

Collateral Damage: Pay-for-Performance Initiatives and Safety-Net Hospitals

A century ago, Abraham Flexner's effort to upgrade medical education closed underresourced schools that provided almost all of the training opportunities available to black and female physicians. Six decades passed before medical schools—prodded by the civil rights and women's movements—began restoring opportunities for these groups.

In this issue, Gilman and colleagues (1) document an analogous side effect of Medicare's pay-for-performance (P4P) initiatives: They divert money from underresourced safety-net hospitals that are mainstays of care in many minority communities to hospitals serving more affluent patients.

This worrisome finding is not unexpected. Even excellent physicians who care for disadvantaged patients often score low on P4P metrics (2). Patients from disadvantaged backgrounds rate their hospital care lower (3) and have higher readmission rates (4).

Medicare's P4P program, which does not adjust for patients' socioeconomic status, assumes that bonuses and penalties will prod substandard providers to improve or see their patients migrate to higher-quality options. However, when quality problems are due to a hospital's financial distress and patients cannot go elsewhere, penalizing low scorers may well punish patients and exacerbate quality disparities. Prescribing a starvation diet for safety-net hospitals that are strapped for cash and are quality challenged makes no sense unless the goal is to close them.

Gilman and colleagues' analysis suggests that the amount diverted from safety-net hospitals so far is modest. However, their estimate is conservative, even modest sums may be important for low-margin institutions, and the safety-net penalty seems certain to grow. To categorize hospitals as "safety-net," Gilman and colleagues relied on standard but flawed Centers for Medicare & Medicaid Services metrics that heavily weight a hospital's Medicaid load. Hospitals caring for many nursing home residents dually covered by Medicare and Medicaid could meet this standard, even if the hospitals shunned the uninsured. True safety-net hospitals do more than care for Medicaid patients. They have an open-door policy for patients lacking access to care elsewhere, including patients with unprofitable diagnoses, such as mental illness; racial and linguistic minorities; and, most importantly, the uninsured. Because Gilman and colleagues' safety-net group was diluted by facilities for which that label was inappropriate, they probably understated P4P's financial toll on true safety-net providers. Moreover, both Medicare and private insurers are set to ratchet up P4P rewards and penalties, even as safety-net hospitals struggle with reductions mandated by the Patient Protection and Affordable

Care Act in the disproportionate share funding on which they have long relied.

Are P4P's benefits worth the risk? The evidence is surprisingly slim. A few small, randomized, controlled trials in outpatient settings have shown improvement on surrogate measures, but most have found no improvement, and none have demonstrated reductions in death or disability rates.

No randomized, controlled trials have examined P4P's effect on inpatient care, but the 2 large-scale, quasi-experimental studies of hospital P4P yielded disappointment. In the Centers for Medicare & Medicaid Services' Premier Hospital Quality Incentive Demonstration, the process-based quality indicators at the 200 P4P hospitals initially improved faster than at control hospitals, but the differences subsequently evaporated and patient outcomes did not improve at all (5). An initial evaluation of an English hospital P4P program suggested a mortality benefit (6), but a longer-term follow-up found no improvement (7). Disturbingly, another English study suggested that incentives to shorten surgical queues worsened myocardial infarction mortality rates (8). Diverting resources to elective surgery cases may have distracted from emergency care.

After Medicare began penalizing hospitals for high 30-day readmission rates for patients with pneumonia, heart failure, and acute myocardial infarction, readmissions for those conditions decreased by 1% to 2%. However, the all-cause readmission rate did not budge. Moreover, hospitals' "gaming" of the readmission metric probably accounts for much of the slight improvement for the 3 target conditions. "Observation" stays, which Medicare does not count as readmissions, have increased sharply and now account for approximately 10% of all hospital stays occurring within 30 days of discharge. Recently discharged patients' emergency department visit rates (without admission) have also spiked.

Safety-net providers spend less on administration and may lack the administrative resources to keep up in the gaming arms race—the health care equivalent of teaching to the test. Expensive computer systems and consultants facilitate aggressive upcoding and meticulous documentation of comorbid conditions that can exaggerate patients' severity of illness and inflate risk-adjusted quality scores. Costly administrative efforts can also ferret out exceptions and ensure that boxes are checked (for example, "Yes, we counseled against smoking"), increasing process-of-care scores, regardless of whether they improve care.

With vast sums at stake, gaming has become ubiquitous. Researchers who used external data to audit Medicare Advantage plans' "safe prescribing" quality

reports discovered that 95% of plans overstate their scores, pushing down the rankings of the 5% who told the truth (9). At a hospital where we worked, the administration responded to poor scores for risk-adjusted mortality rates by hiring coding consultants, even as it was laying off clinical staff. Within 6 months, risk-adjusted mortality rates improved to better than expected rates and Medicare reimbursement climbed \$3 million. Soon after, administrators made physician attendance at coding-coaching sessions mandatory.

