

PRE-PUBLICATION DRAFT

Please do not quote, cite or copy without notification to the author.

**CONTRACTING FOR PROVIDER QUALITY:
THEN, NOW AND P4P**

By Alice G. Gosfield, Esq.

Alice G. Gosfield and Associates, P.C.
2309 Delancey Place
Philadelphia, PA 19103
215-735-2384
215-735-4778
agosfield@gosfield.com
www.gosfield.com

**“Accepted for publication in the Health Law Handbook, 2004 Edition.
Alice G. Gosfield, Editor, (c) West, A Thomson Company.
A complete copy of the Health Law Handbook is available from West by calling
1-800-382-9352 or online at www.west.thomson.com.”**

- TABLE OF CONTENTS -

Chapter 3

CONTRACTING FOR PROVIDER QUALITY: THEN, NOW AND P4P

- 3.1** Quality Policy Evolution
- 3.2** P4P in Brief
 - 3.2.1** Bridges to Excellence
 - 3.2.2** Integrated Health Association
 - 3.2.3** Leapfrog
 - 3.2.4** CMS Physicians
 - 3.2.5** CMS Hospitals
 - 3.2.6** Others
- 3.3** Existing Financial Incentives
- 3.4** Provider Contracts
 - 3.4.1.** Network Configuration
 - 3.4.2.** Access and Availability
 - 3.4.3.** Record Keeping and Access to Information
 - 3.4.4.** Communication Clauses
 - 3.4.5.** Medical Management
 - 3.4.6** Formularies
 - 3.4.7.** Other Clauses and Overview
- 3.5** Perspectives
 - 3.5.1** Quality Impacts
 - 3.5.2** Financial Implications
 - 3.5.3** Contract Concerns
- 3.6** Conclusion

CONTRACTING FOR PROVIDER QUALITY: THEN, NOW AND P4P
Alice G. Gosfield

To create significant leaps in the quality of care by recognizing and rewarding health care providers who demonstrate that they have implemented comprehensive solutions in the management of patients and deliver, safe, timely, effective, efficient, equitable and patient-centered care
---Bridges to Excellence Mission Statement

To mobilize large purchasers to alert the healthcare industry that big leaps in patient safety and customer value will be recognized and rewarded with preferential use and other intensified market reinforcements.
--The Leapfrog Group

Leap – v. – to pass over at a bound
-- The Oxford English Dictionary

With their reference to robust, athletic imagery, two of the most prominent efforts today to propel health care quality demonstrate an exigency to their missions that reflects lagging provider behavior. Drawing on the six principles for quality set forth in the Institute of Medicine’s landmark study, Crossing the Quality Chasm¹, both the Bridges to Excellence program and the Leapfrog Group are among a developing cadre of efforts springing up all over America to reward improved provider quality. These programs come in a variety of iterations. Some focus on leveraging employer dollars directly to providers to change behavior. Some originate with health plans. Some are government sponsored. All exist in a world in which contracts define the obligations of providers in relationship to who pays them. Yet, astonishingly, at least to an attorney, some of these new “pay for performance” (P4P) programs have no contractual link from the payor source to the provider whose responsive behavior is sought. In other words, the provider has no legally actionable claim that money is owed him based on the performance he understood was the goal of the very incentive the program is about. In other settings, there is a contractual link with the payment source, but the provider waives any right to dispute the payment. These confounding approaches to motivating changed provider behavior so as to yield specific health care results are set against a backdrop of longstanding contractual relationships in the basic managed care contract between the provider and the plan.

¹ Corrigan et al, CROSSING THE QUALITY CHASM, Institute of Medicine, National Academy Press, Washington, D.C., 2001.

Given these new developments it is time to reexamine the contractual context for health care quality with particular attention to providers. There are other quality leverage points on the health care delivery continuum²: (1) where the purchaser mandates certain quality-relevant behavior from the plan which seeks to be chosen as the insurer the purchaser will make available to its covered population; (2) where the plan ‘tiers’ its provider-hospitals by developing health insurance products that group hospitals into segments based on the cost or efficiency of care they deliver and then steers patients to choose the preferred providers through lower premiums or cost sharing;³ (3) within physician groups where the employment compensation reflects quality concerns;⁴ and (4) where the purchaser motivates the enrollee to exercise more cost effective or quality relevant behavior as in General Motor’s program which pays its covered population to use better providers.⁵

Although most of these initiatives arise within the context of managed care – whether traditional HMOs, PPOs or hybrid combinations – not all of them depend on the presence of a managed care contract. Still, the P4P programs tend to turn on identified patient populations, so managed care programs with better mechanisms to identify patient targets are typically, but not exclusively, the focus of these efforts. Moreover, it cannot be denied that our attention to quality of care has heightened considerably with the advent of managed care, not only because its incentives led to criticism of increased health care risk from their emphasis on cost containment, but also because managed care systems combining payment and delivery offered the first widespread opportunities to develop meaningful data about health care performance.

To understand the mechanisms and import of the P4P programs, this article (1) examines the evolution of the concept of ‘quality’ which led to these new payment mechanisms and what they are leaping to accomplish; (2) offers a snapshot of the P4P

² I would like to thank Peter Lee, CEO of the Pacific Business Group on Health for his taxonomy of these leverage points.

³ Reportedly these programs have met substantial provider resistance and have raised concerns regarding quality. Some payors have abandoned an exclusively cost driven approach to the tiers and have allowed some hospitals to be in preferred tiers if they produce certain quality behaviors, but some of these programs have been abandoned in some communities where they were initiated. Still, in light of increasing premiums, they continue to draw interest from health plans as another cost containment strategy. See, Mays, Claxton and Strunk, “Tiered-Provider Networks: Patients Face Cost-Choice Trade-Offs”, Issue Brief No 71 (Nov. 2003) Center for Studying Health System Change, <http://www.hschange.com/CONTENT/627/> Still, some payors – notably Aetna and United – are rolling out products which combine much more narrowly restricted, better performing provider networks with consumer incentives to use these networks, reportedly at the instigation of employers. “Innovative Products offer Narrow provider networks Targeted to High Cost Diseases”, *Managed Care Week*, October 6, 2003.

⁴ This is an area around which there has been very little research. See, Robinson, “Theory and Practice in the Design of Physician Payment Incentives,” 79 *The Milbank Quarterly*, 149 (2001).

⁵ See Norris, “Health costs squeeze everyone: Employers and employees struggle with rising costs,” *Detroit Free Press*, Nov 26, 2001. (http://www.freepress.com/money/business/benes26_20011126.htm)

world and the variety of its implementation techniques in relationship to providers; (3) considers how those initiatives relate to diverse financial incentives that have proliferated and remain in play; (4) looks at traditional elements of provider contracting which co-exist with these programs; and (5) concludes with an assessment and some speculation about whether these undertakings can produce the bounding results their sponsors seek given the contractual setting in which they are deployed.

3.1 Quality Policy Evolution

Over thirty years, the progressing concern for health care quality in this country reflects both persistent themes and shifting orientations. Ten years ago, the Institute of Medicine noted that the literature documents efforts over at least the last eighty years to place concepts of quality assessment and assurance in health care into operational frameworks.⁶ Avedis Donabedian, widely regarded as the modern father of quality analysis, drew attention in 1966 to focusing assessment of quality on three aspects of health care: (1) structure or inputs – the context in which actual services are provided, including the physical environment, the qualifications of personnel, and the organizational structure of the institution under review; (2) process – how care is delivered, including technical competence and appropriateness of the services to the condition; and (3) outcomes or end results – what happened to the patient as a result of care, i.e., improvement, deterioration, no change. From a public policy perspective though, because Medicare had emerged in 1966 as a significant factor in the federal budget, quality and cost were joined as dual if not sometimes paradoxical policy foci.

In 1972, when the last HR 1 generated the first widespread Medicare reform, Congress enacted the Professional Standards Review Organization (PSRO) program, to provide a unified system for Medicare and Medicaid to assure the delivery of appropriate, cost effective care in accordance with professionally developed norms, criteria and standards to be applied by 54 physician organizations around the country.⁷ Before then, the primary mechanism of assuring quality in Medicare lay in Conditions of Participation -- the entry requirements for hospitals to be eligible to receive Medicare dollars. The Joint Commission on Accreditation of Hospitals got its major grip on the hospital quality sector nationally when it was recognized in law in a unique role, so that on the basis of a private Joint Commission survey a hospital would be deemed qualified to participate in the Medicare program.⁸ The only condition which could not be deemed was the utilization review standard which was oriented around length of stay,

⁶ Lohr ed., *MEDICARE: A STRATEGY FOR QUALITY ASSURANCE*, Vol I, National Academy Press, Washington, D.C., 1990, p.45.

⁷ For a discussion of the development of the PSRO program see, Gosfield, *PSROs: THE LAW AND THE HEALTH CONSUMER*, Ballinger Publishing Co., Cambridge MA, 1975.

⁸ 42 USC ' 1395bb

and therefore cost, an aspect of review which was not part of the Joint Commission approach.⁹ Six years later, questions had already been raised about what it meant to give this power to a private organization, and the concept of validation surveys was introduced, giving the Secretary of HEW the right to come in to do a resurvey to be sure what the Joint Commission had found was still a valid basis for a hospital to remain in the Medicare program.¹⁰ But, when in the same legislation, he introduced the amendment creating PSROs, Senator Bennett observed:

In addition to the rapidly rising cost of health care, a problem exists with respect to the quality of that care. The Committee on Finance held two extensive series of hearings on health care in 1970....During the course of those hearings, disturbing testimony was heard bearing on the quality of health care. We heard practicing physicians testify to the effect that in many areas of the country a good deal of unnecessary and avoidable surgery was being performed and excessive and inappropriate health care services provided. We learned of significant variations between sections of the country in the lengths of hospitalization for similar patients having a given illness.¹¹

His concerns are unfortunately quite contemporary today. Still, in the early 1970's public policy about health care quality focused primarily on the mechanisms in the Medicare program to assure some measure of quality.

In 1974, the Institute of Medicine, which had recently been chartered (in 1970) as part of the 100 year old National Academy of Sciences "to serve as adviser to the nation to improve health," published one of its earliest health care studies. That monograph addressed where and how quality assurance programs functioned in health care; and, since most of health care policy was focused around services within hospitals, hospital quality assurance activities were the locus of assessment. "The primary goal of a quality assurance system should be to make health care more effective in bettering the health status and satisfaction of a population, within the resources which society and individuals have chosen to spend for that care."¹² While the goal was unassailable as social policy, there was no widely agreed upon definition of health care quality. The

⁹ Id and 1395x(e)(6)

¹⁰ 42 USC ' 1395bb(2)

¹¹ 118 Cong Rec. S418 (Jan 25, 1972)

¹² Institute of Medicine, *Advancing the Quality of Health Care*, National Academy Press, Washington, D.C., 1974, pp 1-2. Between 1974 and 1989, the Institute of Medicine focused relatively little on quality. The reports on quality during that period included "Assessing Quality in Health Care: an Evaluation Report" (Nov. 1976), "Improving the Quality of Care in Nursing Homes" (March 1986). Besides, the sporadic reports cited here, it was not until 1999 that the Institute of Medicine began to focus consistently on quality of care issues.

fluidity, variety, and sweep of the then extant definitions of quality was so broad that in its Medicare Quality Strategy work fifteen years later, the IOM formally analyzed and reviewed about 100 definitions of quality.¹³

Beginning in the mid-1980's, influential policy journals began to pay more attention to quality. In Spring, 1988 Health Affairs published a themed issue focused on quality issues ranging from the quality of medical evidence, to involving consumers in quality, to competition and quality, to how HCFA would promote quality, to literally defining quality in medical care. Clearly policy pundits were concerned about quality, but no real public debate over the issues or legislative action, or even significant voluntary initiatives captured the public imagination. It was not until the widespread implementation of techniques of managed care that the public policy debate over quality would reach far and wide.

Instead, the late 1980's and early 1990's saw the advent of "value purchasing". Rising health care costs increased attention to controls on utilization. The early 1970's utilization review – evaluating care already rendered to deny payment -- gave way to more proactive "utilization management."¹⁴ Major corporate purchasers of health care were beginning to struggle with producing more value in their own consumer products. This was the era of the burgeoning dominance of the Japanese automakers and the big three automakers found themselves challenged to make "Quality Job 1." But quality at what cost? American industry could not compete without improving quality at a fair price. The watchword of major segments of the American economy, which were also seeing fast increasing health care costs, became "value".¹⁵

Transferring what they were experiencing in their core businesses to this progressively draining part of their budgets, purchasers began to make it clear that what they wanted and therefore plans and providers should seek to provide was controlled costs, albeit using appropriate care processes that demonstrated reasonable health care outcomes, but the value would, more and more, have to be demonstrated in data. To be able to assure themselves of value, the purchasers would offer their employees fewer health insurance choices, which would increase competition among those seeking to be the health care delivery vehicle of choice. To increase the value to them, purchasers would try to choose the scope of services and providers whom they would pay for health care which meant they would have to have some data on which to

¹³ IOM, Medicare Strategy, n.6, Vol II., pp. 116-139

¹⁴ See, Gray ed., CONTROLLING COSTS AND CHANGING PATIENT CARE? THE ROLE OF UTILIZATION MANAGEMENT, National Academy Press, Washington D.C., 1989

¹⁵ See, Gosfield, "Value Purchasing in Medicare Law: Precursor to Health Reform", XX Amer J of Law and Med 169 (1994); Gosfield, "Value Purchasing and Effectiveness: Legal Implications", HEALTH LAW HANDBOOK, 1991 ed., Clark Boardman Callaghan, New York, 1991, pp. 185-217

compare the options. It was during this period of time that report cards and other manifestations of performance measures began to emerge as significant factors in health care.¹⁶ Still, though, the primary concern was escalating health care costs.

By 1985 Business Week had its cover story headlined as “Health care: the cost crisis is over – but what about quality?” Paul Ellwood, widely regarded as one of the godfathers of managed care in this country, picked up the banner for the ‘alternative delivery systems’ which he saw as under attack over the “more is better” debates in health care.¹⁷ He noted that until that point most of the creativity in developing managed care models had gone into methods to compensate physicians, market services on the basis of price, reduced hospitalization and improved access and other elements of service that enhanced patient satisfaction. Citing the remarkable but then as yet unrealized potential of organizations that had responsibility for large, defined populations, comprehensiveness of care, prepayment and common medical records to take on quality in a meaningful way, he articulated a prescription for managed care quality: (1) Quality of care should be identified as a separate highly valued function within the organization; (2) Clinical quality and patient service are legitimately inseparable since dissatisfied patients do not comply with clinical advice; (3) Traditional distinctions around structure, process and outcomes are inapplicable and simplistic in this context; (4) Quality assurance system development ought to command at least 1% of an organization’s revenues; (5) Non group structures need to make an even greater investment in quality than do groups to overcome the disadvantages of ‘dispersion of physicians’; (6) Some of the national organizations need to make major investments to develop the next generation of quality assurance systems. His prescription was not filled. In fact, well into the late 80’s most managed care organizations remained mired in the creative endeavors he had noted absorbed their energies. For their failure to embrace quality as a core mission, plans paid a price in managed care backlash.

