

Should Medicare's Value-Based Pricing Be Adjusted for Social Risk Factors? The Role of Research Evidence in Policy Deliberations

Steven Sheingold

US Department of Health and Human Services

Karen E. Joynt Maddox

Washington University

Rachael B. Zuckerman

Nancy De Lew

US Department of Health and Human Services

Arnold M. Epstein

Harvard University

Abstract Health services researchers now have enhanced opportunities for overcoming past obstacles to their work being relevant and useful for policy makers. Understanding the policy process itself is essential to taking advantage of these opportunities. Public policy making can be a complex process, and even the highest-quality evidence is only one of many factors considered and balanced during deliberations. We use a recent research report focused on social risk factors and provider performance as an example of the role evidence plays in policy deliberations. The report addressed the question of whether Medicare's value-based purchasing programs should be adjusted for patients' social risk factors. Despite high-quality and consistent results about the relationship between performance and social risk factors, policy deliberations did not result in a simple recommendation about making adjustments. Thus, the report offers a window into the role of research evidence in the development of policy.

Keywords Medicare, evidence, policy

The idea that health policy decisions should be informed by the best evidence available is hardly controversial (Califf et al. 2016). At this time, health services researchers are in a better position than ever to inform the policy-making process with the near real-time availability of big data and the ability to analyze it at the desktop. Nonetheless, researchers must learn the lessons from a past in which their findings have often failed to have a

Journal of Health Politics, Policy and Law, Vol. 43, No. 3, June 2018
DOI 10.1215/03616878-4366160 © 2018 by Duke University Press

large impact (Davis, Gross, and Clancy 2012). Indeed, there is often a gap between what research results suggest and the policies actually implemented (Brownson, Chiqui, and Stamatakis 2009).

Disconnects between research and policy occur for several reasons (Davis, Gross, and Clancy 2012; Cairney and Oliver 2017). First, academic research agendas are not always attuned to the information needs of policy makers. Second, research questions or results may not be fully consistent with the policy issues being deliberated. Third, the timing for release of research results is not always well synchronized with the policy process.

Understanding the policy process and environment is essential to addressing these issues (Gluck 2015; Cairney 2015; Cairney and Oliver 2017; Oliver, Lorenc, and Innvaer 2014). In both the public and private arenas, making important decisions can be a complex and multifaceted process. While results from high-quality research are critical to making good decisions, other factors, including values and the policy context, play key roles and must be integrated with research results (Hardee, Wright, and Spicehandler 2015). An additional source of complexity is that program and policy decisions are seldom binary; decision makers typically might evaluate multiple courses of action by balancing policy considerations, evidence, and values. Put another way, often there is no simple pathway between a *p*-value and a policy. Both in developing research and in presenting results, it is important that researchers fully understand the decision-making environment in which their products will be considered (Oliver, Lorenc, and Innvaer 2014).

A recent report by the assistant secretary for planning and evaluation (ASPE) of the US Department of Health and Human Services (HHS) provides a useful example of research, high-quality evidence, and policy decision making (HHS 2016). The report examined the relationship between social risk factors and performance measures and considered policy alternatives that might account for these factors in Medicare's value-based purchasing programs. In this article, we describe a generic framework for thinking about evidence-informed decision making processes and use the ASPE report as a case study.

Overview of the Case Study

Over the past ten years, Medicare has increasingly moved toward value-based purchasing, in which fee-for-service payments are supplemented to incentivize higher-quality and lower-cost care. The transition of Medicare from a payment system largely based on quantity to a system that holds

plans and providers financially accountable for cost and quality intensified a debate about whether performance measures should be adjusted for patients' socioeconomic status (SES). A growing body of literature had already demonstrated that individuals with low SES received lower quality of care on reported measures and that providers treating a large share of such patients were more likely to be penalized under Medicare's value-based purchasing programs (HHS 2016; Desai et al. 2016; Gilman et al. 2015; Rajaram et al. 2015; Joynt and Jha 2013). The National Quality Forum, the organization that endorses quality measures, assembled a technical panel to examine the issue and make recommendations concerning sociodemographic adjustment of measures (National Quality Forum 2014). One side of this debate claimed that reporting quality measures and basing payment on them without adjustment for SES factors would financially harm safety net providers and could reduce access for vulnerable patients. Others were concerned that any adjustment for these factors would set lower-quality standards for patients with lower SES and reduce incentives for the providers to improve quality.

In October 2014, Congress passed the Improving Medicare Post-Acute Care Transformation (IMPACT) Act (H.R. 4994), which indicated its intent to better inform the debate with research. Section 2(d)(1) of the IMPACT Act calls for the secretary of HHS, acting through the ASPE, to conduct a study evaluating the effect of individuals' SES on quality measures and measures of resource use under the Medicare program. ASPE was required by the IMPACT Act to deliver the first of two reports by October 2016. The key objectives of this first report to Congress were to develop and analyze the best evidence and evaluate policy alternatives for the Centers for Medicare and Medicaid Services regarding access to new data, accounting for socioeconomic factors in quality and resource measures, and the adjustment of value-based payments for these factors.

To examine issues related to SES, one first needs a definition of this concept. Therefore, ASPE commissioned the National Academies of Science, Engineering, and Medicine (National Academies) to define and conceptualize SES for the purposes of the IMPACT studies (Committee on Accounting for SES in Medicare Payment Programs 2016). The National Academies added external expertise and credibility to the discussion of a challenging topic. It convened a committee of experts in the field and conducted an extensive literature review. Based on the information collected, it concluded that the appropriate framework is social risk factors instead of SES. Social risk factors include socioeconomic position, race, ethnicity, gender, social context, and community context. ASPE's report to Congress also adopted the social risk framework.

ASPE decided to conduct the research internally and had many advantages for overcoming past obstacles to using research to inform policy. First, the legislation clearly defined the research and policy questions. Second, a well-qualified research team had rapid access to the data needed for the research. Third, because the research was conducted internally, continual interaction between the researchers and decision makers within HHS and the White House was possible. This allowed for frequent feedback to carefully align policy and research questions and to modify the way results were presented. The team was also able to brief congressional staff as results became available and receive their input. Despite being of considerable depth and breadth, the research was completed well in advance of the statutory October 2016 due date. External panels of experts in the field, constituted to provide input and review the work, were generally very supportive. Internal policy discussions and final report drafting required some additional time, and the report was transmitted to the Congress in December 2016.

