

# Health Affairs

At the Intersection of Health, Health Care and Policy

Cite this article as:

John Peabody, Riti Shimkhada, Stella Quimbo, Jhiedon Florentino, Marife Bacate,  
Charles E. McCulloch and Orville Solon  
Financial Incentives And Measurement Improved Physicians' Quality Of Care In The  
Philippines  
*Health Affairs*, 30, no.4 (2011):773-781

doi: 10.1377/hlthaff.2009.0782

The online version of this article, along with updated information and services, is  
available at:

<http://content.healthaffairs.org/content/30/4/773.full.html>

**For Reprints, Links & Permissions:**

[http://healthaffairs.org/1340\\_reprints.php](http://healthaffairs.org/1340_reprints.php)

**E-mail Alerts :** <http://content.healthaffairs.org/subscriptions/etoc.dtl>

**To Subscribe:** <http://content.healthaffairs.org/subscriptions/online.shtml>

*Health Affairs* is published monthly by Project HOPE at 7500 Old Georgetown Road, Suite 600, Bethesda, MD 20814-6133. Copyright © 2011 by Project HOPE - The People-to-People Health Foundation. As provided by United States copyright law (Title 17, U.S. Code), no part of *Health Affairs* may be reproduced, displayed, or transmitted in any form or by any means, electronic or mechanical, including photocopying or by information storage or retrieval systems, without prior written permission from the Publisher. All rights reserved.

Not for commercial use or unauthorized distribution

By John Peabody, Riti Shimkhada, Stella Quimbo, Jhiedon Florentino, Marife Bacate, Charles E. McCulloch, and Orville Solon

# Financial Incentives And Measurement Improved Physicians' Quality Of Care In The Philippines

DOI: 10.1377/hlthaff.2009.0782  
HEALTH AFFAIRS 30,  
NO. 4 (2011): 773-781  
©2011 Project HOPE—  
The People-to-People Health  
Foundation, Inc.

**ABSTRACT** The merits of using financial incentives to improve clinical quality have much appeal, yet few studies have rigorously assessed the potential benefits. The uncertainty surrounding assessments of quality can lead to poor policy decisions, possibly resulting in increased cost with little or no quality improvement, or missed opportunities to improve care. We conducted an experiment involving physicians in thirty Philippine hospitals that overcomes many of the limitations of previous studies. We measured clinical performance and then examined whether modest bonuses equal to about 5 percent of a physician's salary, as well as system-level incentives that increased compensation to hospitals and across groups of physicians, led to improvements in the quality of care. We found that both the bonus and system-level incentives improved scores in a quality measurement system used in our study by ten percentage points. Our findings suggest that when careful measurement is combined with the types of incentives we studied, there may be a larger impact on quality than previously recognized.

**John Peabody** (peabody@psg.ucsf.edu) is a professor at the University of California, San Francisco, and the University of California, Los Angeles, and chief medical officer at Sg2.

**Riti Shimkhada** is a senior research associate in the Department of Health Services, University of California, Los Angeles.

**Stella Quimbo** is a professor in the School of Economics, University of the Philippines, in Quezon City.

**Jhiedon Florentino** is a consultant at the Health Policy Development Program, US Agency for International Development, in Manila, the Philippines.

**Marife Bacate** is a consultant at the Asian Development Bank in Manila.

**Charles E. McCulloch** is a professor and head of the Division of Biostatistics at the University of California, San Francisco.

**Orville Solon** is a professor in the School of Economics, University of the Philippines.

Consumers, payers, providers, and policy makers regularly express concerns about low-quality health care,<sup>1</sup> prompting an ongoing search for new ways to improve the quality of care.<sup>2</sup> Unfortunately, at least so far, this effort has not met with much success.

Pay-for-performance is perhaps the most commonly used or widely considered incentive scheme.<sup>3-9</sup> Under pay-for-performance schemes, financial incentives are given to providers to meet certain performance standards. These are usually process measures, such as following evidence-based treatment protocols or population outcome targets—for example, the proportion of hypertensive patients who are adequately controlling their blood pressure. Other financial and nonfinancial incentives, including system-level rewards such as budget increases, individual or group feedback in the form of peer comparisons,

and public recognition have also been used.<sup>10-12</sup>

These different types of incentives, which link individual and system-level rewards to measures of work performance, appear to be associated with only modest improvements (or none at all) in quality of care and patient outcomes.<sup>13-17</sup> Moreover, even the most rigorous pay-for-performance studies have been plagued by selection bias, wherein providers who adopt the incentives and participate in the studies may well be the ones most likely to respond and improve their clinical practice. As a result, the studies may give an inflated sense of the effectiveness of the incentives in improving provider performance.<sup>14,15,17</sup>