The theory of “managed care” had been to create incentives and systems that would literally manage care to contain costs. By merging payment and delivery, health maintenance organizations offered radically different incentives to behavior. Gatekeepers, restricted networks, and changed compensation, primarily to capitation models, were all new techniques that were intended to produce different care delivery patterns. That was the point. The dilemmas this created, by comparison with traditional fee for service indemnity insured care, were succinctly articulated by a federal court judge.

¹⁶ See, Gosfield, “Health Care Report Cards: Quality in the Public’s Cross Hairs”, *HEALTH LAW HANDBOOK*, 2000 ed., WestGroup, New York, 2000, pp. 501-542

¹⁷ Ellwood and Paul, “But what about quality?”, Health Affairs, (Spring 1986) pp. 135-139.

A health maintenance organization offers for a fixed fee, as much medical care as the patient needs. Providers using traditional fee-for-service methods, by contrast, charge for each procedure. Each method creates an unfortunate incentive: a physician receiving a fee for each service has an incentive to run up the bill by furnishing unnecessary care, and an HMO has an incentive to skimp on care (once patients have signed up and paid) in order to save costs. Each incentive encounters countervailing forces: patients or insurers on their behalf, resist paying the bills for unnecessary services, and HMOs must afford adequate care if they are to attract patients. HMOs also have reason to deliver excellent preventive medicine. Prevention may reduce the need for costly services later. Competition among the many providers of health care, and between the principal methods of charging for the care, affords additional protection to consumers.¹⁸

The perverse incentives of fee for service medicine were tempered by utilization review programs which did not claim an explicit focus on quality since they reflected insurance payment principles.¹⁹ Safeguarding against the perverse under-utilization incentives of managed care was the reason for the creation of the National Committee for Quality Assurance (NCQA), the managed care accrediting organization.²⁰ The theory was that by evaluating organizations ensuring they had an infrastructure that incorporated continuous quality improvement, health plans could distinguish themselves within a market on the basis of quality.

The early NCQA accreditation standards focused on quality management and improvement, utilization management, credentialing, members' rights and responsibilities, preventive health services and medical records. An entirely separate parallel program also grew up in the form of HEDIS – the Health Plan and Employer Data and Information Set. As managed care became more prevalent nationally, large purchasers and plans sought to find some way to limit the multiplicity of requests to which plans had to respond in order to get business, while simultaneously giving purchasers some methodology upon which to compare plans. Although it began as a program launched by a benefit consultant and a group of large HMOs, by 1993 it was turned over to NCQA.

There were some 65 measures in HEDIS 2.0, the first NCQA sponsored version of it addressing issues such as member access and satisfaction, utilization, finance and

¹⁸ Anderson v. Humana 24 F 3d 889 (7th Cir 1994)

¹⁹ See Gosfield, "Is Less Really More?: Utilization Management in the 1990s," HEALTH LAW HANDBOOK, 1996 ed., Clark Boardman Callaghan, pp 89-122.

²⁰ For a broad consideration of concerns regarding managed care quality in the mid 1990's see Gosfield, GUIDE TO KEY LEGAL ISSUES IN MANAGED CARE QUALITY, Faulkner and Gray, NYC, 1996

health plan management. These choices reflected the then current priorities in HMO management that Paul Ellwood had noted ten years earlier. The quality measures per se included the extent of delivery of preventive services, prenatal care, treatment of acute and chronic illness (measured by asthma inpatient admission rates and performance of diabetic retinal exams). It was not terribly clinically focused. Its limitations reflected the feasibility of actually measuring anything. The measures were selected based on (1) relevance and value to employers who were paying the freight; (2) reasonable ability of plans to develop and provide the requested data; (3) and potential impact on improving processes of care. Even so, market need for this type of information was manifest since, by 1996, 59% of employers with 10,000 or more employees were using HEDIS. More than 330 health plans were reporting HEDIS data. Still, though, not all health plans were accredited, not all health plans reported HEDIS data, and neither was a guarantee that no quality problems would occur.

By this time, the major health reform which had been expected in the earliest days of the Clinton administration had failed. The resulting market chaos produced a variety of quality-relevant reactions which also played into new notions of quality.²¹ Confusing to purchasers, providers and consumers alike, the boundaries between the functions of the players in health care became blurred. Large health care systems coalesced from smaller community based providers. Some of them obtained licenses to operate their own HMOs. Employers shifted to self-insurance mechanisms becoming more like the plans to which they had previously just paid premiums. Insurers and health systems purchased physician practices. Publicly traded companies bought physician practice assets, as well as hospitals, and managed them on a national basis. Even HMOs were bought and sold nationally creating ever larger MCOs. The corporatization and consolidation of health care at all levels permeated the market. The sheer complexity of delivering health care services intensified with more entities transacting business at more points along the health care continuum. From new kinds of independent practitioners to new types of vendors of supplies and services, to increased subspecialization of physicians, the elements of the delivery system which had to be marshaled even in a single episode of patient care were more diverse, far flung, and diffused. The potential for quality problems, both within organizations and from redundancy and continuity problems among entities, increased with this complexity.

Contemporaneously, there was new concern for the lack of systematically applied science in a health care industry which spent extraordinary sums on the development of new and expensive techniques and technologies which paradoxically, were used widely when not yet proven effective even as proven developments were not timely brought to bear for most Americans. What to do? Congress entered the fray

²¹ See, Gosfield, "The New Playing Field", 41 St. Louis Univ. L.J 869 (1997)

with the creation of the Agency for Health Care Policy and Research and its clinical practice guidelines development program. That initiative was directly connected to the new physician reimbursement system in Medicare. An annual volume performance standard which would allow the size of the budget for Medicare physician payment to be linked to quantification of the services that ought to appropriately be provided to the Medicare patient population.²² Concerns for what was appropriate also turned on widespread unexplained variation in the health care services delivered around the country. John E. Wennberg MD and his colleagues had shown repeatedly over time that where he lived reflected more about the volume of services a patient would receive than what science might dictate.²³ This unexplained variability had both quality and cost implications. The Agency turned to the IOM to help it determine how to develop guidelines which would reflect, as the statute required “the best available research and professional judgment.”²⁴ Whether care to be provided was appropriate, necessary, effective, helpful, or just plain worth paying for was a central issue.

Quite apart from Medicare, the legislation called for government sponsored clinical practice guidelines which would take into account three priorities: (1) improve methods of prevention diagnosis, treatment and clinical management for the benefit of a significant number of individuals; (2) reduce clinically significant variations among physicians in the particular services and procedures utilized in making diagnoses and providing treatments; (3) reduce clinically significant variations in the outcomes of health care services and procedures.²⁵ Although it eventually published 18 guidelines, throughout its existence, the Agency was challenged in political wrangles and budget fights. Its impact through the guidelines program was more on the side of establishing appropriate standards for the creation of guidelines than in moving quality in meaningful ways by guidelines development alone.²⁶ Still, the legislation did fix the concept of more science based decision-making in the policy zeitgeist, and spurred a trend that is only today seeing a real flourishing.

Related to the variability work of Wennberg, others, primarily including David M. Eddy MD PhD in working on AHCPH guidelines (as well as otherwise²⁷) began to call the question regarding the quality of evidence that could be said to demonstrate the

²² Gosfield, “Clinical Practice Guidelines and The Law: Applications and Implications,” *HEALTH LAW HANDBOOK*, 1994 ed., Clark Boardman Callaghan, pp 65-100.

²³ Wennberg, “Dealing with Medical Practice Variations: A Proposal for Action,” *Health Affairs* (1984), pp 6-32.

²⁴ 42 USC ' 299b-1(b)

²⁵ 42 USC ' 299b-3(a)(2)

²⁶ See Gosfield, *KEY LEGAL ISSUES*, n.20, pp 191-203.

²⁷ See Eddy, “Clinical Decision-Making: From Theory to Practice,” a series of twenty-two articles from Jan. 1990 through Feb. 1996 in the *Journal of the American Medical Association*.

best research and professional judgment based upon it. Generally credited with coining the phrase “evidence-based medicine”, Eddy’s work began to refine and discipline efforts at the evaluation of care to reflect what scientific research could substantiate as useful. But his efforts and those of the Cochran Collaborative in England were hardly granted sweeping endorsement in health care delivery and financing in this country.

In the absence of real and widespread evidence as the foundation for care delivery and payment policies, the late 1990’s saw a far more intense focus on whether managed care denied patients purportedly medically necessary services in the interests of cost, both in policy discussions as well as in lawsuits where patients had been harmed. When the ERISA statute foreclosed to many aggrieved plaintiffs a litigation avenue to redress their harm, disgruntled patients and their physicians who felt their professional judgment both impinged upon and thwarted, joined to confront legislatively the problems of ‘drive by deliveries’, one day mastectomies and a host of other clinical restrictions – topics which state legislatures had never seen before. In response, many states adopted managed care reform legislation.²⁸ The result was a sort of public hysteria over the techniques of managed care which in many ways found its apotheosis in 1997 in the spontaneous applause that erupted in movie theaters all over the country when Helen Hunt in the movie “As Good as It Gets” cursed her chronically ill son’s HMO coverage and deified the caring, HMO resistant physician who eventually treats him.

In 1996, responding to managed care backlash, President Clinton convened an Advisory Commission on Consumer Protection and Quality in the Health Care Industry. By the next year a major policy statement was put forth in “Why the Quality of US Health Care Can Be Improved?”²⁹ By 1998, responsible policy analysts sought to dispel the notion that fee for service medicine was good and managed care was bad. In a highly influential paper in 1998, from the National Quality Roundtable, came the following observation:

At its best, health care in the United States is superb. Unfortunately, it is often not at its best. Problems in health care quality are serious and extensive; they occur in all delivery systems and financing mechanisms. Americans bear a great burden of harm because of these problems, a burden that is measured in lost lives, reduced functioning, and wasted resources. Collectively these problems call for urgent action.³⁰

²⁸ See Roth, Chapter 6 in Gosfield, KEY LEGAL ISSUES, n.20, pp 133-157.

²⁹ Schuster, McGlynn, Brook, Report to the National Coalition on Health Care, Rand Corporation, (Oct. 1997), www.nchc.org/emerge/quality.html

³⁰ Chassin, Galvin et al, “The Urgent Need to Improve Health Care Quality”, 280 JAMA 1000 (Sept. 16, 1998)

Donabedian's structure process-outcomes triad was insufficient to get at the real problems of health care which now were characterized as ones of misuse, overuse and underuse. And then, in 2000 the Institute of Medicine published *To Err is Human*, the seminal work on medical errors which through its sensational projection of needless deaths galvanized attention to problems in health care like nothing had previously. Even the Harvard study, which had pointed out ten years earlier the extent to which events which merited malpractice lawsuits occurred in hospitals,³¹ did not produce the reaction the errors study did. That patient safety from errors was a component of the broader quality problems was now joined as a major policy issue.

That science and evidence ought to be the bedrock of health care returned with a vengeance in the debates over autologous bone marrow transplants for breast cancer and the lawsuits brought to obtain insurance payment for them, which were just another part of the challenges to the restrictions managed care imposed. Heartrending stories of patient deaths from denial of this and other unproven techniques were litigated over and over when there was almost no evidence of the efficacy of the treatment. When screening mammographies should be paid for was another battleground. The evidence based issues have found their most recent high profile manifestation in the response to the data over the cancer risks from hormone replacement therapy. What the evidence showed and how to apply it was now not only squarely in the consumer zeitgeist, it was garnering more than mere policy attention.

The 2001 IOM study *Crossing The Quality Chasm* propounded principles for a new health system for the 21st century and unlike prior reports has seen its critical values furthered in significant ways. It is this study which underlies the principles in many of the P4P programs. The fundamental values to be manifest in this new system are that health care should be:

Safe—avoiding injuries to patients from the care that is intended to help them;

Effective—providing services based on scientific knowledge to all who could benefit and refraining from providing services to those not likely to benefit (avoiding underuse and overuse respectively);

Patient-centered – providing care that is respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient values guide all clinical decisions;

³¹ Leape, Brennan et al, "Patients, Doctors and Lawyers: Medical Inquiry, Malpractice Litigation and Patient Compensation in New York," A Report by the Harvard Medical Practice Study to the State of New York, 1990.

Timely – reducing waits and sometimes harmful delays for both those who receive and those who give care;

Efficient – avoiding waste, including waste of equipment, supplies, ideas and energies;

Equitable – providing care that does not vary in quality because of the personal characteristics such as gender, ethnicity, geographic location and socio-economic status.³²

At long last, a set of common, simple principles that can create a quality driven health care system! Yet to get from where we are today to a national health care system which employs these principles at its core is a major undertaking. Given the time lags that my brief, idiosyncratic review of the history of quality policy in this country demonstrates, no wonder major purchasers feel the need to force an energetic, concerted push to leap over the increments of change – an evolution demonstrably so glacially paced that we find ourselves more than thirty years later facing the same concerns which motivated Senator Bennett in 1972 ; and still the overuse/misuse issues he cited are juxtaposed against our new understanding that Americans get only 55% of the health care services the evidence would indicate they should.³³

Why have we failed to make real our unquestionably more sophisticated understanding of what is necessary to improve the quality of our health care? Many have argued that it is the lack of a business case for quality. “Who will pay for the infrastructure we need to do this work? How can we take the time to advance from where we are when we have so little time already to devote to actual care because of the exigencies of the payment system in which we operate?” And so, the P4P programs seek to propel needed change to more science, more safety, and more patient centeredness made known in a more transparent system where performance can be seen, compared and judged based on public reporting. They will do this by paying for results, paying for new processes and systems, or otherwise using purchasing power to compel a market response.

3.2 P4P in Brief

Paying for performance is not an unknown concept in other settings. It is so straightforward as to be almost startling to be considered so innovative in health care. Why would anyone pay the same amount for vastly different grades of products and services? We don't pay that way for gasoline, perfume, clothing or cars. In most of

³² Corrigan et al, *CROSSING THE QUALITY CHASM*, n.1, p. 5-6.

³³ McGlynn et al, “The Quality of Health Care Delivered to Adults in the United States,” 548 *NEJM* 2635 (Jan 26, 2003).

health care we pay the same amount for poor, mediocre and excellent quality care. At the same time, there is not enough high quality care nor enough evidence-based medicine actually delivered to patients.

While pay for performance seems utterly self-evident in some ways, it has now come to be seen as critical to a business case for quality that primarily will motivate physicians. Some of the P4P programs continue to talk about “realigning the incentives” to the players. At its most simplistic P4P entails paying some additional amount of money to physicians or hospitals for certain behavior. However, as we will see below, not all of the P4P initiatives involve only enhanced payments; some have other motivational techniques, too. Most have either a regional focus or are pilot projects. Many involve multiple purchasers and multiple health plans in a community. Perhaps the two highest profile programs are those of Bridges to Excellence (BTE) which comes from the purchaser community and the California initiative stimulated by the Integrated Healthcare Association involving six health plans.

