network
Most relevant features of quality measurement	• Measures for which specific actions can improve care (process of care)	• Patient clinical outcomes, and patient experience • Intuitive process of care measures (e.g., hand-washing)	• First generation programs: underuse; screening; process of care • Second generation: clinical quality, patient outcomes, patient experience	• Typical: structural (e.g., accreditation, reported to standards body, claims-based process of care) • Aspirational: clinical quality, patient outcomes, patient experience

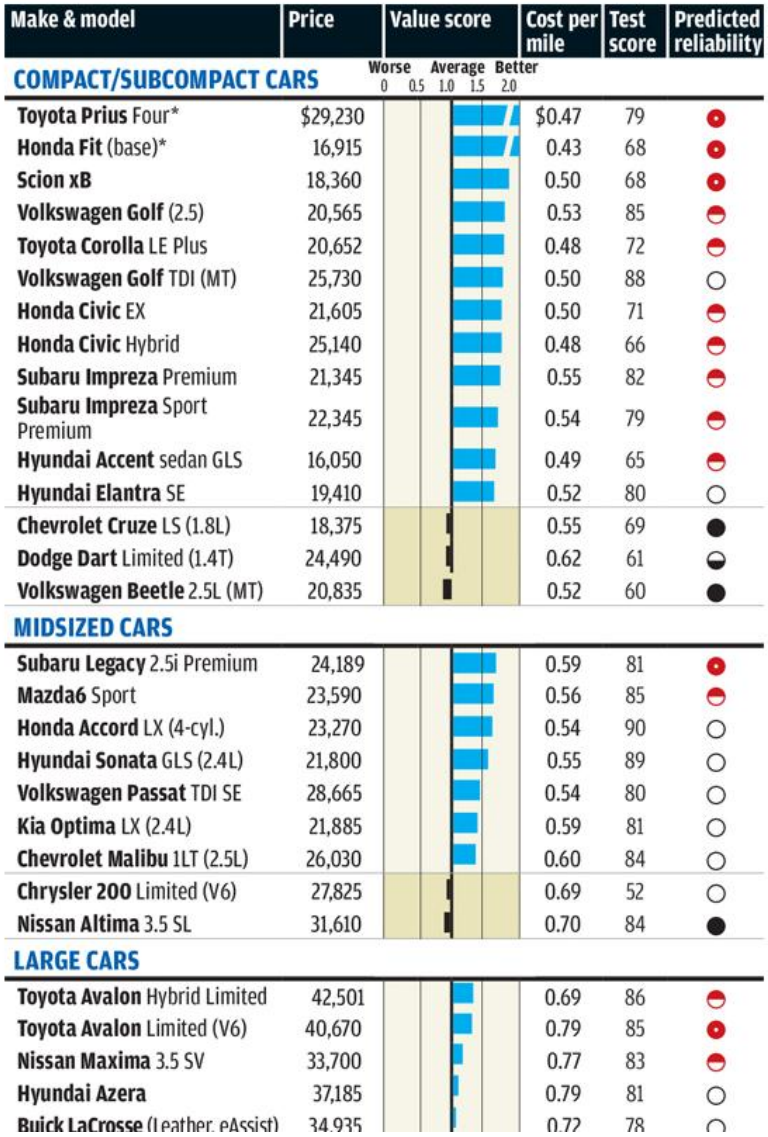
Aspect of combining quality and cost measures	Internal efficiency improvement	Public Reporting	Pay-for-performance (value-based purchasing)	Network design
Most relevant features of cost measurement	• Production costs • Profit margins • "Waste" (e.g., delayed diagnosis or delayed discharge) • Inefficient substitution (e.g., overuse of high cost drugs)	• Patients' financial exposure (out-of-pocket spending) • Actual (unstandardized) prices • Economic practice style (provider cost-consciousness, awareness) • Therapeutic versus complication costs	• Cost to the payer that are tied to a discrete clinical event (episode) • Cost profiling based on actual (unstandardized) prices; most likely in private programs with negotiated prices • Utilization profiling based on average or adjusted (standardized) prices; most likely in public programs with fixed (regulated) prices	• Cost to the payer that are tied to a discrete clinical event (episode) • Unstandardized prices reflect the payer's financial exposure
End product of efficiency or value measurement	• Aggregated performance by roles and responsibilities (service line)	• Easily understood aggregated measures (e.g., grades or star ratings) for efficiency, quality, and cost	• Financial adjustments for providers (bonuses or penalties)	• Tiers for providers (ranging from high efficiency (preferred tier) to low efficiency non-preferred tier)
Downside risk to providers	• Missed opportunities; waste in production process	• Lost referrals or members	• Net revenue reduction	• Lost volume and revenues

