

Outcomes-Based Compensation: Pay-for-Performance Design Principles

*4th Annual Disease Management
Outcomes Summit
November 11 – 14, 2004
Rancho Mirage, California*

Table of Contents

Foreword	i
Conference Participants	ii
Outcomes-based Compensation: A Physician's Perspective	iv
Section 1. Why Physician Payment Must Change	1
Section 2. Pay-for-Performance: A Definition	4
Section 3. Pay-for-Performance Design Principles: Assumptions ...	6
Assumption #1	6
Assumption #2	6
Assumption #3	7
Assumption #4	7
Section 4. Design Principles	9
Design Principle Category #1: Payment Structure	9
Design Principle 1.1: Accountability Level	9
Design Principle 1.2: Distribution of Financial Incentives ..	10
Design Principle 1.3: Financial Incentive Type	11
Design Principle 1.4: Frequency of Assessments and Incentive Distribution	11
Design Principle Category #2: Transparency	12
Design Principle 2.1: Public Disclosure of Method	13
Design Principle 2.2: Disclosure of Results	13
Design Principle Category #3: Metrics	14
Design Principle 3.1: Measurement Level	14
Design Principle 3.2: Metric Attributes	16

Table of Contents

Design Principle 3.3: Metric Domain	18
Design Principle 3.4: Range of Metrics	19
Design Principle Category #4: Evaluation	20
Design Principle 4.1: Need of Evaluation	20
Design Principle Category #5: Community Participation ...	21
Design Principle 5.1: Community-wide Participation in Program Development	21
Design Principle 5.2: Patient Participation	22
Design Principle Category #6: Fairness	22
Design Principle 6.1: Methods to Maximize Fairness ...	23
Design Principle 6.2: Sample Size	24
Section 5. Conclusions	25
References	27
Summary of Pay-for-Performance Design Principles	29

* **Outcomes-Based Compensation: Pay-for-Performance Design Principles** can also be accessed online at www.rewardingquality.com. Visitors to the site can download the complete document, view it by section or research additional information on pay-for-performance.

This summit was planned and conducted by faculty members of the Johns Hopkins Outcomes Evaluation Program and the executive leadership of American Healthways. Accordingly, both organizations attest to the validity of the process leading to the Consensus Statement of the participants as presented in this document. However, the consensus reached at the Summit does not necessarily reflect the opinions or practice of either the faculty of Johns Hopkins or that of American Healthways. Neither Johns Hopkins nor American Healthways is responsible for the implementation of these guidelines or protocols beyond their own institutions.

F o r e w o r d

Outcomes-based compensation is upon us, and it is imperative that physicians' voices be heard. If physicians in the United States want to transform the health care system into one that produces outcomes rather than merely producing services, a logical step is to make a declaration on how the incentives used to promote quality and outcomes should be designed.

This document provides a set of design principles to guide the development of pay-for-performance programs. Each principle was carefully developed for review by a 23-member steering committee that convened regularly over a five-month period. The committee was composed of practicing physicians, health plan administrators, disease management administrators, large employer purchasers, an actuary and academic researchers.

In November 2004, a conference of 250 physicians and medical managers from across the United States was convened in Rancho Mirage, CA. This groundbreaking conference was the first time that physicians and medical managers provided a consensus statement on how outcomes-based compensation arrangements should be developed in order to align health care toward evidence-based medicine, affordability and public accountability for how resources are used.

The three objectives of the consensus conference were:

- (1) To review and revise the consensus document draft developed by the steering committee,
- (2) To articulate physician preferences for the specific wording of the design principles, and
- (3) To elicit physician perspectives on the ideal PFP program.

Conference participants were asked to consider how PFP should be crafted in order to have an impact on outcomes for the largest number of patients possible. At its conclusion, the conference generated a blueprint for the design of pay-for-performance programs. We believe that these design principles, reflecting the thoughtful deliberations of a large group of physicians, can have an important impact on future physician payment policy. The principles that were produced are intended as a guideline for any organization interested in developing a pay-for-performance program, or as a framework for studying programs currently in place.

Conference Participants

Abraham A. Alecozay, M.D.
Judith Ameriks, M.D.
John E. Anderson, M.D. *
George Andrews, M.D.,
MBA, CPE
Linda Bahrke, RN, BSN,
MAOM
Timothy S. Bailey, M.D.
Jack Barry, M.D.
Jeffrey Beardmore, M.D.
David Berry, M.D.
Jim Bishop, M.D.
Daniel Bonbrisco, M.D.
Charles H. Booras, M.D. *
William Bossany
Wendy Bottomley
John K. Bower, M.D.
James Bracikowski, M.D.
Frederick L. Brancati, M.D.,
MHS
Dwayne Broussard, M.D.
Sheila Brown
Theodis Buggs, M.D.
James Burks, M.D.
Nial Castle, D.O.
Billy Chacko, M.D.
Larry K. Chase, M.D.
Michael Chesner, M.D.
Peter Choy, M.D.
Richard S. Chung, M.D. *
John J. Cogan, M.D.
Richard Colucci, M.D.
James S. Coxe III, M.D.
David Crosby
M. Olubunmi Dada, M.D.,
Ph.D.
William Dang, M.D.

Barbara Davis
Christine Del Gaizo, M.D.
Gregory B. Diette, M.D. *
Lonnie Divack, M.D.
John Dixon, M.D., FACC,
FAHA, FSCAI
Louella Dougherty
Daniel F. Driscoll, M.D.
Linda Dunbar *
Dennis Robbins, Ph.D.,
MPH
Donald Robinson
Mary E. Edwards
Eugene Eisenberg, M.D.
Bernie E. Elliott, M.D.
Laurence A. Engelberg,
M.D.
Peter Fagan
Kim Fecher
Christopher B. Forrest,
M.D., Ph.D *
Judith Frampton, R.N., MBA
Ron Y. Fujimoto, D.O.
George Gamblin, M.D.
John S. Georgakopoulos,
D.O.
James K. Geraughty, M.D. *
William R. Gold, M.D.
Michael Gold
Carlos J. Gonzalez, M.D.
Michael Goran, M.D. *
Sheila Gould, R.N. BC
CPHQ
Stephen Grayson, D.O.
Thomas J. Hamilton, D.O.
Shirley Heintz
John Heryer, M.D.

Douglas Hiza, M.D. *
Joel Hoffman, FCA, ASA,
MAAA *
Hugh Hood, M.D., F.A.C.P.
Jamie D. Howard, M.D.
Roger K. Howe, M.D., MMM
Jeryl Janson
Herman Jenich
Pete Keim, M.D.
Matt Kelliher
Rhonda Ketterling, M.D.
Marjorie King, M.D., FACC
Sid King, M.D.
E. John Kleinholz, M.D.
Robert Klugman, M.D.
Nat Kongtahworn
Allan M. Korn, M.D., FACP
Leonard J. Kryston, M.D.
Ganesh Kumar, M.D.
Richard Kuroghlian
Sandy Kurtz, M.D.
Paul D. Wardlaw, M.D.
Winston Watkins, M.D.
Kenneth A. LaBresh, M.D.
S. L. Lampkin IV, M.D.
Richard Lee-Ching, M.D.
Marc S. Levitt, M.D.
Whitney Limm, M. D.
Della M. Lin, M. D.
Steven M. Lum, M.D.
Sanford R. Mallin, M.D.
Bruce Mann, M.D.
Susan E. Manning, R.N.
Leonard Mastbaum, M.D.,
FACE
Janet Maurer, M.D.
Paul McGinnis, M.D.