3.2.1 Bridges to Excellence

The rationale for focusing on physicians is explicitly stated in the BTE materials. Among the three primary quality management-cost containing strategies identified by BTE as care coordination, disease management and information technologies, no one solution has emerged as the optimal efficient and effective model to improve quality. BTE notes further that care coordination and disease management have limitations because there are no incentives without capitation, there is a lack of financial and non-financial implementation capacity and there are no current IT standards to which incentives can be attached. Therefore, “creating an incentive system based on linking outcomes and process improvements to lower costs and improve quality is the best approach the payors can take in this stage of the continuing evolution of the healthcare system.”³⁴

This program began as a relatively loose series of principles to be applied by purchasers seeking to change costly behavior in communities where they were paying for substantial amounts of care. It has morphed into a detailed, regionally focused, clearly elaborated full-fledged program that is incorporated and has its own governance structure which draws in local leadership in each region. BTE is guided by three operating principles intended to “realign **everyone’s** incentives around higher quality.”³⁵

- “Re-engineering care processes to reduce mistakes or require investments, for which purchasers should create incentives;

³⁴ “New Employer – Existing region implementation toolkit,” Bridges to Excellence, www.bridgestoexcellence.org.

³⁵ Id at page 5.

- Significant reductions in defects (misuse, underuse, overuse) or reduce the waste and inefficiencies in the healthcare system today;
- Increased accountability and quality improvements will be encouraged by the release of comparative provider performance data delivered to consumers in a compelling way.”

The program is affirmatively marketed to those employers in a region who recognize that “(1) the market is not delivering optimal care; (2) that patient, provider and purchaser interests can be at odds; (3) that purchasers have little way of knowing how effectively and efficiently their healthcare dollars are being used; (4) that there is little or no accountability at the individual provider level for the quality of care delivered; and (5) that the purchaser has not yet created a robust business case for better quality of care by recognizing and differentially paying providers that can demonstrate better outcomes.” Employers who accept those tenets are asked to finance change in three realms within the target community: (1) diabetes care; (2) physician office processes; and (3) cardiac care. There are payments to physicians and rewards to consumers who are actively engaged. The design of the program is intended to generate and conform to the six principles of the *Quality Chasm* study.

The rationale for selecting diabetes care -- the most evolved aspect of the program to date -- is explicitly articulated in terms of both cost and quality: only 25% of diabetics receive appropriate hemoglobin tests (HbA1c) on an annual basis. Up to three-quarters of diabetics did not receive recommended care from their provider in the middle range state from 1997-99; and over 50% of diabetic patients did not perform recommended self-care during the same period of time. The cost of diabetic care has gone up dramatically in the past few years. From 1997-2002, national diabetes medical care costs as identified by BTE more than doubled from \$44 billion to \$92 billion dollars, and the average cost per diabetes patient as of 2002 was \$13,243 by contrast with the average cost per non-diabetic patient of \$2,560. And still these patients are not receiving necessary and appropriate care. Research and published literature show a \$350 per diabetic patient per year savings opportunity. It is on this basis that BTE takes the position that “to move the market we need to focus our leverage on engaging consumers and engaging providers.”³⁶

This Diabetes Care Link (DCL) initiative offers to physicians who are certified by the NCQA Diabetes Physician Recognition Program (DPRP), \$100 per diabetic patient per year and the cost of the certification application will be paid as well. In touting improvements in diabetic care as a result of the NCQA certification, BTE notes that from 1997-2002 DPRP applicants improved the average rate of diabetes patients who

³⁶ Id. Page 7.

had hemoglobin A1c levels of less than 7 from 25% to 45%, their low-density lipoprotein cholesterol (LDL-C) levels below 100 mg/dl rose from 17% to 45% and the rate of diabetes patients monitored for kidney disease rose from 60% to 83%. This is asserted to be significantly higher compliance than average.

To be eligible for the payment, primary care physicians and endocrinologists must have at least 10 diabetic patients combined across plans. The fact that this program involves multiple health plans in a community distinguishes it from the plan generated programs, so that the incentive is the same across the chronic population. This eliminates one of the barriers to change in health plan sponsored programs where a plan with relatively little market share has difficulty capturing the attention of physicians who have widely dispersed payor mixes in their practices. In marketing the program to physicians in Louisville, BTE noted that groups could join if together they had an average of at least 10 diabetic patients across the group or a total of 70 patients treated by the group based on inpatient and ambulatory claim and encounter data supplied by the participating health plans. With the launch of the program each physician would receive a count of his or her diabetic patients. The marketing materials specifically say there is no appeal of the count in the first year, but such an appeal might be possible in the future depending on the first year's experience.

To become certified, a DPRP physician must meet goals for specific measures thereby earning points for achieving those goals. For example there are outcomes measures as in the proportion of patients with HbA1c below 8% (goal: 55%) and those with HbA1c above 9.5% (goal: less than 21%). There are separate measures for pediatric patients. There are measures common to both a one year and three year DPRP certification; and then for the three year certification there are additional measures.

It is expected that it takes an hour to complete the NCQA application form and 15 minutes to abstract a single medical record to obtain the data on the patient, depending on the nature of the medical records the physician maintains. A criticism of this mechanism is that all the data is self-reported by the physicians. Reportedly "random audits of applications to NCQA will occur to ensure the validity of the data submitted."³⁷ Interestingly, BTE states that some physicians could earn up to \$20,000 in bonuses under the program, but this would mean they had 200 diabetic patients in the practice. To abstract the data for those patients even at merely 15 minutes a chart would entail 50 hours of work devoted just to producing this data.

Consumers are to be motivated by the opportunity to earn CareRewards when they engage in effective self-management techniques through a specifically designed website, customized to display the employer's logo. The patients in the program earn points for various aspects of self-care – appropriately frequent blood glucose

³⁷ For more information on the NCQA diabetes certification program, see <http://www.ncqa.org/dprp/dprpmain.htm>.

monitoring, appropriate drug compliance, maintaining a four times a week exercise program and weight, getting an annual eye exam, discussing foot care with the physician, having blood pressure and cholesterol checked at an office visit, having their hemoglobin A1c checked, and maintaining an appropriate hemoglobin level. Where the patient is seeing a physician who is NCQA certified, presumably the physician will be performing the services that reward the patient. Depending on how many points are earned over a six month period, the patient would get a bonus from the employer of coupons worth \$5, \$15 or \$35 dollars which can be used to purchase diabetic treatment supplies and lifestyle products from a dedicated vendor. If the patient is not seeing a physician who would do this on his own, the patient becomes his own stimulant to proper quality care.

It is expected that the employer would pay \$150 per patient per year, but would realize savings of \$350 per diabetic patient per year. Although these seem very small amounts on a per patient basis, assuming 300 covered diabetic employees or dependents, and that 50% of them go to certified physicians and 50% of the patients qualify for the consumer rewards made available, projected net return on investment is \$52,500 for the 300 diabetic patients. In November, 2003, BTE paid its first bonus payments to physicians in Cincinnati and Louisville. The program had been launched there the previous April involving General Electric, Proctor and Gamble, Verizon, UPS and Ford. The average payment was claimed to be in the thousands. The largest payment made was \$7500.

The Cardiac Care Link (CCL) program will not be implemented until 2004 and then in Albany and Schenectady. It also draws on an NCQA recognition program for heart and stroke treatment, which measures similar goals and outcomes for heart and stroke patients including LDLs tested and controlled, blood pressure tested and controlled, use of aspirin and smoking cessation advice. It is estimated that if all the measures are met there would be an annual savings per patient of \$390. The elements of this program were not as well elaborated as of the writing of this article.

The third component – the Physician Office Link (POL) program – is intended to move physicians' offices more in the direction of *Quality Chasm* principles and to reduce defects that cost money and harm patients. Unlike the diabetes care program which is targeted at PCPs and endocrinologists, the POL program targets all physicians and all patients in the region. The primary sources of savings are projected to come from disease management, computerized physician order entry, electronic health records, and ambulatory care sensitive hospital programs and care coordination.

POL focuses on three aspects of physician offices including clinical information systems, patient education and support, and care management. Awarding points for 35 separate measures, the POL assessment looks at the use of basic registries with follow up, use of electronic registries and systems for prescribing, checking efficacy and safety,

and retrieving results with follow up, elements of partial electronic medical records or an advanced electronic medical record. In support of patients POL evaluates provision of educational resources, referrals for risk factors and chronic conditions and quality measurement and improvement. In care management, the program looks at care of chronic conditions in terms of identification of process management and resources to assist with medication compliance, appointments and the like. POL applicants are also evaluated in terms of how often they create preventable admissions to the hospital and then how they care for high risk medical conditions.

The physician payment for complying with any or all of the three modules in the first year is \$50 per covered patient per year. In year two the payment goes down to \$20 if only one module is in place but remains the same if all three modules are present. To motivate the most comprehensive implementation, in the third year the payment goes down again to \$10 for only one module, to \$30 for two and remains at \$50 only if all three modules are in place. This program is linked to the DCL and CCL-condition-specific bonuses by withholding 20% of the POL bonus until the practice meets one or both of the condition specific programs. The physician gets the full POL bonus plus an extra \$80 for each diabetic and cardiac patient when meeting CCL and DCL. Projecting a 3 PCP practice with 1000 patients covered by the program of whom 3.5% are diabetic and 2.5% are cardiac patients, the practice would receive a total of \$54,800: \$40,000 for meeting the POL measures and \$14,800 for meeting the DCL and CCL measures. Unlike the DCL chart abstraction estimates, though, there are no estimates of either the time or cost of implementing the systems in the POL measures. Implementation of the POL program was slightly delayed until December 2003, so there were no results to report at this writing.

This fairly elaborate approach to moving physician practices in specified markets turns on several critical contractual relationships, but none of them involves the providers themselves. Employers enter into a contract as a BTE participant which obligates the employer to provide money for the general contractor for data, MEDSTAT, and for the physician incentives. The CareRewards program for the consumer incentives is optional. BTE itself has an agreement with MEDSTAT. The employer, participating health plan and MEDSTAT enter into a data use, confidentiality and indemnification agreement. MEDSTAT and the plan enter into a HIPAA Business Associate Agreement. Indemnifications abound. The physicians, whose behavior change is sought, have no privity of contract with any of the parties with respect to the anticipated payment. They participate with the health plans which continue to pay them under those contracts, and the health plans and the employers have their contractual relationships. There may also be disease management programs in place at the applicable health plans, including for diabetes and cardiac care. There are contracts implementing those relationships and programs as well.

While BTE has explicitly stated that “Incentives have to be compelling enough that physicians cannot afford to ignore them” it is too early to know whether this program has found that sweet spot. It is also interesting to speculate on how broad a leap this approach represents with its focus on two basic conditions. The rigor of evaluation, though, has been built into the program and there are explicit measures of whether it will be found to have succeeded. Still, the essential features that motivate physician behavior to improved quality lie in meeting the NCQA recognition standards, being paid to do so, and having the recognition publicized to consumers. This is the same model that was used to stimulate HMO accreditation by NCQA. There, when major employers like Xerox, IBM and the Big Three automakers said they would not offer HMOs that were not accredited, plans changed their behavior in significant ways in order to get that business. Whether physicians will be so motivated will turn on more complex factors. Whether consumers will make any changes in selecting physicians based on the NCQA recognition is also not yet known. Whether recognized physicians experience adverse selection by becoming more popular with more expensive patients in a capitated payment system is another unexplained potential unintended effect, and then whether the BTE payments are sufficient to offset the cost of the resulting more expensive patient population is unknown.

3.2.2 Integrated Health Association

By its own statement “the Integrated Healthcare Association (IHA) is composed of top decision-makers from the major health care stakeholder groups in California. Its mission is to promote the continuing evolution of integrated health care, supported by financial mechanisms that align incentives of purchasers, payors, and providers, as the best means to achieve positive outcomes for the patient and the general public.”³⁸ Its participants include representatives from the leadership of health plans, physician groups, and health systems, plus at large academic, purchaser, pharmaceutical industry and consumer representatives, involved in policy development and special projects around integrated health care and managed care.

In January, 2002, the IHA got agreement from six health plans in the region to use a common set of performance measures to evaluate providers and offer enhanced payment. Although the plans agreed to use common measures, they did not agree on how to apply them in terms of payment. Among the measures, cardiac care is an issue in common with BTE, and although both IHA and BTE target office processes, there the similarities end. The IHA model is keyed far more toward the types of HEDIS measures that have long been in play in California where not only does managed care

³⁸ www.ihha.org

dominate the payment models, but the physician groups perform under delegation many of the functions that plans do themselves in other markets.³⁹

The clinical measures include percent of patients receiving childhood immunizations, breast cancer screening, cervical cancer screening, use of inhaled corticosteroids in asthmatic patients, patients discharged from the hospital with acute myocardial infarction, coronary artery bypass graft (CABG) or percutaneous transluminal coronary angioplasty (PTCA) with evidence of LDL-C screening, and diabetic patients enrolled for at least a year who had evidence of HbA1c screening. Assessment includes measures of patient satisfaction with specialty care, timely access to care, doctor-patient communication and overall ratings of care. Groups are also measured on their investment in information technology including integrated clinical electronic data sets and mechanisms to support clinical decision-making at the point of care. The performance on the measures is then weighted.

All of the plans tier the physicians and key enhanced payment to their performance percentile by comparison with other physicians. By contrast with BTE, this means whether any physician or physician group will realize real dollar benefit is unknown going in, and no group has the ability to influence other groups with whom they are compared. To the extent physicians are measured and rewarded individually, large groups will have more influence than small ones, but still, the ultimate amount of the pay for the comparative performance is speculative at best.

The payment enhancements among the plans for performance are plan-specific. Although the IHA program was touted as entailing common measures among all six plans, the implementation among the participating plans is quite variable. Aetna will pay up to 3.5% of the capitation payment, based on all the common measures weighed 50% on clinical, 40% on patient satisfaction and 10% on IT (which is the weighting most of the plans have adopted). Groups must have 12,000 Aetna member months in 2003, and have signed delegation agreements (meaning they perform these functions at the group level rather than relying on the plan to do so) for patient management, claims payment and credentialing. If the physicians meet the 75th percentile and up on each measure they earn up to the full bonus amount. None of this is in the physician group contract with the plan.

Blue Cross has two programs: The first, the Quality Incentive Program adds pap smears as a clinical measure and does not measure for IT investment. In addition, 72.5% of the bonus they offer is keyed to non P4P measures including grievances, audit, UM

³⁹ NCQA created its physician organization certification program primarily to meet the need in California for physician groups to be able to be approved to perform delegated functions to lower administrative burdens on them of having to comply with disparate demands from multiple plans. See www.ncqa.org/Programs/Accreditation/Certification/poc/pocftp.htm. It has long been the case in California that physician groups are quite large and that single plans do not dominate their payor mix as they do in other markets. This is one of the reasons that the common measures among six plans were seen as an advance in terms of administrative burdens on the physician groups.

requirements and some others. The physicians are paid up to \$4.50 per member per month with a \$2.50 additional bonus for generic prescribing depending on the percentile of their performance (e.g., at the 20th, 40th, 60th and 80th). Contract amendments are used to specify the payment parameters and obligations. The second, Quality Incentive III, offers \$2.00 per member per month depending on the percentile, calculated differently from the other program and is explicitly decoupled from the contract with the physician group. New groups and existing groups with no other incentive plan are automatically enrolled. Existing groups that have some incentive payment receive pro-rated payments as they transition to the new model.