Appendix A: Applications outside of health care that combine indicators of quality and cost

There are a number of efforts outside of health care that link measures of products' cost and quality to measure "value." These include efforts by Consumer Reports to measure the value of automobiles and US News and Weekly Report's rankings of the value of colleges. Consumer Reports' 2014 rankings of the "Best and Worst Cars for the Money" assigns cars a "value-score" based on the combination of a "road test score" (i.e., quality), a composite rated on a 100 point scale, a predicted reliability score (assessing how well the car is likely to hold up given the reliability of recent models), and 5-year ownership costs. This value score is then displayed alongside the car's price and the component factors comprising the value score (ownership costs, road test score, and predicted reliability). The rankings also use a fudge factor to "not recommend" cars as a result of poor reliability. They are able to do this because the ratings are concerned primarily with identifying highly recommended cars: it is therefore not concerned about precise measurements of poor value, and instead focuses on precisely measuring the highest value cars. Value rankings are then displayed within classes of cars (e.g. compact cars, midsized cars, luxury cars, small sport utility vehicles, midsized sport utility vehicles, etc.). See **Appendix Exhibit 1** for how these data are displayed. The model used to calculate value rankings is not specified. Consumer Reports' automobile value ratings have a close analogue to efficiency-tiering in health care. These programs are frequently designed to direct patients towards higher efficiency providers (through a "designation" program) rather than directing patients away from lower efficiency providers and report efficiency for different "classes" of physicians (i.e., different specialties).

US News and World Reports compiles a list of the highest value colleges and universities in their "Best Value Schools."[85] To do this, they first assess school quality as a composite of the student selectivity, college graduation rates, assessment from peer institutions, faculty resources (i.e., class size), financial resources, and alumni giving. They then create a value score as the weighted combination of three factors: 1) 60% is for the ratio of quality to cost (including tuition, room and board, books, and other expenses), discounted according to the average need-based scholarship; 2) an unspecified percentage is based on the percentage of students who received need-based aid; 3) an unspecified percentage is based the percentage of total costs that are discounted. Similar to Consumer Reports' automobile rankings, US News ranks the value of schools within different categories (e.g. national universities, regional universities, national liberal arts colleges, and regional liberal arts colleges). When showing the rankings of schools based on value, US News and World Reports displays some of the component parts side-by-side (percentage of students receiving need-based grants and the average cost after receiving grants based on need), but not all of the parameters going into the value calculation (see **Appendix Exhibit 2**). This model for combining quality and cost to measure value is similar to the unconditional model with a large weight towards cost.