Despite the apparent advantages for linking evidence to policy, the resulting deliberations did not yield simple policy outcomes with regard to whether or not to adjust value-based purchasing programs for social risk. For health services researchers in general, and particularly for those who believe that statistical estimates with the expected sign and a “good” *p*-value should result in an obvious policy direction, the IMPACT report offers a window into the role of research evidence in the development of policy. We explore these issues in more detail in the following sections.

Program/Policy Decision Making Using Evidence: A Conceptual Model

There are several models of the ways that research and other factors can influence policy (Peterson 2017; Gold 2009). Well-known models of the policy process begin with problem identification and describe the pathways by which evidence affects policy formulation. For the purposes of this article, we emphasize the evidence assessment and deliberative phases of the process. That is, we focus directly on the use of evidence in deliberating policy alternatives once the problem and the need for a decision is recognized—even if that decision may ultimately be to maintain status quo.

Figure 1 summarizes the process of moving from a recognized need to addressing an issue to a policy decision. Such a process involves integrating multiple factors into a decision, including evidence, organizational culture, and the external policy context. It is important to note that the model

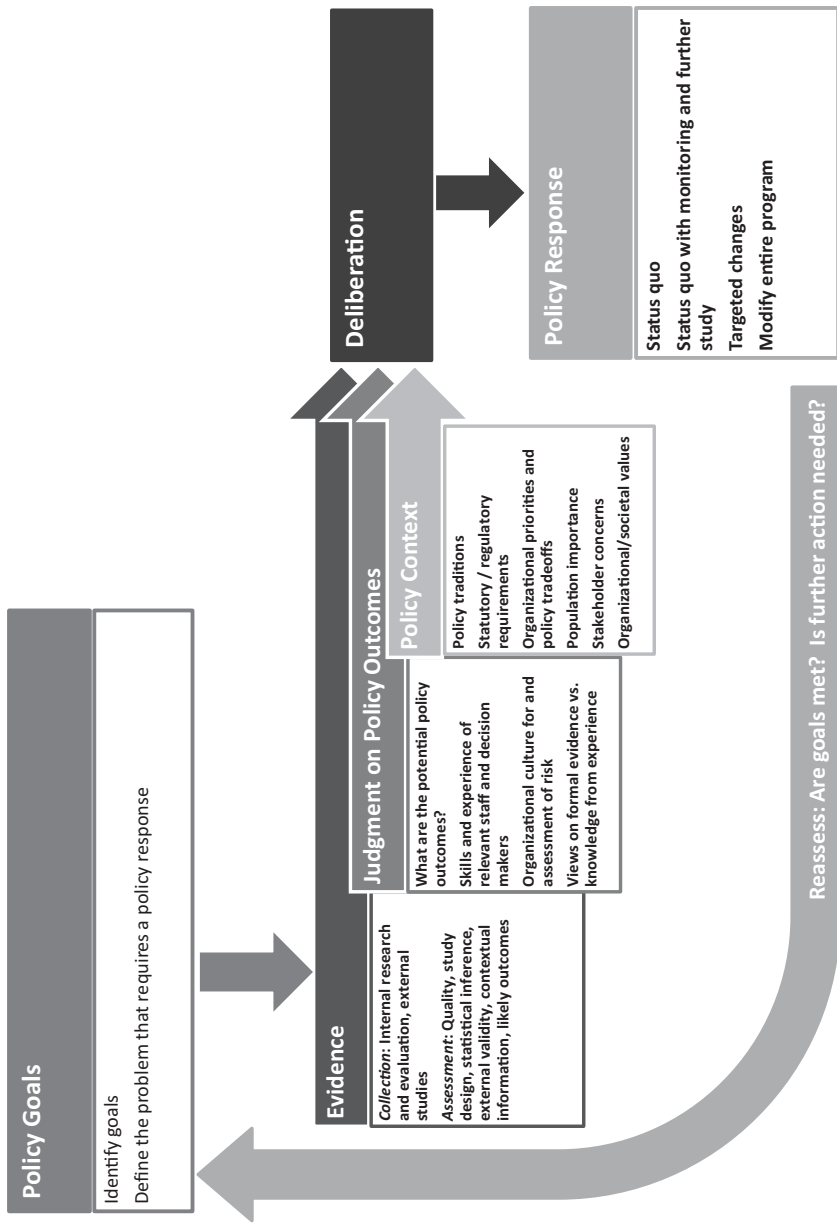


Figure 1 Conceptual Framework for Evidence-Informed Decision Making

is not intended to suggest factors that have been considered or should be considered in every policy decision. Based on the literature and the authors' experience, it describes factors that may be considered appropriate for specific policy deliberations.

Policy Goals

The first step in evidence-informed decision making for policy is defining the policy goals. Whether for a new policy or modifying an existing one, these should be clearly delineated along with the problem that requires the policy response in the first place.

Evidence

Evidence Collection. Evidence collection should proceed broadly and often requires a combination of internal research along with external studies. One important component of evidence collection is commonly referred to as the “environmental scan,” which provides a broad overview of the available evidence. Such a scan identifies information gaps and helps formulate the research questions.

Researchers should recognize that policy makers may consider valuable a variety of information beyond typical research (Cairney 2015). Decision makers often seek a range of “evidence” that includes not only research studies but also information from stakeholders, gray literature, case studies, narratives, and anecdotes (Hardee, Wright, and Spicehandler 2015).

Evidence Assessment. While researchers strive to produce the highest-quality work possible, conveying a careful assessment of the relevant body of research to policy makers is critical. Evidence assessment should include evaluation of the quality of the evidence, including assessment of study design, statistical inference, and external validity. One controversial issue for researchers in presenting and interpreting research results for policy makers is how much weight to put on the p -value. The size of the p -value is often held up as the key criterion in evaluating the importance and significance of results and may influence what research gets published in high-level journals. More important, the p -value is often interpreted as a “walk-off” criterion for decision making. In this approach, a p -value less than a commonly used threshold, such as 0.05, is interpreted as a statistically significant program/policy effect, while a p -value greater than that level is interpreted as not statistically significant. These findings can often be overstated—statistically significant is often wrongly conflated with the

existence of an important program effect, while not statistically significant (by the p -value criteria) is assumed to mean that there is no program effect. Either overstatement might inappropriately influence policy decisions.

