Commentary

Using bioinformatics to keep “pay for performance” fair: a modest proposal

SHERMAN C. STEIN, M.D.

University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

The essayist Jonathan Swift was distressed by the inhumane living conditions of the Irish lower classes. He had written a series of pamphlets, *The Drapier's Letters*, in which he requested help from their English landlords. Appalled at the lack of response to his writings and by the worsening starvation, he wrote “A Modest Proposal” in 1729. In this satirical essay, Swift used savage irony to secure a response from the English overlords. He sarcastically proposed that the poor sell their children to the English to be eaten, thus simultaneously reducing the population and enriching the survivors.

I may be overstating my case by likening the condition of organized neurosurgery to that of the Irish poor, and in so doing I do not wish in any way to demean their suffering. I possess neither Swift's moral outrage nor his wit, and I shall not stress the parallel too far. Nevertheless, both groups share a sense of helplessness in the face of contending forces. In the case of neurosurgery, among those forces are pressure to improve outcomes while simultaneously reducing costs.⁴

The most recent development is the effort to establish pay for performance (P4P) standards, in which an individual physician's reimbursement would be tied to patient outcomes, as judged by the physician's ability to reach performance targets. These targets relate to both the quality and cost of care. Following a proposal by the Institute of Medicine in 2001,²² P4P has been gaining traction. The Medicare Payment Advisory Commission, in its March 2005 report to Congress, recommended implementation of P4P in Medicare, and P4P plans by private insurers are burgeoning.^{7,27} Although P4P was originally limited to rating group practices and Health Maintenance Organizations, its scope has expanded to cover individual physicians; neurosurgeons will be affected in the near future.⁹ Some third-party payers have announced plans to make physician “report cards” public.³²

The definitions and metrics for quality of care are well established, all using a quantitative estimation of the success of an individual medical intervention.^{12,14,15,17,25} A collection instrument is constructed to organize data, as provided by the patient's insurance claim and supplemented as needed by the provider. The patient's health outcome is given a numerical value, and (in some systems) is adjusted for the processes of care, structural and efficiency measures, patient preferences, and satisfaction with care delivery. The result is compared with a predetermined norm, and the treating physician is graded accordingly. How are the benchmarks determined? Most sponsors assemble expert panels to establish practice guidelines by consensus. Popular evaluation processes such as the Health Plan Employer Data and Information Set (HEDIS)²⁶ limit the guidelines to those based primarily on process, structural, and efficiency measures. The resulting rating is calculated on how closely the outcome follows the guidelines; the algorithm is often proprietary and not made public. As currently practiced by a number of P4P programs, the provider's score accounts for as much as 20% of his or her income.⁹

The P4P program has the potential to inflict great harm on the neurosurgical profession. It was first introduced in industry and commerce to reduce overhead costs and maximize return on investment.²⁸ These same motivations are evident in the promotion of P4P for the healthcare field,^{11,34} although the effort has been somewhat tempered in recent years by a recognition that quality improvement must be emphasized.² There is an inherent conflict between improving quality and reducing cost in healthcare.²⁴ There can be little doubt on which side of this conflict third-party payers stand. Some investigators have even pointed to an underlying ethical dilemma in quality improvement; restrictions implemented in the name of cost containment represent medical experimentation without informed consent.²³

Will P4P work in medicine? There is little evidence that

it works in industry.³⁰ Indeed, one could argue that P4P for top executives is behind the many recent corporate scandals. There is the risk that such programs emphasize ratings at the expense of patient needs, teamwork, the reporting of deficiencies, and opportunities for improvement.¹⁰ Physicians attempting to meet targets will be more inclined to avoid certain patients and perhaps stretch indications to treat others.