Without accounting for this selection bias and other shortcomings in study design, the benefits of an incentive program aimed at improving clinical practice can be incorrectly interpreted, leading to either failure to adopt policies that

might lead to real improvements in quality, or, at the opposite end, adoption of policies that are ineffective and costly or even have unintended consequences.<sup>18</sup>

A second challenge in evaluating pay-for-performance strategies has been the difficulty in finding measures that are impervious to “gaming,” which occurs when providers focus on improving only in areas that are being monitored in order to give the appearance of success, while neglecting other (unmonitored) aspects of quality. Gaming is particularly problematic when measures of performance are too limited or narrowly defined.<sup>19</sup>

Finally, many incentive performance studies suffer because different facilities treat patients with varying health status or different health problems, or receive revenue from payers with different payment structures and incentives. The presence of variation in case-mix or patient health or of multiple payers makes it hard to control for the effects of such differences on performance and hard to interpret comparisons over time or across facilities.<sup>20,21</sup>

We had the opportunity to overcome each of these problems in the Philippine Child Health Experiment, known in-country as the Quality Improvement Demonstration Study. In this study we experimentally introduced two incentive-based policy interventions and compared them to control sites, where the intervention was not being used.

### Overview Of The Philippine Health Care System

In the Philippines, health care is delivered by both public and private doctors and facilities. Public facilities, which include primary health care clinics and hospitals, are mostly under the jurisdiction of local government units. The Philippine Department of Health is the principal government agency that formulates national health policies and programs and guides the development of local health systems.

Children in rural areas of the Philippines, such as the Visayas, a large island group in central Philippines around the Visayan Sea, have a high burden of disease. According to the 2003 National Demographic and Health Survey conducted by the Department of Health, the mortality rate for children younger than age five is 52 per 1,000 in rural areas. That is much higher than the overall national average of 34 per 1,000, or the average for all countries in the World Health Organization’s Western Pacific Region of 31 per 1,000. During 2000–03, diarrheal disease and pneumonia accounted for 25 percent of all deaths among children younger than age

five.<sup>22</sup> Approximately 12 percent of children in rural areas have symptoms of acute respiratory illness.<sup>23</sup>

The Philippine Health Insurance Corporation (PhilHealth), the largest insurance carrier in the Philippines, administers the government-sponsored National Health Insurance Program. It provides the equivalent of first-dollar coverage for public and private inpatient services, including room and board, diagnostic procedures, drugs, surgical and other treatment procedures, and professional fees. The program covers employees who pay through wage deductions in the formal sector—medium-size and large businesses and government, retirees and pensioners, and self-employed individuals who pay their own premiums. Subsidized coverage is extended to poor households through PhilHealth’s Indigent Program.

### Study Data And Methods

**STUDY SETTING** We provide a summary of our methods here; a more detailed description of the methods is available in the online Appendix.<sup>24</sup> The subject of our study was a policy experiment that evaluated the affect of two interventions on physician practices, health behavior, and the health status of children younger than age five in the Philippines. These interventions are financed and implemented by PhilHealth.

Under the experiment, physicians in hospitals randomly selected as “bonus intervention” sites could directly receive extra pay based on average clinical competence scores of randomly selected physicians, facility caseload, and average patient satisfaction. Hospitals randomly selected as “expanded insurance intervention” sites received greater revenue in the form of PhilHealth insurance benefits covering 100 percent of costs for ordinary cases of common conditions such as pneumonia and diarrhea. In this latter set of incentives directed at hospitals, the incentive for physicians was indirect. The study sites were from ten provinces located in the Visayas and Mindanao regions of the Philippines. If the province was slated for health reform and if it had a district hospital, that hospital was selected as a site. Thirty district-level government hospitals were so identified and selected.

Districts were sorted into groups of three, based on basic supply and demand characteristics such as population size, average income, labor-force participation rate, functional literacy, infant mortality rate, and percentage of the population with insurance. Within each matched group, one site was randomly chosen as the expanded insurance intervention site, one as the bonus intervention site, and the third as

the control site with no intervention.

**DATA COLLECTION** The data were from two sources: clinical performance vignette assessments (described below) and a physician survey. Physicians practicing at the thirty district-level government hospitals were the target population and were surveyed every six months starting at baseline—the beginning of the study—and extending through three-and-a-half years after the intervention. Baseline data were collected in 2003, policy interventions were introduced in 2004, and physician assessments continued until 2007.