A recent statement by the American Statistical Association is clear that “scientific conclusions and business or policy decisions should not be based only on whether a p value passes a specific threshold” (Wasserstein and Lazar 2016). Whether decision makers would accept p -values that are higher than 0.05 may depend on the specific values of the organization (see below), the context of the particular program and its potential outcomes, and how decision makers view available anecdotal or contextual information. For example, decision makers might accept a higher p -value for programs with a potential for positive outcomes and whose worst case is no effect, while setting stricter standards for programs whose worst case is negative outcomes or harms. Researchers must understand these factors and carefully explain results in the context of the appropriate meanings of p -value as described above. Researchers should also note the difference between statistical significance and magnitude (e.g., clinical significance) as appropriate for given situations.

Decision makers also need information on how much of the variation in key outcomes can be explained by the policy or program effect. First, it is possible for statistical models to exhibit a highly significant (low p -value) effect of one or more factors but leave much of the variation in outcome unexplained. This would mean that predictions from the model would be imprecise—that is, the prediction interval would be wide. More important, unexplained variation in outcomes may be as policy relevant as the explained variation. For example, was such unexplained variation random or due to unmeasured but important variables that might explain differences in performance? Finally, it is also important for researchers to convey information to policy makers about the causes of the variability. Much can be learned about the generalizability of findings from a randomized controlled trial by detecting and quantifying variation in program impacts (Raudenbush and Bloom 2015). For example, policy may be better informed by not only learning whether accountable care organizations reduce patient care costs on average but also contrasting the characteristics of these organizations that were successful with those that did not perform as well.

Judgment on Policy Outcomes

Based on the evidence, decision makers will make judgments about the implications for policy: Is there a problem to address, and how large? What

are the implications of the findings for how current programs or policies are performing? What do these results tell them about the magnitude of risks associated with the status quo versus alternative courses of action?

Even the highest-quality research leaves uncertainties with regard to outcomes. The culture of each organization and the skills, preferences, and experiences of its decision makers are all likely to have a significant effect on how evidence is interpreted and used for decision making (Oxman et al. 2009; Liang et al. 2012). Decision makers' current environment will almost surely influence the way they interpret the likely outcomes from a given body of evidence. As noted by Cairney (2015: 42), attention of policy makers to particular kinds of evidence can change rapidly "as events prompt policymakers to shift their focus quickly, or ambiguity and uncertainty contributes to shifting attention to different policy images."

Lastly, organizations differ in terms of their attitudes and preference toward risk under various conditions, that is, the extent to which they will foster experimentation and accept the risk of failure to seek the possibility of significant gains. The same set of evidence might lead one decision maker/organization to believe that a policy is worth pursuing to produce desirable effects but another to find it too risky.

Policy Context

Once decision makers have assessed the research evidence, they must also consider the broader policy context in reaching a decision. Policy context may be influenced by both internal and external factors, such as statutory and regulatory requirements, budget considerations, and organizational priorities and values. Stakeholder positions and concerns can also play a role in how policy alternatives are evaluated.

For public policy, statutory and regulatory requirements are often critical to choosing among policy alternatives. These requirements can set bounds and provide directives on both policy options and their implementation. For example, program officials often cannot eliminate policies or programs authorized by statute, even if the evidence suggests they have negative consequences. Similarly, statutes can allow program officials to modify programs via regulatory change, but limit the scope of those modifications.

Budgetary limitations often play a significant role in decision making. Limitations in funding, staffing, and other administrative resources may affect an organization's ability to maintain or expand some programs, even those with strong evidence of success. Then other policy factors, such as the importance to the population served and consistency with other organizational goals and policies, become critical.

Particularly for public agencies, policy traditions and precedents may have a large impact on decision making. Policy traditions may be the way an organization has historically chosen among alternatives and preferences for reaching and implementing decisions; it may also be embedded in the underlying statute of a program. Most policy decisions have several types of effects: the impact on the program, the impact on the population served, and the precedent it might set for future decisions. In particular, decisions that depart from policy traditions without strong rationale may set precedents that could affect future decisions.

A good example of the power of a precedent is how payment is determined for hospitals in Medicare's prospective payment system. The Medicare statute has generally established a single rate for all providers. Differences in costs that are considered legitimate, and beyond the immediate control of providers, are recognized by adjustments to the single national rates (e.g., for teaching and local wage cost for hospitals, local practice cost variation for physicians) rather than creating separate rates for each of the various provider groups. Policy alternatives that create separate payment groups for providers would need to have a strong rationale for departing from this tradition, and the government has generally been reluctant to do so.

Potential Trade-offs between Policy Goals. Policy makers also pay attention to how any decision complements or conflicts with other organizational priorities. For a hypothetical example, an agency may have a stated objective of training the most disadvantaged individuals, but well-designed program evaluations suggest that it is more cost effective to target less disadvantaged individuals. In this case, strong evidence may suggest one policy direction, but the ultimate decision may be different, depending on how the agency assesses the trade-offs among its objectives. Undoubtedly, the values and beliefs of the organization and those of its decision makers would play a key role. Likewise, the evidence would play a key role by shedding light on the magnitude of potential trade-offs.

Deliberation and Decision

Ultimately, a decision must be reached—and such a decision can vary in its departure from the status quo. Decision makers may weigh the evidence within the context of all of the factors discussed above and decide that the best way to meet the policy goal at hand is to maintain the status quo for now and couple this with monitoring and further study. Alternatively, they

may decide on a course of action such as wholesale program modifications or more incremental and targeted changes.

Applying the Framework to a Case Study: Should Value-Based Payment Account for Social Risk Factors?