John S. Melish, M.D.
Burnell J. Mellema, M.D.
Robin Merlino, M.D.
Michael L. Miccolo, M.D.
Peggy Mohoric
Wendy Monroe
Edwin Montell, M.D.
Rian Montgomery, M.D.
Carlisle Moore, RN
Allen J. Naftilan, M.D., PhD.
Robert Neaderthal, M.D.
C. Warren Neel, Ph.D.
Alan C. Nelsen, M.D.
J. Darryl Nesmith, M.D.
Jeff A. Odiet, M.D.
Kimberly O'Keefe
William C. O'Neil
Kenneth W. Ordene, M.D.
John Oujiri, M.D.
Donald E. Paynter, M.D.
Pat Pingitore, M.D.
Mary G. Pixler, M.D.
Glenn Pomerantz, M.D.
Thomas J. Quam, M.D.,
MBA
Lawrence I. Rand, M.D.
Keith Rau, M.D.
John Reitano, M.D.
Ronald M. Robinson, M.D.
Marsha Rogers, M.D.
Brenda Rogers-Grays, M.D.
James L. Rosenzweig, M.D.
Sharon Rudnick, MSW
John Rule, M.D.
John Saalwaechter, M.D.,
MBA
Patricia R. Salber, M.D.,
MBA
John Santalo

H. Scott Sarran, M.D.
Marie A. Savard, M.D.
Sidna M. Scheitel, M.D.,
MPH *
James Schrenker, M.D.
Gerald Schulman, M.D.
Robert E. Sevier, M.D.
Sheldon Sheps, M.D.
Allen T. Sherwood, M.D.
Irena Shperling, M.D.
Dhana Shrestha, M.D.
Dennis Shuman, M.D., Ph.D.
Jeffrey L. Simmons, MD
Bryan Sitzmann, M.D.
Karen Smoot
James W. Snyder, M.D.
Alan Sokolow, M.D. *
Stephen A. South, M.D.
FACE
Michele Spagna
Donald Stangler, M.D.
Marcia J. Stark
Alton Steiner, M.D.
Randall Suzuka, M.D.
Michael Swarin, D.O., MPH
Victor G. Villagra, M.D.*
Cynthia Villareal-Taylor,
M.D.
Dirk Wales, M.D.
James Wang, M.D.
Raymond L. Webster, M.D.,
MPH
Mark D. Westbrook, M.D.
Arthur Williams, D.O.
Myra Williams
Michael Wittenberg, DPM
Nancy Wolfe-Sidberry,
M.S.,M.D.
James Woodburn, M.D., MS

Janet S. Wright, M.D.
Junping Yang, M.D., Ph.D.
Laraine Yapo, BSN
Vincent Zuccarello

** Steering Committee members*

Outcomes-based Compensation: A Physician's Perspective

Optimal delivery of preventive services and management of chronic diseases for my patients have been my passion for the past decade. Throughout the literature, and even in our own practice, extensive gaps have been documented between what is recognized as optimal care and what is recorded as actual performance. The emergence of evidence-based medicine and best practice models has been enormously helpful to my colleagues and me by shifting our focus toward providing those services that can lead to the best possible outcomes.

However, I believe that both our medical practice and the larger health care system contain several barriers to achieving the best health for my patients. One of the most obvious is the fee-for-service payment system within which our practice operates.

There are some positives and some negatives when practicing in a fee-for-service environment. Cost containment programs initiated by the practice are less likely to be viewed suspiciously by our patients. Patients appreciate being offered appropriate preventive services and being evaluated and treated for acute illnesses. However, follow-up for chronic diseases by telephone or e-mail communication is not a billable encounter, even though use of non-visit care modalities is convenient for the patients. This is revenue "lost" when you are getting paid only for office visits. Incentives need to be realigned to promote value-added care.

The paradigm shift that pay-for-performance presents to the physician community holds the potential for revolutionizing how we are paid. More importantly, it can improve patient care. The incentives could foster practice changes such as an electronic infrastructure in the form of registries and electronic medical records with decision support that would improve quality of patient care. In addition, allied health staff could be utilized to assist in sending out patient reminders for preventive services and routine chronic disease management. Cultural acceptance by clinicians to redesign care delivery to include such things as interim care by telephone and e-mail to improve chronic disease care could become a reality.

Although I did not know what to expect when I was asked to participate in the planning of this conference, it has been a rewarding experience. It has given me, and I hope will give all of us, the opportunity to voice our opinions regarding new modes of payment. Too often managers and non-clinical personnel make profound decisions regarding how we practice medicine. I hope this conference allows us to shape future payment policies in ways that those of us who actually see patients believe will work best.

— Sidna M. Scheitel, MD, MPH, Mayo Clinic

Section 1. Why Physician Payment Must Change

"The social obligation for best practice is part of the commodity the physician sells."

KJ Arrow, 1963

"Quality problems are everywhere, affecting many patients. Between the health care we have and the care we could have lies not just a gap, but a chasm."

Institute of Medicine, 2001

"There are many mechanisms for paying physicians; some are good and some are bad. The three worst are fee-for-service, capitation and salary."

JC Robinson, 2001

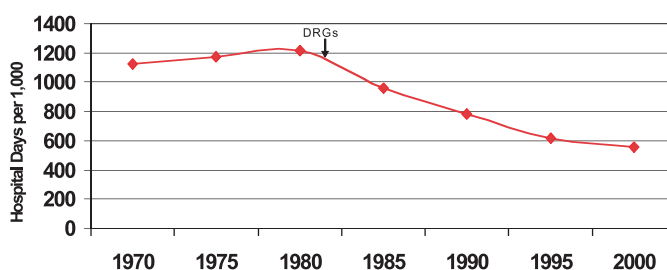
Americans think their health care is the best in the world. A large biomedical research enterprise produces a steady output of new pharmaceutical, medical and surgical products and procedures. Technological diffusion is rapid, giving the United States an incredibly resource-rich health care system. However, availability of technology does not equate with excellence in quality of care. A growing body of empirical evidence has documented gaps between how health care should be delivered to achieve the best possible outcomes and how it is actually delivered (Schuster, McGlynn et al. 1998; Institute of Medicine 2001; Fisher E. S., Wennberg et al. 2003; McGlynn, Asch et al. 2003). These gaps are so large that a 2001 panel of experts convened by the Institute of Medicine called it a "quality chasm" (Institute of Medicine 2001).

Our inability to consistently produce health care of the highest quality results from failures throughout the health care system. It is not merely a problem of getting health professionals "to do the right thing." Change needs to occur at multiple levels: (1) among health professionals to ensure they have current and accurate knowledge, skills and expertise; (2) at the group or team level to facilitate integration of services across practitioners and time; (3) among health delivery organizations so that the necessary infrastructure, such as clinical information systems, is available; and (4) at the larger environmental level to address regulatory, coverage and payment policies (Shortell SM 2004). Transforming health care cannot be done by targeting just one of these levels; we need to align system change at all four levels.

One of the most powerful levers for modifying the organization of health care is through alterations in provider payment policy. Medicare's prospective payment system is an excellent

case in point. In the 1980s, hospital payment for Medicare switched from a retrospective, cost-based payment methodology to a prospective, per-case method. This new payment system, also called Diagnosis Related Groups (DRGs), removed incentives for keeping patients in the hospital. The per-case DRG method gave hospitals a lump sum for every patient hospitalized for a specific condition, regardless of how long they stayed. Hospitals' responses were dramatic and quick (see Figure page 5). Mean length of stay in the nation decreased, leading to fewer hospital days, while ambulatory services and post-acute care increased in importance. Because DRGs dealt with inpatient services only, its impact on cost

Figure. Impact of DRGs on Hospital Days in the US, 1970-2000. (Data Source: National Hospital Discharge Survey)



containment for all health care sectors was minimal. Much of health care delivery shifted from inpatient to outpatient care. The DRG example suggests (1) that health care organizations can be dramatically altered in response to payment and (2) both intended (i.e., lower inpatient costs) and unintended effects (i.e., increased outpatient costs as a result of the DRGs) may result from health care financing reforms that occur in isolation.

