HEALTH & MEDICINE

More Volume than Value in Health Quality Measures?

Efforts to measure treatment quality need to be more transparent, not more complex.

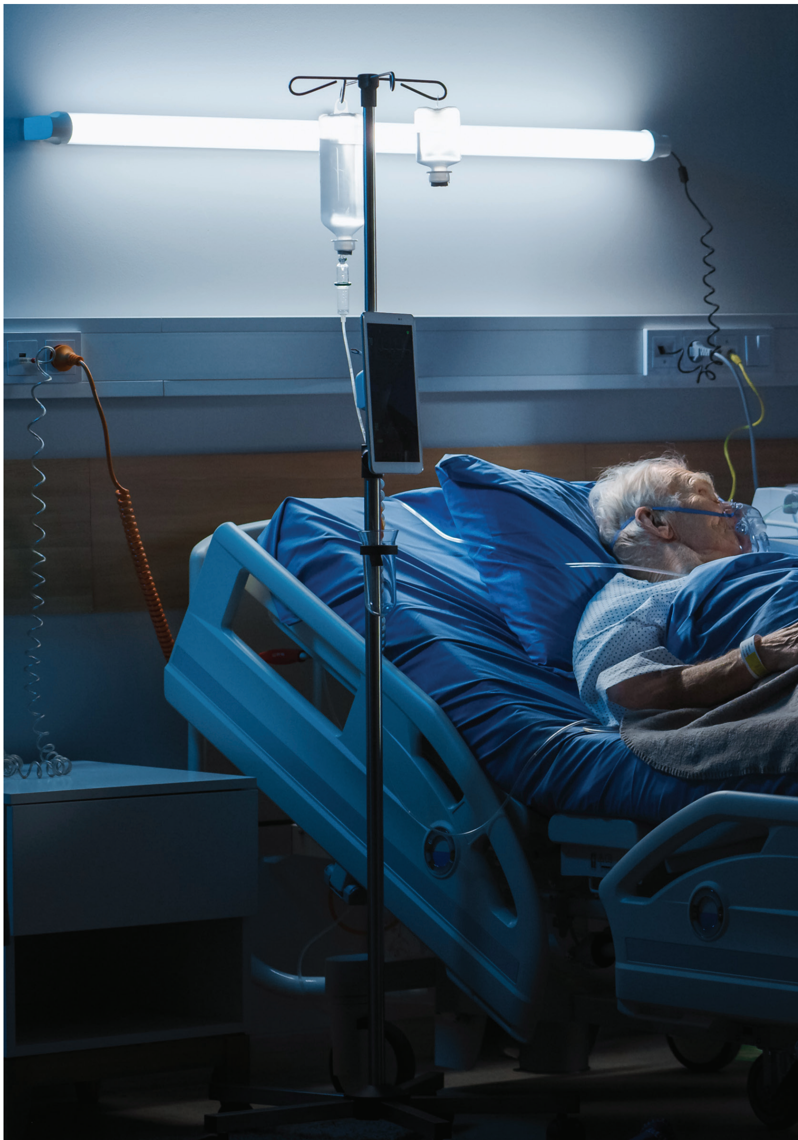
BY TOM MILLER

Quality measurement in healthcare policy struggles with two particularly chronic pre-existing conditions in U.S. politics: aversion to the appearance of "discriminatory" decision making and resistance to accountability for overtly restricting marginally beneficial care. Most quality measures find it hard to reward doing less, not more. They are biased toward yielding above-average ratings (that is, producing the Lake Wobegon effect) for most healthcare providers. However, they still can succeed politically for decision makers by deflecting blame while hitting closer to the sweet spot of discretionary power without accountability.

Most current measures of care quality and value also remain imperfect, limited, and contentious in more technical scientific and economic terms. But their greatest vulnerabilities in our mixed private–public healthcare system remain political.