As outlined above, ASPE's report to Congress examined the relationships between several social risk factors, quality measures, and program performance for providers who treat a disproportionate share of socially at-risk patients. Specifically, social risk factors included dual enrollment in Medicare and Medicaid as a marker for low income, residence in a low-income area, Black race, Hispanic ethnicity, and residence in a rural area. The report evaluated nine Medicare value-based payment programs that were currently operational or defined in statute and incorporated quality and/or efficiency metrics into payment: the Hospital Readmissions Reduction Program (HRRP), Hospital Value-Based Purchasing Program, Hospital-Acquired Condition Reduction Program, Medicare Advantage Quality Star Rating Program, Medicare Shared Savings Program, Physician Value-Based Payment Modifier Program, End-Stage Renal Disease Quality Incentive Program, Skilled Nursing Facility Value-Based Purchasing Program, and Home Health Value-Based Purchasing Program.

The report was based on three areas of research and policy analysis: the effect of social risk factors on performance measures at the beneficiary and provider levels, the simulation of impacts of value-based purchasing programs on providers, and the evaluation of policy alternatives. An original objective for the Report was to make recommendations concerning whether quality measures, and value-based payment should be adjusted for social risk factors and, if so, what policy mechanism should be chosen.

Evidence: The Impact of Social Risk Factors on Quality Measures

The primary research was conducted internally by ASPE staff. The methods and results were vetted through both external statistical consultants and technical expert panels. The results at both at the beneficiary and provider levels were strong and consistent across all the programs studied:

1. Beneficiaries with social risk factors had poorer performance on process measures (e.g., cancer screening), clinical outcome measures (e.g., diabetes control, readmissions), safety (e.g., infection rates), and patient experience measures (e.g., communication from doctors and nurses), as well as higher resource use (e.g., higher spending per

hospital admission episode). This was true even when comparing beneficiaries within the same hospital, health plan, accountable care organization, physician group, or facility. Dual enrollment was typically the most powerful predictor of poor performance among those social risk factors examined. For the most part, these findings persisted after risk adjustment, across care settings, measure types, and programs, and were moderate in size.

2. In every care setting examined, providers that disproportionately cared for beneficiaries with social risk factors tended to perform worse than their peers on quality measures. Some of these differences were driven by beneficiary mix, but some of the difference persisted even after adjusting for beneficiary characteristics. The single exception was that accountable care organizations with a high proportion of dually enrolled beneficiaries were more likely to share in savings under the Medicare Shared Savings Program, despite slightly worse quality scores.

Table 1 summarizes patient-level results for HRRP. Dual eligibility for Medicare and Medicaid, a proxy for low income status, is the most powerful predictor among social risk factors. The raw odds of readmission are 24 percent higher for a dually enrolled individual relative to all other beneficiaries. The clinical risk adjusters used in HRRP reduce these odds by nearly half, to 13 percent greater. Further adjusting for other social risk factors and hospital characteristics reduces these odds to 10 percent.

Table 2 summarizes these results based on safety net status of the hospital. The raw odds of readmission for a patient in a safety net hospital were 9–20 percent higher than in other hospitals. Even after clinical risk adjustment, the odds ratios were 9–14 percent higher. Including patient social risk and hospital characteristics, the odds remained 5–9 percent higher.

Results for the eight other payment programs were consistent with the HRRP (HHS 2016). Dually eligible beneficiaries and safety net providers scored worse on most quality indicators used in value-based purchasing, even after any adjustments made (e.g., clinical risk) in calculating bonuses and penalties. For example, physician practices that treated larger shares of dually eligible beneficiaries were more likely to be penalized under the Physician Value-Based Payment Modifier Program.

It is important to note that these analyses cannot fully determine why patterns exist. Beneficiaries may have poorer outcomes due to higher levels of medical risk, worse living environments, greater challenges in adherence and lifestyle, bias, or discrimination. Providers may have poorer

Table 1 Patient-Level Odds of Readmission by Social Risk Factor

Social risk factor	Odds ratio for social risk factor		
	Unadjusted	Adjusted for comorbidities (age, kidney failure, heart failure, etc.)	Adjusted for comorbidities and other social risk factors
Dual status	1.24*	1.13*	1.10*
Low-income ZIP code	1.00*	1.00*	1.00
Black	1.10*	1.09*	1.01
Hispanic	1.10*	1.04	0.96
Urban	1.05*	1.04*	1.05*

*Significant at $p < 0.05$.**Table 2** Odds of Readmission by Safety-Net Hospitals

Readmission Cohort	Odds ratio for safety-net status (highest 20 percent of Disproportionate Share Hospitals index)		
	Unadjusted	Adjusted for comorbidities (age, kidney failure, heart failure, etc.)	Adjusted for comorbidities, patient social risk factors, and hospital characteristics
Acute Myocardial Infarction	1.20*	1.14*	1.09*
Heart Failure	1.12*	1.10*	1.06*
Pneumonia	1.12*	1.09*	1.05*
THA/TKA	1.09*	1.09*	1.09*
COPD	1.18*	1.12*	1.06*

*Significant at $p < 0.05$.

performance due to having fewer resources, more challenging clinical workloads, lower levels of community support, or worse quality. Many of these factors, for both beneficiaries and providers, are not easily measured with current data. Thus, assessing explained versus unexplained differences was also an issue. In the readmissions example, the 5–9 percent higher odds of readmission that remains after adjusting for the patient and hospital characteristics available to the researchers is unexplained, meaning that one-third to one-half of the raw odds of readmission remains

unexplained. Other methods of variance decomposition were employed yielding similar results. These include fixed effects to evaluate unspecified patient-, hospital-, and area-level influences, but a significant amount of variation remained even after these analyses.

The Financial Impact on Providers

The second stage of the research was to determine the actual financial impact on safety net providers of the Medicare value-based payment programs. The payment programs assess penalties and bonuses based on a variety of quality and performance measures. The process of converting performance on these measures to penalties and bonuses uses a variety of statistical and formula-based approaches, as well as applying upper and lower bounds specified in statutes and regulations. Depending on the specific method employed in each program, differences in financial consequences for providers may not be of the same magnitude as the differences in the performance measures themselves. In this second phase of the IMPACT work, program impacts were simulated to provide additional information for policy deliberations.