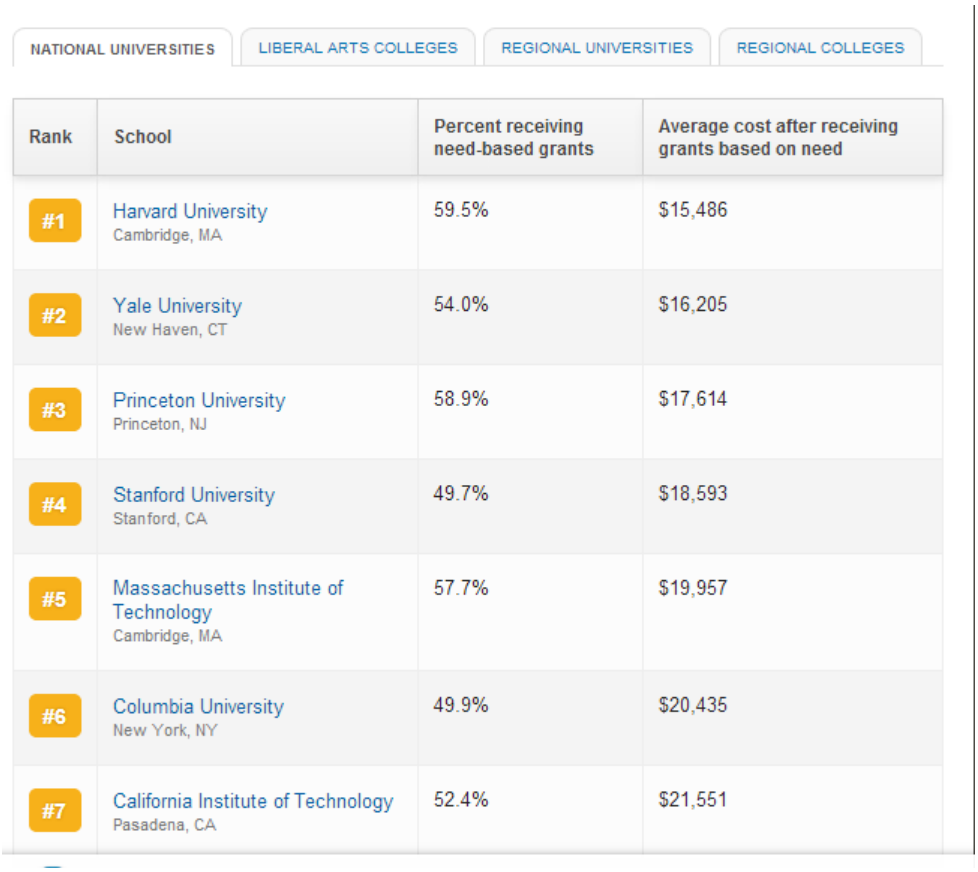
Appendix A Exhibit 1. Screen shot from Consumer Reports 2014 "Best and Worst Cars for the Money"



Make & model	Price	Value score (Worse 0 0.5 Average 1.0 1.5 Better 2.0)	Cost per mile	Test score	Predicted reliability
COMPACT/SUBCOMPACT CARS					
Toyota Prius Four*	$29,230		$0.47	79	
Honda Fit (base)*	16,915		0.43	68	
Scion xB	18,360		0.50	68	
Volkswagen Golf (2.5)	20,565		0.53	85	
Toyota Corolla LE Plus	20,652		0.48	72	
Volkswagen Golf TDI (MT)	25,730		0.50	88	
Honda Civic EX	21,605		0.50	71	
Honda Civic Hybrid	25,140		0.48	66	
Subaru Impreza Premium	21,345		0.55	82	
Subaru Impreza Sport Premium	22,345		0.54	79	
Hyundai Accent sedan GLS	16,050		0.49	65	
Hyundai Elantra SE	19,410		0.52	80	
Chevrolet Cruze LS (1.8L)	18,375		0.55	69	
Dodge Dart Limited (1.4T)	24,490		0.62	61	
Volkswagen Beetle 2.5L (MT)	20,835		0.52	60	
MIDSIZED CARS					
Subaru Legacy 2.5i Premium	24,189		0.59	81	
Mazda6 Sport	23,590		0.56	85	
Honda Accord LX (4-cyl.)	23,270		0.54	90	
Hyundai Sonata GLS (2.4L)	21,800		0.55	89	
Volkswagen Passat TDI SE	28,665		0.54	80	
Kia Optima LX (2.4L)	21,885		0.59	81	
Chevrolet Malibu 1LT (2.5L)	26,030		0.60	84	
Chrysler 200 Limited (V6)	27,825		0.69	52	
Nissan Altima 3.5 SL	31,610		0.70	84	
LARGE CARS					
Toyota Avalon Hybrid Limited	42,501		0.69	86	
Toyota Avalon Limited (V6)	40,670		0.79	85	
Nissan Maxima 3.5 SV	33,700		0.77	83	
Hyundai Azera	37,185		0.79	81	
Buick LaCrosse (Leather, eAssist)	34,935		0.72	78	