CIGNA uses the IHA common measures and adds more for encounter data which demonstrates annual thresholds of procedures in lab, radiology and professional outpatient of \$.05 per member per month with a maximum of \$.15 per member per month; and additional \$.10 per member per month for maintaining panels open to CIGNA members. There is a minimum of \$1.60 per member per month for the top performing groups in all components measured in this program. Apparently CIGNA has never put bonuses in their contracts and are not adding these either.

HealthNet's maximum bonus is \$2.25 per member per month plus another 10% bonus if the group has an individual physician bonus plan. Each of the measures is weighted differently and comprises a specified percentage of the maximum bonus amount. Groups must be contracted for the program for both 2003 and 2004. Bonus payments will not be set off against risk pools and an amendment must be signed affirmatively removing the incentive program from their contracts.

In a unique spin on performance, Pacificare pays the physicians an increased bonus if the hospitals at which they practice perform more effectively; so, Pacificare has added four additional hospital measures and six other clinical measures. It uses Leapfrog participation and measures, with performance targets for this aspect of the program. In essence this tiers the hospitals to the physicians rather than to the consumers: a physician under this mechanism has a personal incentive to use the better hospitals. They offer a \$2.00 per member per month maximum if the patients are commercial only, with an additional dollar for the Medicare patients. Performance is pro rated by percentile and is documented in the contracts. Here, though, the physicians waive their right to dispute any of the payments made under the program.

3.2.3 Leapfrog

The Leapfrog Group is not actually a payment program. Consisting of some 145 public and private organizations that provide health care benefits, The Leapfrog Group (TLG) hopes that by harnessing the buying power of the 34 million consumers its purchaser members represent, it can foster speedier developments in the market to patient safety and quality, primarily in hospitals. TLG has promulgated standards for three aspects of hospital quality -- Computerized Physician Order Entry (CPOE), ICU Staffing with Intensivists, and Evidence-Based Hospital Referral -- which it hopes to see

rewarded by preferential contracting and improved payment by its members which will result from applying its six purchasing principles.

TLG expects its purchaser members to (1) inform and educate employees about the importance of comparing the performance of health care providers and assist them in understanding how to use such measures to make informed health care choices. To do so purchasers will aggregate available validated performance information on their major providers into (2) comparative value ratings for their covered populations. (3) The real impact will come from purchasers using two or more of three types of incentives to reward delivery systems with higher value and “will annually increase their intensity until they prove sufficient to motivate widespread and substantial annual performance improvement among their major providers.” The first recommended technique is (a) driving patient volume by several methods including promoting by selection and deselection the preferred more value producing providers; through consumer economic incentives, whether in higher co-pays if the patient goes to lower value providers or otherwise; and, by consumer decision support through easily understandable performance valuation. The second differentiation is through (b) price variations including value-based bonuses, or rebates and/or risk adjustments; and the third incentive is through (c) public recognition.

The fourth purchasing principle is to (4) focus on discrete forward leaps in patient safety. Here the selection of the three safety topics – CPOE, intensivists and evidence based hospital referral – are to be targeted in selecting providers and plans that also emphasize these specific leaps in their proven performance. The fifth principle is to (5) hold health plans directly accountable for applying these principles through nationally standardized Leapfrog questions in health plan RFPs, heavily weighted scoring criteria, “robust health plan performance incentives” and other methods consistent with Leapfrog. Finally, in selecting benefit consultants and brokers, purchasers will (6) create strong incentives for them to incorporate Leapfrog principles in both their advice to their clients and in their own standard tools for assessing health plans and delivery systems.

The first step in the comparative performance information has been the voluntary hospital survey as to progress in meeting the three standards. To explain the standards more fully: (1) the intensivist standard is based on the fact that mortality rates are significantly lower in hospitals where intensive care units are managed exclusively by board certified intensivists in closed units where patients are cared for by teams of clinicians who work together in fine tuning routine ICU care as well as responding to emergencies. These physicians, board certified in specially identified areas (e.g., emergency medicine, critical care, anesthesia, pediatrics or surgery with subspecialty training in critical care), must be present in the ICU during daytime hours and at least 95% of other times must be able return to the ICU in five minutes and/or arrange for a certified non-physician responder to reach ICU patients within five minutes.

(2) As to the evidence based hospital referrals (EHR), patients with high risk conditions ought to be treated for elective admissions at hospitals with characteristics shown to be associated with better outcomes. The first six conditions selected for attention are coronary artery bypass graft (CABG), percutaneous coronary intervention, abdominal aortic aneurysm repair, pancreatic resection, esophagectomy and high risk deliveries involving low expected birth weights, short gestational age or pre-natal diagnosis of major congenital abnormality. To meet the standards a hospital must have a specified minimum volume of the procedures or admissions per year and meet outcome and process standards too. Outcome measures are risk adjusted mortality for CABG and percutaneous coronary interventions. Although process measures are present for all of the conditions they are somewhat less emphasized.

(3) CPOE systems are electronic prescribing systems that intercept medication errors when they most commonly occur – at the moment of prescribing. The costs in both health care expenditures and mortality associated with adverse drug events (ADEs) are significant. According to the IOM medication errors alone contribute to 7,000 deaths annually and \$2 billion per year in hospital costs.⁴⁰

To meet TLG's CPOE standard, a hospital has to assure that 75% of physicians enter hospital medication orders through a CPOE system; demonstrate that its inpatient CPOE system can alert physicians to at least 50% of common prescribing errors using a customized testing protocol which was still being developed in December 2003; and require that physicians electronically document a reason for overriding a CPOE intercept.

When it published the standard TLG was really pushing the industry since fewer than 2% of hospitals have these systems in place. One reason is the very significant costs. TLG cites the Brigham and Women's upfront expense of \$1.9 million and \$500,000 annual maintenance but also cites savings to the hospital of between \$5 and \$10 million annually. No wonder purchaser engagement to drive this expenditure is particularly emphasized.

Hospitals have now begun reporting their progress. As of October 2003, 1012 hospitals had voluntarily reported data, a 60% increase over the previous year. In addition, as the size of the pool of hospitals reporting grows, the percentage reporting as adhering to TLG standards goes down. There is speculation this may be because the first volunteers were those hospitals who had less risk of reporting because they were already engaged in these initiatives. The previous year, of the 637 hospitals reporting, 5% had fully implemented CPOE and another 25% expected to reach the standard

⁴⁰ Kohn, Corrigan, Donaldson, ed., TO ERR IS HUMAN, National Academy Press, Washington, D.C., 2000.

within a year. Under the more recent survey 3% had an operational CPOE system in place and 13.5% were on target to have a system in place by 2005. These are hardly sweeping responses, and are self-reported at that. CPOE has been controversial because of the expense in putting it in place and the difficulty of getting physicians engaged to use it. Still, the results on intensivists are only somewhat better. In 2002 21% of reporting hospitals had an intensivist on duty 8 hours a day and 16% more said they were on target to have that standard met in the next year. In the next year though, 17.6% of hospitals met the 8 hour standard and 4.8% said they were on target for compliance in a year.⁴¹ Like BTE – and many of the same purchasers are involved in both initiatives – TLG uses a regional roll out strategy. ⁴² TLG says it has 22 regional initiatives to integrate community-wide, multi-stakeholder collaboration into its national initiative.

Also, like the NCQA physician recognition programs that are core elements of the BTE physician incentives, the Leapfrog hospital data is self-reported. MEDSTAT collects, analyzes and supports the data submissions, and the results are posted on the Leapfrog website. Hospitals are reported as “fully implemented”, “good early stage effort”, “willing to report publicly”, “did not disclose” or “does not apply” for each of the measures. “Fully implemented”, however, does not mean what one might expect. Fully implemented means at least 75% of prescribers use the CPOE system. Fully implemented on the ICU module means truly complies with the standards as they are published. The EHR standard is scored by giving credit for each of the factors on the volume, outcomes, and process measures; so it is reported slightly differently.

Purchasers and plans are beginning to use the Leapfrog standards in payment. In one of the earliest reported movements in that direction Empire Blue Cross and Blue Shield, a major payor in one of the regional roll out communities, says it will pay hospitals who meet TLG standards a 4% bonus in their payment. But there is no contract between Empire and the hospitals that establishes this. When questioned, the general counsel stated, “We will do what it says on our website.”⁴³

3.2.4 CMS Physicians

⁴¹ Morrissey, “Adding up quality standards,” *Modern Healthcare*, October 20, 2003, p. 8.

⁴² In fact, in terms of a People Magazine perspective, the similarities and overlaps are not surprising since one man, Robert S. Galvin, MD, MBA; Director, Global Healthcare at General Electric has been instrumental in the creation of both initiatives.

⁴³ Personal communication to author 8-13-2003.

The Medicare program has also gotten on the P4P bandwagon in physician demonstration projects, albeit pursuant to statutory mandate.⁴⁴ In the Physician Group Practice (PGP) Demonstration, Medicare will test a hybrid payment methodology for paying physician-driven organizations that combines Medicare fee for service with a bonus pool derived from savings achieved through improvements in the management of patient care and services. In projects to involve six groups of at least 200 physicians⁴⁵, the focus is chronically ill Medicare beneficiaries. A three year demonstration project, PGP purports to combine fee for service with “new” financial incentives that are “more in line” with those used by managed care organizations and commercial payors.⁴⁶ In addition to the bonus pool, annual performance targets will be established for each participating group equal to the average Part A and Part B expenditures of beneficiaries assigned to the group, adjusted for health status and expenditure growth. Measured against that baseline, savings will be evaluated. Bonuses are only paid if the participating group demonstrates Medicare savings and achieves quality goals. The savings are measured by comparison with other groups in the market.

The participants to whom the patients are assigned are unique primary care physicians, meaning each patient is assigned to one physician who is credited both with the performance and the resulting payment, even if more than one PCP takes care of the enrolled patient. The project uses 25% withholds of the bonus amount and an annual settlement. The final project settlement comes a full year after the end of the third year of demonstration. The most a PGP can get paid is 15% of target Medicare expenditures for the assigned population. A savings of 2% is the threshold and no bonus is earned if that is the only level of savings produced, even if the process measures are met. The target measures selected come primarily from HEDIS and Medicare Health Care Quality Improvement Program (HCQIP) and include such typical chronic care measures as eye exams, HgA1c tests, and lipid profiles for diabetics, mammograms, chest x-ray and EKG within 3 months after initial diagnosis of CHF, flu and pneumonia vaccines, office visits every 6 months for patients with any of 4 chronic conditions.

⁴⁴ The PGP program was mandated by ' 412 of the Benefits Improvement and Protection Act of 2000. The goals are to encourage coordination of Part A and Part B services; promote efficiency by investment in administrative structure and care processes and reward physicians for improving health outcomes. There is another physician focused quality initiative in the Doctor Office Quality project which is a 3 state pilot project working through QIOs (the old PROs) to define quality of care for chronic disease and preventive services in physician offices and test non-financial incentives for physicians to participate in quality improvement. In Iowa, California and New York, generalist physicians from practices of varying sizes will participate. The non-financial incentives being tested include CME credits for participating physicians, public recognition and potential reductions in malpractice insurance. CMS, DOQ Fact Sheet, March, 2003.

⁴⁵ See, Hawryluk, “Medicare experiments with quality incentive programs,” AMNews (Nov 4., 2002) p.7

⁴⁶ All descriptions of the PGP program come from the CMS website, cms.hhs.gov/healthplans/research/PgpDemo.asp

Determination of performance on those standards will be based on claims data. The program's designers assert that the relevant data can be found in claims themselves which by Medicare law must include procedure and diagnosis codes. It is expected that there will be applied both threshold measures of quality performance as well as measures of improvement over time for the agreed upon targets for the groups.

The 2002 description of the construction of the bonus pool, the assigning of targets, the eligibility of groups, the way the incentives will work, and the enrollment of the beneficiaries is a 211 page document. For all of its methodological detail, there is little in the program description that states a sure physician financial reward for performance. What if the comparison group also improves over time? Because the project runs on a three year cycle, if the group saves in the first year and earns its bonus and then does not meet the targets in the second year and runs a deficit, and because the withhold on the bonus payments is only 25%, the group could find itself in deficit at the end of the three years. Although this convoluted approach to payment draws some of its elements from existing Medicare programs like the Medicare+ Choice Quality Assessment and Performance Improvement (QAPI) and HCQIP programs, its complexity is noteworthy by comparison with the non-governmental P4P programs reviewed here. And this is not the only CMS demonstration project to combine physician payment with quality.

CMS has launched yet another experiment in New Jersey involving eight hospitals which will pay physicians bonuses from their Medicare Part A payments – a classic and admittedly “gainsharing” model.⁴⁷ What is most astounding is that this government sponsored pilot sets up a reward mechanism which if the hospitals implemented voluntarily on their own would be illegal according to the Office of the Inspector General under its rejection of gainsharing models where physicians are paid by the hospital to reduce care provided there, even if off of a baseline of overuse.⁴⁸

The designer of this initiative was one of the creators of the original DRG based payment model which operated in New Jersey before it was extended nationally. He claims the project's cornerstone is an incentive program that evaluates physician performance according to costs per case adjusted for severity and case mix. The costs, however, are the hospital's costs and not the physicians'. Part of the advertised appeal of the program is that it removes managed care entities from the equation. One of the participating hospital managers says that physicians on average can expect to gain as

⁴⁷ Becker, “New Jersey experiment”, *Modern Healthcare*, Dec 1, 2003, pp 6-14.

⁴⁸ Office of Inspector General, “Gainsharing Arrangements and CMPs for Hospital Payments to Physicians to Reduce or Limit Services to Beneficiaries” (July 1999) <http://oig.hhs.gov/frdalrt/gainsharing>. For a broader consideration of the implications of this position to quality initiatives, see Gosfield, “Making Quality Happen: In Search of Legal Weightlessness,” *HEALTH LAW HANDBOOK*, 2002 ed., WestGroup, pp. 31-43. For a broader consideration of fraud and abuse enforcement and quality, see Gosfield, “The Quality/Compliance Nexus: Moving to Programmatic Integration,” *AGG Notes* (July 2003) www.gosfield.com/notes/index.html.

much as \$340 per Medicare admission in incentive payments. Top performing doctors, it is claimed, could reap as much as 25% more in Medicare fees in addition to recouping income lost because of reduced lengths of stay and ordering fewer diagnostic tests. Echoing the basic PGP model, though, although they can opt out in the first year, if the hospitals do not achieve 2% savings in the second year of the program they have to make up the difference. Physician participation is purely voluntary and the program expects them to be evaluated individually rather than by groups. Like classic gainsharing, this model gives the physicians part of the savings of the hospital's expenditures. Its connection to more purely pay for quality physician performance is unclear and the physician's direct payment continues to come from the entities with which he is otherwise contracted and other payors which surely incentivizes him differently. While this program purports to be about P4P, the only performance that is described as being measured is the savings which produce the bonuses to the physicians although there are reportedly some safeguards for quality.