Across programs, the simulation results were directionally consistent with the findings from the quality measures, but the magnitude of financial effects was not as consistent. Safety net providers were more likely to face financial penalties across all five operational Medicare value-based purchasing programs in which penalties are assessed, including programs in the hospital, physician group, and dialysis facility settings. Plans serving duals were less likely to receive bonuses in Medicare Advantage. The differences in these financial impacts were generally smaller than the observed differences in the performance measures themselves.

Table 3 provides an example for the HRRP. Penalties for safety net and other hospitals were generally similar, despite the differences in readmission rates. Apparently, the exact methods of converting readmission rates into penalties compress them considerably. As discussed below, the contrast in these two sets of findings for the HRRP, which also exist in varying degrees in the other programs, can influence how the results are interpreted and how trade-offs with other policy objectives are navigated in the decision-making process.

Accounting for Social Risk in Payment: Impact of Alternatives

Several different options to account for social risk were considered and simulated, including direct adjustment for social risk factors, comparing

Table 3 Current HRRP Penalties by Safety Net Status

Hospital group	Hospitals penalized under the HRRP, 2015 (%)	Average HRRP penalty	
		% base diagnosis-related group (DRG) payments	US dollars
All hospitals	81	0.46	\$158,000
Safety-net hospitals (top 20 percent of Disproportionate Share Hospitals)	87	0.48	\$191,000
All other hospitals	80	0.45	\$150,000

providers only to those similar in terms of measures of patients' social risk (stratification), and conditioning any changes to penalties and bonuses on attainment and/or improvement. In general, the impact of implementing these options was small, although it varied somewhat among the payment programs. For example, the impact for the HRRP of these different approaches was to reduce average penalties for safety net hospitals approximately \$20,000–45,000.

To Adjust or Not to Adjust: Judgments on Policy Outcomes and Policy Context

The results described above were consistent across programs and performance measures and had a clear direction. It would be understandable if someone reviewing these data thought they created a strong rationale for policies that account for social risk. However, as foreshadowed by the conceptual model described above, a series of different factors played a role in how the evidence was interpreted and contextualized, especially with regard to weighing trade-offs among policy objectives.

Judgments on Policy Outcomes

Despite the consistency of the research findings, their application was controversial. Members of the leadership team differed in factors such as their experience with past payment policies, their expectations about the likelihood of unanticipated consequences, and their sensitivity to stakeholder concerns. Reflecting these differences, some members of the team thought there was a strong rationale for risk adjustment, whereas other

members viewed the impacts as small enough to merit careful monitoring in moving forward but not significant enough to take action.

Policy Context

A range of policy objectives came into play as the leadership team evaluated policy alternatives. These policy objectives included (1) protecting safety net providers from unfair financial harm, (2) preserving payment equity by accounting for cost/quality factors beyond the control of providers, (3) providing strong incentives to both improve overall quality of health care and reduce disparities, and (4) maintaining transparency of information. Moreover, the secretary of HHS had recently established goals for the share of Medicare payment that would flow through alternative payment models (Burwell 2015). Thus, encouraging providers to join new delivery system models that provide patient-centered, coordinated care across the full range of services was an additional objective.

Policy makers pay close attention to the way Medicare payment policy impacts providers that disproportionately serve low-income or vulnerable populations. Maintaining payment equity in a way that reflects the greater clinical or access needs of low-income or vulnerable populations has been a central tenant across Medicare's prospective payment systems. Policies such as the disproportionate share adjustment to hospital inpatient rates and the accommodations provided for many rural hospitals recognize the unique challenges that providers face related to their patient population and environment. Similar models could be used to account for social risk in payments or quality measures. The rationale for these policies to achieve payment equity is that they are necessary when deviations from a national standard (such as a group of providers whose costs are above national payment rate) are beyond providers' control but that providers should be financially accountable for variations resulting from factors considered within their control.

Although it was one of the policy objectives, existing data did not fully clarify the extent of disparities that were and were not under providers' control. The IMPACT research used measures of explained and unexplained variation as a guide. The explained variation decomposed into within provider and between provider variations, with the former used as a proxy for the portion of the estimated impact that was beyond providers' control. The assumption was that each provider rendered the same care to all of its patients, so differences in outcomes were due to differences between patients, particularly their social risk profiles. This proxy is imperfect at

best, as it is quite possible that there were differences in care provided to patients within the same setting. Furthermore, it is not possible to assess whether the unexplained variation was due to quality differences that could be addressed by providers or unmeasured factors beyond their control.

These uncertainties have potential implications for assessing adjustment options. For example, direct risk adjustment for the social risk factor(s) based on the statistical models would reflect explained variation and carry less risk of accounting for factors within the providers' control. Simulations revealed that direct risk adjustment would have a small impact on the financial consequences in most programs, meaning this alternative might not sufficiently protect safety net providers. Stratification simulations showed a larger protective effect, but this option might be accounting for factors within the providers' control.

Maintaining incentives for quality improvement was also important. There is growing evidence that value-based payment has resulted in quality improvements, particularly for providers that faced the largest penalties when programs began (Zuckerman et al. 2016). So, an important consideration was assuring that the financial consequences of these programs did not reduce safety net providers' ability to care for patients while also maintaining the incentives for continued quality improvement. Providers serving socially at risk patients clearly fared worse in value-based pricing programs, but the overall adverse financial impact was limited. For some members of the leadership team, the positive response to program incentives, combined with less threatening financial impacts, seemed to support taking a more conservative approach. That is, maintaining the status quo while monitoring current programs to ensure that any one of them, or the cumulative impact of several, did not significantly or unfairly impact safety net providers.

Another key policy objective in current law and policy implementation is transforming the delivery system to improve value. Currently, innovative payment models are being tested that are designed to encourage providers to be responsible for providing high-quality, well-coordinated care across the full range of health care services. Under these models, providers are financially accountable for both cost and quality. One concern about accounting for social risk is that reducing financial consequences for some providers would lessen incentives for them to join the new delivery models. To the extent that the most productive way to improve care for patients at social risk is creating networks of providers to coordinate care, reducing the incentive to enter these models might be harmful to socially at risk patients.