**VIGNETTES AND PHYSICIAN SURVEYS** We used clinical performance vignettes to measure changes in performance associated with the policies.<sup>25–30</sup> Vignettes use a prescribed scenario to simulate a clinical encounter and judge clinical competence by asking the physician to state what he or she would do in response to the information presented in ten different scenarios. We used clinical performance vignettes because they have been validated against actor patients; this validation is regarded as the gold standard for such comparisons and is widely used.<sup>26,29,31,32</sup> These clinical performance vignettes required the physician to have an entire set of clinical skills in five domains: taking a history, doing a physical examination, ordering tests, making a diagnosis, and prescribing treatment. Differences between these and other kinds of vignettes are described in the Appendix.<sup>24,33</sup>

All physicians who took care of children at least 25 percent of the time in each hospital were registered for the study. Each semester (a period of six months), three randomly selected physicians from each study hospital participated in three vignettes, one for each of the three medical conditions (dermatitis, diarrhea, and pneumonia) selected for the study. A list of the vignette cases for each condition is shown in the Appendix.<sup>24</sup> The conditions and the order were all randomly assigned.

Each clinical performance vignette was scored by two trained abstractors blinded to the physician's identity and the intervention. The number of correct items formed the basis for a physician's vignette score. Correct responses were aggregated to obtain a final summary score.<sup>26</sup> The scores for all three vignettes per physician were averaged, producing an average vignette score, which was used as the measure of quality of care for that facility at that time.

The vignettes were accompanied by a survey, which collected data on physician demographics, physician education and training, practice characteristics, clinic characteristics, and physician income, to account for differences in physician composition within and between

sites and for any shifts in this composition over time. Patient volumes were also measured to determine whether there were shifts in volumes over time.

**DEMONSTRATION STUDY INTERVENTIONS** A primary aim of the demonstration study was to examine the effect of bonus payments on the quality of care at baseline and at postintervention intervals in intervention sites compared to control sites. Another study aim was to determine the effect on health care services and outcomes—before and after the intervention—of expanded insurance benefits to the hospital. The target for the interventions in all sites, which remain in place, is all children younger than age five.

Each quarter we computed the quality metric<sup>34</sup> for each hospital assigned to either intervention. Based on the audited financial reports of Phil-Health and the survey data, total bonus payments represented approximately 5 percent of total doctor salaries in qualifying bonus intervention site hospitals. Details on the calculation of the bonus amount are given in the Appendix.<sup>24</sup>

Physicians took the clinical performance vignettes regardless of which intervention site they were in. Vignette scores, overall and by skill domain, were reported individually to the participating physicians; the head of the hospital received aggregate scores of all of the physicians.

**CONCEPTUAL FRAMEWORK** We hypothesized that observed improvements in physician vignette scores would be a product of behavioral responses to varying incentives. By having two separate interventions and control sites, as well as a distinct subgroup of physicians who had taken vignettes multiple times, we were able to disaggregate the observed changes in quality. Our randomized study design introduced a conceptual framework that explored four effects—bonuses, expanded insurance, repeat test taking, and dissemination effects—and estimated where and to what extent they affected quality performance.

The bonus incentive effect occurred only in the bonus intervention sites; it was estimated by the improvement in vignette scores between baseline and postintervention compared to changes that occurred over the same period in the control sites.

Whereas physicians in the bonus intervention group might have improved vignette performance over time because of the direct effect of bonus incentives, physicians in the insurance intervention group would improve quality performance by responding to indirect financial rewards for improved quality at the hospital level. Physicians at insurance expansion intervention sites knew that bringing a greater number of insured patients into their facility translated into

greater reimbursement for the hospital. We estimated this effect by examining whether greater reimbursement to hospitals that were able to draw in more insured patients improved quality.

In control sites, where no incentives were offered, quality was also regularly measured and scores were fed back to physicians, their hospital chiefs, and policy makers. This act of measurement and feedback affected all three arms of the study, and any subsequent improvement was due to awareness and consequent motivation to perform better. This effect was estimated by looking at the change in control site scores over time.

Lastly, there is the question of whether repeat test taking confers some advantage that alters scores instead of reflecting any actual change in performance. The possibility of a repeat-test-taking effect can be evaluated by comparing repeat test takers with one-time testers, which we did in our analysis. The full regression model is shown in the Appendix, along with details on the variables used.<sup>24</sup>

**DATA ANALYSIS** We calculated basic descriptive statistics for physicians and for physician performance. We combined the six rounds of semi-annual vignette data to estimate the effects of the interventions on the average vignette scores before and after the intervention. We controlled for clustering at the hospital level and effects of smaller versus larger facilities, and, to estimate potential learning effects, we estimated a model that accounts for repeated test taking.