Source: http://www.consumerreports.org/cro/2012/05/best-new-car-values/index.htm

Appendix A Exhibit 2. Screen shot from US News and World Reports 2014 "Best Value Schools"



NATIONAL UNIVERSITIES | LIBERAL ARTS COLLEGES | REGIONAL UNIVERSITIES | REGIONAL COLLEGES

Rank	School	Percent receiving need-based grants	Average cost after receiving grants based on need
#1	Harvard University Cambridge, MA	59.5%	$15,486
#2	Yale University New Haven, CT	54.0%	$16,205
#3	Princeton University Princeton, NJ	58.9%	$17,614
#4	Stanford University Stanford, CA	49.7%	$18,593
#5	Massachusetts Institute of Technology Cambridge, MA	57.7%	$19,957
#6	Columbia University New York, NY	49.9%	$20,435
#7	California Institute of Technology Pasadena, CA	52.4%	$21,551

Source: http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-universities/best-value/spp+50

Appendix B. NQF Linking Cost and Quality Expert Panel Roster

Joyce DuBow (Co-Chair)
AARP, Washington, DC

Carole Flamm, MD, MPH (Co-Chair)
Blue Cross and Blue Shield Association, Chicago, IL

Peter Almenoff, MD FCCP
Department of Veterans Affairs, Kansas City, KS

Steven Asch, MD, MPH
Center for Innovation to Implementation, Palo Alto VA, Palo Alto, CA

Larry Becker
Xerox Corporation, Rochester, NY

David Cohen, MD, MSc
Saint Luke's Mid America Heart Institute, Kansas City, MO

Mary Cramer, MBA, CPHQ
Massachusetts General Hospital / Massachusetts General Physician Organization, Boston, MA

Christine Goeschel, ScD, MPA, MPS, RN, FAAN
MedStar Health, Columbia, MD

Donald Likosky, PhD
University of Michigan, Ann Arbor, MI

Timothy Lowe, PhD, MSW
Premier, Inc., Charlotte, NC

Catherine MacLean, MD, PhD
University of California Los Angeles, David Geffen School of Medicine, Los Angeles, CA

Jack Needleman, PhD
University of California Los Angeles, Fielding School of Public Health, Los Angeles, CA

Steven Pantilat, MD, FAAHPM
UCSF Palliative Care Leadership Center, San Franciso, CA