3.2.5 CMS Hospitals

Although the New Jersey experiment, binds the hospitals and physicians together if the physicians want any bonus, it is in the CMS Hospital Quality Incentive Project with Premier Inc, a nationwide organization of not for profit hospitals, that Medicare would reward participating hospitals directly for their quality performance.⁴⁹ Premier was reportedly selected because through its database of hospitals, it has the ability to track and report quality data for 34 quality measures for each of the hospitals. Top performing (read this to mean tiered or normatively compared) hospitals will receive bonuses based on their performance on evidence based measures including for inpatients with heart attack, heart failure, pneumonia, coronary artery bypass graft and hip and knee replacements.

Each of those conditions has between four and eight separate measures that together determine the condition score. For example, heart failure admissions will be evaluated to find whether there has been left ventricular function assessment, detailed

⁴⁹ <http://www.cms.hhs.gov/quality/hospital/PremierFactSheet.pdf>. The obverse is now also on the horizon. CMS with the American Hospital Association, the Federation of American Hospitals and the American Association of Medical Colleges, launched a "voluntary" hospital quality reporting initiative in 2003. National Voluntary Hospital Reporting Initiatives, [Fact Sheet](#) (Oct. 8, 2003) www.cms.gov/quality/hospital/TQIFactSheet.pdf The primary thrust of this program is to provide public information about hospital quality. The first measures to be reported are ten measures endorsed by the Joint Commission, CMS and the National Quality Forum which address three conditions: acute myocardial infarction (aspirin and beta blocker at arrival, aspirin and beta blocker at discharge, ACE inhibitor for left ventricular systolic function); heart failure (left ventricular function assessment, ACE inhibitor for left ventricular systolic function); and pneumonia (initial antibiotic timing, pneumococcal vaccination, oxygenation assessment). Of the 4100 eligible short term acute hospitals, by June 19, 2003, only 415 of those eligible had actually reported. By December 23, 2003, the AHA claimed 2,456 hospitals were "participating." [www.hospitalconnect.com/aha/key_issues/patient_safely/initiative_map/initiative_map.html] Congress has now raised the stakes. What was voluntary now has financial risk. In ' 501 of H.R.1 (PL 108-173), ' 1886 of the Social Security Act was amended to provide that for each of the fiscal years 2005 through 2007, any hospital which does not report will have its Medicare payment update reduced by 0.4% (42 USC ' 1345nn (b)(3)(B))

discharge instructions, ACE inhibitors and smoking cessation advice or counseling. The hip and knee replacement admissions will be examined for prophylactic antibiotic received within 1 hour prior to surgical incision, prophylactic antibiotic selection for surgical patients, prophylactic antibiotics discontinued within 24 hours after completion of surgery, post operative hemorrhage or hematoma, post operative physiologic and metabolic derangement, and readmissions within 30 days post discharge.

Each participating hospital will be scored on each measure and the scores rolled up into deciles of performance. Hospitals in the top 20% of quality over all will be paid a financial reward. Hospitals in the top 10% for a given diagnosis will get a 2% bonus of their Medicare payments for the condition (one half of what Empire will pay to its Leapfrogging institutions), and the next lower decile will be paid 1%. Again, the program will operate over a three year timeframe and for all of the participating hospitals Medicare expects the bonuses to cost \$7 million annually and \$21 million overall. The hospitals also have to improve over a baseline and the lowest performing hospitals that choose to participate will have their subsequent DRG payments lowered for the clinical conditions on which they performed poorly. Premier expected that 300 hospitals would participate and 280 have enrolled.⁵⁰ Given the \$7 million dollars to be paid out, then, the top 20% or 56 hospitals would share some of the money for their overall quality performance, and 28 of the hospitals would share from the same pool for their performance on each of the six conditions. One Ohio hospital system projected that they could net \$2 million dollars over the three years if they were among the top performers.⁵¹

The outcomes of both the physician and hospital initiatives in the Medicare program are not yet known. Described as a nascent but generally positive development by some of the pioneers in quality measurement, evidence based medicine, managed care and health care quality policy nationally⁵², those experts also said that “Measured against the magnitude of the problem, however, these efforts have barely begun to achieve critical mass and momentum...The available measures are less than perfect, but the CMS-Premier demonstration and the National Quality Forum hospital initiative show that we have adequate tools to accelerate the pace of change.....[P]ayment for performance should become a top national priority and Medicare payments should lead in this effort, with an immediate priority for hospital care.”⁵³

3.2.6 Others

⁵⁰ <http://www.premierinc.org/all/informatics/qualitydemo/index.html>

⁵¹ Powell, “Medicare offers more to Summa”, *Beacon Journal*, October 2, 2003, <http://www.premierinc.org/all/informatics/qualitydemo/news/03-oct/summa.htm>

⁵² including Donald Berwick of the Institute for Healthcare Improvement, David Eddy, Paul Elwood, Elizabeth McGlynn, John Wennberg, and others

⁵³ “Paying for Performance: Medicare Should Lead” 22 *Health Affairs* pp 8-10 (Nov. 2003)

Although the programs just described are the broadest in geographic sweep, they are by no means the whole story on the P4P phenomenon. Going back to the mid-90's occasional health plans sporadically financially recognized HEDIS performance among their physicians. Over time more and more health plans have instigated their own versions of P4P and regional health care business coalitions have entered the fray on their own.

Blue Cross of California began some recognition of physician performance as early as 1994. While this insurer is now part of the IHA efforts, it has had its own programs in place as well. In 2001 they announced a plan to pay physicians up to a 10% bonus for patient satisfaction and preventive care measures. They have since gone further to pay based on three additional factors: a medical groups' own internal quality and clinical performance measurement system and its award disbursement system to its own physicians; the number of patients who transfer to other physicians because of dissatisfaction and how often members disagree with a medical decision or complain about the care received. This Quality Score Card enhanced the Quality Measurement and Bonus System to award points for items which can include waiting times for appointments, peer and staff reviews, as well as performance in managing chronic diseases.⁵⁴ By September, 2003, Blue Cross had paid \$28 million under this program. Of the 160 physician groups eligible for the incentives, about half received them.⁵⁵ The California market, however, does not have the exclusive franchise on plan sponsored P4P roll outs.

Harvard Pilgrim in Boston agreed to pay Partners Health Care, the largest hospital system in the market with 4,000 affiliated physicians, an increased fee for service rate—with a withhold -- for providing care that demonstrated improved HEDIS type measures. On the hospital side there are standards related to patient safety, specifically concerning prescription drugs and disease management. Although Boston is a BTE and Leapfrog community, this program predates either of those initiatives.⁵⁶

Minnesota, also a hotbed of managed care, has had its own share of P4P in play. As far back as 1998, the Buyers Health Care Action Group simply established an unvarnished awards program for those physician groups who would seek to be evaluated for their improved quality performance. The first prize winners get a \$100,000 unrestricted cash grant and the three second place winners an additional \$50,000 each.⁵⁷

⁵⁴ “Blue Cross of California Announces Restructure Physician Quality Initiatives,” July 10, 2001, www.hillphysicians.com/About_Hill_Physicians/Hill_in_the_.../Quality_Initiatives.htm

⁵⁵ “Blue Cross of California Has Paid \$28 Million to Providers Under Incentive Plan”, (Pondel, Los Angeles Daily News, 9/10/2003) www.californiahealthline.org/members/basecontent.asp?contentid=49731

⁵⁶ Eaton, “Health Plans Rewarding Doctors For Meeting Quality Standards Rather Than Controlling Costs”, Physicians' Financial News (Sept 30, 2001), www.pfnpublishing.com/quality.htm

But this is not truly *pay for quality*. More recently Blue Cross and Blue Shield of Minnesota, the largest payor in the state, announced a slightly different take on P4P in its 'Pay for Performance' Clinic Incentive program which will reward clinics for proven outcomes on pre-agreed measures. For the contract year July 2003 to July 2004 under one of their contract products, the new Recognizing Excellence and Assisting Smokers to Quit will replace all the other quality incentives and, based on clinic size and number of Blue Cross members, participating clinics could earn rewards ranging from \$12,500 to \$500,000. The clinics will self report measures in five provider specialties: internal medicine, family practice, pediatrics, OB/GYN and cardiology. The groups report two measurements per specialty focusing on disease specific results for chronic conditions such as diabetes, pediatric asthma or hypertension. Generic prescription patterns are another measurement per specialty area and tobacco dependence metrics including tobacco status and documented advice to quit are also counted. Metrics pertaining to descriptive clinic information, including such items as languages spoken, are also reported. According to the insurer itself, the clinics who realize the rewards must demonstrate superior care over time.⁵⁸

Their competition in the same state includes an Outcomes Recognition Program (ORP) with Payment for Quality (PFQ). HealthPartners ties an annual financial bonus to primary care physician performance, the threshold for which is adjusted upward annually. To participate, a primary care group must have 2,500 HealthPartners members under a risk contract, having been part of the network for at least two years. The 22 clinics eligible for the ORP awards in 2002 care for more than 90% of the plan's members and can get rewards ranging from \$63,8000 to \$250,000 depending on the number of patients enrolled and targets reached. In 2002, 13 of the 22 groups received a total of \$347,400. Patient satisfaction accounted for 30% of the bonus in 2002 but will be 55% in 2003.

On the chronic care measures, HealthPartners does not pay for having performed a service, but for having produced an outcome; for example on diabetes a minimum of 20% of the members with diabetes had to have HbA1c at less than 8, LDL less than 130/mg/dl and blood pressure under 130/85 mm Hg. They had to be non-smokers and those over 40 had to be taking aspirin daily. The level was increased to 30% for 2003. There are coronary artery disease targets, preventive services and tobacco cessation targets.

In addition the Payment for Quality program goes beyond what we have now seen as a typical emphasis on primary care and chronic conditions to entail agreements

⁵⁷ Aquilina, "Buyers Reward Physicians for Quality", (Feb 2001) www.qipphysician.com/cgo-bin/article.cgi?article_id=885

⁵⁸ "Blue Cross Launches New 'Pay for Performance' Clinic Incentive Programs," (Sept 16, 2003) ClariNews, Story from Blue Cross and Blue Shield of Minnesota via BizWire

with all high volume cardiology and emergency medical groups, 70% of orthopedists and ENT specialists and 40% of OB/GYNs. For 2003 the plan expects to pay as much as \$8.5 million under this program with about \$4.5 million to primary care, \$1 million to specialty groups and \$3 million to hospitals. For 2004, the specialty and hospital payments are expected to grow by 30-50%. The participating groups with the plan establish baseline and goals to earn monies set aside as a pool to be awarded depending on results.⁵⁹

Physician focused P4P programs involving a variety of payment enhancements have been in place in Buffalo, New York under the auspices of Independent Health, an HMO; in Hawaii sponsored by the local Blues plan PPO; by Anthem in New Hampshire; Blue Cross Blue Shield of Tennessee; Priority Health, an HMO in Michigan; CIGNA in Atlanta; Highmark in Pittsburgh; and Anthem in Virginia, among others.⁶⁰ Hospital models are also increasingly in evidence. Anthem has written quality language into its 2003 hospital contracts that impact 54% of all Anthem admissions across nine states.⁶¹ In Virginia, the Anthem hospital program involves 13 large facilities that provide broad ranging cardiovascular services. Those hospitals account for a little more than 50% of Anthem member's annual discharges in Virginia. The performance goals are linked to indicators published by CMS, JCAHO, Leapfrog and AHRQ. The outcomes measures draw on JCAHO ORYX core measures bolstered by cardiac catheterization performance indicators linked to the American College of Cardiology National Cardiovascular Data Registry. Patient safety accounts for 30% of the reward determination, health outcomes is 55% and patient satisfaction survey results account for 15%. Unique features of this program include the formal collaboration of the ACC to provide technical and educational support to the hospitals, a contract with the Virginia QIO as the Patient Safety Organization which validates the data submitted and another contract with Virginia Health Information, an all payer data repository, for web-based data submission.

From the business coalition side even where a coalition participates in Leapfrog or even, BTE, as in the Central Florida Health Care Coalition, they may establish their own P4P programs. CFHCC has implemented a physician P4P program which targets performance benchmarks on ten conditions, tiers the physician performance and pays more and faster to the top tiered 'platinum' physicians, along with lowered administrative burdens of reportedly allowing them to go off formulary as well as providing them with free PDAs to facilitate prescribing. The second tier does not get all

⁵⁹ Reece, "Program Links Rewards to Outcomes," (Nov 2003) QIPhysician.com

⁶⁰ See Maguire, "As they struggle to improve quality, HMOs try a new incentive: bonuses", (June 2001), ACP-ASIM Online, www.acponline.org/journals/news/jun01/bonuses.htm; Sipkoff, "Pay for Performance Programs Spreading, Adding to Work Load," (October 2003) www.qualityindicator.com/cgi-bin/article.cgi?article_id=1524; personal communication 8-26-2003, from Kim Suarez, Director Provider Network Management of Priority Health

⁶¹ Personal communication 9-14-2003 from Samuel Nussbaum MD, Medical Director, Anthem.

the benefits but gets some enhanced payment.⁶² Coalitions all over the country including in big cities and smaller communities like Rockford, Illinois, and Dayton, Ohio are also trying P4P approaches.

The variety of these programs is significant as a manifestation that there is no clarity regarding what will work best. The range of complexity and simplicity is also impressive. One of the truths of this phenomenon though, is that in no instance yet seen is the entire payment model based on P4P. As a result, whatever is induced by P4P is living along side the pre-existing financial incentives. Those incentives, therefore, bear scrutiny in light of cross currents and mixed messages they may send in this new world.

3.3 Existing Financial Incentives

P4P programs are an add-on. They add an enhancement on top of existing financial incentives, the goals and influences of which may not be consistent with the P4P desired behavior. Still further, the varieties of financial incentives at work on physicians and hospitals can vary even within a single market. The interplay between P4P efforts and what already exists is unknown.

The first and classic payment incentive to change care delivery patterns was capitation – actuarially determined monthly payment to the physician or his group for each patient assigned to his panel. Payment is made whether the patient ever enters the office or has any contact with the physician. The incentive drives toward less services because the physician only receives one payment for the patient and must manage care for the entire month for the money received, balancing the needs of each individual patient against the rest for whom he is responsible. Stop loss insurance protects against the risk of catastrophic cases.

Even today, though, there is very little that can be said definitively about the quality implications of capitation. Some argue that it motivates greater emphasis on prevention and early detection because a healthier panel of patients is less expensive to the physician. But not much has been proven over the last thirty years about the quality incentives of capitation. The California Association of Health Plans cites capitation’s “clinical empowerment” to coordinate care and manage acute and chronic illnesses.⁶³ But the fundamental information that drives the construction of the capitation amount turns on standard actuarial principles which consider past utilization behavior projected forward, whether those patterns entailed good care, misuse, overuse or

⁶² Prager, “Coalition proposes pay based on quality,” *AMNews*, June 19, 2000, www.ama-assn.org/amednews/2000/06/19/prsa0619.htm and Sipkoff, “Employer, Health Plans to Pay Physicians for Quality Efforts” (March 2002), www.qipphysician.com/cgi-bin/article.cgi?article_id=1123

⁶³ “Capitation and Quality in Health Care,” http://www.calhealthplans.com/ib_capitation_quality.htm.

underuse. Capitation is in no way a clinically driven payment model and has been described as a “meat cleaver approach.”⁶⁴

As early as 1991, there were concerns regarding the impact of these financial incentives. Clearly, physicians experienced the impact of the incentives differently from the way the managed care entity regarded them. In 1991, an article surveyed HMO managers regarding two components of typical payment incentives including (1) withholds – money reserved against the expensive ordering practices of physicians – and (2) bonuses paid when behavior is appropriate. At that time, approximately two-thirds of all HMOs used withhold accounts to evaluate the extent to which primary care physicians referred to specialists, to hospitals, and for drugs and laboratory testing. The amount of the withholds was generally not returned to the physicians if there was a deficit. Primary care physicians who overspent might incur additional penalties.