Table 4 summarizes the different policy options and objectives. The table was presented to key HHS decision makers as part of regular briefings

Table 4 Trade-offs among Policy Goals

Policy criteria	No adjustment	Direct adjustment	Stratification	Rewarding improvement	Targeted incentive payments
Protects safety net providers from unfair financial stress					
Adjusts only for what is beyond provider control					
Encourages reduction in disparities in access, quality, and outcomes					
Promotes transparency to facilitate consumer choice					
Does not discourage providers from joining Alternative Payment Models					

concerning the study's results and analysis of policy alternatives. Given all of the factors described above, the cells of the table were left empty so that individual decision makers could use their own weighting system to determine how they would evaluate the trade-offs among these objectives.

The Report's Policy Considerations

In fulfilling the IMPACT Act's mandate, the report to Congress included a number of "policy considerations" in lieu of recommendations. These considerations addressed the general question about whether to account for social risk in quality and cost measures used in value-based purchasing. The focus was on two separate ways in which measures are used: for reporting purposes and in calculating penalties and bonuses. These considerations were embedded in a three-part strategy:

1. *Performance on quality and outcomes should be measured and reported specifically for beneficiaries with social risk factors.* Doing so would allow policy makers and clinical leaders to identify, track, and address disparities in care.
2. *High-, fair quality standards should be set for all beneficiaries.* Whether the "fairest" standard is one that does or does not adjust for social risk will depend on the type of measure and how the considerations outlined earlier apply to that particular measure. Additionally, all measures should be studied to determine whether accounting for frailty, medical complexity, functional status, or other factors might improve their ability to fairly and accurately assess provider performance.
3. *Reward and support better outcomes.* Meeting quality standards, particularly for outcome measures, may be harder for beneficiaries with social risk factors, who face specific challenges to achieving good health outcomes. Therefore, value-based purchasing programs should do the following:
 - a. Provide specific payment adjustments to reward achievement and/or improvement for beneficiaries with social risk factors. Leveraging the power of value-based purchasing to provide specific payment adjustments to reward providers for successfully achieving high quality and/or good health outcomes in beneficiaries with social risk factors may provide important incentives to focus on these individuals and help offset any real or perceived disincentives to caring for them.

- b. Where feasible, provide targeted support for providers who disproportionately serve them. Providing targeted support, for example, through quality improvement programs designed specifically for beneficiaries with social risk factors, is also critical to ensuring that all beneficiaries can have the best health outcomes possible. Another key component of support is ensuring that current base payments are adequate to support high-quality care for beneficiaries with social risk factors.

Did the Policy Considerations Reflect the Evidence?

The IMPACT report to Congress represented cutting-edge research that yielded strong and consistent results to inform the policy process. The subsequent policy deliberations sought to maximize the criteria listed in Table 4 by crafting broad policy considerations.

At first glance, some might take the view that evidence did not significantly affect the decision-making process. After all, a seemingly high-quality research effort did not result in a simple answer to a rather straightforward policy question: should value-based purchasing programs account for social risk factors? As we have attempted to describe in this article, the policy considerations arose from a deliberative process that carefully weighed all of the evidence with a number of other relevant policy considerations and contexts.

A strong case can be made that the evidence indeed heavily influenced the policy deliberations in two ways. First, the small financial impacts on safety net providers of the current programs and potential policy changes provided some basis for a more cautious, or “watchful-waiting,” policy rather than immediately accounting for social risk factors. This approach was bolstered by concerns that accounting for social risk might reduce providers’ incentives to improve quality and engage in new delivery models. The results do indicate that these providers may face larger financial problems as value-based purchasing expands, a problem that should be monitored with an eye toward policy modification as needed.

The second and perhaps most influential research finding was the magnitude and consistency of the relationship between social risk factors and a variety of cost and quality measures across all of the payment programs. These results suggest socially at-risk patients receive systematically worse-quality care with subsequent poor outcomes. It is apparent that there is a much larger and more serious quality problem that should be addressed to meet overall health system quality goals. Simply making

direct adjustments for social risk to measures or payments might provide some short-run financial relief to some providers but would do very little to address the core problem. In brief, the policy considerations reflect a holistic rather than a Band-Aid approach to addressing quality issues.

Lessons Learned

Perhaps the overall lesson in the IMPACT report experience is that high-quality evidence is important to policy consideration but represents only one of many factors used in making a decision. It is also useful to distinguish between the research question and the policy question as we consider evidence-informed decision making. In this case the research question was clear: is there a relationship between social risk factors and performance on the metrics in Medicare payment programs at the patient and provider level? The empirical answer was yes, there was a clear and consistent relationship across all of the programs studied. The original policy question addressed by the report was how we should account for social risk factors in Medicare's value-based purchasing programs. The policy answer is yet to be fully formulated. It involves a careful balancing of the evidence provided, evidence that will be forthcoming in a future report, and the many other policy factors described in this article.

The related lesson learned is that good evidence can illustrate a problem in a way that can help reformulate policy questions. In this case, there was consensus that a very important policy question should be, how do we improve the systematically lower-quality care that patients at social risk receive compared to other patients?

Another lesson that researchers can learn from is that occasionally policy making follows its own timetable due to factors unrelated to evidence. Many stakeholders may play a role in defining a problem and movement toward possible policy solutions. Again, the social risk and value-based purchasing issue provides an excellent example. In 2013, prior to availability of data on the actual financial impact of HRRP, the Medicare Payment Advisory Commission (MedPAC) registered its concern about safety net hospitals by recommending that the penalties be calculated based on peer group comparisons or "stratification" (MedPAC 2013). The peer groups would be based on hospitals' share of low-income Medicare beneficiaries. Since members of Congress were hearing concerns from hospitals in their districts, the MedPAC recommendation was drafted into proposed legislation. The Twenty-First Century Cures Act passed in December 2016 provided a convenient legislative vehicle for the proposal. Ironically, the bill passed just weeks before delivery of the ASPE report that Congress had requested to inform

the legislative efforts. The report did not recommend this approach over others, and its evidence suggests the peer group approach's actual beneficial financial impact for safety net providers will be very small. These events illustrate that stakeholder concerns and the timing of legislative vehicles can sometimes require action before the relevant evidence is available.