QUALMS ABOUT QALYS

The most recent illustration might be observed in the "Protecting Health Care for All Patients Act of 2023," HR 485, which moved rapidly through House committees last spring but now is stalled, awaiting action by the full House of Representatives. The bill is cosponsored by two House committee chairs, Cathy McMorris Rogers (R–WA) of the Energy and Commerce Committee and Jason Smith (R–MO) of the Ways and Means Committee. The legislation would prohibit the use of quality-adjusted life year (QALY) and similar measures for coverage and payment determinations (usually on comparative effectiveness grounds) under all



TOM MILLER is a senior fellow at the American Enterprise Institute.

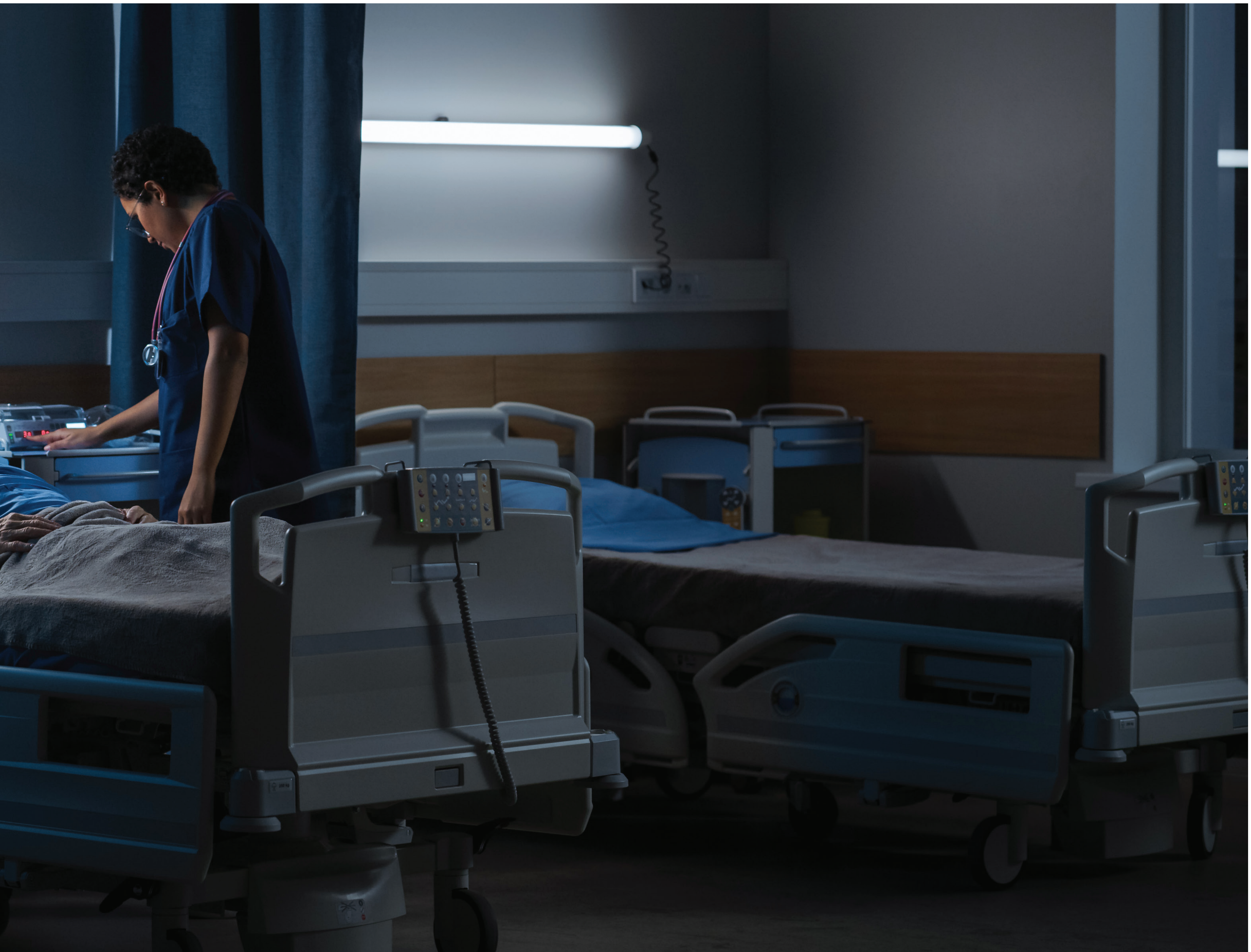
federal and state healthcare programs.