Kimberly Rask, MD, PhD
Alliant Health Solutions, Atlanta, GA

Iyah Romm
Massachusetts Health Policy Commission, Boston, MA

Matthew Rousculp, PhD, MPH
GlaxoSmithKline, Research Triangle Park, NC

Dennis Scanlon, PhD
The Pennsylvania State University, State College, PA

Jeremiah Schuur, MD, MHS, FACEP
Brigham and Women's Hospital, Boston, MA

Jeffrey Silber, MD, PhD
Children's Hospital of Philadelphia, Philadelphia, PA

Alan Speir, MD
Inova Health System, Falls Church, VA

Joe Stephansky
Michigan Health & Hospital Association, Okemos, MI

Herbert Wong, PhD
Agency for Healthcare Research and Quality, Rockville, MD

Gregory Wozniak, PhD
American Medical Association, Chicago, IL

Gary Young, JD, PhD
Northeastern University, Boston, MA

References

1. Crawford S. Linking cost and quality expert panel orientation. Paper presented at: NQF 2014.
2. Cutler DM. Where are the health care entrepreneurs? The failure of organizational innovation in health care. In: Lerner J, Stern S, eds. *Innovation Policy and the Economy.* Vol 11. Chicago: University of Chicago Press; 2011:1-28.
3. The watch on the job. *New England Journal of Medicine.* 1969;281(14):792-793.
4. Thornton TN, Leonard RC. Experimental comparison of effectiveness and efficiency of three nursing approaches. *Nursing Research.* 1964;13:122-125.
5. Hussey PS, de Vries H, Romley J, et al. A systematic review of health care efficiency measures. *Health Services Research.* 2009;44(3):784-805.
6. National Quality Measures Clearinghouse. Varieties of measures in NQMC. *Tutorials on Quality Measures* http://www.qualitymeasures.ahrq.gov/tutorial/varieties.aspx. Accessed 04/14/2014.
7. Korenstein D, Falk R, Howell E, Bishop T, Keyhani S. Overuse of health care services in the United States: An understudied problem. *Archives of Internal Medicine.* 2012;172(2):171-178.
8. Burke LA, Ryan AM. The complex relationship between cost and quality in US health care. *The Virtual Mentor.* 2014;16(2):124-130.
9. Doyle JJ, Graves JA, Gruber J, Kleiner S. Do high-cost hospitals deliver better care? Evidence from ambulance referral patterns. www.nber.org March 2012.
10. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Annals of Internal Medicine.* 2003;138(4):273-287.
11. Silber JH, Kaestner R, Even-Shoshan O, Wang Y, Bressler LJ. Aggressive treatment style and surgical outcomes. *Health Services Research.* 2010;45(6 Pt 2):1872-1892.
12. Kaestner R, Silber J. Evidence on the Efficacy of Inpatient Spending on Medicare Patients. *Millbank Quarterly*. 2010; 88(4):560-594.
13. Gao J, Moran E, Almenoff PL, Render ML, Campbell J, Jha AK. Variations in efficiency and the relationship to quality of care in the veterans health system. *Health Affairs (Millwood).* 2011;30(4):655-663.
14. NQF. *Measurement framework: Evaluating efficiency across patient-focused episodes of care.* Washington D.C.: NQF;2009.
15. Institute of Medicine. *Crossing the quality chasm: The IOM health care quality initiative.* Washington, D.C.: Institute of Medicine;2001.
16. Casalino LP, Nicholson S, Gans DN, Hammons T, Morra D, Karrison T, Levinson W. What does it cost physician practices to interact with health insurance plans? *Health Affairs (Millwood).* 2009; 28(4):w533-543.
17. Sood N, Ghosh A, Escarse J. *The effect of health care cost growth on the U.S. economy*. United States Department of Health, 2007.
18. Leatherman S, Berwick D, Iles D, Lewin L, Davidoff F, Nolan T, Bisognano M. The business case for quality: case studies and an analysis. *Health Affairs (Millwood).* 2003;22(2):17-30.
19. Ryan AM. Effects of the Premier Hospital Quality Incentive Demonstration on Medicare patient mortality and cost. *Health Services Research.* 2009;44(3):821-842.
20. Robinson JC. The end of managed care. *Journal of the American Medical Association.* 2001;285(20):2622-2628.
21. VanLare JM, Blum JD, Conway PH. Linking performance with payment: implementing the Physician Value-Based Payment Modifier. *Journal of the American Medical Association.* 2012;308(20):2089-2090.