**LIMITATIONS** This study, even with its randomized design, had a number of limitations. The first is that we did not link changes in quality performance to changes in health outcomes in patients. Although we know that vignette performance is associated with clinical outcomes, including patient satisfaction,<sup>35,36</sup> assessment of health status changes among patients would strengthen the evidence of the relationship between incentives and quality.

Another limitation is that we did not survey physicians to identify what they perceived the incentives to be in each intervention arm. Although we can posit that financial incentives in bonus sites and indirect system incentives in insurance expansion sites led to quality improvement, physician perceptions would provide a clearer picture of the actual drivers of behavioral change.

Lastly, an evaluation of the changes in the insurance status of patients over time relative to performance would strengthen the evidence that system-level effects led to improvements in performance in insurance expansion sites.

## Study Results

The total study population consisted of 617 physicians. Comparing physician characteristics, we found that the two intervention groups and control sites were statistically similar at baseline except that physicians in the insurance expansion group were more likely to be women. There were no significant differences in facility characteristics between the intervention and control sites (Exhibit 1; a more detailed version of Exhibit 1 is available in the Appendix).<sup>24</sup> We also compared physician characteristics at the end of the study and among qualifying facilities, and we found no difference in their composition.

At baseline, vignette scores were also similar among the three sites. Across the five domains of the clinical evaluation, physicians performed the best on ordering tests and gradually declined in performance as they progressed through the clinical evaluation, achieving the lowest scores in diagnosis and treatment. There was no difference by individual skill domains among sites in history taking, physical examination, ordering imaging studies and laboratory tests, and diagnosis and treatment (see Exhibit 2; a more detailed version of Exhibit 2 is available in the Appendix).<sup>24</sup>

**DETERMINANTS OF QUALITY BONUS EFFECT** To estimate the effects of the bonus incentive on quality after controlling for physician characteristics, we evaluated whether the average vignette scores in bonus sites improved over time to a greater extent than in control sites (Exhibit 3). At baseline and again at six months after the intervention period, there were no significant score improvements in bonus sites ( $p = 0.64$  and  $p = 0.61$ , respectively).

At twelve months after the intervention, there were significant improvements in scores in the bonus sites ( $p = 0.006$ ). At eighteen months after the intervention, improvements continued in the bonus sites compared to baseline and compared to control sites ( $p = 0.01$ ).

Improvements in the bonus sites continued over time: at thirty-six months after the intervention, bonus sites were 9.7 percentage points higher than baseline ( $p < 0.001$ ). When we disaggregated the overall scorers by domain, we observed that there were gains across all five (results not shown).

By measuring the average number of monthly inpatients, we found that there was no substantive change in patient volumes in the hospitals that qualified for bonuses (329) versus those that did not (347). As expected, in the insurance expansion sites there was an increase in the number of patients seen after insurance benefits were expanded.

**SYSTEM EFFECTS** To assess system-level incen-



**EXHIBIT 1**
**Physician And Facility Characteristics, By Intervention Type At Baseline, Philippines, 2003**

	Mean				Difference in means	
	All sites	Insurance expansion sites	Bonus sites	Control sites	Insurance expansion vs. control	Bonus vs. control
<b>PHYSICIAN</b>						
Age (years)	41.53	42.36	41.23	41.10	1.26	0.14
Male (%)	33.71	10.71	43.33	45.16	34.47***	1.83
Specialty, pediatrician (%)	28.09	28.57	36.67	19.35	9.21	17.31
<b>FACILITY</b>						
Accreditation by PHIC (%)	100	100	100	100	0	0
Number of beds	51.67	55.00	24.86	47.50	7.50	5.00
Number of physicians	11.30	12.60	10.50	10.80	1.80	-0.30
<b>PATIENT</b>						
Average travel time (minutes)	32.51	31.02	31.02	33.90	-2.87*	-1.49
Average household income (PhP)	58,650	58,626	58,884	58,442	183.69	441.45

**SOURCE** Authors' analysis of study data. **NOTES** PHIC is Philippine Health Insurance Corporation. PhP is Philippine pesos. \* $p < 0.10$  \*\*\* $p < 0.01$

tives, we compared insurance expansion to control sites. As in the previous comparison, there were no significant differences at baseline. There was also no improvement in vignette scores six months after the intervention was introduced ( $p = 0.79$ ). However, at the twelve-month, post-intervention assessment, vignette scores in insurance expansion sites significantly improved and were greater than scores in the control sites ( $p = 0.02$ ).