At that time, most HMOs also used some form of bonus payment if there were surpluses in the account and the physicians met some standard of productivity. Approximately one half of HMOs paid primary care physicians by capitation, two-fifths paid by fee-for-service and 13% paid on a salary basis. About 60% of the respondents to the survey used a withhold account reserving from 11% – 20% of the income. Of those using withholds, 25% placed the individual physician at risk rather than the group. Almost two-thirds of HMOs had risk pools with 50 or more physicians as members. Twelve percent (12%) had smaller risk pools. Most managers thought that the withhold effect would only be meaningful if it was above 5%. Four-fifths of the HMO managers believed there would be a noticeable impact at a level of withholding between 5% and 30%.

Not surprisingly, given the business interests of the HMOs as noted by Ellwood, the level of withholding at which managers would express concern about “appropriateness” of physicians’ quality judgment was higher than the level at which there would be an impact. Obviously, this was a reflection of the overuse problem that motivated the use of HMOs in the first place. The effect of bonuses was seen to be the same. As the authors observe, “managers apparently believe that financial effects are symmetric. Whether an action rewards an income already expected or provides additional income is irrelevant.”⁶⁵

That capitation reduces care delivered is unquestionable even where the change is directed by large physician groups themselves at financial risk. In 1995, two different studies examined physician groups with large volumes of capitated dollars in their payer mix. The first looked at ninety-four physician groups, all contracting with a

⁶⁴ Wehrwein, “The March of Capitation: Reversed or Just Delayed?”, Managed Care. (Nov 1997) (http://www.managedcaremag.com/archives/9711/9711.capitation_part1.shtml.)

⁶⁵ Hillman, Pauly, Kerman and Martinek, “HMO Managers’: Views on Financial Incentives and Quality”, Health Affairs (Winter 1991), pp 207-219.

single, large, network-model health maintenance organization in California. The second looked at six large groups managing full capitation including hospital care. The physician managed entities were at least as aggressive as plans generally, if not more so.⁶⁶ Of potentially greater significance to how P4P can affect treatment behavior in a context of capitation incentives was the contemporaneous study of 108 managed care organizations enrolling 33.5 million individuals about half of them in HMOs, the rest in PPOs.⁶⁷ Taking into account and differentiating among group and staff models, IPA and network models and PPOs⁶⁸, the study found that financial incentives only were apparently seen as inadequate to assure the desired change in care patterns. Virtually all of the entities – 89% - used some formal utilization management program, and most used more than one including preadmission review of non-emergency admissions, concurrent review of hospital stays, discharge planning, ambulatory care review, and retrospective claims review. So much for the vaunted faith in the new financial incentives to do their work.

By 1998, concern for the effects of the financial incentives was in full flower. Studies showed fifty-seven percent (57%) of primary care physicians reported they felt pressure from the managed care organization to limit referrals and 17% said they believed such pressure compromised patient care. Another 75% felt pressure to see more patients per day, and 24% felt such pressure compromised patient care. Not surprisingly, the authors concluded that “incentives that depend on the quality of care and patients’ satisfaction are associated with greater job satisfaction among physicians.”⁶⁹

For hospitals, the equivalent shift in payment came in 1983 when Medicare changed from cost-based payment to a single per admission payment in accordance with the patient’s diagnosis. Although diagnoses were grouped (diagnosis related

⁶⁶ Kerr, et al., “Managed Care and Capitation in California: How Do Physicians at Financial Risk Control Their Own Utilization,” 123, *Ann. Of Int. Med.* 500 (Oct. 1, 1995), and Goldfarb, “Physicians in Control of the Capitated Dollar: Do Unto Others....,” *id.*, at 546. Another review of the same phenomenon has come to a similar conclusion where the arrangement extends to global capitation, assigning to the physicians responsibility to buy hospital care while managing medical services. See also Robinson & Casalino, “The Growth of Medical Groups Paid Through Capitation in California,” 333 *NEJM* 1681 (Dec. 21, 1995), and Winslow, “Doctors Cut Costs More When They Own Managed Care Practices, Share Profits,” *Wall St. J.*, Dec. 21, 1995 at B8.

⁶⁷ Physician Payment Review Commission, *Arrangements Between Managed Care Plans and Physicians: Results from a revised and shortened version of the same study.* See Gold, et al., “A National Survey of the Arrangements Managed Care Plans Make with Physicians,” 333 *NEJM* 1678 (Dec 21, 1995).

⁶⁸ The study did not rigorously define or differentiate the explicit characteristics of each model. In selecting the samples, they distinguished HMOs and PPOs and then within those distinctions they also differentiated product lines – among traditional HMO, open-ended HMO product, PPO without a primary care gate keeper, PPO with a primary care gatekeeper, exclusive provider organization, and other point-of-service product. *Id.* at 24-25.

⁶⁹ Grumbach, et al. “Primary Care Physicians’ Experience of Financial Incentives in Managed Care Systems”, 339 *NEJM*, 339: 1516 (Nov 19, 1998).

groups – “DRGs”) by resource consumption, the intent was to make hospitals become more efficient in their utilization of resources within the single DRG payment. Like the stop loss coverage which is generally available in capitation contracts for physicians to cover catastrophic cases, the DRG program offers outlier payments for medically appropriate, longer lengths of stay and higher intensity of services for some admissions. The old PRO program, now the QIO program offers safeguards for the incentives that this program would otherwise create with respect to premature discharges and under utilization.⁷⁰ Like capitation though, DRGs reflect nothing clinical. The groups were created based on resource consumption at that time, whether it reflected overuse, underuse, misuse or stellar care. There has been very little discussion of the impact of the DRG payment system on quality of care. Rather, safety issues, the role of nurses in the hospital⁷¹ and some very recent concerns about perverse financial incentives that penalize hospitals for providing good care⁷² have been discussed in the context of the quality debates. There has been nowhere near as much focus on the quality affects of hospital payment systems by comparison with the attention to physician payment systems effects on quality.

As managed care organizations have struggled to meet purchaser demands, and contain costs while they solidify their relationships with physicians and cushion themselves against public opprobrium, other forms of payment have emerged. Contact capitation was said to arise because of “a failure by traditional capitation to produce expected results.”⁷³ Here, the physician specialist is paid pro rata for the share of the patients he is managing, based on a separate per member per month payment for each specialty. The physician gets credit only when there is initial contact with a patient. He is then paid monthly for a year. Although the theory posits adjustments for specific procedures and diagnoses, and even subpools for selected procedures, most of these programs have become neither that sophisticated, nor particularly widespread. The intent was to insulate individual physicians from expenses generated by over-utilizing colleagues by putting more clinical control in the hands of specialists, who need not continually seek referral authorizations from a primary care gatekeeper; but because the specialist has to see the patient to get paid, it blunts the perverse incentive of getting paid for not seeing the patient under pure capitation. Still though, the per member/per month rate is actuarially and not clinically driven.⁷⁴ Another version of this approach

⁷⁰ 42 USC § 1320c et seq. See Gosfield, “Utilization Management, Quality Assurance and Practice Guidelines, Ch 25, AHLA, *HEALTH LAW PRACTICE GUIDE*, WestGroup, pp 25-29 to 25-55.

⁷¹ Page, ed. *KEEPING PATIENTS SAFE: TRANSFORMING THE WORK ENVIRONMENT OF NURSES*, National Academy Press, 2003.

⁷² Abelson, “Hospitals Say They’re Penalized by Medicare for Improving Care,” *NY Times*, Dec. 5, 2003, p1, col 2.

⁷³ Frank, “Contact Capitation for Specialists.” (June 1999) <http://www.cost-quality.com/5,2art3.htm>

⁷⁴ See Carlson, “Contact Capitation and Its Cousins”. *Managed Care* (Nov 1997) http://www.managedcaremag.com/archive/9711/9711.capitation_paut2.shtml

is episode of care payments. Here, the plan pays on the basis of episodes of care for cardiology (or other specialty defined scope of services). First it sets aside an amount of money which will be available to pay for cardiology care (usually the previous year's cardiology expenses) and then determines how much will be paid for each episode of care (ninety days), irrespective of the scope of the services provided (whether an EKG or cardiac catheterization or hospitalization for a heart attack). The amount available to pay for each episode depends on how many episodes are triggered – the volume of utilization – during the previous quarter reviewed.⁷⁵ An impact on utilization will only be realized, though, if the specialists involved respond to this diffuse, generalized incentive to think about how much of the targeted services they provide along with their unaffiliated cardiologist brethren. In addition, the episodes are often triggered by the PCP's referral rather than the cardiologist's behavior.

Global capitation systems were developed for hospital systems and PHOs where the hospitals and physicians would manage care together. These arrangements typically have involved risk sharing with the physicians who get the added reward of a share of surplus if they manage utilization effectively; but the physicians are also responsible for deficits that may be generated. The hospital system takes on all the risk of the insurer, including adverse selection. These contracts were often negotiated to be paid as a percent of the insurer's premium dollar, and the hospital system took on the insurer's pricing risk. Absent pricing control and data and systems to effectively monitor and manage care, these have been high risk contracts that generated multi-million dollar losses. The real pitfalls of this approach could be seen in the collapse of the Allegheny Health System which had been operating under some significant percent of premium contracts which paid them 78% of the premium dollar to manage all of the care. Allegheny's cost of care would have required 95% of premium to breakeven. Although there are some global capitation contracts paid as a percent of premium that remain in place, this is nowhere near as widespread a phenomenon as was initially anticipated.

Similar to hospital DRGs, some commercial payors have offered case rates where there is a single payment over a period of time to pay for a surgical procedure, the admission of the patient and the subsequent related services. In essence, this is a longer extension of the traditional global surgery concept where one payment covers the surgery and predefined post operative related services for up to ninety days. For example, some payors offer a single rate for high cost procedures and all of the follow up care. A patient is admitted for hip replacement. The physicians are paid one fee to include the operation, the follow up skilled nursing care, the further durable medical equipment provided and the rehabilitation services. This approach gives the physicians risk of effective medical management, but not actuarial risk for marketing to a high risk

⁷⁵ Uhlman, "Insurer's payment plan giving cardiologists a jolt," Philadelphia Inquirer, July 24, 1995 at G1.

population or other adverse selection risks. Sometimes the hospital and the physician are bound together under a case rate, but more frequently they are not.⁷⁶

This brief review of some of the major approaches to financial incentives cannot convey adequately the ebb and flow, claims and challenges associated with the full panoply of risk sharing arrangements, utilization management mechanisms and other financial incentive techniques that have flourished and withered in the last fifteen years. Most of the incentives reviewed, however, were created to motivate very broad changes in delivering care and not to motivate specific clinical behavior. Still though, there have been instances in the past of some very directed incentives in this regard. For example, a payor which pays a higher rate for a vaginal delivery after Caesarean (VBAC) than it will pay for a repeat Caesarean⁷⁷ wants obstetricians to do more VBACs. Prior to the current iteration of P4P programs, some payors had begun to experiment with bonuses, even in a fee for service payment system for meeting quality targets for administering beta blockers after a heart attack or prescribing ACE inhibitors for congestive heart failure.⁷⁸

While the quality impacts of these financial incentives remain murky, what is clear is that there is not universal satisfaction with any of them. The interrelationship between pure payment based incentives and medical management/utilization management programs has also remained a conundrum even after the 1995 Mathematica study showed that MCOs have not trusted payment incentives alone to produce change and truly ‘manage’ care. Still further, the RICO lawsuits filed against large HMOs like Aetna, CIGNA and others, which are beginning to produce major settlements⁷⁹ can be said to be challenging the fundamental incentives that managed care has stood for: “We’re not challenging the managed care system, says [one of the plaintiffs’ attorneys]. “We are challenging the misuse and the secret nature of these incentives to such an extent that patients do not receive care they need, or care that’s been represented to them that they would receive and that they believed they were paying for when they became a member of an HMO.”⁸⁰ The design of the P4P programs can withstand those attacks, but whether P4P programs produce the desired results will turn not only on the mechanisms of the financial incentives and whether providers will respond to them, but also on the legal obligations among the parties.

3.4 Provider Contracts

⁷⁶ See Carlson, n.74

⁷⁷ Burling, “Insurer seeks to cut back on Caesareans,” Philadelphia Inquirer July 24, 1995 at G1.

⁷⁸ Carlson, “Physician Financial Incentives: Another UM Tool Bites the Dust”, Managed Care. (June 2000) <http://www.managedcaremag.com/archives/0006/0006.incentives.html>

⁷⁹ In re Managed Care Litigation, S.D., Fla., MDL No. 1334.

⁸⁰ Carlson, n.78.

Most P4P programs are implemented without even contractual obligations from the payor to the providers. Still, the provider has quality relevant obligations in the contract he does have. The interrelationship between what is a legal obligation and a desired additional payment may also prove to be a source of tension.

The diversity among these contracts is remarkable considering in how many ways the managed care relationships are fundamentally similar. Moreover, despite variability in the detail of these documents, which may range from typical primary care gatekeeper contracts of ten (10) to twelve (12) pages of text to large integrated delivery system arrangements that may run up to sixty (60) or seventy (70) pages of text plus exhibits, there is virtually none in which the entire transaction is contained in the contract itself. Not only do many of these contracts structurally rely on exhibits, practically all managed care contracts require those who contract with them to adhere to other texts that are external to the contract. Examples include provider manuals, peer review and credentialing systems, quality assurance programs, subscriber grievance mechanisms, customer satisfaction controls, utilization review procedures, miscellaneous policies, protocols and guidelines, bylaws, rules and regulations and other idiosyncratic mechanisms upon which any particular plan might rely. Many of these documents are relevant to quality, but most of the provisions are intended to give life to the theory of an organized delivery system which responds to the financial and medical management techniques which define managed care as distinct from pure “pay the bill” indemnity insurance.

The essence of the contract turns on the basic service obligations and specified payment for fulfilling the contract. The financial clauses do not merely address the amount of payment the provider can expect. Financial terms also include beneficiary hold harmless provisions, and the provider’s right to charge for non-covered services and for co-payments and deductibles.

Although not the basic point of most of these financial provisions, some do have some quality implications. For example, if a patient is repeatedly charged for non-covered services, this could discourage the patient from seeking necessary services because the financial risks are becoming unbearable. Similarly, if a physician uses charging for non-covered services as a loophole to unbundle his charges, then the physician may be circumventing the financial incentive the plan intends to impose.