The Institute of Medicine has recognized the need to transform physician payment. In its now famous report, *Crossing the Quality Chasm* (IOM 2001), the following recommendations regarding physician payment were made:

- fair payment should be given for good clinical management,
- providers should have the opportunity to share in the benefits of quality improvement,
- purchasers should have the opportunity to recognize quality differences in health care and direct decisions accordingly,
- financial incentives should align with implementation of care processes based on best practices and the achievement of better patient outcomes and
- payment should promote better coordination of care.

Current physician payment systems are not designed to promote quality or better outcomes. Both theory and history support this claim. Fee-for-service is essentially pay-for-

production and offers rewards for seeing more patients, generating more services (whether appropriate or inappropriate care) and upcoding procedures and diagnoses. When used within an environment in which consumers have nearly all their costs covered by insurance, fee-for-service can lead to large increases in health care expenditures. Conceptually, capitated payments should enhance efficiency of health care production and reduce provider and patient demand for services. However, some have argued that capitation causes stinting on care, under-use, quality problems and risk selection. Salaried payment reduces incentives for productivity and is basically a pay-for-time method. None of these forms of provider payment aligns compensation with outcomes. New methods for paying physicians are needed so that doctors are appropriately rewarded for providing high-quality care and promoting better outcomes for their patients.

Section 2. Pay-for-Performance: A Definition

Health care organizations have already begun to pay physicians for meeting quality standards (Casalino L., Gillies et al. 2003; Strunk BC 2004). These forms of compensation are usually called "pay-for-performance (PFP)," because physicians or their practices are given financial incentives for achieving certain quality targets. Recent reports suggest that just 1% to 2% of physician compensation in PFP programs is from incentive pay for quality (Kralewski, Rich et al. 2000; Casalino L., Gillies et al. 2003). Even so, the word from the "market" is that the number of physicians and amount of money that will be involved in some form of PFP in the near future will be substantial (Epstein, Lee et al. 2004).

Patient Vignette: Uncompensated Chronic Care

"Jim is a 53-year-old male with a history of Type 2 Diabetes. He has repeatedly missed scheduled appointments, and over three years his A1C increased from 6.8 on maximum oral agents to 11.3. Only after extensive discussions, many of which occurred over the telephone, did he agree to start on a long-acting insulin at bedtime. Because of his phobia of needles, my staff and I spent a great deal of time between office visits talking with him on the telephone. I would review his glucoses on a weekly basis, adjust his insulin over the phone and calm his fears about changes in vision that would ultimately resolve. Those 10 to 15 minutes a week were "uncompensated encounters," yet they were critical to helping him self-manage his condition. Over three months his A1C decreased to 6.6. He is excited about how well he feels and has a renewed commitment to caring for his diabetes, even offering to speak to other "difficult" patients I might have. I am pleased with his outcome, but it would have been nice to have been compensated for the result not accurately reflected in billing on a fee-for-service basis only."

John E. Anderson, MD
Internal Medicine

Although research projects recently funded by Robert Wood Johnson Foundation and Centers for Medicare and Medicaid Services will provide results on the effects of alternative PFP models, none has been published yet. In general, there is very little research on the effects of PFP methods. We have no empirical information on a wide range of topics related to PFP: how large should payments be, should payments be made to individual physicians or groups, what metrics should be used for the payments, what effects on quality and outcomes do these payments have, and what other changes in the health care system should be made to enhance and reinforce the effectiveness of the PFP model. Thus, PFP is an unproven method of physician payment and could be considered "experimental." This lack of documented evidence of benefit suggests the need for evaluations of PFP interventions.

PFP is not an all-encompassing solution for improving quality. It is one method among a wide array of approaches targeted at different levels of the health care system (the milieu, organizations, groups and individual practitioners) and can be combined with non-financial methods as well.

For the purposes of this document, we define pay-for-performance in the following way:

DEFINITION: PAY-FOR-PERFORMANCE

The use of incentives to encourage and reinforce the delivery of evidence-based practices and health care system transformation that promote better outcomes as efficiently as possible.

We use the term incentives to denote reinforcers. A reinforcer is anything that alters the chances that a behavior occurs (Town, R. 2004), or at the organizational level, that a structural change occurs. In our definition, we therefore include positive reinforcers (e.g., bonus payments), negative reinforcers (e.g., withhold distributions), punishments (e.g., physicians pay a penalty for not meeting target levels for quality or outcome measures) and non-financial mechanisms (e.g., making results of quality assessments publicly available).

The definition indicates that PFP should reinforce high levels of performance while encouraging lower performers to improve. It also makes clear that it is not enough to link incentives with the production of care. PFP is designed to promote better outcomes—that is, better health, functional status and well-being.

Section 3. Pay-for-Performance Design Principles: Assumptions

Four assumptions guided the developed of the PFP design principles. These included: (1) unit of analysis and payment, (2) requirement for investment of new resources or redistribution of existing resources, (3) patient incentives will not be addressed, and (4) the percentage of income at risk will not be addressed.

Assumption #1: Our focus is on incentives targeted at individual physicians and physician organizations.

PFP can be implemented at many levels of the health care system: integrated delivery systems, health plans, hospitals, physician organizations and individual physician practices. For the purposes of this document, our emphasis is on community-based physicians, which is where most individuals get most of their care, and "physician organizations." Both single-specialty and multi-specialty group practice organizational models were considered to be within the scope of physician organizations.

Assumption #2: PFP will require investment of health care resources.

PFP is a tool to improve quality and outcomes. The early stages of PFP programs are almost certain to entail the allocation of new, or a redistribution of existing health care resources. Focusing patient care on the production of outcomes rather than merely producing services will require physicians and their organizations to develop new care management capacity. This may include, for example, disease registries, electronic information systems capable of producing quality and outcome metrics and the addition of nurses and other practice staff to implement these new processes. We expect that any return on investment generated from PFP programs will take several years and will result from better outcomes and reduced complications, which result in

Health Plan Perspective: Who should bear the costs of PFP?

"We have recently begun paying some of our providers bonuses for achieving a variety of quality and outcome targets. We recognized early in the design process of our PFP program that additional funds would be necessary for the practices to enhance their chronic care services and to develop the information systems necessary for state-of-the-art evidence-based medicine. As a purchaser of health care services, we chose to bear the burden of these new costs equally with the physician practices. It is my strong belief that in the next few years, this new "investment" will more than pay for itself with reduced high-cost services."

James K. Geraughty, MD
Senior Medical Director

lower downstream demand for high-cost specialty and hospital services.

Assumption #3: Patient incentives will not be addressed, although we recognize that patient participation in the health care process is critically important to enhancing outcomes.

Better outcomes are not achieved solely by applying high-quality medical care during office visits. Patients themselves play a central role in improving their health. They choose whether and how to participate in the care delivery process, as well as whether and how to implement treatment programs. Physicians and plans have a responsibility for engaging all individuals in health care, not just those who access services on their own accord.