The improvement also persisted at each subsequent period compared to baseline. At thirty-six months it was 9.1 percentage points greater than baseline ( $p = 0.001$ ) and was statistically indistinguishable from the improvement in the bonus sites (see also Exhibit 3). Improvements

occurred throughout all domains among the insurance expansion sites, although the improvement in the treatment domain was not statistically significant (results not shown).

As noted above, we also looked for repeat testing effects, dissemination effects, and possible effects from changes in the hospital workforce. We found that repeat test taking did not improve vignette scores ( $p = 0.09$ ), and there were no significant interactions with the interventions.

The only evidence for a dissemination effect occurred in the control sites in the last assessment period, thirty-six months after intervention. Vignette scores improved compared to baseline ( $p = 0.03$ ), but the magnitude of the improvement was not as large as the increases

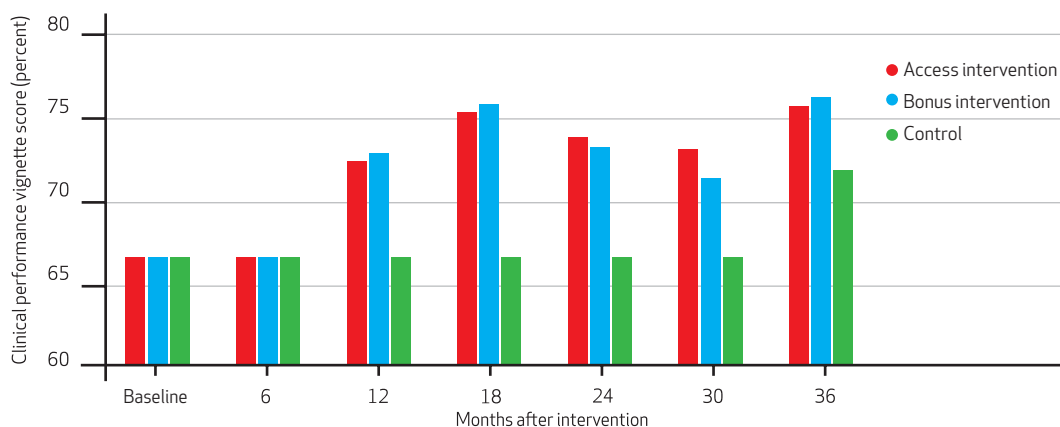
**EXHIBIT 2**
**Clinical Performance Vignette Scores, By Intervention Type And Domain At Baseline, Philippines, 2003**

Vignette score	Mean				Difference in means	
	All sites	Insurance expansion sites	Bonus sites	Control sites	Insurance expansion vs. control <sup>a</sup>	Bonus vs. control <sup>a</sup>
Overall	53.21	53.90	54.10	51.70	2.20	2.40
History taking	56.33	57.20	57.62	54.32	2.90	3.30
Physical exam	57.73	57.75	58.84	56.63	1.11	2.22
Test ordering	64.14	65.18	64.63	62.72	2.45	1.91
Diagnosis	46.10	46.92	47.41	44.09	2.84	3.32
Treatment	38.17	39.23	37.64	37.73	1.50	0.09

**SOURCE** Authors' analysis of study data. <sup>a</sup>No differences were significant at the 0.10 level.

## EXHIBIT 3

Predicted Average Vignette Scores, By Intervention Site For Each Assessment Period, Philippines, 2003–07



**SOURCE** Authors' analysis of study data. **NOTE** A version of the exhibit showing the standard error of the estimate is available in the online Appendix (Note 24 in text).

that occurred in the insurance expansion and bonus sites during the same period.

Some improvements in quality might have been due to changes in the physician workforce, which averaged 10 percent with a 3–27 percent range among hospitals, during the thirty-six months of the study. We tested for such effects by comparing at baseline the doctors who stayed with those who subsequently left the hospital and by comparing the original doctors with new physicians postintervention. Neither comparison showed any significant differences, which indicates that staff selection did not contribute to quality improvements.

### Discussion

We found that clinical performance vignette scores—a measure of the quality of care—were significantly improved by bonus incentives directed toward the individual doctor in a pay-for-performance program and by indirect system-level expanded insurance incentives that improved access to care and overall reimbursement. By the last assessment period, quality had improved approximately ten percentage points in both intervention groups.