The co-payments and deductibles which are required to be collected at the time services are rendered are part of the utilization controls which also have implications for quality. If the patient has “first dollar coverage”, then he has no financial liability in seeking services and will likely seek more of them. If the patient feels in his wallet the decision to seek medical care, perhaps he will think twice before triggering the incursion of additional costs. Obviously, the extent to which a patient is discouraged from seeking care can be a significant quality influence. Similarly, the stop-loss

coverage in a capitated or risk based transaction as in the federal regulations controlling physician incentive plans⁸¹ can influence the services the patient receives. Still, these issues are primarily financial and not directly linked to the quality of care. Besides these ubiquitous elements, most provider contracts in HMOs and their variants include at least eight (8) categories of contract provisions which relate more directly to quality, even while they address other business needs of the plan.

3.4.1. Network Configuration

In the earliest days of managed care, the use of a limited, selected network of providers was one of the principal tenets. Although many would argue that the original promise of managed care to channel patients to a restricted, quality-vetted network of providers has been thwarted by the phenomenal emphasis on broad provider choice, maintenance of the selected network is still critical. Credentialing and selection of providers, as well as where they are located, that reflect specifically the MCO's interest in maintaining an appropriate network configuration both to assure patient access and to weed out identifiable malactors.

So, the first quality initiative of any MCO is credentialing. This screening mechanism was initially viewed as one of the most critical differences between managed care and traditional indemnity insured care. The managed care industry as a whole was, at least in its infancy, relatively lax about performing the credentialing functions. In fact, many plans simply required that a physician have hospital privileges. This was inadequate for NCQA accreditation which considered verified, timely data an important component of quality assurance. As more credentialing requirements have proliferated, there has been significant rumbling in the managed care industry about the paper chase quality of these programs. This has led to the creation of an industry of credentials verification organizations (CVOs) which, in fact, NCQA certifies directly. CVOs are supposed to facilitate one-stop shopping so that a physician group need not recreate repeatedly the credentialing information it provides to substantiate the quality bona fides of their physicians.⁸²

Typically, the physician application requires qualification data: licensure, medical school graduation, internship, residency, and subsequent work experience. Hospital affiliations are requested and practice characteristics are elicited as well: office hours, office locations, special language abilities, and often the extent to which ancillary personnel are used to provide services. Questions as to whether the physician has ever been disciplined by any other authorities (whether hospitals, licensing agencies, board

⁸¹ 42 USC ' 1395nn (e) (3) (B); 42 USC ' 1395w-22(j) (4); 42 CFR ' 417.479(d); 42 CFR ' 422.208, 210 and Gosfield, "Making Quality Happen," n 48 at pp 26-31.

⁸² www.ncqa.org/Programs/Accreditation/Certification/cvo/cvotext.htm

specialties, or the federal government) must be answered. As part of the application, the physician consents to the plan's investigation of the information he has provided. In fact, HMOs are specifically mentioned within the scope of protected entities engaged in peer review under the Health Care Quality Improvement Act.

As part of the credentialing processes, many MCOs affirmatively require the physicians to notify them within a defined period of time (within five (5) days, within then (10) days) of any major changes in the information that was provided as to licensure, hospital medical staff privileges, legal authority to prescribe drugs, sanctions from the federal government and those matters which were addressed in the initial application.

While most of the historical focus has been on physician credentialing, more MCOs are moving in the direction of facility and institution "credentialing" too. Some of them are getting a boost in this endeavor by virtue of recent NCQA requirements to go beyond evaluating the bona fides of hospitals, nursing homes and home health agencies prior to permitting them to join the provider network in terms of accreditation, but to have a plan for getting patient safety and quality data from hospitals.⁸³

3.4.2. Access and Availability

Network configuration relates to the accessibility and availability of care. It involves both assuring the location of providers to offer a measure of convenience, as well as safeguarding the right mix of primary care and specialty providers. Typical contract features include the obligations that the providers keep their offices at the locations where they practiced when they entered into the relationship or notify the MCO in advance if they relocate, or close an office. In any of these situations, the MCO may be faced with a hole in the fabric of the network. Consequently, providers are often even obligated to notify the MCO six (6) months in advance of a planned retirement so that the MCO can replace them in an orderly way. The more popular the physician or the medical group, or the more specialized the terminating practitioner, the more difficult it is for the plan to transfer patients without major disruption.

Because patient access to physicians is an essential aspect of comprehensive health insurance coverage, the provider contract usually requires physician availability twenty-four (24) hours a day, seven (7) days a week, three hundred and sixty-five (365) days a year. Some contracts also go so far as to require physicians, in particular

⁸³ CR 11 requires NCQA-accredited MCOs to evaluate institutional providers in their networks. Q12, effective July 2004, requires MCOs to have a plan to get safety and quality data which can include, among other options, requiring hospitals to respond to the Leapfrog Survey. NCQA 2004 MCO Standards and Guidelines, NCQA, Washington, D.C.

primary care physicians, to maintain appointment availability in accordance with pre-established standards (e.g., emergency appointments within two (2) hours, urgent appointments within forty-eight (48) hours, routine visits within three (3) weeks). To meet these requirements, physicians also have to make arrangements for peer coverage on weekends, evenings and vacations. Because of the other quality controls and hold harmless aspects of the arrangements, where the MCO insists on these obligations with the directly contracted physician, some are also imposed with regard to physicians who provide on-call coverage. Many MCOs will not allow coverage by someone who has not executed a participating physician agreement as well.

3.4.3. Record Keeping and Access to Information

Assuring the bona fides of the physician practice activities is an essential aspect of managing quality, since, except in a setting where the physicians are the direct employees of the MCO, it does not itself deliver care; so access to the records documenting what has been done is a critical element of the MCO's ability to manage anything. Consequently, plans require that providers report data to them.

Many MCOs not only specify that medical records be maintained in accordance with general standards of the community, but also in accordance with the MCO's own standards, including reports to and from referral providers, discharge summaries, and records of emergency care provided to the member. In a very few MCOs nationally, physician offices may be directly online with the MCO. The MCO may have direct access to encounter data as each patient is seen. That is unusual. Most MCOs specify forms and formats and require regular encounter data reporting even when it is not necessary for payment as in capitation. Most MCOs reserve the right to inspect both clinical and financial records pertaining to services provided to their members. Depending on the payment mechanism, some also require access to the practices' book to ascertain their overall financial health. A medical group in precarious financial straits is likely to change their practice behavior or even to go out of business after the plan has already paid for care in advance, through capitation or a similar technique.

Record review is not the only mechanism by which the MCO gathers information. For NCQA compliance and for some MCOs who merely choose to do so, MCOs perform an onsite visit to physician offices to be sure that services are provided in a setting as claimed in the application. Many HMOs access the National Practitioner Data Bank themselves, but PPOs often require that the physician produce his own NPDB report because they are not HMOs and do not provide care, as a peer review entity must, to gain access to the NPDB information.

3.4.4. Communication Clauses

The sobriquet of “gag clauses” has been bestowed on what are three types of contractual provisions. The first is the anti-disparagement clause which prohibits the physician from communicating with the member in a manner which is disparaging of the reputation of the MCO. Provider criticism of the extent to which the MCO restricts services or otherwise implicates quality has been a battleground for some years. These provisions are a noteworthy manifestation of some of the profound tensions in the peculiar tri-partite relationship between the MCO, the physician and the patient. They demonstrate that some MCOs recognize that those on whom they must depend so completely to fulfill their contractual obligations, namely the physicians, are in many instances relatively unwilling entrants onto the playing field and recalcitrant while engaged. These MCOs know that they are at some risk of the physician maligning them to the primary customer who is the patient. The plans seek to have a chilling effect on the physician’s tendency to denounce managed care in general and potentially them in specific.

The second type of provision is one intended to safeguard the confidentiality of the plan’s trade secrets. They incorporate very broad confidentiality clauses which are intended to preclude from physician disclosure anything they consider proprietary. Very often these include both methods of business and payment rates. In the past, some physicians interpreted this confidentiality clause to mean that they could not disclose the financial incentives under which they were operating. Many plans in the late 1990s added clauses specifically allowing discussion of the general financial incentives, but not the actual payment amounts.

The third communication relevant clause entails the implication that physicians are not permitted to describe to the patient treatment options that may not be covered by the patient’s plan. This controversy was generated by an article by Stephanie Woolhandler, M.D. and David Himmelstein, MD in the *New England Journal of Medicine*, December 21, 1995, in which it was claimed that some MCOs prohibited physicians from describing all treatment options to a patient until the course of treatment had been approved by the utilization management program. The result was an enormous outcry when, in fact, the truth was there was one Permanente Medical Group which had imposed such a rule. Nonetheless, both in state law as part of many state managed care reform movements and in voluntary contracting initiatives, most MCOs now explicitly acknowledge the right of the physician to discuss any medical options with the patient and to not feel encumbered in his desire to appeal the financial determinations of the plan.⁸⁴

3.4.5. Medical Management

⁸⁴ For a view of both sides of this debate, see Uhlman, “Are Doctors Being Paid to Insure That You Don’t Get the Health Care You Need?” *Philadelphia Inquirer*, December 21, 1995.

The medical management system is the core characteristic which defines managed care as distinct from other insurance mechanisms. It encompasses typically a number of components and appeals from their judgments (1) quality review/assurance/improvement program; (2) utilization management systems; (3) delegation; and (4) the subscriber grievance system. Through their own quality mechanisms, MCOs seek to employ organized and systematic efforts to identify continually opportunities to improve care. Typical features include peer evaluations of care, specific focused studies and evaluations of pre-selected indicators of quality. Where a plan is NCQA accredited, it is obligated to engage in these activities to maintain its status. NCQA accredited plans contractually obligate providers to adhere to plan mandates for data they will submit to HEDIS, onsite inspections or virtually anything else they believe is necessary to maintain accreditation.

Primary care gatekeeping was essentially designed as a utilization management technique – no expensive services including specialty care would be allowed unless prior authorized by the one physician who would know the most about the patient. Yet, gatekeeping and its significant quality value was never well-explained or advertised by the managed care industry: in the increasingly complex delivery system, ideally the gatekeeper could shepherd the patient through the process assuring continuity, coordination, reduced redundancy and seamless access. The advent of “concierge care,” where patients pay a premium to have their physician perform this interface with the rest of the system is an ironic demonstration that MCOs missed the boat on this point. This positive potential was rarely realized as plan marketing staff was often fearful of emphasizing these controls. In fact, in many ways, without this clearly articulated quality rationale, gatekeeping was perceived as what plans needed it to be – blunt force cost control through limited access. Consumers saw it as turning Dr. Welby into Dr. No.

Benefit determinations are made through utilization management, which may, as we have seen, also incorporate pre-admission review, prior authorizations, approval of emergency services after they are delivered out-of-network or out of the assigned physician’s office, concurrent evaluation of care and post-delivery follow-up. Providers are obligated to conform to the processes, dictates and decisions of the systems even when it is often difficult to discover the rules of the game. Failure to adhere has been a basis for termination from networks and such a battleground in the MCO backlash skirmishes that many state laws now prohibit terminating a physician who appeals MCO utilization management decisions.

As providers have become more and more sophisticated in their relationships with plans and the plans increasingly devolve more and more financial risk onto their provider arrangements, the providers themselves have sought to have delegated to them the right to perform these functions. Delegation raises its own quality concerns under the contract. The standards by which a provider will be judged to be eligible to

perform these essential functions is as important from a quality perspective as the basic activity delegated. In addition, a triggering event that will cause the plan to withdraw the provider's right to perform delegated functions is also at issue. Fundamental to both concerns, however, are the standards used by the provider, which may be identical with, slightly variant from, or otherwise divergent from those the plan would use itself.

NCQA regards the delegation process as a critical component of its accreditation and scrutinizes carefully the accountability mechanisms in place to link those performing delegated quality assurance and utilization management back into the plan's legal ambit. Contracts between providers and plans that are accredited must speak in specific terms to these matters. Further, NCQA differentiates clearly between the functional responsibilities of these reviews and the ultimate accountability for their outcomes. Consequently, NCQA takes the position that certain functions essential to quality control are never delegable. Still further, NCQA has set-up a program that will certify physician organizations as qualified and competent to perform delegable functions.⁸⁵

Subscriber grievance mechanisms also relate directly to quality. Because in most plan arrangements the patient is restricted to obtaining care from the assigned physician pool, and in a gatekeeper model must obtain all care only after contact with a gatekeeper, dissatisfaction with a gatekeeper can impinge on quality. Whether primary or specialty care, if the patient cannot get proper access, or is uncomfortable in the doctor-patient relationship, and, therefore, not forthcoming about matters that relate directly to the physician's understanding of the patient's clinical circumstances, then the quality of care the patient receives will be diminished. Usually providers are contractually required to respond to grievances, and in many instances are obligated to do so by regulation evaluating physicians for the number of grievances they generate and how they respond in the grievance repeal process is also information plans use in their physician quality evaluation.

3.4.6 Formularies

One of the battlefields in managed care, including arguments about the extent to which it implicates quality, is the use of formularies – lists of approved and covered drugs. Physicians must manage their prescribing patterns to conform to a plan's specific formulary which may reflect more of the business deal the manufacturer has struck with the plan (or its pharmacy benefit manager) than a clinical foundation. The ability to appeal a formulary's strictures is important to physicians and patients who

⁸⁵ See n. 82.

find far more prevalent consumer-driven cost initiatives in this aspect of care than anywhere else in the system. The P4P programs which focus on drugs prescribed (e.g. ACE inhibitors, beta blockers) have these co-existing mechanisms to contend with.

3.4.7. Other Clauses and Overview

MCO contracts typically contain anti-discrimination clauses, post-termination obligations, and indemnification clauses, all of which could be said to relate to quality at least tangentially. Anti-discrimination clauses prohibit physicians from discriminating in the treatment of the plan members on the basis of sex, race, religion, national origin, and other basic civil rights type concerns, as well as on the basis of the member's clinical needs or source of payment. Carried to its extremes, these clauses could be seen to implicate the very changes managed care was intended to produce. For example, most indemnity insurers do not pay for screening examinations or wellness services, unless mandated by state law. Managed care contracts encourage such care. Physicians providing those types of services to managed care subscribers, are, of course, discriminating on the basis of source of payment.

Post-termination clauses variably require the physician and/or hospital to continue services during an inpatient stay, or to complete a course of treatment, even after the Agreement terminates. Without such a requirement, the plan would be at risk of being unable either to meet its contractual obligation to its employer and subscriber customers, or to ensure continuity of care.

But other than the pure medical management programs imposed by MCOs, there are few provisions in the typical MCO provider contract which address quality that even have relevance to current P4P initiatives. The litany of contracted-for provider obligations create administrative demands on physicians in credentialing requirements and record keeping, but otherwise have little connection to P4P initiatives which create their own administrative burdens. In the medical management programs though, one can imagine potential conflict where a typical UM decision denies payment for some service the physician believes is necessary to achieve the outcome goals set in the NCQA DPRP program certification, for example: How does the physician judge his risk here and how much do the plans coordinate their basic systems with these P4P add-ons? These issues are not addressed; and, where there is no contract between the physician and the P4P program, the physician has no way of knowing the answer. Where the plan creates its own P4P program and puts it into the contract, there is no statement I have seen about whether UM can trump P4P. Whether the potential for the tensions becomes a reality, or even whether the potential alone will chill physician enthusiasm, has yet to be seen.