22. Ryan A, Blustein J. Making the best of hospital pay for performance. *The New England Journal of Medicine.* 2012;366(17):1557-1559.
23. Ryan AM, Damberg CL. What can the past of pay-for-performance tell us about the future of Value-Based Purchasing in Medicare? *Healthcare.* 2013;1(1-2):42-49.
24. Fisher ES, McClellan MB, Safran DG. Building the path to accountable care. *New England Journal of Medicine.* 2011;365(26):2445-2447.
25. *Variation in Health Care Spending: Target Decision Making, Not Geography.* The National Academies Press; 2013.
26. Robinson JC. Hospital tiers in health insurance: Balancing consumer choice with financial incentives. *Health Affairs (Millwood).* 2003;Suppl Web Exclusives:W3-135-146.
27. Obamacarefacts.com. Types of Health Insurance Plans. 2014; http://obamacarefacts.com/insurance-exchange/health-insurance-plans.php Accessed April 14, 2014, 2014.
28. Wall JK. Amid Obamacare's changes, WellPoint keeps old playbook. *The Dose.* July 25, 2013 www.ibj.com: Indianapolis Business Journal; 2013.
29. Robinson JC, Brown TT. Increases in consumer cost sharing redirect patient volumes and reduce hospital prices for orthopedic surgery. *Health Affairs (Millwood).* 2013;32(8):1392-1397.
30. Sennett C, Starkey K. *Measuring and improving efficiency in Health Care: Report from an ABIM foundation/IOM meeting.* Philadelphia: ABIM Foundation;2006.
31. Krumholz HM, Keenan PS, Brush JE, Jr., et al. Standards for measures used for public reporting of efficiency in health care: A scientific statement from the American Heart Association Interdisciplinary Council on Quality of Care and Outcomes Research and the American College of Cardiology Foundation. *Circulation.* 2008;118(18):1885-1893.
32. Timbie JW, Normand SL. A comparison of methods for combining quality and efficiency performance measures: Profiling the value of hospital care following acute myocardial infarction. *Statistics in Medicine.* 2008;27(9):1351-1370.
33. Tompkins CP, Higgins AR, Ritter GA. Measuring outcomes and efficiency in Medicare Value-Based Purchasing. *Health Affairs (Millwood).* 2009;28(2):w251-261.
34. Song Z, Safran DG, Landon BE, He Y, Mechanic RE, Day MP, Chernew ME. The 'Alternative Quality Contract,' based on a global budget, lowered medical spending and improved quality-Supplemental Materials. *Health Affairs (Millwood).* 2012;31(8):1885-1894.
35. Kessler D, McClellan M. *The effects of hospital ownership on medical productivity.* Cambridge, MA: National Bureau of Economic Research;2001.
36. Hollingsworth B. The measurement of efficiency and productivity of health care delivery. *Health Economics.* 2008;17(10):1107-1128.
37. Rosko MD. Measuring technical efficiency in health care organizations. *Journal of Medical Systems.* 1990;14(5):307-322.
38. Meltzer DO, Chung JW. The population value of quality indicator reporting: A framework for prioritizing health care performance measures. *Health Affairs (Millwood).* 2014;33(1):132-139.
39. Berenson RA. *Moving payment from volume to value: What role for performance measurement?* Urban Institute;2010.
40. Massachusetts Division of Health Care Finance and Policy. *Massachusetts health care cost trends: Price variation in health care services.* Boston, MA 2011.
41. The Leapfrog Group. Survival predictor (surgical mortality). Washington D.C.: Academy Health; 2011.
42. Krumholz HM, Wang Y, Mattera JA, Wang Y, Han LF, Ingber MJ, Roman S, Normand SL. An administrative claims model suitable for profiling hospital performance based on 30-day