Conclusions

By focusing on the IMPACT report to Congress as a case study, we provide a framework to illuminate the use of research evidence in policy decision making. In particular, we highlight the complex set of factors considered with research results in an evidence-informed policy process.

In summary, we found large differences in performance on quality measures when we compared safety-net and non-safety-net providers but more modest differences in payment effects. Numerous policy assessments and trade-offs were considered. Our consensus was that addressing the social risk issue with a narrow set of adjustments for quality measures or payment would not really address the underlying problem. Rather, we recommend a wide-ranging set of policy considerations to address broad disparities in the processes and outcomes of care. Evidence played a key role in the deliberations.

In large organizations such as HHS there are multiple components with differing functions. The responsibilities attached to these parts of the organization may influence how their decision makers weigh the research evidence and all of the other relevant factors discussed in this article in terms of formulating policy preferences. Thus, it is possible that the same body of research results in different preferences across the organization, meaning that finding internal consensus during policy deliberations is critical. But one would need to look no further than the authors of this article to witness such heterogeneity in views. We represent an interdisciplinary team that participated fully in the planning, research, and deliberations that led to the report's policy considerations. We reflect different degrees of government, academic, and clinical experience. We entered the process with differing views on the appropriateness of adjusting quality measures for social risk. It is safe to say that, after months of analyzing the data together, we still had some differences in our policy preferences. But that may make the point of this article more clearly than anything else: our different background and experiences lead us to interpret the implications for policy of all of the research results somewhat differently. In this case, however, consensus policy considerations that we were all comfortable with did emerge.

■ ■ ■

Steven Sheingold is an economist and director of the Division of Health Financing Policy in the Office of the Assistant Secretary of Planning and Evaluation in the US Department of Health and Human Services. He conducts economic and policy research on the health care markets as well as the Medicare program, including payment policy, the role of social risk in Medicare value-based purchasing programs, the Hospital Readmissions Reduction Program, Part D, and implementation of the new physician payment systems mandated by the Medicare Access and CHIP Reauthorization Act of 2015.

steven.sheingold@hhs.gov

Karen E. Joynt Maddox is a practicing cardiologist at Barnes-Jewish Hospital and assistant professor at Washington University School of Medicine and School of Social Work, as well as a health policy adviser for BJC Healthcare, a fifteen-hospital system in Missouri and Illinois. She served from 2014 to 2016 as senior adviser in the Office of Health Policy, US Department of Health and Human Services. She is a health services and health policy researcher whose main areas of focus are quality measurement, value-based and alternative payment models, and disparities in care.

Rachael B. Zuckerman is a social scientist analyst in the Office of the Assistant Secretary of Planning and Evaluation in the US Department of Health and Human Services. She conducts policy research on the Medicare program, including current projects on the role of social risk in Medicare value-based purchasing programs, the Hospital Readmissions Reduction Program, and implementation of the new physician payment systems mandated by the Medicare Access and CHIP Reauthorization Act of 2015.

Nancy De Lew is the associate deputy assistant secretary for health policy in the Office of the Assistant Secretary for Planning and Evaluation at the US Department of Health and Human Services. She serves as the senior career official in the Office of Health Policy, where she leads a team who apply their skills in policy development, strategic planning, research, and evaluation to some of the department's most challenging health policy problems. She provides executive leadership and coordination on a broad range of health care financing, coverage, access, public health, and quality issues.

Arnold M. Epstein is the John H. Foster Professor of Health Policy and Management and chair of the Department of Health Policy and Management at Harvard Chan School of Public Health. His research interests focus on quality of care and access to care. During 1993–94 he served in the Clinton administration and in 2014–16 he served as deputy assistant secretary in the US Department of Health and Human Services. He received the Distinguished Investigator Award from AcademyHealth in 2015. He is an associate editor at the *New England Journal of Medicine* and is member of the National Academy of Medicine and the Association of American Physicians.

References

- Brownson, Ross C., Jamie F. Chiqui, and Katherine A. Stamatakis. 2009. "Understanding Evidence-Based Public Health Policy." *American Journal of Public Health* 99, no. 9: 1576–83. doi: 10.2105/ajph.2008.156224.
- Burwell, Sylvia M. 2015. "Setting Value-Based Payment Goals—HHS Efforts to Improve U.S. Health Care." *New England Journal of Medicine* 372, no. 10: 897–99. doi: 10.1056/NEJMp1500445.
- Cairney, Paul. 2015. *The Politics of Evidence Based Policy Making*. London: Palgrave Macmillan.
- Cairney, Paul, and Kathryn Oliver. 2017. "Evidence-Based Policymaking Is Not like Evidence-Based Medicine, So How Far Should You Go to Bridge the Divide between Evidence and Policy?" *Health Research Policy and Systems* 15, no. 1: 35. doi: 10.1186/s12961-017-0192-x.
- Califf, Robert M., Melissa A. Robb, Andrew B. Bindman, Josephine P. Briggs, Francis S. Collins, Patrick H. Conway, Trinkia S. Coster, Francesca E. Cunningham, Nancy De Lew, Karen B. DeSalvo, Christine Dymek, Victor J. Dzau, Rachael L. Fleurence, Richard G. Frank, J. Michael Gaziano, Petra Kaufmann, Michael Lauer, Peter W. Marks, J. Michael McGinnis, Chesley Richards, Joe V. Selby, David J. Shulkin, Jeffrey Shuren, Andrew M. Slavitt, Scott R. Smith, B. Vindell Washington, P. Jon White, Janet Woodcock, Jonathan Woodson, and Rachel E. Sherman. 2016. "Transforming Evidence Generation to Support Health and Health Care Decisions." *New England Journal of Medicine* 375, no. 24: 2395–2400. doi: 10.1056/NEJMs1610128.
- Committee on Accounting for SES in Medicare Payment Programs. 2016. *Accounting for Social Risk Factors in Medicare Payment: Identifying Social Risk Factors*. Washington, DC: National Academies of Sciences, Engineering, and Medicine.
- Davis, Matthew M., Cary P. Gross, and Carolyn M. Clancy. 2012. "Building a Bridge to Somewhere Better: Linking Health Care Research and Health Policy." *Health Services Research* 47, no. 1, pt. 2: 329–36. doi: 10.1111/j.1475-6773.2011.01373.x.
- Desai, Nihar R., Joseph S. Ross, Ji Young Kwon, Jeph Herrin, Kumar Dharmarajan, Susannah M. Bernheim, Harlan M. Krumholz, and Leora I. Horwitz. 2016. "Association between Hospital Penalty Status under the Hospital Readmission Reduction Program and Readmission Rates for Target and Nontarget Conditions." *Journal of the American Medical Association* 316, no. 24: 2647–56. doi: 10.1001/jama.2016.18533.
- Gilman, Matlin, E. Kathleen Adams, Jason M. Hockenberry, Arnold S. Milstein, Ira B. Wilson, and Edmund R. Becker. 2015a. "Safety-Net Hospitals More Likely than Other Hospitals to Fare Poorly under Medicare's Value-Based Purchasing." *Health Affairs* 34, no. 3: 398–405. doi: 10.1377/hlthaff.2014.1059.
- Gluck, Michael E. 2015. "Moving Health Services Research into Policy and Practice: Lessons from Inside and Outside the Health Sector." *AcademyHealth*.
- Gold, Marsha. 2009. "Pathways to the Use of Health Services Research in Policy." *Health Services Research* 44, no. 4: 1111–36. doi: 10.1111/j.1475-6773.2009.00958.x.