Measuring quality is hard under the best of circumstances. Hospitals that look good according to one risk-adjustment algorithm often seem downright dangerous according to another. For instance, in the analysis by Gilman and colleagues, safety-net hospitals scored better on mortality despite scoring poorly on process-of-care measures. The use of quality metrics for financial reward or punishment can render the clinical data needed for real quality improvement no more accurate than a tax return.

The recent sustainable growth rate fix, which will tie 18% of Medicare physician fees to cost and quality performance metrics, will intensify P4P pressures, paperwork, and confusion (given the complexity of figuring out which of a patient's many physicians is culpable for excessive costs or quality deficits).

Tethering physicians' rewards to box checking and redundant documentation risks both substituting insurers' priorities for patients' goals and demoralizing physicians. Pay for performance can crowd out intrinsic motivation that keeps us doing good work even when no one is looking. A growing body of behavioral economics research indicates that when preexisting motivation is high, monetary incentives often undermine performance on complex cognitive tasks—particularly when incentives are contingent on specific task performance or are associated with surveillance, deadlines, or threats (10).

Paying for quality has strong intuitive appeal. However, as with other medical interventions, intuition may mislead, and adopting everywhere policies that have been proven nowhere puts millions at risk for unintended side effects.

Steffie Woolhandler, MD, MPH

David U. Himmelstein, MD

City University of New York School of Public Health at
Hunter College
New York, New York

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-1393.

Requests for Single Reprints: Steffie Woolhandler, MD, MPH, City University of New York School of Public Health at Hunter College, 255 West 90th Street, New York, NY 10024; e-mail, swoolhan@hunter.cuny.edu.

Current author addresses are available at www.annals.org.

Ann Intern Med. 2015;163:473-474. doi:10.7326/M15-1393

References

1. Gilman M, Hockenberry JM, Adams EK, Milstein AS, Wilson IB, Becker ER. The financial effect of value-based purchasing and the Hospital Readmissions Reduction Program on safety-net hospitals in 2014. A cohort study. *Ann Intern Med.* 2015;163:427-36. doi:10.7326/M14-2813
2. Hong CS, Atlas SJ, Chang Y, Subramanian SV, Ashburner JM, Barry MJ, et al. Relationship between patient panel characteristics and primary care physician clinical performance rankings. *JAMA.* 2010;304:1107-13. [PMID: 20823437] doi:10.1001/jama.2010.1287
3. McFarland DC, Ornstein KA, Holcombe RF. Demographic factors and hospital size predict patient satisfaction variance-implications for hospital value-based purchasing. *J Hosp Med.* 2015. [PMID: 25940305] doi:10.1002/jhm.2371
4. Kind AJ, Jencks S, Brock J, Yu M, Bartels C, Ehlenbach W, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. *Ann Intern Med.* 2014;161:765-74. [PMID: 25437404] doi:10.7326/M13-2946
5. Jha AK, Joynt KE, Orav EJ, Epstein AM. The long-term effect of premier pay for performance on patient outcomes. *N Engl J Med.* 2012;366:1606-15. [PMID: 22455751] doi:10.1056/NEJMsa1112351
6. Sutton M, Nikolova S, Boaden R, Lester H, McDonald R, Roland M. Reduced mortality with hospital pay for performance in England. *N Engl J Med.* 2012;367:1821-8. [PMID: 23134382] doi:10.1056/NEJMsa1114951
7. Kristensen SR, Meacock R, Turner AJ, Boaden R, McDonald R, Roland M, et al. Long-term effect of hospital pay for performance on mortality in England. *N Engl J Med.* 2014;371:540-8. [PMID: 25099578] doi:10.1056/NEJMoa1400962
8. Propper C, Burgess S, Gossage D. Competition and quality: evidence from the NHS internal market 1991-9. *Economic J.* 2008;118:138-70.
9. Cooper AL, Kazis LE, Dore DD, Mor V, Trivedi AN. Underreporting high-risk prescribing among Medicare Advantage plans: a cross-sectional analysis. *Ann Intern Med.* 2013;159:456-62. [PMID: 24081286] doi:10.7326/0003-4819-159-7-201310010-00005
10. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull.* 1999;125:627-68. [PMID: 10589297]

Annals of Internal Medicine

Current Author Addresses: Drs. Woolhandler and Himmelstein: City University of New York School of Public Health at Hunter College, 255 West 90th Street, New York, NY 10024.