mortality rates among patients with an acute myocardial infarction. *Circulation.* 2006;113(13):1683-1692.
43. Burack JH, Impellizzeri P, Homel P, Cunningham JN. Public reporting of surgical mortality: A survey of New York State cardiothoracic surgeons. *Annals of Thoracic Surgery.* 1999;68(4):1195-1200.
44. Silber JH, Rosenbaum PR, Ross RN, Ludwig JM, Wang W, Niknam BA, Mukherjee N, Saynisch PA, Even-Shoshan O, Kelz RR, Fleisher LA. Template matching for auditing hospital cost and quality. *Health Services Research.* 2014;epub ahead of print.
45. Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry: 2012 Annual Data Report. Besthesda, MD: 2013. http://www.cff.org/livingwithcf/qualityimprovement/patientregistryreport/ Accessed 10/29/2014.
46. Virginia Cardiac Surgery Quality Initiative. 2014; http://www.vcsqi.org. Accessed 10/29/2014.
47. Centers for Medicare and Medicaid Services. How to Obtain the 2012 QRUR and the 2012 IEP PQRS Performance Report. Baltimore, MD: 2014. http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/ReportTemplate.html Accessed 10/29/2014.
48. Birkmeyer JD, Kerr EA, Dimick JB. Appendix F Commissioned Paper: Improving the Quality of Quality Measurement. In: Institute of Medicine of the National Academies. *Performance Measurement: Accelerating Improvement*. Washington, D.C. 2014.
49. Mehrotra A, Hussey PS, Milstein A, Hibbard JH. Comsumers' and providers' responses to public cost reports, and how to raise the likelihood of achieving desired results. *Health Affairs (Millwood)*. 2012; 31(4):843-51.
50. Epstein AJ. Do cardiac surgery report cards reduce mortality? Assessing the evidence. *Medical Care Research and Review*. 2006; 63(4):403-26.
51. Kolstad JT, Chernew ME. Quality and consumer decision making in the market for health insurance and health care services. *Medical Care Research and Review*. 2009; 66(1)suppl 28S-52S.
52. Hibbard J, Sofaer S. *Best practices in public reporting No. 1: How to effectively present health care performance data to consumers.* Rockville, MD: Agency for Healthcare Research and Quality;May 2010.
53. Porter ME. What is value in health care? *New England Journal of Medicine*. 2010; 363:2477-81.
54. Deshpande PR, Rajan S, Sudeepthi BL, Nazir CPA. Patient-reported outcomes : A new era in clinical research. *Perspectives in Clinical Research*. 2011; 2(4): 137-44.
55. Manary MP, Boulding W, Staelin R, Glickman SW. The patient experience and health outcomes. *The New England Journal of Medicine*. 2013; 368: 201-3.
56. Rosenthal MB, Fernandopulle R, Song HSR, Landon B. Paying for quality: Providers' incentives for quality improvement. *Health Affairs (Millwood)*. 2004; 23(2): 127-41.
57. Doran T, Kontopantelis E, Reeves D, Sutton M, Ryan A. Setting performance targets in pay for performance programmes : What can we learn from QOF ? *BMJ*. 2014; 348: g1595.
58. King AL. *'Physician Transparency' movement advances thanks to New York Attorney General.* ABA Health eSource; 2008; 4(6). http://www.americanbar.org/newsletter/publications/aba_health_esource_home/King.html#_ftn2 Accessed 10/29/2014.
59. Sinaiko AD, Rosenthal MB. The impact of tiered physician networks on patient choices. *Health Services Research*. 2014: 49(4): 1348-63.
60. NQF. *Composite performance measure evaluation guidance.* Washington, DC: NQF;2013.
61. Aetna. *Aexcel performance network designation measurement methodology.* Aetna;2014.