The primary stated rationale for the bill is to avert discrimination against the elderly, disabled, or terminally ill in resource allocations and price determinations. This is not a baseless concern because less-refined reliance on increased life expectancy alone could undervalue qualitative improvement in life for those already suffering from impaired health and well-being (in other words, chronic diseases or longer-term disabilities). Still, proper estimates of QALYs would direct health resources to more broadly beneficial uses, but the legislation would block that.

In response to what seemed like imminent passage of the legislation last April, a squad of nine health policy researchers assembled in a *Health Affairs* article to temper political qualms about QALYs and related measurement tools. University of Washington School of Medicine professor Sean Sullivan et al. argued that other alternatives could address concerns about discrimination and facilitate explicit consideration of both quality and length of life in comparative-effectiveness assessments. They mentioned at least three newer measurements—Healthy Years in Total, Life Years Gained, and Generalized Risk-Adjusted QALY—as methods to avoid discriminating against the disabled and devaluing extensions of their lives. The field of regulatory and reimbursement review in healthcare already extends to other measures like Disability-Adjusted Life Years and Value of a Statistical Life Year gained. Further efforts seem likely to ensure that every gain in care quality can be recognized for reimbursement in some manner.

In its most basic form, QALY claims to measure the value of health outcomes to the people who experience them. It considers both the length of life gained from a treatment and the quality of that gained time, and produces a single number that then can be compared across different types of treatments. But the increased lifespan and its quality are measured first in terms of how people feel in specific states of health (generally ranging from perfect health to death, although a case can be made that some states are worse than death), the time they live in those various states of health, and then multiplying the two measures. Further weighting modifications can involve determining utility values by methods such as time tradeoffs, standard gambles, or visual analog scales. Medical costs for particular treatments may be combined with QALYs to estimate the cost-per-QALY for a given intervention.

As QALY-style measures evolved over the last 50 years, so did the list of potential objections to them. For example, well before



GORODENKOFF / GETTY IMAGES

the latest proposed legislation, the 2010 Affordable Care Act prohibited the Patient-Centered Outcomes Research Institute, an influential government-sponsored organization that funds comparative-effectiveness research, from using cost-per-QALY benchmarks.

The litany of QALY criticisms includes: excluding broader concepts of well-being from measurements of "health"; assuming that people's preferences between time and quality of life do not change; relying on a public rather than a patient perspective of utilities for various states of health; excluding non-health benefits that produce positive externalities and social utility; and failing in neutral QALYs to consider patients' burden of disease, their condition severity, or other equity factors. Critics also argue that QALYs do not differentiate between overall improvement in health state and prevention of decline, and they discriminate against the elderly with shorter life expectancy.

MODULATING MONEY AND MEDICINE

The empirical precision behind these assorted measures is prone to overstatement, but they really are aimed more at clouding harder choices on the margins and keeping key parties tolerably happy. Most disputes over quality measures that make a difference ultimately reflect the key battleground in health policy: who gets paid and how much. When regulatory or reimbursement officials decide that a treatment adds QALYs for a given population, that alone does not determine how much it is worth. For that, many decision makers turn to the concept of Value of a Statistical Life (VSL). The approach relies on evidence of tradeoffs in markets between risk and money to generate estimates of the benefits of risk reduction efforts (in other words, to monetize them). Government policy for such interventions (including health treatments and the policies governing them) is supposed to reflect society's willingness to pay for such benefits. Policies to enhance or maintain human life differ in their effectiveness, and "buying" an extra year or life may not be worth the same across the board.

