

Measuring and Improving Healthcare Value

Drew Guerra

Key Takeaways:

1. Value is often defined as a composite of many components, including the appropriateness, quality, cost, and equity of care.
2. There is significant disagreement regarding how best to measure value. However, numerous public and private actors have established sets of value metrics across care settings. The “best” metrics are typically evidence-based and built upon large enough sample sizes to allow for statistically significant comparisons.
3. Commonly cited challenges in measuring value include metric selection, data collection and reporting, benchmarking and risk adjustment, patient attribution, and restrictions on provider autonomy.
4. Insurers, employers, and other purchasers have attempted to incentivize and improve value through a variety of levers, including alternative payment models, narrow networks, centers of excellence, employer-based clinics, and behavioral interventions. There is an ongoing debate as to the efficacy of several of these levers.
5. Key themes across each lever include identifying variations in value across providers and health systems, directing higher volumes to high-value care, cutting out low-value care, and incentivizing and enabling providers to improve the value of care they deliver.

Defining Value

Before we can attempt to improve value, we must define it. Many argue this is not straightforward – value is inherently subjective, as any definition must prioritize potential components of value to arrive at a composite definition, which will differ across stakeholders. Commonly cited components of value include the appropriateness, quality, and cost of care. More recently, equity has been raised as a value component. Notably, some of these categories are often combined to arrive at integrated components of value, such as cost-effectiveness. Disagreements primarily arise over differences in perspective, priorities, and power dynamics. First, different stakeholders’ definitions of value are informed by their organizational perspective and personal priorities. Additionally, which stakeholders *should* define value and which stakeholders ultimately have the final say often differ, and the latter is heavily influenced by capital allocations, information asymmetries, market power, and regulations.

Despite stakeholders' efforts to arrive at definitions of value that incorporate diverse viewpoints, there are still significant differences of opinion, which influence how value is measured and how those measurements are used to incentive higher-value care.

Measuring Value

Defining value is of limited value if it is not measurable. For example, we might agree that value is best defined as the cost-effectiveness of any given treatment – but if we cannot measure the intervention's cost or effectiveness, we have no means of evaluation or comparison.

Measurement introduces a new set of challenges. Specifically, we must consider what metrics to use, how to capture required data, and whether selected metrics are truly indicative of value. Proponents of measurement argue that these hurdles can be overcome, but opponents point to flaws in existing metrics, ongoing data, attribution, and benchmarking challenges, and restrictions on provider autonomy as evidence that measuring value is difficult, if not impossible.

Value Metrics Selection Criteria

Notably, nearly all stakeholders agree that the best metrics follow several key guidelines, including:

- Evidence-based and connected along the care pathway to desired outcomes and/or cost reductions
- Large enough samples that allow for valid comparisons
- Risk-adjustable, where possible, to account for variation across patient populations

Metrics may also be differentiated as measuring process, structure, or outcomes:

- Outcome metrics are direct, but can be difficult to evaluate given sample size limitations
- Process metrics are often used as proxies for outcomes metrics and allow for closer to real-time responsiveness, but not all are meaningfully linked to positive outcomes
- Structural metrics are also used as proxies and are considered longer lasting but slower to respond than process metrics. They should also be linked to desired outcomes

Common Value Metrics & Panels

- Medicare's Merit-Based Incentive Payment System Quality Measures (MIPS)
- Medicaid Adult and Child Core Set
- Healthcare Effectiveness Data and Information Set (HEDIS)
- Consumer Assessment of Healthcare Providers and Systems (CAHPS)
- Core Quality Measures Collaborative
- ORYX (Joint Commission)

Finally, different components of value vary in difficulty of measurement. Cost (or, at least price) is easy. So is patient satisfaction: just ask. Equity is not particular hard in theory, but we lack the data in many areas. Finally, quality is often hardest to measure accurately and meaningfully.