It has been emphasized that the quality of care can be improved only in a nonpunitive environment, in which policy changes are coordinated among all the stakeholders.^{5,6,13} Rating physicians, especially if the scores are made public, may have a number of unintended consequences, among them reduction in quality, rather than its improvement.³³ The innovation and broadening of therapeutic indications for established procedures to high-risk patients will undoubtedly suffer.¹⁰ Of course, P4P has already been tried in medicine: it failed in Indonesia,¹⁶ and exhibited mixed results in Italy and Great Britain.^{1,3}

What about the unique situation of neurosurgery? We are a specialty with a small number of providers, each of whom performs a relatively low volume of procedures. A single adverse outcome could affect one's rating for years, thus a low rating for a neurosurgeon in an underserved area could imperil local healthcare delivery. Clinical research and innovation will inevitably suffer, as may clinical judgment itself. It is one thing for HEDIS to set standards for beta blockers after myocardial infarction, a situation in which there is little controversy. Treatment of cervical radiculopathy is quite another matter, as evidence-based guidelines are lacking. One can imagine a scenario in which practicing neurosurgeons are so risk averse that they will limit the types of procedures they perform and restrict treatment to only the youngest and healthiest patients. Once they begin to keep one eye on the ratings of their competitors and adjust their practices to maximize their own scores, it will truly be a Swiftian setting.

It is unlikely that physicians can avert P4P. Nor, if past experience is any indicator, will we be able to alter its implementation or the rules of the game.²⁹ What we should and must have a say in are the outcomes data and their analysis. We must exploit bioinformatics to study outcomes and benchmarks independent of the third party payers. Only then can our patients and we be assured that quality is being emphasized over profits.

Dr. Harbaugh, who has unparalleled credentials in the field of outcomes and quality assurance in neurosurgery,^{19,20} has recommended a voluntary outcomes database derived from the American Board of Neurological Surgery case log.¹⁸ Neurolog is an Internet-based data collection instrument that can be adapted readily for outcomes reporting. Nevertheless, it could be argued (and probably will be by government agencies and third-party payers) that this system is time-consuming and prone to compliance and accuracy issues. The insurance industry is likely to rely primarily on claims data, which it already collects. My suggestion is simply that, whatever the source of the information, it be assembled in a standard format and entered into a national, federally supervised database. The data can be made anonymous, with references to patient, physician, and hospital identities removed. The information should then be made publicly available for all investigators. We can independently apply metaanalytic tools^{8,21,31} to pool data for out-

comes research, decision-making, and cost-effectiveness studies. This will permit us to monitor and, when necessary, credibly to challenge standards imposed from without.

CONCLUSIONS

On the question of whether the powers that be will give us access to such data, I am reminded again of Swift's essay. Denied access to education, few Irish could even read "A Modest Proposal." By contrast, the essay was widely read by the English, who found it extraordinarily amusing. Nothing changed.