A notable finding, however, is that performance improvement ultimately occurred even in control sites after thirty-six months. This observation raises the intriguing possibility that there was also a real delayed dissemination and feedback effect, fomented by the bonuses or possibly the expanded insurance. More information and further monitoring of sites are needed before any such conclusion can be drawn from this study.

This study overcomes several methodological limitations found in other studies on incentive schemes that have made it difficult to distinguish between pure financial incentives that benefit the provider and incentives that benefit the system. The direct monetary incentive was modest—about 5 percent of a physician's salary—and accounted for only a marginal increase in PhilHealth payments, yet it was large enough to stimulate a significant improvement in performance. Physicians in insurance expansion sites also improved in quality performance to a similar magnitude as did the physicians in bonus sites.

Although all sites (including controls) had feedback and public disclosure, only in the insurance expansion site did hospitals, and not just the doctors, have the opportunity to benefit financially. Their higher scores may have helped them draw in more insured patients and, with them, a greater opportunity to bill a reliable payer.

Alternatively, insurance expansion site performance may have improved because the incentive benefited the entire team of providers, including the hospital. The similar benefits of the direct and the indirect incentives underscore the importance of considering system-level effects that may drive individual behavior.

A body of literature suggests that both monetary and nonmonetary incentives lead clinicians to reflect on their relative performance and encourage the adoption of improvements.<sup>16,37,38</sup> For instance, quality performance feedback (measurement and feedback), a basic nonmonetary incentive, has been shown to improve clinical performance.<sup>35</sup>

There are several reasons that quality may have improved in this study besides physicians' responding to the incentives. At the behest of PhilHealth and to facilitate generalization and eliminate any possible training effects, no training or remediation was offered. Similarly, there might have been a volume effect, with more patients going to the hospitals that faced incentives, but when we checked for changes in inpatient census among the bonus sites, we found none.

Our results also indicate that there was no performance difference among those doctors

who completed the vignettes more often than those who only did them once, and thus no likely effect of repeat test taking or gaming over time.

In summary, this experimental study provides evidence that pay-for-performance has a significant effect on clinical performance. Moreover similar quality effects may be possible through indirect financial incentives that operate at the system level. These effects on quality affected performance earlier and to a greater degree than measurement and feedback of performance alone. ■

The Quality Improvement Demonstration Study is funded by the US National Institutes of Health through an R01 grant (No. HD042117).

## NOTES

- 1 Jamison DT. Investing in health. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al., editors. *Disease control priorities in developing countries*. 2nd ed. New York (NY): Oxford University Press; 2006. p. 3–36.
- 2 American Academy of Pediatrics Steering Committee on Quality Improvement and Management, American Academy of Pediatrics Committee on Practice and Ambulatory Medicine, Hodgson ES, Simpson L, Lannon CM. Principles for the development and use of quality measures. *Pediatrics*. 2008;121(2):411–8.
- 3 Epstein AM. Performance reports on quality: prototypes, prospects, and problems. *N Engl J Med*. 1995;333:57–61.
- 4 Epstein AM. Public release of performance data. *JAMA*. 2000;283(14):1884–86.
- 5 Epstein AM, Lee TH, Hamel MB. Paying physicians for high-quality care. *N Engl J Med*. 2004;350:406–10.
- 6 Kuhmerker K, Hartman T. Pay-for-performance in state Medicaid programs: a survey of state Medicaid directors and programs. New York (NY): Commonwealth Fund; 2007.
- 7 Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. *JAMA*. 2000;283:1866–74.
- 8 Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA* 2005;294(14):1788–93.
- 9 Trude S, Au M, Christianson JB. Health plan pay-for-performance strategies. *Am J Manag Care*. 2006;12(9):537–42.
- 10 Baker DW, Cebul RD. Should consumers trust hospital quality report cards? *JAMA*. 2002;287(24):3206–7.
- 11 Bufalino V, Peterson ED, Krumholz HM, Burke GL, LaBresh KA, Jones DW, et al. Nonfinancial incentives for quality: a policy statement from the American Heart Association. *Circulation*. 2007;115(3):398–401.
- 12 Shaller D, Sofaer S, Findlay S, Hibbard J, Lansky D, Delbanco S. Consumers and quality-driven health care: a call to action. *Health Aff (Millwood)*. 2003;22(2):95–101.
- 13 Campbell S, Reeves D, Kontopantelis E, Middleton E, Sibbald B, Roland M. Quality of primary care in England and the introduction of pay for performance. *N Engl J Med*. 2007;357:181–90.
- 14 Epstein AM. Pay for performance at the tipping point. *N Engl J Med*. 2007;356(5):515–17.
- 15 Glickman SW, Ou FS, DeLong ER, Roe MT, Lytle BL, Mulgund J, et al. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA*. 2007;297(21):2373–80.
- 16 Lindenaer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, et al. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med*. 2007;356:486–96.
- 17 Rosenthal MB, Frank RG. What is the empirical basis for paying for quality in health care? *Med Care Res Rev*. 2006;63(2):135–57.
- 18 Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay for performance on the quality of primary care in England. *N Engl J Med*. 2009;361(4):368–78.
- 19 Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann Intern Med*. 2006;145:265–72.
- 20 Epstein AM. Paying for performance in the United States and abroad. *N Engl J Med*. 2006;355(4):406–8.
- 21 Landon BE, Normand SL, Blumenthal D, Daley J. Physician clinical performance assessment: prospects and barriers. *JAMA*. 2003;290(9):1183–9.
- 22 World Health Organization. *World health statistics 2006*. Geneva: WHO; 2006.
- 23 Philippines Department of Health. *National objectives for health, Philippines, 1999–2004*. Manila: DOH; 1999.
- 24 To access the Appendix, click on the Appendix link in the box to the right of the article online.
- 25 Dresselhaus TR, Peabody JW, Lee M, Wang MM, Luck J. Measuring compliance with preventive care guidelines: standardized patients clinical vignettes and the medical record. *J Gen Intern Med*. 2000;15(11):782–8.
- 26 Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715–22.
- 27 Peabody JW, Luck J, Glassman P, Jain S. Measuring what we want to measure: using vignettes in clinical education [Abstract]. *J Gen Intern Med*. 2002;17(1 suppl):232.
- 28 Dresselhaus TR, Peabody JW, Luck J, Bertenthal D. An evaluation of vignettes for predicting variation in the quality of preventive care. *J Gen Intern Med*. 2004;19(10):1013–18.
- 29 Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med*. 2004;141(10):771–80.
- 30 Peabody JW, Tozija F, Munoz JA, Nordyke RJ, Luck J. Using vignettes