Data Collection, Reporting, and Standardization

Although data collection, reporting, and standards are improving, there are significant technical challenges to data collection. Primary difficulties include¹:

- Lack of standardization across EMR / EHR systems
- Prevalence of small and incomplete data sets
- Potential competitive disincentives to share data with other healthcare stakeholders

- Challenges with data aggregation across different stakeholders

Additionally, there are significant data collection costs. First, collection and reporting (and the estimated 785 hours per physician they require) contribute to physician burnout. Additionally, there is a significant monetary cost – an estimated \$15.4 billion in 2016.ⁱⁱ These costs are exacerbated by the myriad metrics that providers must report (more than 1,700 alone for Medicare), and limited overlap across different metrics panels (estimated to be as little as 20%).ⁱⁱⁱ

Attribution and Benchmarking

Critics of existing measurement techniques contend that attribution at the individual provider level is both technically difficult and potentially misleading. For example, should a PCP who has recently begun seeing a patient with numerous monthly specialty appointments be the patient's attributed provider? Typically, the PCP has little influence over the value delivered by patients' other care providers. Other methods include attributing patients to all of their providers or the provider that bills for the largest percentage of claims, but both also have evident limitations. Additionally, critics point to the difficulties inherent in benchmarking provider performance. Existing measurement systems often use both prior provider data and risk adjustment to account for differences between provider panels, but both have limitations. First, providers may 'sandbag' their initial performance to demonstrate improvements more easily on future assessments. Additionally, risk adjustment methods may be insufficiently nuanced, easily 'gamed', and/or not accurate enough to account for variations in provider panels. This can lead to significant variation in revenue that is not reflective of value or performance.^{iv}

However, proponents of measuring provider value and performance point to improvements in risk adjustment models and learnings from early efforts to implement benchmarks. Although no attribution system or benchmarking method is perfect, they contend that long-run improvements have and will continue to expand opportunities to measure and incentivize value.

Provider Autonomy

Finally, many providers express frustration that an overreliance on metrics reduces providers' autonomy to diagnose their patients' needs and vary treatment methods based on individual differences.^v In fact, metrics may distort providers' behaviors away from optimal care delivery. In support of this view, Gupta et al. argue that evidence-based medicine should incorporate physicians' clinical expertise and patient values, in addition to clinical best practices.^{vi}

Influencing Value as a Purchaser

Finally, if we can sufficiently overcome the challenges inherent in defining and measuring value, we are left with a final obstacle. How do we *improve* value?

Insurers and employers have historically played active roles in attempting to influence and improve healthcare value. Beginning in the 1990s and 2000s, government and commercial insurers began to experiment with alternative payment models in an effort to shift incentives from volume to value. Even before that, purchasers started to use preferred and narrow networks to encourage their members to see providers perceived as providing higher-value care. In recent years, purchasers have become increasingly creative in identifying and deploying levers to improve the value of care provided to their enrollees and employees. This section will explore these efforts, their outcomes, and related academic literature.

Alternative Payment Models

Both public and commercial payers have experimented widely with alternative payment models (APMs). These models are typically structured to shift risk (and potential financial reward) from the purchaser to the provider, to incentivize care that is more cost-effective and higher quality. APMs are often considered on a spectrum, from “entry-level” pay-for-reporting and pay-for-performance models that tie financial incentives to data reporting and/or quality metrics to fully capitated models that typically pay a risk-adjusted per member, per month amount. One-sided and two-sided risk models, bundled payments, and global budgets are also well-trod APMs, although the APMs that shift the most risk to providers (typically global budgets and capitated models) are generally the least used to date. As APMs have been used extensively, there is significant literature studying their impact.

Proponents of APMs point to several studies that document substantial cost savings without declines in quality, including evaluations of BCBS Massachusetts’ Alternative Quality Contract and the Medicare Shared Savings Program.^{vii,viii} Additionally, some studies provide evidence of quality improvements, including a study by Navathe et al. that found Hawaii’s new capitated primary care payment system led to small but statistically significant improvements in quality.^{ix} However, opponents counter that the evidence is much more mixed.^x An analysis of CMMI’s Bundled Payments for Care Improvement initiative found no meaningful reduction in cost across any of the 5 medical bundles.^{xi} The Government Accountability Office’s initial evaluation of the Hospital Value-based Purchasing Program implemented as part of the ACA found that the new financial incentives had no impact on quality, as did an earlier assessment of Medicare’s Premier Hospital Quality Incentive Demonstration.^{xii,xiii} A meta-analysis of 69 studies assessing pay-for-performance programs across the United Kingdom and United States found little evidence of quality improvement, and that the few marginal gains were limited to process measures only.^{xiv} Additionally, some argue against the presence of selection biases in many APMs where program participation is voluntary. In these programs, providers that elect to enroll are presumably better positioned than other providers to be ‘successful’ and so may be better prepared to deliver higher-quality and/or lower-cost care.