For example, VSL appears to vary by income, with an elasticity in the United States ranging from about 0.5 to 0.7. In essence, being wealthier equates to being more willing to buy what might make one "healthier." Other complications in determining the comparative value or effectiveness of different treatments via VSL metrics include how a policy to reduce risk of one type may increase risk of another type or create incentives for individuals to compensate by reducing their other efforts to reduce risk. Moreover, the unintended consequences of some risk reduction policies may add extra costs that end up increasing fatalities, or they may decrease income available to finance other types of health and safety expenditures (including some non-healthcare investments that improve lifetime health outcomes). More simplistic VSL measures may neglect individual preferences for reduced morbidity—and not just mortality—risk.

One can try to merge VSL and QALY metrics by calculating the value of a statistical life year (VSLY) saved for more incremental mortality improvements. But it might be worth considering geographic or behavioral adjustments in measuring VSL. After all, in places like Casablanca, "human life is cheap" and perhaps, too, in certain states, at least for those covered at pre-ACA Medicaid income levels only. (The effects of state policies with liberal abortion laws versus states with liberal gun ownership laws tend to offset each other politically, if not in body counts.)

Another more personalized VSL might consider whether the previous actions of individuals speak more loudly to how much they value their own lives. After all, if one's lifetime habits and practices (e.g., tobacco use, sedentary lifestyle) clearly indicate little concern for future health, why should everyone else pay a much higher premium to extend it? We might call this a health behavior quality-adjusted life year.

The main tradeoff in easing qualms about QALY involves how these sorts of adjustments can dilute their effectiveness in making cost effectiveness comparisons across broader treatment options. "Discrimination" need not always be treated as invidious when it instead is used to make evidence-based distinctions in advancing more efficient and effective use of limited public resources, albeit tempered on the back end for overriding policy considerations involving equity and target efficiency. QALY provides limited tools that by themselves cannot address fully the larger problems of our political system in failing to account for the full costs of spending other people's money, assess the public's willingness to pay, maintain the consistency of choices over time, and—most of all—accept responsibility for making those choices more transparently.

OPPORTUNITY COSTS AND COGNITIVE DISSONANCE

Of course, the current backlash over QALY and similar measures to assess the comparative effectiveness and value of health treatment alternatives may be less about scientific accuracy or elusive concepts of social equity than more mundane political calculations. The cognitive dissonance of House Republicans insisting on more budgetary austerity in the aggregate while recoiling from applying tools to trim health expenditures involving specific products and services of lower marginal value is more of a chronic preexisting condition first contracted by encounters with public opinion snapshots. It may even be aggravated by other would-be sellers of those items who are quite adept at couching their actual objections to lower prices in more high-minded rhetoric. Further irony arises from the unwillingness of those who dodge responsibility for making tougher choices to delegate them instead with fewer pre-set mandates and prohibitions, but better information, to consumer markets.

In any case, opportunity costs can be disguised, but they eventually must be paid, even if indirectly and over longer time horizons. Reasonable discount rates reveal that paying them later will cost more. By ignoring the variance in the value of different states of health, we will also pass up gains from changes in treatment that can cure or improve our health.

Consider some of the rhetorical paradoxes that coexist too comfortably within much of our contemporary health policy debate. For example, policymakers and commentators lament that Americans spend too much on health care, except when we spend too little. (Perhaps we do both at the same time!) Prices and costs for current health services and health insurance are too high, yet the value of a chance for better health is said to be "priceless." The healthcare sector is an engine of growth and innovation, yet it also threatens to swallow the rest of our economy and bankrupt younger and future generations.