to compare the quality of clinical care variation in economically divergent countries. *Health Serv Res.* 2004;39(6):1951–70.

- 31** Hrisos S, Eccles MP, Francis JJ, Dickinson HO, Kaner EF, Beyer F, et al. Are there valid proxy measures of clinical behaviour? A systematic review. *Implement Sci.* 2009;4:37.
- 32** Glassman PA, Luck J, O’Gara EM, Peabody JW. Using standardized patients to measure quality: evidence from the literature and a prospective study. *Jt Comm J Qual Improv.* 2000;26(11):644–53.
- 33** Das J, Hammer JS, Leonard KL. The quality of medical advice in low and middle income countries: 2008. Washington (DC): World Bank; 2008. (World Bank Policy Research, Working Paper No. 4501).
- 34** Solon O, Woo K, Quimbo SA, Shimkhada R, Peabody JW, Florentino J, et al. A novel method for measuring health care system performance: experience from QIDS in the Philippines. *Health Policy Plan.* 2009;24(3):167–74.
- 35** Peabody JW, Shimkhada R, Florentino J, Quimbo S, Solon O. The relationship of different levels of quality on cost and satisfaction: evidence from the QIDS study. *Med Care.* 2010;48(1):25–30.
- 36** Peabody JW, Nordyke RJ, Tozija F, Luck J, Muñoz JA, Sunderland A, et al. Quality of care and its impact on population health: a cross-sectional study from Macedonia. *Soc Sci Med.* 2006;62(9):2216–24.
- 37** Jamtvedt G, Young JM, Kristoffersen DT, O’Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev.* 2006 Apr 19;2:CD000259.
- 38** Pham HH, Schrag D, O’Malley AS, Wu B, Bach PB. Care patterns in Medicare and their implications for pay for performance. *New Engl J Med.* 2007;356:1130–9.

## ABOUT THE AUTHORS: JOHN PEABODY, RITI SHIMKHADA, STELLA QUIMBO, JHIEDON FLORENTINO, MARIFE BACATE, CHARLES E. MCCULLOCH & ORVILLE SOLON



**John Peabody** is a professor at the University of California, Los Angeles and San Francisco.

As they explain in this issue of *Health Affairs*, John Peabody and Orville Solon, co-principal investigators, led an investigation of how two models of pay-for-performance worked in Philippine hospitals. Peabody reports that the Philippine experience shows that “detailed clinical measurement rapidly improves performance” by physicians, “apparently more than financial incentives.”