The role of patient self-management as an integral element of health care quality and optimal outcomes has received a good deal of attention in the research literature and the lay press. However, a thoughtful framework for using incentives either through benefit designs or other mechanisms to encourage healthful patient self-management (not just cost-shifting) has not been fully developed. Past experience has demonstrated that financial factors such as differential premium contributions, higher co-payment or restricting access to services results in lower utilization, but these measures also may worsen health outcomes (Soumerai 2004). The growing popularity of consumer-directed plans also places a premium on patient health care purchasing behavior without clear evidence regarding its long-term impact on quality and health outcomes.

This document does not address strategies for providing incentives to patients to align their actions with better outcomes. We concluded that this is a topic so large and complex that both it and physician incentives could not be dealt with adequately at the same consensus conference.

Assumption #4: The amount of physician or practice income affected by PFP is not addressed by our design principles.

The design principles do not address the specific percentage of income that should be performance-based. In one study, bonus payments had little effect on physician organizations' use of care management programs, largely because they were perceived as too small to influence decision-making (Casalino L., Gillies et al. 2003). The specific level of physician or practice income that effects substantive change without unintended negative consequences, such as shunning high-risk patients from physician practices, is unknown.

Supporting the government's proposal for a bold experiment in quality, family practitioners in the United Kingdom voted affirmatively for a new pay-for-performance bonus payment system. Up to 20% of physician income is based on meeting targets for clinical indicators, practice organization indicators and patient experiences obtained from surveys (Roland 2004). The impact of this level of bonus payment on promoting quality and on other perhaps unforeseen changes in patient care is unknown and currently being evaluated. Moreover, we acknowledge that any PFP program must operate within the framework of the statutes and regulations, such as anti-trust and anti-kick-back laws, governing physician payments.

Section 4. Design Principles

This section describes the design principles, which are grouped into six categories: (1) Payment Structure, (2) Transparency, (3) Metrics, (4) Evaluation, (5) Community and Patient Participation, and (6) Fairness.

The conference included an initial session in which each principle and respective design options were discussed in terms of meaning, clarity and comprehensiveness of the description. A second session was devoted to formulating specific wording for each principle. The conference concluded with a consensus process involving all attendees during which final versions of the principles were produced.

The 15 design principles produced at the consensus conference are reproduced below. After each one is stated, we provide a discussion of important and in some cases critical considerations offered by conference participants as they formulated the final wording of the principle.

Design Principle Category #1: Payment Structure

There are several considerations regarding how to structure the incentives used in pay-for-performance. This category addresses four discrete principles related to the structure of payments.

Design Principle 1.1: Accountability Level

The accountability level for physician pay-for-performance may be either the individual physician or practice. Incentives targeted at physicians are more likely than those targeted at practices to change physician behavior; however, practice incentives may have a bigger impact on altering the infrastructure needed to provide high-quality care. Some have argued that within a practice, multiple physicians manage any given patient, so there should not be a single accountable physician designated. Instead, all treating physicians within a practice should benefit from the outcomes and resultant compensation. Health professionals often share both patients and resources, such as clinical information systems and ancillary personnel within a practice. Alternatively, making multiple physicians accountable for a single patient may dilute the impact of the incentive and diffuse a sense of accountability.

Regarding accountability level, the consensus conference participants made the following declaration:

Physician organizations that directly interface with payers should be the accountable entity in PFP programs.

Discussion: Although payments may be distributed to physician organizations, practices are strongly encouraged to disburse funds to individual physicians based on their performance within the group. Maximal effectiveness of PFP programs will be achieved when both physician organizations and individual physicians within those groups are accountable for measurable patient outcomes. PFP programs targeted at physician organizations may be more likely to alter structural aspects of the practice milieu than those targeted at individual physicians.

Design Principle 1.2: Distribution of Financial Incentives

PFP may take the form of a variable component of payment that is added to base compensation, which is unaffected by the PFP formula. How to distribute the incentives, both positive and negative, is a decision that designers must make. One approach is to give all entities some form of positive incentive with increasing amounts linked to better performance. In this scenario, even the low-performing entities would receive some amount of variable payments, albeit not as large as their higher performing counterparts. Providing incentives to most or all physician organizations in the PFP program may be necessary to encourage continued participation among the lower performers.

A second approach is to give positive incentives to entities meeting certain target thresholds for quality or outcomes, and the specific amount may or may not be graded by performance. For some, this approach may be appealing because only "excellent quality" is rewarded. The threshold approach to distributing incentives assumes that physicians will be motivated to make substantial changes in their clinical actions and/or practice infrastructure as a result of the possibility of a reward. It is unclear whether this holds true for average and low performers, who may not perceive the incentives to be within their grasp.

Regarding distribution of incentives, the consensus conference participants made the following declaration:

PFP programs should provide variable incentives to physician organizations that meet certain target thresholds or demonstrate a clear improvement over baseline performance levels.

Discussion: Thus, the threshold approach should be applied to high-performing physician organizations to reward excellence. It should also be applied to organizations showing some minimum level of improvement over baseline to reward substantive improvements in quality.

Design Principle 1.3: Financial Incentive Type

According to recent surveys, several types of rewards have been used in the early PFP programs. Most common are bonuses to groups and individual physicians, tiered co-payments with higher performing providers having lower patient co-payments, payment rates tied to performance and quality infrastructure grants (Bailit Health Purchasing 2002; Strunk BC 2004). Bonus payment may be used as a positive (extra funds) or a negative (expected funds not distributed) incentive. Another approach is to modify a conversion factor used in a fee-for-service system or the capitation rate, higher for better performers and lower for those in the poorer spectrum. The co-payment structure can be used to encourage (by lowering or eliminating them) or to discourage patients from using a physician or practice based on its quality.

There is suggestive evidence that negative incentives may be more powerful stimuli to induce behavior change among physicians than positive incentives. A "quality withhold" is a form of negative incentive. A portion of physician income is set aside pending the achievement of certain quality targets.

Regarding the type of incentive that PFP programs should use, the consensus conference participants made the following declaration:

PFP programs should be based on positive financial incentives.

Discussion: Negative financial incentives should be avoided. In the early phases of PFP programs, developing adequate levels of provider buy-in to the process will be critical to program success. Negative incentives would discourage providers from participating. In addition, loss of income associated with negative incentives can create serious gaps in the flow of practice revenue.

Design Principle 1.4: Frequency of Assessments and Incentive Distribution

There is a necessary lag between the end of the assessment period and availability of data because of the time it takes to collect data and produce the quality metrics. The impact of incentives is likely to be stronger when they are applied close in time to the clinical activity.

Distributions with very long lags are less likely to allow physicians to make timely adjustments in their behavior and practice. Timeliness of assessments and distributions is counterbalanced by the larger payments that may result from longer intervals. Furthermore, frequency of assessment must be weighed against the administrative burden on payers, physicians and practices associated with generating the metrics. In many cases, the assessment interval is determined by the definition of the metric.

Regarding the frequency of assessments and incentive distribution, the consensus conference participants made the following declaration:

Metric assessments and payments should be made as frequently as possible in order to better align rewards to actual performance. Results of assessments should be reported and payments provided to the physicians involved as soon as possible after the close of the measurement period.

Discussion: Frequent assessments in the absence of electronic medical records could place a large burden on practices and payers. Until EMRs or other means of securing performance data are commonplace, the timing of payments under PFP should depend on the technical capabilities of the providers and health plans involved.