- Hardee, Karen, Kelsey Wright, and Joanne Spicehandler. 2015. "Family Planning Policy, Program and Practice Decisionmaking: The Role of Research Evidence and Other Factors." Working paper/technical report. Washington, DC: Population Council. doi: 10.13140/RG.2.1.1379.8883.
- HHS (US Department of Health and Human Services). 2016. *Report to Congress: Social Risk Factors and Medicare's Value Based Purchasing Programs*. Washington, DC: Assistant Secretary for Planning and Evaluation, HHS.
- Joynt, Karen E., and Ashish K. Jha. 2013. "Characteristics of Hospitals Receiving Penalties under the Hospital Readmissions Reduction Program." *Journal of the American Medical Association* 309, no. 4: 342–43. doi: 10.1001/jama.2012.94856.
- Lavis, John N., Andrew D. Oxman, Simon Lewin, and Atle Fretheim. 2009. "SUPPORT Tools for Evidence-Informed Health Policymaking (STP) 1: What Is Evidence-Informed Policymaking?" *Health Research Policy and Systems* 7 (suppl. 1): S1. doi: 10.1186/1478-4505-7-s1-s1.
- Liang, Zhanming, Peter F. Howard, Sandra G. Leggat, and Gregory Murphy. 2012. "A Framework to Improve Evidence-Informed Decision-Making in Health Service Management." *Australian Health Review* 36, no. 3: 284–89. doi: 10.1071/ah11051.
- MedPAC (Medicare Payment Advisory Commission). 2013. *Report to the Congress: Medicare and the Health Care Delivery System*. Washington, DC: MedPAC.
- National Quality Forum. 2014. "Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors: A Technical Report." August 15. www.qualityforum.org/Publications/2014/08/Risk_Adjustment_for_Socioeconomic_Status_or_Other_Sociodemographic_Factors.aspx.
- Oliver, Kathryn, Theo Lorenc, and Simon Innvaer. 2014. "New Directions in Evidence-Based Policy Research: A Critical Analysis of the Literature." *Health Research Policy and Systems* 12, no. 34. doi: 10.1186/1478-4505-12-34.
- Peterson, Mark A. 2017. "When and How Health Policy Research Matters: A Framework for Analysis." Paper presented at the *Journal of Health Politics, Policy and Law*/NYU Robert F. Wagner Graduate School of Public Services Conference on Policy Analysis and the Politics of Health Policy: Scholarship, Knowledge Translation, and Policymaking, New York, NY, May 2.
- Rajaram, Ravi, Jeanette W. Chung, Christine V. Kinnier, Cynthia Barnard, Sanjay Mohanty, Emily S. Pavey, Megan C. McHugh, and Karl Y. Bilimoria. 2015. "Hospital Characteristics Associated with Penalties in the Centers for Medicare and Medicaid Services Hospital-Acquired Condition Reduction Program." *Journal of the American Medical Association* 314, no. 4: 375–83. doi: 10.1001/jama.2015.8609.
- Raudenbush, Stephen W., and Howard S. Bloom. 2015. "Learning about and from a Distribution of Program Impacts Using Multisite Trials." *American Journal of Evaluation* 36, no. 4: 475–99.
- Wasserstein, Ronald L., and Nicole A. Lazar. 2016. "The ASA's Statement on p-Values: Context, Process, and Purpose." *American Statistician* 70, no. 2: 129–33. doi: 10.1080/00031305.2016.1154108.
- Zuckerman, Rachael B., Steven H. Sheingold, E. John Orav, Joel Ruhter, and Arnold M. Epstein. 2016. "Readmissions, Observation, and the Hospital Readmissions Reduction Program." *New England Journal of Medicine* 374, no. 16: 1543–51. doi: 10.1056/NEJMsa1513024.

Appendix: Resources Related to ASPE's Report to Congress

ASPE's Report to Congress:

aspe.hhs.gov/pdf-report/report-congress-social-risk-factors-and-performance-under-medicares-value-based-purchasing-programs

Statute Authorizing the Report:

www.congress.gov/bill/113th-congress/house-bill/4994

NAM Report:

www.nap.edu/catalog/23635/accounting-for-social-risk-factors-in-medicare-payment

Four Constituent Reports:

www.nap.edu/catalog/21858/accounting-for-social-risk-factors-in-medicare-payment-identifying-social

www.nap.edu/catalog/23513/accounting-for-social-risk-factors-in-medicare-payment-criteria-factors

www.nap.edu/catalog/23605/accounting-for-social-risk-factors-in-medicare-payment-data

www.nap.edu/catalog/21914/systems-practices-for-the-care-of-socially-at-risk-populations

NAM Summary of Dissemination Meeting for the Report:

www.nap.edu/resource/23635/SES-and-Medicare-Dissemination-Meeting.pdf

National Quality Forum Project on Risk Adjustment for Sociodemographic Factors:

www.qualityforum.org/Risk_Adjustment_SES.aspx