References

- Adinolfi P: Performance-related pay for health service professionals: the Italian experience. **Health Serv Manage Res** 11: 211–220, 1998
- Anonymous: Changing the dial on quality: what will pay-for-performance reimbursement strategies mean for providing better care? **Qual Lett Healthc Lead** 17:2–8, 2005
- Arrowsmith J, French S, Gilman M, et al: Performance-related pay in health care. **J Health Serv Res Policy** 6:114–119, 2001
- Asmus JJ: Why U.S. Neurosurgeons are in trouble: competition, cost containment, and caring. **Surg Neurol** 56:137–138, 2001
- Baker GR, Norton P: Making patients safer! Reducing error in Canadian healthcare. **Healthc Pap** 2:10–31, 2001
- Ball MJ, Douglas JV: IT, patient safety, and quality care. **J Healthc Inf Manag** 16:28–33, 2002
- Benko LB: Blues primed to pay. 29 plans are already paying for performance. **Mod Healthc** 35:16, 2005
- Berman NG, Parker RA: Meta-analysis: neither quick nor easy. **BMC Med Res Methodol** 2:10, 2002
- Bertagna M, Gliklich R: P4P for specialists gains velocity. **AANS Bulletin** 14(1):18–19, 2005
- Berwick DM: The toxicity of pay for performance. **Qual Manag Health Care** 4:27–33, 1995
- Bledsoe DR, Leisy WB, Rodeghero JA: Tying physician incentive pay to performance. **Healthc Financ Manage** 49:40–44, 1995
- Blumenthal D: Part 1: Quality of care—what is it? **N Engl J Med** 335:891–894, 1996
- Bonney E, Baker GR: Current strategies to improve patient safety in Canada: an overview of federal and provincial initiatives. **Healthc Q** 7:36–41, 2004
- Brook RH, McGlynn EA, Cleary PD: Quality of health care. Part 2: measuring quality of care. **N Engl J Med** 335:966–970, 1996
- Chassin MR: Quality of health care. Part 3: improving the quality of care. **N Engl J Med** 335:1060–1063, 1996
- Chernichovsky D, Bayulken C: A pay-for-performance system for civil service doctors: the Indonesian experiment. **Soc Sci Med** 41:155–161, 1995
- Donabedian A: An assessment of the criteria approach to medical care evaluation. **Hosp Med Staff** 5:1–9, 1976
- Harbaugh RE: Characterizing quality. The opportunity of outcomes research. **AANS Bulletin** 14(1):14–16, 2005
- Harbaugh RE: History of outcomes measures in neurosurgery. **Neurosurg Clin N Am** 12:217–221, 2001
- Harbaugh RE: Quality assurance in neurosurgery: United States concepts. **Acta Neurochir Suppl** 78:53–58, 2001
- Hasselblad V, McCrory DC: Meta-analytic tools for medical decision making: a practical guide. **Med Decis Making** 15: 81–96, 1995
- Institute Of Medicine: **Crossing The Quality Chasm. A New Health System for the 21st Century**. Washington, DC: National Academy Press, 2001
- Kofke WA, Rie MA: Research ethics and law of healthcare sys-

Using bioinformatics to keep pay-for-performance fair

- tem quality improvement: the conflict of cost containment and quality. **Crit Care Med** 31 (Suppl 3):S143–S152, 2003
24. McLoughlin V, Leatherman S: Quality or financing: what drives design of the health care system? **Qual Saf Health Care** 12:136–142, 2003
25. Merry MD, Crago MG: The past, present and future of health care quality. Urgent need for innovative, external review processes to protect patients. **Physician Exec** 27:30–35, 2001
26. National Committee for Quality Assurance: **The Health Plan Employer Data and Information Set (HEDIS)**. (<http://www.ncqa.org/Programs/HEDIS/index.htm>) [Accessed 14 August 2005]
27. O'Hare PK: Pay for performance: will your hospital be ready? **Healthc Financ Manage** 59:46–48, 2005
28. Rock RH: Pay for performance: accent on standards and measures. **Compens Rev** 16:15–25, 1984
29. Romano M: Performance anxiety. Is it too late for physicians to set their own rules on pay-for-performance? **Mod Healthc** 35: 6–7, 14, 2005
30. Rynes SL, Gerhart B, Parks L: Personnel psychology: performance evaluation and pay for performance. **Annu Rev Psychol** 56:571–600, 2005
31. Stroup DF, Berlin JA, Morton SC, et al: Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. **JAMA** 283:2008–2012, 2000
32. Terry K: Pay for performance: a double-edged sword. **Med Econ** 82:64–66, 68, 71, 2005
33. Werner RM, Asch DA: The unintended consequences of publicly reporting quality information. **JAMA** 293:1239–1244, 2005
34. Woodson SB: Making the connection between physician performance and pay. **Healthc Financ Manage** 53:39–42, 44, 1999

Manuscript received July 25, 2005.

Accepted in final form September 9, 2005.

Address reprint requests to: Sherman C. Stein, M.D., Department of Neurosurgery, University of Pennsylvania School of Medicine, 310 Spruce Street, Philadelphia, Pennsylvania 19106. email: sherman.stein@uphs.upenn.edu.