Design Principle Category #2: Transparency

Physicians act as patients' agents, helping them, as well as their designated advocates and caregivers, to make medical decisions in the face of uncertain effects and outcomes (Arrow KJ 1963). Physicians' decisions are influenced by their concern for their patients' welfare and health and by their professional norms. Public disclosure of the results of quality and outcome assessments are non-financial incentives that can alter a physician or group's esteem among their peers or patients. Public recognition for quality of care may be a stronger incentive than bonus payments, which in the past have tended to be too small to garner much attention from physicians and practices (Casalino L., Gillies et al. 2003). Either used alone or combined with financial rewards, public disclosure could have powerful effects on modifying the structure and performance of physicians and their practices.

Like many aspects of PFP programs, the effects that public disclosure will have on patients, physicians, the patient-physician relationship and the health care system are not fully known. The argument for disclosure, therefore, is not predicated on a strong research base. It is motivated in part by a general consumerist trend to make more information available to the public.

On an ethical level, patient autonomy argues for transparency. Autonomy refers to the tenet that patients should have all the information they need to make informed decisions. Withholding details regarding the methods used to pay physicians, the providers participating in a PFP program and the results of quality assessments are threats to autonomy. Patients' right to know this information is counterbalanced by physicians' desire to keep information about their professional practice private and confidential. Thus, regarding public disclosure, there is a tension between autonomy of patients and autonomy of physicians. Some research evidence suggests that disclosure of physician incentives to patients does not alter patients' trust in their doctors or insurance companies and may actually have a positive effect on trust (Hall, Dugan et al. 2002).

Public accountability also supports transparency. Society entrusts physicians with powerful prerogatives in the care of patients, a privileged status that must be preserved and strengthened. Transparency of incentives supports continued trust in the profession. Two design principles can be derived from the concept of transparency. The first relates to making the method used to pay physicians or groups transparent to the public. Disclosing the method may also include revealing the identity of the participants in the pay-for-performance program. The second principle relates to disclosure of results of the quality and outcome assessments.

Regarding public disclosure of method, the consensus conference participants made the following declaration:

Design Principle 2.1: Public Disclosure of Method

A list of physician organizations participating in the PFP program, as well as the quality and outcome metrics used in the program, should be disclosed to the public.

Discussion: Disclosure of physician participation in a PFP program sends the public the positive message that the provider is focusing on quality and improving care and that they are willing to be accountable for their performance.

Design Principle 2.2: Disclosure of Results

This design principle addresses transparency of the results produced from the quality and outcome assessments.

Regarding public disclosure of results, the consensus conference participants made the following declaration:

PFP programs should publicly disclose a list of physician organizations who meet quality and outcome target thresholds and those who are demonstrating improvement over time.

Discussion: Ranking of all physician organizations should not be done because of unreliability inherent in conventional statistical methods and the resultant risk of falsely identifying outliers.

There should be a baseline period before public disclosure to provide physicians with opportunities to review, validate and interpret their results. In effect, this preliminary phase would involve disclosure to the physician organization only. A process for validating results and expressing disagreement with the findings should be established in all programs. Once the validity of the quality and outcome assessments has been substantiated, then disclosure of results can proceed.

Some participants strongly felt that disclosure has several potential negative effects. For example, if PFP uses only a limited number of measures, consumers choosing a practice may have incomplete information about the global quality of care delivered by that practice. In such cases, doctors' practices may suffer or be rewarded inappropriately. There were also concerns about potential misuse of such data for the purposes of contracting or in legal cases.

Design Principle Category #3: Metrics

In PFP, the measures used to assess quality and outcomes provide a basis for determining the amount of reward to distribute to providers. Health services researchers have spent years developing the technology necessary to accurately measure quality and outcomes. The field has advanced to a point that there are a sufficient number of metrics with established measurement properties to build credible quality-improvement programs. The design principles in this category build on this knowledge base and add a few considerations that may be unique to the PFP context.

Design Principle 3.1: Measurement Level

For the purposes of quality assessment, three measurement levels can be assessed: structure, process and outcomes.

Structural measures refer to aspects of the health care system that are present before patients and professionals meet. Examples of structural measures include disease registries, electronic information sources, electronic medical records and availability of health educators, social workers and case managers. These measures are easier to obtain than process or outcome measures and place the least data collection burden on the practice. On

the other hand, they have weaker links with outcomes as compared with process measures.

Processes of care refer to what happens when patients and professionals interact. Process quality measures assess the degree to which those interactions conform to evidence-based guidelines of care. Some of these measures have strong empirical evidence for a linkage with outcomes. However, most are disease-specific and cover a very narrow range of clinical activity. Examples of process measures are whether certain lab tests (e.g., A1C and LDL) are checked during a specified interval among patients with particular diseases, patients' assessments of their interactions (also called satisfaction with care), and appropriate administration of certain drugs to patients. Process measures should also include assessments of waste and inefficient practices.

Outcome measures are the intermediate and long-term results of health care, and include changes in health status (both self-assessed health and clinical markers such as organ function), functioning (ability to participate in desired activities, cognition, mobility and self-management) and well-being. Although improving patient outcomes is the most important goal of health care, providers have voiced concerns about being held accountable for those changes in health, functioning and well-being on which their interventions have little direct effect. For many, outcomes factors outside the control of health professionals are critically important determinants. Obtaining outcome information presents the greatest methodological challenges, because many measures require patient report and survey sampling, others rely on laboratory results, and clinical measures such as organ functioning may not be apparent for a number of years.

Outcomes may be intermediate biochemical or physiologic changes (such as LDL level or results from pulmonary function tests) or long-term end-organ effects (such as rate of myocardial infarction). Whereas the latter are arguably more important, it may be more feasible to measure the former.

Ultimately, provider behavior change and system transformation, which are the targets of PFP, should have positive effects on outcomes. However, as these change processes unfold and evolve, it will be essential to have quality assessments done at all three levels of measurement. Structural measures can be used to assess practice infrastructure, process measures tap into practitioner behavior, and physician-patient interactions and outcome assessments provide information on the end results of care processes.

Regarding measurement level, the consensus conference participants made the following declaration:

The metric set used in PFP programs should include a mix of outcome, process and structural measurements.

Discussion: Outcome measures are intentionally listed first, because they are viewed as the most important type of measure for PFP programs. The exact mix of all these measures cannot be specified for every program and will depend on priorities for change and existing technical capabilities.

Design Principle 3.2: Metric Attributes

When selecting specific metrics, PFP designers must choose a number sufficient to cover several clinical processes, but not so many as to engender confusion in the participants regarding how and where they should focus their efforts. Because a large volume of patients may be needed to obtain stable estimates of performance, particularly for disease-specific quality, the balance between comprehensive assessment and respondent burden should be carefully considered.

There are several metric attributes that can be considered during a selection process. The criteria proposed in this document have been adapted from an Institute of Clinical Systems Improvement (ICSI) internal document¹ proposing performance measurement for ICSI member organizations.

First, a candidate measure is more worthy of being used to the extent that the health care structural element, process of care or outcome is common or frequently experienced. We term this criterion **volume**. Improvements in high-volume measures can have a larger impact on the health and health care of patients than those that are lower volume.

Second, the potential impact on health associated with changes in performance is an important consideration. This is called the **gravity** of the measure. For example, measures associated with cancer quality/outcomes have higher gravity than those linked to acne because of cancer's threat to survival. Delays in treatment for life-threatening illnesses have higher gravity than long waiting times for routine ambulatory care. Obstetrical concerns have high gravity because of the many years of health or disability at stake for the newborn.