Finally, opponents of pay-for-performance models often argue that if financial incentives are employed incorrectly, they can distort provider incentives and promote behaviors that decrease overall value, rather than increase it. One common theme raised by Dan Pink and other psychologists differentiates between algorithmic and heuristic tasks. Pink finds that performance on algorithmic tasks, which require minimal creativity, can be improved through financial incentives, but that performance on more creative heuristic tasks, such as care delivery, may be unresponsive to financial rewards such as those offered in some alternative payment models.^{xv}

Preferred and Narrow Networks

Commercial insurers and employers have also created preferred and narrow networks in an attempt to direct their enrollees and employees to higher-value care. A traditional narrow network restricts the providers that a plan member can see, and presumably selects these providers based on an evaluation of cost, quality, and other components of value. The broadest and least restrictive narrow networks are often called preferred provider networks (PPOs), and the most restrictive networks are typically referred to as health maintenance organizations (HMOs) and exclusive provider organizations (EPOs).

First, there is clear evidence that narrow networks reduce health insurance premiums (Dafny et al.). However, there have been only limited studies examining the drivers of these reduced

costs, which are likely to include plans' increased leverage when negotiating with providers, plans' exclusions of higher-cost providers, and a potential selection bias in which low-cost patients are more willing to accept a narrow network in exchange for reduced premium costs. Recently, Gruber and McKnight studied the narrow networks' impact on Massachusetts state employees and found that spending per beneficiary in the narrow network plan fell by nearly 40%, driven by both price and quantity declines, with no significant impact on quality.^{xvi} Proponents of narrow networks can point to the Gruber and McKnight study as evidence that narrow networks are 'all upside, no cost.' Notably, the cost savings they found were exclusively from reduced spending in specialty and hospital settings, both of which are often cited as key opportunities to reduce low-value care.

Opponents of using narrow networks as a means of improving value can point to these plans' inherent restrictions on access. Specifically, narrow networks often exclude the highest-cost providers from their plans to reduce costs, and these high-cost providers may be viewed as the most high-quality and/or prestigious providers in a geography. Additionally, narrow networks may require that members commute significant distances to visit specialists (or even primary care providers), especially if these members live in rural areas. Further, Dafny et al. note that as the number of narrow network plans on the health exchanges increases, there may be some form of resultant death spiral where only the sickest, highest-cost patients are willing to pay for the broader networks and eventually the exchange plans only offer narrow networks.^{xvii} Finally, there are several related network models that are designed to achieve similar improvements in value, including tiered networks, high performance networks, and clinically integrated networks. Some models also aim to identify individual high-value providers and direct referrals to those providers, to improve overall value.

Notably, as a counter to both alternative payment models and narrow networks, some argue that variation at even the individual provider level is relatively small. This perspective posits that cutting the lowest value providers out – or improving their practice methods – will make minimal headway in improving overall value. However, there is some evidence that intra-provider variation is quite large. For example, a study found a 20% cost differential between emergency medicine physicians ranked at the 75th and 25th percentile by spending distribution.^{xviii}

Other Interventions

Recently, purchasers have explored numerous creative measures to improve healthcare value. To avoid unexpectedly high surgical or specialty care costs, some purchasers require or offer an option for their employees to go to high-quality Centers of Excellence (CoEs). These employers leverage the large volume of procedures they can direct to the CoE(s) to extract provider cost concessions, typically in the form of a bundled or capitated payment agreement. In some instances, purchasers may elect to work only with one CoE, further increasing their leverage in price negotiations. Providers including Johns Hopkins Medicine and the Cleveland Clinic have signed onto CoE agreements. These larger CoEs handle significant patient volumes, allowing providers additional 'reps' on more complex cases – and several studies, including a large meta-analysis, found that these additional 'reps' drive holistic quality improvements.^{ix} In the most extreme cases, employers have contracted with prestigious U.S. providers to perform surgeries outside of the United States to avoid high hospital facility fees, pharmaceutical costs, and medical device costs, among others. The employers pay more than typical commercial rates to the U.S. provider, to incentivize them to travel, and also pay employees an incentive fee. Ashley Furniture Industries' benefits manager estimates the total cost for procedures done outside the U.S. is approximately 50% of the cost in the U.S.^{xx}

Evidence on these programs' quality is limited, but concerns have been raised about the primary surgeon's lack of familiarity with the setting they are practicing in and team they are working with, as well as potential complications arising from language barriers.