3.5 Perspectives

That P4P program proliferation manifests profound dissatisfaction with how quality performance has evolved in this country is indisputable. It is also noteworthy that all of these programs segregate and differentiate some physicians and hospitals from others who provide care in the same insurance product, and the differentiation turns on quality factors, an approach which represents a significant departure from the history of American health care payment to date. Still, almost all of these programs are less than five years old, and despite the irrefutable fact that real money in terms of total dollars has been put on the table to change behavior, there are significant questions to be raised about P4P and they center on three primary areas: (1) the quality impacts of these initiatives; (2) their financial implications; and (3) their significance in contract terms.⁸⁶

3.5.1 Quality Impacts

There has been some criticism, almost knee jerk in some respects, that programs that emanate from the plans in particular have not involved physicians enough in their design. Paranoia, however, is not always a sign of mental defect and physicians who have been skeptical and have been encouraged to be so in some quarters, fear that some of the P4P programs are a disguise to provide least care rather than right care. These critics take the position that “the only credible proof of ...goodwill will be if those at the front line of caring for patients have a strong voice in defining what quality is.”⁸⁷ While the actual proof will be in payment and outcomes, most of the P4P programs reviewed here did have significant physician involvement in their development and substantial physician willingness to give them a chance. There are broader policy implications to the quality import of P4P, though.

Almost all of these programs fall into four categories to identify those eligible for rewards – (1) where something not previously sufficiently provided meets a threshold of performance (e.g., the programs which reward numbers of diabetic retinal exams or numbers of patients with specific HbA1c results); (2) where providers are normatively compared on their production of the applicable measures against other participating providers in tiering programs and the better performers are rewarded; (3) where there is a combination of tiering and benchmarking against another comparative standard as in the CMS PGP and Premier Hospital Initiative programs; and (4) where a provider meets a threshold of behavior that entails some achievement on measures and then is paid for each patient with a qualifying condition (e.g., BTE). Although there are a few instances in which providers are scored not only on what they do and the patient outcome but also on negative effects, (e.g., readmission within 30 days, incidence of

⁸⁶ Some of the initial critical thinking on these points was part of a consideration of how to improve the business case for quality for physicians set forth in Gosfield and Reinertsen, *Doing Well by Doing Good: Improving the Business Case for Quality*, (June 2003), http://www.uft-a.com/PDF/uft-a_White_Paper_060103.PDF

⁸⁷ Editorial, “Rewarding quality care: Physicians should set the standard”, *AMNews*, Jan 6, 2003

hematoma), today's P4P programs' thrust is almost exclusively at underuse problems. Virtually everything for which payment is to be made is something which has not been done previously or has not been done enough. There is little in these programs that addresses overuse or misuse. To pay someone to do something which can be measured and counted is much easier as a first step than to pay for more generally appropriate utilization through techniques which do not smack of the blunt incentives of capitation, contact capitation or episodes of care, methods which have been devalued in today's world. As for misuse, payment systems to prevent giving the patient the wrong thing are even more complex to define and design let alone to pay for.

That some aspects of quality can be improved by the approach underlying today's P4P programs seems logical, at least. In fact, physicians offered the opportunity to participate in NCQA's Diabetes Recognition Program have not only been willing to do so without direct financial reward, but have reported that the undertaking on that one condition has improved their care for other conditions as well.⁸⁸ While in today's low hanging fruit environment, paying for recognition in this way, or paying for measures as in the IHA initiative, makes a certain amount of sense, these are neither broad nor comprehensive quality initiatives. The themes of chronic primary care conditions vastly predominate in the design of physician P4P initiatives, when we know there are significant other problems among specialists too. In the hospital setting, most of the attention is to safety and cost savings, again, significant and important to address but not revolutionary in the conceptual sweep of their quality impact. For all their rhetoric about significant and big leaps in quality, most of the P4P sponsors claim merely that their efforts are groundbreaking and that we have to start somewhere to make a real and measurable difference in quality. Of greater concern perhaps than the limited nature of these first steps for payment for quality are the financial implications of P4P.

3.5.2 Financial Implications

The most obvious financial question to be raised is whether the amounts of money in these programs are sufficient to motivate real behavior when the contracts to which the providers must adhere pay them differently and the incentives there are not only not clinical, some argue they are anti-clinical. For the hospitals, the costs to implement Leapfrog measures or to succeed in the CMS program have barely been quantified, nor are the rewards clear. In the programs which tier – whether hospitals or physicians – the speculative nature of the eventual benefit is troubling. The players have no way of knowing whether they will ever actually be rewarded for their efforts since they have no idea how their colleagues in the risk pool will behave. Even if they do know they will be paid, as in BTE, there is another significant question as to whether the margins associated with getting the reward are worth it. Leapfrog begins to come to

⁸⁸ Gray, "As NCQA turns its attention to physicians, performance measures find some ready takers," ACP Observer, December 2003,

grips with this in its explicit advocacy for purchasers to help hospitals implement CPOE given its extraordinary upfront and maintenance costs.

In BTE, there is a stated estimate that culling data from a chart for one program will take 15 minutes. This is a real cost in terms of FTE staff which the physician practice must incur to get the payment especially when the other components beyond diabetes care are added. While increased revenues are evident, it is worthwhile to reflect at least briefly on the margin implications of these programs:

For physicians who have been under-providing the relevant services (most of them), there will be increased revenue for providing the formerly underused services; but the cost to do that entails staff time in rendering the additional services and in some instances may entail additional operating expenses for drugs, equipment, and supplies associated with the once missing services now to be rendered. This increases physician practice expenses. For physicians who may have been over-utilizing in relationship to the measured activities there will be lowered expenses (and decreased revenues in a fee for service context). But, foregoing the expense of unnecessary drugs, equipment, supplies and staff time or increased expenses to produce results may not be at all symmetrical with the increased revenue.

Where public recognition is added to the mix, the resulting potential adverse selection by drawing more of the sicker patients to the recognized providers could have nasty downsides, particularly in a capitated environment. It is unlikely the relatively small bonuses realized by an individual physician per patient in BTE, for example, would offset those costs. Still, though, the monies at issue in the tiered programs can be significant. Hill Physicians Medical Group, an enormous 2,000 physician IPA in San Ramon, California, got \$3.9 million in the first year from all the payers participating in the IHA initiative.⁸⁹ But for all the apparent windfall, the bonus averaged to \$1950 per physician.

Another major question raised by physicians about the bonuses is where they come from. Many physicians, particularly where the plans initiate the programs, think the dollars they are being paid in these contexts are simply being taken out of the same pool of money that has always been available to them, so there is no net financial gain and the payments are meaningless. This is particularly sensitive in California where some of the more aggressive plans have paid below even California market rates. When Blue Cross of California announced its first move to rewarding patient satisfaction, the company was roundly criticized when physicians claimed they were “going broke” trying to provide care where the plan wants to emphasize patient satisfaction and won’t

⁸⁹ Conn, “California group cashes in on quality,” Modern Physician, Dec. 18, 2003.

adequately fund basic care.⁹⁰ This also raises the question of what happens in the interplay between the more familiar, non-clinically driven financial incentives for all their care, to which the physicians must respond under their contracts, and these add-ons. About this we know almost nothing, today.

Other questions also crop up: In the threshold and tiered bonus programs, as the physicians respond to the incentive, once the vast majority of the patients reach the targets, what happens? If the program then shifts to new quality targets, what will happen when the continuing costs of meeting the earlier measures are no longer recognized? As more care becomes “bonusable,” where will the money to create the bonus pools come from? If the bonus pools come from new dollars from purchasers rather than merely shifting dollars among the physicians, will employers and taxpayers really accept increased costs, to get what they thought they were buying all along—i.e. high quality care? Will the focus on a few discrete measures cause physicians to achieve these targets at the expense of other, unmeasured, but perhaps even more important aspects of quality?⁹¹ There are value judgments hidden throughout, and although there is recognition that these programs are transitional, they entail enormous effort on the part of the providers whose behavior change is sought. Moreover, whatever the financial import of P4P, for lawyers the already existing contractual relationships cannot be ignored.

3.5.3 Contract Concerns

BTE, Leapfrog and many of the IHA participants have purposefully chosen to walk away from a contractual foundation for their initiatives. Some of the plan generated programs do entail contract amendments, but in all the P4P programs I have identified, the basic financial incentives and the basic contract terms remain in effect with an addendum which describes the add-ons. The rest of the contract with all of its controls for quality, its medical management programs, and resulting administrative burdens reflecting the plan’s other concerns remain in place. In the CMS programs, the rest of the statutory and regulatory requirements for payment also persist. Moreover, as in the Pacificare contracts and BTE, there is an explicit denial of an appeal of the basic calculation which determines payment. This may prove a weakness. Physicians have long decried bitterly the quality of the data they obtain from plans about their practices. Plan data is described as “junk” and in New Hampshire, one physician in the Anthem P4P program has been quoted as saying they not only pay staff to gather data to report, they pay staff to validate the data they get from the plan as to which patients qualify

⁹⁰ Mitchell, “Quality Bonuses Questioned”, (Aug 27, 2001)
www.advanceforhie.com/pastarticles/aug27_01feature2.html

⁹¹ For another consideration of these issues in the context of a different proposal, see Gosfield and Reinertsen, n. 86.

under the program.⁹² In BTE, the plans report to MEDSTAT which produces the report which the physicians may not challenge. In the Anthem Virginia cardiovascular care program there is a data aggregator, but the contract terms were proprietary and unavailable for review.

There are contractual data concerns from another perspective as well. Virtually all of the data regarding performance in the P4P programs reviewed here are self reported – in the Leapfrog program, BTE through the NCQA recognition programs, and in other settings. There have been noises about validation and audits, but there is a level on which these programs are apparently designed around the expectation that those who volunteer to participate will do so honestly. In the context of improving health care quality this is a good and noble view, but as NCQA learned in establishing data specifications for HEDIS as well as for scoring accreditation, there is much wiggle room in data reporting. The validity of what is reported will remain largely unstudied according to the reviews to date. To the extent there are disagreements, though, without privity of contract between the employers paying for BTE and the physicians, why would physicians be motivated when they have no legal guarantee they will be paid?

Still further, although health plans are participating in BTE, for example, unless they explicitly agree to make their utilization management programs acknowledge the clinical behavior and resulting claims differentials associated with these programs which affirmatively call for more services, what standards will be applied to physician behavior – the same cost cutting criteria that apply for the rest of their programs? And what of MCOs already paying a disease management vendor to affect improved efficiencies for many of the same disease states confronted by P4P programs?⁹³ Is the advent of P4P a manifestation that these programs have had limited success? Do physicians treating patients in these settings get credit for the improved outcomes their patients demonstrate when some success may be more appropriately attributed to the disease management vendor? Does that even matter if what we are primarily concerned about are the results to the patients?

As to the contracts themselves, many of the quality-oriented features in traditional MCO contracts reflect a very different and simplistic view of what physicians ought to do. In these contracts there is precious little recognition of the plan's role, or indeed responsibility, to help the physicians meet their obligations. The provisions directly related to quality address credentialing and the physician's

⁹² See Maguire, "As they struggle to improve quality HMOs try a new incentive: bonuses," *ACP-ASIM Observer* (June 2001) www.acponline.org/journals/news/jun01/bonuses.htm

⁹³ See, Gosfield, "Disease Management, Demand Management and Telemedicine: The Leading Edge of Managing Care," *HEALTH LAW HANDBOOK* (1998 Ed.), WestGroup, pp. 235-266

cooperation with the plan's HEDIS obligations or desires regarding accreditation. Over-reaching confidentiality and anti-disparagement clauses reflect an era far removed from the transparency, evidence based medicine and patient centeredness values of *Crossing the Quality Chasm*. What, after all, are the deep, proprietary plan secrets the physicians must hold confidential? That which evidence would dictate as clinically appropriate is not proprietary. To the extent the plan's programs entail intellectual property, those rights can certainly be protected, but their content is hardly confidential. In their insistence on secrecy, managed care organizations have aligned themselves more with the nuclear power industry than principles of open clinical dialogue. No wonder physicians regard many plan initiatives with skepticism and suspicion. Mostly these clauses are about not revealing the fee schedule or capitation amounts – and toward what end? Would that there was something of major import to safeguard! In my thirty years of practice, I have never seen the wonderful secret that deserves such confidentiality.

Some MCOs who have created their own P4P programs appear to be stepping up to a more active role in facilitating improvement, as in the Minnesota based P4P programs reviewed here. But the contracts I have been able to review, stymied as I was by proprietary concerns there despite offers of confidentiality protection, do not change the basic contractual relationship at all. Most MCO contracts today still reflect an orientation that primarily supports the plan's marketability in network configuration, access and availability requirements, provides for enforcement of the medical management system through termination of those who cannot conform, and imposes record keeping and other administrative burdens in formulary implementation and otherwise. Surely as we learn more from P4P efforts, the basic contractual relationship between plan and provider will have to change to reflect an approach which is designed affirmatively to be more consistent both with the new values and the more sophisticated collaborative interaction between MCO and provider that will be necessary to really affect performance in a broader way.

3.6 Conclusion

The P4P phenomenon is extending far and wide in American health care. Its basic notion of paying differently for evidence based care is unassailable. But these programs still speak in terms of incentives -- carrots and sticks intended to alter behavior. Whether focus on 'incentives' through payment add-ons either creates a real business case for quality or is the most effective way to change behavior is yet to be seen. "We are dangling the potential of new money in front of you. Here are some things you can do to get it. You figure out what else is necessary and incur the expense to make that happen. If you do that, you can have the new money, as long as you don't fight with us about how much you got," seems to characterize the current versions of P4P. In fact, the notion of simply accreting new pieces of payment on top of a system

which is, at its core, not designed to support the desired change, presents both conceptual and perhaps to be seen practical challenges.⁹⁴

In its limited focus today primarily to primary care and within that chronic care, as well as three hospital safety initiatives, P4P does not bespeak Baryshnikovian vaulting. By their existence and in their descriptions, the P4P programs acknowledge that physicians are central to what will be necessary to generate major movement forward on quality. The dilemma for those physicians is how to survive in the rest of their business while responding to these limited opportunities. The trials will hardly be inconsequential. For all of the desire in the industry to propel quality improvement by leaps, the better sports analogy may well lie in the steeplechase – a race to a goal which entails a myriad of hurdles, brambles, watertraps and pitfalls.

“The real challenge is to develop a program which increases physician margins, improves quality and is sustainable over a longitudinal time frame. But for such a program to engage physicians, and not add to their administrative burden, attention must be paid to the disparate forces that prevent physician engagement on quality. For any such solution to have real value, it must (1) preserve and enhance the critical quality-relevant aspects of the doctor-patient relationship; (2) reduce administrative and regulatory burden; (3) propel the best science and improved outcomes, and do all of this in a dynamic and quickly changing health care landscape.”⁹⁵

P4P is an important development in American health care. Its outcomes will offer a significant learning opportunity. But as it is constructed today, it is only transitional as we seek better ways to engage providers in doing the very thing they most want to do anyway – provide the best evidence based health care they can in safe, healing, personal relationships with patients whose lives are improved as a result. That process has only just begun.

⁹⁴ For a different approach which posits using evidence based medicine and clinical practice guidelines as the foundation not just for payment but for many other aspects of health care delivery and finance -- see Gosfield and Reinertsen, n. 86.

⁹⁵ Id at p. 10