¹ Personal communication from R. Dinter and G. Mosser

Third, a measure is more worthy to the extent that there is empirical **evidence** linking changes in the metric with clinically important changes in health, functioning or well-being. Alternatively, for outcome measures, there should be evidence that the application of health care is an important determinant of the outcomes. The metric should be actionable—i.e., specific health care actions associated with the metric lead to better patient outcomes. Beta-blocker use post-MI is a good example of an actionable measure for which there is strong evidence of its linkage to future health outcomes.

Fourth, a proposed measure should assess an aspect of performance for which there is a **gap** between current practice and what can be achieved under optimal circumstances. Metrics for which there is variation, such as rates of diabetic foot exam, are most useful. Those which are nearly uniformly done (e.g., blood pressure checks during routine visits) are less useful.

Fifth, a measure is more suitable for use to the extent that the **prospects** for improvement of the measured performance are good. Not only should there be variation in the metric (what we call a "gap" above), but there should also be no external factor that would preclude improvement in the quality or outcomes among health care units.

Sixth, the measure itself should have an acceptable degree of **reliability, validity and feasibility**. In other words, a measure is a better choice to the extent that experience in its use has shown that it produces consistent results over time and across observers ("reliable"), it is consistently associated with outcomes and related health care metrics ("validity") and methods exist for the efficient and minimally burdensome acquisition of data ("feasibility").

Regarding metric attributes, the consensus conference participants made the following declaration:

The attributes of measures used in PFP programs should include:

- **High volume:** common structural attribute or frequently experienced process/outcome of care,
- **High gravity:** large potential impact on health associated with metric,
- **Strong evidence-basis:** research evidence of linkage between change in measure and outcomes,
- **Gap between current and ideal practice,**
- **Good prospects for quality improvement:** no external factor that would preclude health care entities from closing gaps between current and ideal practice,
- **Measurement reliability:** the metric produces consistent results across time and observers,
- **Measurement validity:** the metric actually measures what it is intended to measure and is clearly defined and
- **Measurement feasibility:** methods or technologies exist for the efficient acquisition of the necessary data.

Discussion: The field of quality and outcome assessment changes rapidly enough that PFP programs should have sufficient flexibility to add new metrics and delete existing ones in a dynamic way. The volume consideration can be applied to a general population or within sub-groups defined by age, disease class or some other attribute. With better electronic information systems, the number of measures that can be included in PFP programs will be greatly increased, because the feasibility criterion will be more commonly satisfied.

Design Principle 3.3: Metric Domain

Once the relative mix of structural, process and outcome measures is determined, PFP designers must select metrics from specific quality domains.

Regarding metric domains, the consensus conference participants made the following declaration:

Metrics for PFP programs should be selected from the following quality and outcomes domains:

- **Patient-centeredness:** captures patients' assessments of their experiences in the care process and interactions with providers,
- **Effectiveness:** measures that are linked to health outcomes in real-world settings,
- **Safety:** measures associated with reduced chances of patient harm and
- **Efficiency:** risk-adjusted assessments of service use and expenditures.

Discussion: This approach uses a modified IOM framework for quality assessment (Institute of Medicine 2001). Measures within these four domains can be selected by type of service (i.e., preventive care, acute care, chronic care, long-term care and palliative care) and/or by type of outcome (e.g., biochemical and physiologic outcomes, end-organ outcomes, functional status and well-being).

Patient-centered measures capture patients' assessments of how their physician or the entire health care team respects their personal values and preferences, are responsive to their needs, provides emotional support or physical comfort and involves family and friends. Examples of these measures include: structure—reports on ease or difficulty accessing providers, ratings of office waits, reports about appointment waits; process—evaluations of interactions with health professionals in terms of the respect given to the patient, trust in the provider and emotional support offered; and outcome—patient-reported assessments of their health, functioning and well-being.

Effectiveness metrics include structure or process measures that are linked to outcomes in real-world settings. Structural metric examples are presence of clinical information systems and disease registries, both of which are associated with improved chronic care outcomes. Process metrics could include conformance with clinical practice known to be linked to improved outcomes. Checking AIC on a semi-annual basis is an example of a process-level effectiveness metric, whereas the actual AIC level is an example of an outcome-level effectiveness metric.

Safety measures include metrics associated with reduced chances of patient harm. Drug-drug, drug-age and drug-disease interactions represent three classes of medication-related safety measures. An example of a drug interaction structural measure is presence of electronic prescribing technology, a process measure would be the actual prescribing of a drug inappropriately, and an outcome measure would identify a change in a specific poor health or functional state associated with unsafe medical practice, such as fall injuries associated with sedative use among elders.

Efficiency is a cross-cutting theme in the IOM quality domains. Efficiency measures refer primarily to assessments of utilization and health care expenditures. We believe that they are necessary for a balanced PFP metric set, but they are certainly not sufficient. As our PFP definition indicates, we suggest designers ought to provide incentives that improve quality and promote outcomes as efficiently as possible. To accomplish this goal, some measurement of resource use is necessary. Efficiency measures must be risk-adjusted (for example, by using a method such as the Johns Hopkins Adjusted Clinical Groups [ACG] Case-Mix System) to control for differences in the morbidity burden of patient populations.

Design Principle 3.4: Range of Metrics

Some existing PFP programs focus on a few (e.g., three or four chronic care indicators) or even one metric (e.g., immunization rates), whereas others have opted to include a much larger number. In California, a consortium of six health plans (Aetna, Blue Cross of California, Blue Shield of California, CIGNA Healthcare of California, HealthNet and PacifiCare) have developed a PFP model that includes about a dozen preventive care, chronic care and patient experience measures. Additional measures can be (and have been) added to this list (for more information see <http://www.iha.org>). This relatively small set of metrics contrasts with a much more comprehensive, yet complex, set of indicators in the new British PFP experiment (Roland 2004). British family practitioners' quality bonuses are

benchmarked against performance across well over 100 indicators that assess care for 10 conditions, each with multiple indicators, as well as practice organization and patient experience metrics.

PFP designers must decide on the number of metrics to use to calculate the amount of the incentives. Smaller numbers are easier to explain and understand, but they may also lead to sub-optimal care for conditions not covered in the metric set. A more comprehensive set of indicators runs the risk of too much complexity, leading to physician and practice confusion during implementation.

Regarding range of metrics, the consensus conference participants made the following declaration:

PFP designers should include a sufficient number of metrics across a spectrum of health promotion activities and disease states so that they provide a balanced view of performance.

Discussion: At the onset, this number should be limited within domains, so they could be focused on transforming specific elements of the system and be accurately measured. Over time, this list needs to be re-evaluated and redefined and expanded to be more comprehensive. It is important for multiple groups to be involved in choosing these metrics. These groups include community physicians, insurers, purchasers and patients. Ultimately, the metrics should be broad in scope and physician-specialty specific.

Design Principle Category #4: Evaluation

Because research on pay-for-performance is scant, the actual effects of this new payment method are quite uncertain. One could argue that this degree of clinical uncertainty renders PFP experimental, which suggests the need for evaluations to assess impact in early adopting health care organizations. On the other hand, program evaluations require methodological expertise and additional funds, which are resources that may be difficult to secure for some organizations. It would also be reasonable to suppose that evaluations are best left for the research community, rather than organizations that implement the programs.

Regarding the need for evaluation, the consensus conference participants made the following declaration:

Design Principle 4.1: Need For Evaluation

Every PFP program should have some level of evaluation. The evaluations should include periodic assessments of intended and unintended impacts on access, costs, quality, health outcomes, physician satisfaction and patient satisfaction.

Discussion: A national database of PFP evaluation results should be established so that organizations implementing PFP programs can share their experiences. Funders of health services research are encouraged to support scientifically rigorous studies of innovative programs.