62. Blue Cross Blue Shield Association. Blue Distinction Centers: An overview. http://www.bcbs.com/why-bcbs/blue-distinction/. Accessed 04/24/2014.
63. Blue Cross Blue Shield Association. Blue Distinction Specialty Care Program: 2013 Program selection criteria for Blue Distinction Centers for cardiac care. http://www.bcbs.com/healthcare-partners/blue-distinction-for-providers/cardiacprogramcriteria.pdf. Accessed 04/24/2014.
64. Blue Cross Blue Shield Association. Blue Distinction Specialty Care Program: 2011-2012 Program selection criteria for Blue Distinction Centers for spine surgery and Blue Distinction Centers for knee and hip replacement. http://www.bcbs.com/healthcare-partners/blue-distinction-for-providers/BDC_SpineKneeHip_Selection_Criteria.pdf. Accessed 04/24/2014.
65. Bailit M, Hughes C, Burns M, Freedman DH. *Share-savings payment arrangements in health care: Six case studies.* New York, NY: The Commonwealth Fund;2012.
66. Blue Cross Blue Shield of Michigan. *2014 Hospital pay-for-performance program: Peer groups 1-4.* Michigan2014.
67. *Blue Cross Blue Shield of Massachusetts: The Alternative Quality Contract.* Blue Cross Blue Shield of Massachusettes; May 2010 2010.
68. Chernew ME, Mechanic RE, Landon BE, Safran DG. Private-payer innovation in Massachusetts: The 'Alternative Quality Contract'. *Health Affairs (Millwood).* 2011;30(1):51-61.
69. Song Z, Safran DG, Landon BE, He Y, Mechanic RE, Day MP, Chernew ME. The 'Alternative Quality Contract,' based on a global budget, lowered medical spending and improved quality-Supplemental Materials. *Health Affairs (Millwood).* 2012;31(8):1885-1894.
70. Christianson JB, Feldman R. Evolution in the Buyers Health Care Action Group purchasing initiative. *Health Affairs (Millwood).* 2002;21(1):76-88.
71. Lyles A, Weiner JP, Shore AD, Christianson J, Solberg LI, Drury P. Cost and quality trends in direct contracting arrangements. *Health Affairs (Millwood).* 2002;21(1):89-102.
72. Cigna. *Cigna care designation and physician quality and cost-efficiency displays: 2013 Methodologies Whitepaper.* Cigna;2013.
73. Cigna. *Collaborative accountable care: CIGNA's approach to Accountable Care Organizations.* Cigna;2011.
74. HealthPartners. Cost and quality ratings:Medical group and hospital ratings. 2014; https://www.healthpartners.com/public/cost-and-quality/index.html. Accessed April 14, 2014.
75. The Agency for Healthcare Research and Quality, The Employer Health Care Alliance Cooperative. Efficiency in health care: What does it mean? How is it measured? How can it be used for Value-Based Purchasing? 2006; Madison, WI.
76. Carroll J. Early Tiered networks encounter many obstacles. *Managed Care*: MediMedia Managed Markets; 2007.
77. Alteras T, Silow-Carroll S. *Value-driven health care purchasing: Case study of the Massachusetts Group Insurance Commission.* The Commonwealth Fund;2007.
78. RTI International. *Accountable care organization 2013 program analysis: Quality performance standards narrative measure specifications.* Baltimore, MD: Centers for Medicare & Medicaid Services;2012.
79. NCQA. Relative Resource Use. 2014; http://www.ncqa.org/tabid/1231/default.aspx. Accessed 04/14/2014, 2014.
80. *Tufts Health Plan Blue Ribbon Value Methodology- Summary.* Tufts Health Plan;2006.
81. UnitedHealthCare. *UnitedHealth Premium Physician Designation Program: Summary methodology.* UnitedHealthCare;2014.
82. UnitedHealthCare. UnitedHealth Premium Physician Designation Program: Detailed Methodology. 2014.

83. United Healthcare Online. Premium Methodology. https://www.unitedhealthcareonline.com/b2c/CmaAction.do?channelId=45dff1ab39f24210VgnVCM100000b640dd0a____ . Accessed 04/24/2014.
84. Speir AM, Kasirajan V, Barnett SD, Fonner E. Additive costs of postoperative complications for isolated coronary artery bypass grafting patients in Virginia. *The Annals of Thoracic Surgery.* 2009;88(1):40-46.
85. Morse R. Best value schools methodology. *US News and World Report*2013. http://www.usnews.com/education/best-colleges/articles/2013/09/09/best-value-schools-methodology

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

ISBN 978-1-933875-79-8