Additionally, some employers have elected to launch their own on-site clinics to provide care for their employees. Providers at these employer-based clinics are either on the company's payroll or on the payroll of a vendor specializing in staffing employer clinics. These clinics have different incentives than typical primary care providers or urgent care centers – and may be able to provide higher value care. Typically, providers at these employer clinics are paid on salary (rather than volume), preventative care is emphasized, paperwork and claims are vastly reduced, appointments are longer, and employee access is improved. Additionally, employee productivity may be positively impacted – employee commutes to the on-site clinic are relatively short and chronic conditions can be managed more proactively. Notably, this option is potentially available to employers of all sizes, as smaller employers may partner together to share the costs of an employer-based clinic. However, as the proliferation of on-site clinics has been relatively recent, few studies have assessed their potential benefits. Convenience to employees, and less lost work time for employers, may be sufficient rationale for these clinics even if there is no additional value with regards to care quality or cost reduction.

Finally, insurers, employers, and others have deployed a host of behavioral interventions. Although not limited to purchasers, successful behavioral interventions range from providing physicians with peer comparison data to instituting EMR best practices and/or reminders.^{xxi,xxii} The former has been done by hospitals and payors – and both seem to be effective in reducing undesired provider behaviors when applied correctly. Additionally, there are myriad opportunities to leverage behavioral economics to design provider incentives, from restructure the timing of providers' bonuses to account for humans' strong aversion to losses, to separating out bonus methods of payment from typical salary deposits to increase their salience.^{xxiii} Finally, although there is limited evidence that data alone drive behavioral change, many purchasers have advocated for additional transparency and data analytics.

However, some have raised concerns that behavioral incentives can be overly blunt tools – specifically, that unless an incentive is sufficiently targeted, it may have adverse and unexpected consequences. In one study, Medicare sent out letters to 5,000 physicians with high prescription rates for quetiapine. However, although physicians reduced their prescribing, this reduction was split evenly amongst patients who were guideline-concordant (and could reasonably be expected to benefit from quetiapine) and patients who were not.^{xxiv} Similarly, rankings at the system level miss significant variation between individual providers and can have large confidence intervals. Rankings at the individual provider level may be more difficult to arrive at, given data and other limitations, but may have greater usefulness in improving provider practice.

Employer Capability and Responsibility

Finally, there is a less common but still important debate regarding the roles that purchasers can and should take on. First, while many commercial payors and some large employers may have the expertise necessary to identify and correctly pull levers that drive improved value, many employer HR functions lack the time, ability, or inclination to tightly manage and coordinate their health care benefits package. Additionally, while employers are incentivized to both keep their employees happy and reduce their healthcare expenditures (which considered in tandem may encourage employers to seek out high-value health care options), their definition of value might

not exactly align with their employees – and in fact might create disutility for their employees. If employers are not well-equipped to promote high-value care and/or their vision of high-value differs from their employees, should employers be responsible for driving improvements in care value? These arguments are often cited by providers in opposition to insurer and employer efforts to influence how providers practice.

Opponents of these arguments – specifically, insurers and employers – are likely to acknowledge the limitations of small employers in defining and incentivizing high value care but point out that these limitations do not extend fully to insurers or larger employers with dedicated healthcare benefits teams. Furthermore, employers are unlikely to define value in a substantially different way from their employees – or they risk employee unhappiness, turnover, and productivity declines. Finally, insurers play a similar role today, and if neither employers nor insurers act as a check on providers’ potential biases, providers act unilaterally in defining value.