Design Principle Category #5: Community Participation

Within a single medical market, if each payer develops and implements a unique set of metrics, the impact of PFP is likely to be limited. Moreover, a consortium of plans in a community that cooperate on the design and implementation of PFP is more likely to affect a large enough share of physician income to produce real change. Metric sets that do not overlap across payers (or purchasers) using PFP will increase the level of confusion among providers regarding on which aspects of clinical care and practice organization they should focus change efforts. Payers include both public (Medicare and Medicaid) and private organizations. By developing a common set of metrics and implementation procedures across payers within a community, specific community priorities can be the focus of providers' attention. Community-wide participation has the potential to transform health care within a geographic region and will give high levels of visibility to the effort. On the other hand, forcing community-wide participation could limit health plan innovation and may slow down the implementation process.

Regarding the need for evaluation, the consensus conference participants made the following declaration:

Design Principle 5.1: Community Participation

Employers, public purchasers, payers and providers serving the same medical market should develop a common set of metrics and measurement procedures.

Discussion: Community-wide participation facilitates statistically valid evaluation of smaller physician practices by capturing a large share of their patient populations in quality assessments. Moreover, fewer resources among physician organizations are required for measurement if a common approach is utilized. Communities should consider developing common data sets that aggregate data across payers, purchasers and providers in order to have a uniform methodology for assessing and reporting performance. A common set of national metrics and implementation procedures would greatly enhance the capacity of communities to coordinate efforts across organizations. Without community-wide participation, PFP faces a high risk of failure due to the large burden that will be placed on physicians and their practices.

Design Principle 5.2: Patient Participation

Quality and outcomes of care cannot improve without the active engagement of patients in health care processes. Their importance to improving quality is often overlooked.

Regarding involving patients in PFP design and implementation processes, the consensus conference participants made the following declaration:

Patients should be involved in PFP program development and assessment.

Discussion: Patients have a critical role in quality improvement as central actors in care processes. Thus, it is logical that patient preferences should be incorporated into the design of PFP systems.

A dissenting view expressed is that involving patients in the design process is logistically problematic and unnecessary for the success of PFP.

Design Principle Category #6: Fairness

This category relates to how fair PFP is to physicians and patients affected by it. Superior PFP systems minimize the likelihood that any provider or patient group will be unjustly impacted by PFP.

Patients do not randomly distribute themselves to their providers. Some professionals care for sicker or more socially complex patient populations than others. Achieving quality and outcome targets for these providers will be more difficult than for those whose patients are healthier. For example, in a study on health plan quality, organizations with higher percentages of minority, rural and low socio-economic patients had lower quality ratings. Once differences in patient mix were accounted for, the quality rankings of some organizations changed substantially (Zaslavsky, Hochheimer et al. 2000), making some "bad apples" look "good."

Regarding methods to maximize fairness, conference participants made the following declaration:

Design Principle 6.1: Methods To Maximize Fairness

PFPP programs should include methods to maximize fairness by addressing differences in patient health status, social complexity and patient adherence.

Discussion: To promote fairness, this design principle states that assessments, and thus payments based on those assessments, should be adjusted for differences in patient mix across providers (i.e., risk adjusted). Risk adjustment will minimize the effects of risk selection that may be unfair to both patients and their providers. Another concern is that if PFPP assessments are done using patient populations that are too small to provide valid results, results that may be publicly disclosed would unjustly penalize or reward providers.

Patient Vignette: The Difficult Patient

"Howard L. is a 38-year-old male with BMI of 25 who came to my office as a new patient in late 2003 complaining of weight loss, polyuria and nocturia. His random glucose was 374 and his dipstick urinalysis was negative for ketones. Treatment for diabetes was initiated. Laboratory tests were obtained the next day revealing an A1C of 14.0, with LDL of 131, HDL-34 and triglycerides 364. His urine microalbumin was elevated at 38.4. At his follow-up appointment one week later, a great deal of time was spent with Mr. L. educating him about primary prevention, diabetes and self-management, metabolic syndrome and dyslipidemia. He was advised that he absolutely must discontinue all nicotine use and reduce alcohol to no more than seven standard drinks per week. A referral was made for comprehensive diabetic education and medical nutrition therapy in addition to a referral to ophthalmology for a dilated retinal examination. A comprehensive foot examination was performed, demonstrating a normal quantitative somatosensory threshold. Both pneumococcal and influenza vaccinations were provided and low-dose aspirin therapy was initiated. Wellbutrin was prescribed to assist with smoking cessation.

This office visit required 42 minutes of face-to-face time.

In spite of multiple efforts by telephone and mail, Mr. L. did not return for follow-up until 13 months later. He never went for monitoring laboratories, the dilated retina exam, diabetic education or medical nutrition therapy. He had long since run out of medications and was no longer monitoring his plasma glucose. We had a long talk about the importance of forming a therapeutic patient-physician relationship, adherence with prescribed therapies and follow-up in addition to the importance of controlling his multiple risk factors as quickly as possible. His repeat laboratories demonstrated terrible glycemic control with an A1C of 12.3 and diabetic dyslipidemia. This office visit required 23 minutes of face-to-face time.

Based on these lab "outcomes," one could conclude that I was doing a terrible job of managing this patient and should not receive a positive incentive under a pay-for-performance arrangement. Pay for performance must have risk adjustment and other systems in place to reward dedicated physicians who are willing to go the extra mile for the problem patient such as Mr. L. so that these folks do not end up being shuffled from one practice to another."

Charles H. Booras, MD
Family Medicine

Regarding sample size, conference participants made the following declaration:

Design Principle 6.2: Sample Size

PFP assessments should be done using patient samples that are large enough to produce statistically meaningful results.

Discussion: If an adequate sample size cannot be achieved using data from the physician organization only, results could be pooled across reporting units in order to gain sufficient sample size. For small practices, statistical reporting may be unreasonable and alternative methods for assessment may be needed. No physician organization should be excluded from PFP programs because of the size of its patient population.

Section 5. Conclusions

There is growing interest in changing physician compensation to promote better quality and patient outcomes. Some organizations have already begun to offer physician organizations bonus payments, better contracts and other financial rewards to better align payment with quality. Additionally, public recognition of both good and poor quality is being used and offers a powerful supplement to financial incentives. All these programs have been developed without substantive input from physicians. This document fills this information gap.

Both positive and potentially negative outcomes may result from pay-for-performance programs. For example, there may be a change in the holistic, patient-oriented approach to patient care, if health care is delivered by managing the metric rather than managing the patient. Specifically, some of the potentially negative impacts include:

- Disincentives for physicians to practice in areas with patient populations that have high levels of health care needs or social complexity;
- Less attention to patients' psychosocial needs due to an increased biomedical orientation of health professionals;
- Fragmentation of care that could result from management of metrics rather than management of patients;
- Physician concern that PFP is being done to decrease their income;
- If co-payment tiering is used, access could worsen for patients whose physicians are underperforming and as a result have higher co-payments;
- Disclosure of PFP participation or results could have deleterious effects on the doctor-patient relationship among physicians who are poor performers;
- Loss of the art of medicine because of a preoccupation with charting, flow charts and other forms of documentation;
- Poorer quality of care for conditions not included in the incentive system;
- Higher practice administrative costs entailed in generation of the PFP metrics;

Physician Testimonial: Maintaining the "Art" of Medicine

After beginning participation in a pay-for-performance program, I have noticed that I am spending more time during a patient encounter reviewing flow sheets and checklists. Although I know that this process will provide better evidence-based care for my patients, I fear that it is distracting from the personal interaction. PFP, by distracting from the human interface and hindering a physician's ability to individualize care, may damage the "art" of medicine.