Peabody is deputy director of the Policy and Prevention Health Group in the Institute for Global Health/Global Health Sciences as well as a professor both in the Epidemiology and Biostatistics and Medicine Divisions at the University of California, San Francisco, and the School of Public Health at the University of California, Los Angeles. His

research interests include health-sector reform, quality of care, finance, and cross-system comparisons. Peabody is also the chief medical officer and senior vice president of Sg2, a health care analytics and intelligence firm.

Peabody has a medical degree from the University of California, San Francisco; a diploma in tropical medicine and hygiene from the University of London; and a doctorate in public policy from the Frederick S. Pardee RAND Graduate School.



**Riti Shimkhada** is a senior research associate in the Department of Health Services, University of California, Los Angeles.

Riti Shimkhada, program manager of the Philippine study, is a senior research associate in the Department of Health Services, University of California, Los Angeles. Her research interests include the interface between

epidemiology and policy making, evidence-based policy making, health impact assessment, and health forecasting. She holds a doctorate in epidemiology from the University of California, Los Angeles.



**Stella Quimbo** is a professor at the University of the Philippines.

Stella Quimbo is a professor in the School of Economics, University of the Philippines. Her research interests are in health economics, particularly hospital pricing and health insurance design, and the effect of incentives on physician and patient behavior. She has experience in survey design, implementation, and data analysis. Her doctorate is from the School of Economics, University of the Philippines.



**Jhiedon Florentino** is a consultant at the Health Policy Development Program, USAID.

Jhiedon Florentino is a consultant at the USAID Health Policy Development Program. His research interests are health and regulatory economics and data analysis. He earned his master's degree at the School of Economics, University of the Philippines.



**Marife Bacate** is a consultant at the Asian Development Bank.

Marife Bacate is a consultant at the Asian Development Bank whose

research interests include health and trade economics. Her master's degree in economics is from the School of Economics, University of the Philippines.



**Charles E. McCulloch** is a professor at the University of California, San Francisco.

Charles McCulloch is a professor and head of the Division of Biostatistics at the University of California, San Francisco. His research interests include statistical methods for longitudinal data analysis, generalized linear mixed models, and latent class models.

McCulloch is a fellow of the American Statistical Association and an elected member of the International Statistical Institute. He holds a doctorate in statistics from Cornell University, in

addition to a master's degree in math sciences from the Johns Hopkins University.



**Orville Solon** is a professor at the University of the Philippines.

Solon is a professor in the School of Economics, University of the Philippines, and is recognized as the nation's expert in health care reform. He also heads the US Agency for International Development (USAID) Health Policy Development Program, which supports policy and financing to promote family health. His research interests include health reform, care delivery, and financing. Solon has a doctorate in economics from the School of Economics, University of the Philippines.

## Errata

### **PRONOVOST ET AL., APRIL 2011, P. 573**

The acknowledgment for coauthor Richard Lilford should have contained the following statement: Richard Lilford was funded by the National Institute for Health Research (NIHR) through the Collaborations for Leadership in Applied Health Research and Care for Birmingham and Black Country (CLAHRC-BBC) program. The views expressed in this article are not necessarily those of the NIHR; the Department of Health; the University of Birmingham; or the CLAHRC-BBC.

### **CLASSEN ET AL., APRIL 2011, P. 585**

This article contained several errors. First, in the final paragraph under

“Study Results,” the Patient Safety Indicators method had a sensitivity of 5.8, not 8.5 as shown. Also, in Exhibit 3, the values for “Pulmonary/VTE” in severity level I should have been 1, not 2, leading to a total of 16, not 17. These errors do not affect the article’s findings and conclusions. The text and Exhibit 3 have been corrected online.

### **SMITH ET AL., APRIL 2011, P. 646,**

**P. 652, P. 654** This article contained several errors. First, the fifth sentence in the abstract (p. 646) should have stated that pharmacists resolved nearly 80 percent of drug therapy problems, not nearly 83 percent. Next, Michael P. Starkowski’s tenure as commissioner of the Connecticut Department of Social

Services ended in April 2011. This should have been reflected in the biographical information on pp. 646 and 654. In addition, the notes to Exhibit 3 (p. 652) should have referred readers to Note 18 in text, not Note 19 in text.

### **PEABODY ET AL., APRIL 2011, P. 773,**

**P. 781** Information about these authors contained an error. John Peabody is chief medical officer at Sg2. This information was omitted from his biography on p. 773 and was erroneously attributed to a coauthor on p. 781.

### **WYNIA ET AL., FEBRUARY 2011, P. 267**

On p. 267 of this article, first paragraph under “Study Data And Methods,” the word “psychologists” should be “psychiatrists.”