It may seem somewhat contradictory to proclaim the many bountiful benefits of our healthcare system, yet at the same time suggest that we subsidize it too much and "everyone else" must be spending too much. Policy advocacy by many academics and healthcare interest groups remains biased toward using national politics to preempt more resources for healthcare spending than individuals seem willing to pay with their own money. (Pause briefly to reflect on just who the ultimate beneficiaries really are!). These efforts are reinforced by many right-leaning critics of current policy who default to the view that high levels of overspending on healthcare are better than government rationing.

Age-driven demand for more healthcare carries just-in-time inconsistencies once one examines the distribution of political benefits and burdens across generations.

Admittedly, few people have lost money betting on the capacity of officeholders and their supporters to entertain several conflicting positions and perceptions simultaneously, as long as they do not have to be combined in the same sentence at the same time. The truncated challenges of choosing "your money or your life" in healthcare spending transactions evoke echoes of a much-older comedic skit by self-styled skinflint Jack Benny. In fact, modern bands of political robber-barons may be after both! Or at least much more of your money, and most of your independent healthcare choices in life.

Doubling down on our health spending bets and debts later in life still can claim academic foundations, illustrated best in a 2007 study by Robert Hall and Charles Jones. They concluded that their sophisticated modeling supported the proposition that spending 30 percent or more of GDP on health by 2050 would maximize Americans' social welfare. In other words, stepping on the healthcare accelerator, not the brake, is really the way to go.

Hall and Jones argued that as we grow richer and richer, we run out of interesting things (and greater quantities of them) that we want to consume, at least on the margin. Meanwhile, the most valuable thing we have left to purchase becomes more time on the clock to live. (Increased doubts and fears of the afterlife appear to reinforce each other.) They dressed this up in the more technical econometric jargon of the marginal utility of further non-health consumption falling faster than the marginal utility of additional health spending, as people get richer and the value of life (in statistical years, and perhaps even nanoseconds) rises.

Embedded within this one-sided model are a few givens that remain quite questionable. The law of diminishing returns apparently applies only to non-health consumption, but not to the next dollar of spending on healthcare. Extending life is equated with consuming more healthcare services, rather than in investing most in those health-enhancing tools and resources that yield the best returns, including ones beyond the billable reach of the medical sector. Institutional arrangements and public policies that distort the relationship between our respective healthcare spending and health-enhancing production functions are essentially ignored in such pro-health-spending assumptions.

A stronger analysis of the factors behind preventable mortality by J. Michael McGinnis, Pamela Williams-Russo, and James Knickman in 2002 concluded that perhaps only 10–15 percent of early deaths could be avoided by better availability or quality of medical care, whereas 40 percent are due to behavior patterns that could be modified by preventive interventions, while genetic (30 percent) and social (15 percent) determinants account for most of the rest. Of course, buying another year at the back end of an expiring life might be limited to a different menu of less-effective but more-expensive medical interventions.

It may well be the case that as people get very rich, the most valuable channel for their spending is to purchase additional years of life. Nevertheless, one should not confuse the upper tail of the income-driven health-consumption distribution with its middle and lowest tiers.

Age-driven demand for more healthcare carries its own just-in-time inconsistencies once one examines the distribution of political benefits and burdens across generations. For example, at the end of their excursion through the statistical contortions of QALY, Hall and Jones add that the quality-of-life numbers they reviewed imply that "a 65-year-old would give up 82 percent of her consumption, and an 85-year-old would give up 87 percent of her consumption, to have the health status of a 20-year-old." Perhaps the more accurate way to interpret those preferences is that the 65-year-old would be willing to give up 82 percent of the consumption (and future income) of a genuine 20-year-old. The 85-year-old may not only want the body of a 20-year-old but hopes the latter visits her more often, even if his economic prospects are limited to cleaning the pool or cutting the lawn to help pay off multigenerational debts.