Sid King, MD
Internal Medicine

- Incentives for physicians and organizations to "cherry-pick" the easiest patients to manage and
- Poorly run programs may discourage physicians from participating.

It is incumbent on PFP designers to build mechanisms that monitor the effects of the program on patient access, practice burden, quality and outcomes for conditions not targeted by the PFP formula. Unintended negative effects, if detected, should prompt reassessment, and potentially a redesign, of the PFP program.

If incentives are sufficiently powerful to modify clinician behavior and practice structure, some possible positive effects include (adapted in part from Roland 2004):

- Better access to and delivery of preventive services;
- Potential for reducing waste and inefficiency;
- Increased use of electronic information systems, including medical records and disease registries;
- Stronger connections with community resources that patients may call on to enhance chronic care self-management;
- Improved primary care management of chronic disease with more practices specializing in chronic care;
- Better quality; and
- Improved outcomes.

The relative balance between positive and negative effects needs to be carefully monitored by designers and evaluators of PFP programs. The ongoing input and feedback of physicians will be critical to determining the future success or failure of PFP.

References

- Arrow KJ (1963). "Uncertainty and the welfare economics of medical care." *Am Econ Rev*: 941-973.
- Bailit Health Purchasing, L. (2002). Provider incentive models for improving quality of care. Prepared for the National Health Care Purchasing Institute: An Initiative of the Robert Wood Johnson Foundation.
- Casalino L., R. R. Gillies, et al. (2003). "External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases." *Jama* **289**(4): 434-41.
- Epstein, A. M., T. H. Lee, et al. (2004). "Paying physicians for high-quality care." *N Engl J Med* **350**(4): 406-10.
- Fisher E. S., D. E. Wennberg, et al. (2003). "The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care." *Ann Intern Med* **138**(4): 273-87.
- Hall, M. A., E. Dugan, et al. (2002). "How disclosing HMO physician incentives affects trust. Not all cost-minimizing physician incentives are ethically troubling to patients." *Health Aff (Millwood)* **21**(2): 197-206.
- Hofer, T. P., R. A. Hayward, et al. (1999). "The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease." *Jama* **281**(22): 2098-105.
- Institute of Medicine (2001). *Crossing the quality chasm: a new health system for the 21st century*. Washington, D.C., National Academy Press.
- Kralewski, J. E., E. C. Rich, et al. (2000). "The effects of medical group practice and physician payment methods on costs of care." *Health Serv Res* **35**(3): 591-613.
- McGlynn, E. A., S. M. Asch, et al. (2003). "The quality of health care delivered to adults in the United States." *N Engl J Med* **348**(26): 2635-45.
- Roland, M. (2004). "Linking physicians' pay to the quality of care—a major experiment in the United kingdom." *N Engl J Med* **351**(14): 1448-54.

- Schuster, M. A., E. A. McGlynn, et al. (1998). "How good is the quality of health care in the United States?" *Milbank Q* **76**(4): 517-63, 509.
- Shortell SM (2004). "Increasing value: a research agenda for addressing the managerial and organizational challenges facing health care delivery in the United States." *Med Care Res Rev* **61**(3): 12S-30S.
- Soumerai, S. B. (2004). "Benefits and risks of increasing restrictions on access to costly drugs in Medicaid." *Health Aff (Millwood)* **23**(1): 135-46.
- Strunk BC, H. R. (2004). Paying for quality: health plans try carrots instead of sticks. Issue brief: findings from HSC. Washington, D.C.: 1-4.
- Town R, K. J., Dowd B (2004). "Assessing the influence of incentives on physicians and medical groups." *Med Care Res Rev* **61**(3): 80S-118Ss.
- Zaslavsky, A. M., J. N. Hochheimer, et al. (2000). "Impact of sociodemographic case mix on the HEDIS measures of health plan quality." *Med Care* **38**(10): 981-92.

Summary Of Pay-Per-Performance Design Principles

Design Category #1: Payment Structure		
1.1	Accountability Level	Physician organizations that directly interface with payers should be the accountable entity in PFP programs.
1.2	Distribution of Incentives	PFP programs should provide variable incentives to physician organizations that meet certain target thresholds or demonstrate a clear improvement over baseline performance levels.
1.3	Incentive Type	PFP programs should be based on positive financial incentives.
1.4	Frequency of Assessments and Incentive Distribution	Metric assessments and payments should be made as frequently as possible in order to better align rewards to actual performance. Results of assessments should be reported and payments provided to the physicians involved as soon as possible after the close of the measurement period.
Design Category #2: Transparency		
2.1	Public Disclosure of Method	A list of physician organizations participating in the PFP program, as well as the quality and outcome metrics used in the program, should be disclosed to the public.
2.2	Public Disclosure of Results	PFP programs should publicly disclose a list of physician organizations who meet quality and outcome target thresholds and those who are demonstrating improvement over time.
Design Category #3: Metrics		
3.1	Measurement Level	The metric set used in PFP programs should include a mix of outcome, process and structural measurements.
3.2	Metric Attributes	<p>The attributes of measures used in PFP programs should include:</p> <ul style="list-style-type: none"> • High volume: common structural attribute or frequently experienced process/outcome of care, • High gravity: large potential impact on health associated with metric, • Strong evidence-basis: research evidence of linkage between change in measure and outcomes, • Gap between current and ideal practice, • Good prospects for quality improvement: no external factor that would preclude health care entities from closing gaps between current and ideal practice, • Measurement reliability: the metric produces consistent results across time and observers, • Measurement validity: the metric actually measures what it is intended to measure and is clearly defined and • Measurement feasibility: methods or technologies exist for the efficient acquisition of the necessary data.

Summary Of Pay-Per-Performance Design Principles

Design Category #3: Metrics		
3.3	Metric Domain	<p>Metrics for PFP programs should be selected from the following quality and outcomes domains:</p> <ul style="list-style-type: none"> • Patient-centeredness: captures patients' assessments of their experiences in the care process and interactions with providers, • Effectiveness: measures that are linked to health outcomes in real-world settings, • Safety: measures associated with reduced chances of patient harm and • Efficiency: risk-adjusted assessments of service use and expenditures.
3.4	Range of Metrics	PFP designers should include a sufficient number of metrics across a spectrum of health promotion activities and disease states so that they provide a balanced view of performance.
Design Category #4: Evaluation		
4.1	Need for Evaluation	Every PFP program should have some level of evaluation. The evaluations should include periodic assessments of intended and unintended impacts on access, costs, quality, health outcomes, physician satisfaction and patient satisfaction.
Design Category #5: Community and Patient Participation		
5.1	Community-Wide Participation in Program Development	Employers, public purchasers, payers and providers serving the same medical market should develop a common set of metrics and measurement procedures.
5.2	Patient Participation	Patients should be involved in PFP program development and assessment.
Design Category #6: Fairness		
6.1	Methods to Maximize Fairness	PFP programs should include methods to maximize fairness by addressing differences in patient health status, social complexity and patient adherence.
6.2	Sample Size	PFP assessments should be done using patient samples that are large enough to produce statistically meaningful results.

We encourage your input and feedback.

Please direct all correspondence to: *info@amhealthways.com* or

American Healthways, Inc.

3841 Green Hills Village Drive

Nashville, TN 37215

(800) 327-3822

www.americanhealthways.com



3841 Green Hills Village Drive
Nashville, TN 37215
1-800-327-3822

V.1 ©2005 American Healthways, Inc.