Discussion Questions

1. How do you define value? What are you prioritizing and deprioritizing in that definition?
2. Do you believe we can measure value accurately enough to act on it?
3. What are the 1-2 greatest challenges you believe remain to accurately measure value?
4. What metrics do you believe best measure value? What metrics are least representative of value? Do we track too many metrics today?
5. Which levers do you believe are most effective in improving healthcare value?
6. What insurers or employers are using these levers effectively? What purchasers are being the most creative in their approach?

References

ⁱ “Measuring the quality of physician care.” Agency for Healthcare Research and Quality. 2019-09.

ⁱⁱ Cheney, Christopher. “AMGA endorses 14 metrics for quality and value reporting.” Health Leaders Media. 2018-02.

ⁱⁱⁱ “Doctors are so much more than quality metrics.” Medical Economics. 2016-03.

^{iv} Mackenzie, Andrew et al. “Comparing model error between a standard risk adjustment model and a disease-specific risk adjustment model.” Society of Actuaries. 2020-12.

^v “Doctors are so much more than quality metrics.” Medical Economics. 2016-03.

^{vi} Gupta, Dipti et al. “Value-based reimbursement: Impact of curtailing physician autonomy in medical decision making.” AJR, journal of the American Roentgen Ray Society. 2016-02.

- ^{vii} Song, Zirui et al. “Health care spending, utilization, and quality 8 years into global payment.” *New England Journal of Medicine*. 2019-07.
- ^{viii} McWilliams, J. Michael et al. “Medicare spending after 3 years of the Medicare shared savings program.” *New England Journal of Medicine*. 2018-09.
- ^{ix} Navathe, Amol et al. “Association between the implementation of a population-based primary care payment system and achievement on quality measures in Hawaii.” *JAMA*. 2019-07.
- ^x Burns, Lawton & Mark Pauly. “Transformation of the healthcare industry: Curb your enthusiasm?” *The Milbank Quarterly*. 2018.
- ^{xi} Maddox et al. “Evaluation of Medicare’s bundled payments initiative for medical conditions.” *New England Journal of Medicine*. 2018-07.
- ^{xii} “Hospital value-based purchasing: Initial results show modest effects on Medicare payments and no apparent change in quality-of-care trends.” *Government Accountability Office*. 2015-10.
- ^{xiii} Jha, Ashish et al. “The long-term effect of premier pay-for-performance on patient outcomes.” *New England Journal of Medicine*. 2012-04.
- ^{xiv} Mendelson, Aaron et al. “The effects of pay-for-performance programs on health, health care use, and process of care: A systematic review.” *Annals of Internal Medicine*. 2017-03.
- ^{xv} “Daniel Pink on incentives and the two types of motivation.” *Farnam Street*. 2016-08.
- ^{xvi} Gruber, Jonathan & Robin McKnight. “Controlling health care costs through limited network insurance plans: Evidence from Massachusetts state employees.” *American Economic Journal: Economic Policy*. 2016-05.
- ^{xvii} Dafny et al. “Narrow networks on the health insurance marketplaces: Prevalence, pricing, and the cost of network breadth.” *Health Affairs*. 2017-09.
- ^{xviii} Van Parys, Jessica. “Variation in provider practice styles within and across emergency departments.” *PLOS ONE*. 2016-08.
- ^{xix} Halm, Ethan et al. “Is volume related to outcome in health care? A systematic review and methodologic critique of the literature.” *Annals of Internal Medicine*. 2002-09.
- ^{xx} Galewitz, Phil. “A Mexican hospital, an American surgeon, and a \$5,000 check (yes, a check).” *The New York Times*. 2019-08.
- ^{xxi} Navathe, Amol & Ezekiel Emanuel. “Physician peer comparisons as a nonfinancial strategy to improve the value of care.” *JAMA*. 2016-11.
- ^{xxii} Meeker, Daniella et al. “Effect of behavioral interventions on inappropriate antibiotic prescribing among primary care practices.” *JAMA*. 2016-02.
- ^{xxiii} Emanuel, Ezekiel et al. “Using behavioral economics to design physician incentives that deliver high-value care.” *Annals of Internal Medicine*. 2015-11.
- ^{xxiv} Liao, Joshua & Amol Navathe. “Nudging physicians to reduce quetiapine prescribing using Medicare letters.” *JAMA Psychiatry*. 2018-10.