THE CASE FOR MODEST BUT TRANSPARENT MEASUREMENT

So why bother to overcome untempered aversion to any public assessment of the comparative returns from investments and other spending allocations for different types of medical treatments and technologies? Why study how effectively and efficiently particular healthcare providers tend to dispense them? Poisoning the informational well will limit the potential of other consumers seeking value-based indicators for healthcare choices to draw upon it more deeply in the future. If we do not readjust the current trend lines for health spending growth, we still may have to ration. But that future rationing will involve the rest of our income and wealth that is not devoted to healthcare, particularly when projected growth rates for the former are not guaranteed to overcome the appetite of the latter.

We underinvest in what, earlier in life, might keep us healthier longer and overinvest, later in life, in trying to hold at bay conditions we inherited, neglected, or nurtured.

One way to reach a better balance is to maintain the broader distinction between public spending and private spending. The scope and scale of the former should be, and occasionally is, limited by the key considerations of how it contributes to the public good, minimizes opportunity costs, and justifies its encroachments on our respective pursuits of happiness in the rest of civil society. On the other hand, the level of one's own private spending for health (the unsubsidized and non-mandated kind) should be no one else's business, with the usual disclaimers for force and fraud. For better or for worse, in sickness and in health, till death do we depart.

In other words, spend as much of your own money, or that of friends and family, as you want on healthcare. Just do not touch up the rest of us for even more in higher taxes and interest payments on the public debt to extend and expand the excessive reach of Medicare, Medicaid, private insurance subsidies, regulatory barriers, and the next round of health entitlements—at least without checking to assess how much it matters, based on available evidence.

Our political financing of healthcare remains fundamentally driven by two distortions, or mismatches, related to different time horizons and changing (time-inconsistent) preferences.

First, older Americans, who vote more often and engage in the political process more effectively, would rather spend the money of younger workers than more of their own resources. They personally experience, and politically extract, the greatest differential between the apparent point-of-service cost of the care they receive and its life-extending net present value to them. They buy care at a politically discounted price but value it at an undiscounted rate. Although one's early behavior and decisions might suggest a much lower personal valuation, at the time, of one's many statistical life years ahead, those remaining years are revalued and marked to market as they become less plentiful.

Second, opportunities for the most effective and longest-lasting investments in future health arise early in life (initial child development, primary education, and formation of behavioral habits). However, their relative rates of return diminish over time. Eventually, the installed base of one's depreciating health makes more attractive other costlier and more intensive health interventions, hopefully delivered "just in time" as ultimately the last lottery tickets on sale.

Hence, we underinvest in what, earlier in life, might keep us healthier longer and overinvest, later in life, in trying to reverse, or hold at bay, the health conditions we have mostly either inherited, neglected, or nurtured. Human nature's tendencies to try to ignore the distantly unavoidable, renegotiate terms after the fact, and purchase upgrades with someone else's money, are welcomed and augmented by the "earned entitlement" façade of our mixed welfare state.

One should never underestimate how difficult it is to confront even the initial outlines of these conflicts and contradictions in the name of "health reform." Hence, the better advice for health policy and politics is to aim lower if you hope to hit higher targets. As a more achievable goal, we should hope that quality measures at least become more transparent in revealing not just those stated targets, but also the underlying policy preferences, value judgments, and biases of the decision makers purporting to reach them. Defaulting back to more random, inconsistent, or ambiguous criteria only encourages the abuse and increased pursuit of even less accountable discretionary power and limits efforts to challenge or constrain it.

READINGS

- "Alternatives to the QALY for Comparative Effectiveness Research," by Sean Sullivan, Darius Lakdawalla, Beth Devine, et al. *Health Affairs Forefront*, April 21, 2023.
- "Income Elasticities and Global Value of a Statistical Life," by W. Kip Viscusi and Clayton Masterman. *Journal of Benefit–Cost Analysis* 8(2): 226–250 (2017).
- "The Case for More Active Attention to Health Promotion," by J. Michael McGinnis, Pamela Williams-Russo, and James Knickman. *Health Affairs* 21(2): 78–93 (2002).
- "The Value of Life and the Rise in Health Spending," by Robert Hall and Charles Jones. *Quarterly Journal of Economics* 122(1): 39–72 (2007).