SUPPORTING STATEMENT
Multi-site Coordinated Evaluation of the Impact of Quality-based payment Strategies
for the Agency for Healthcare Research and Quality

A. JUSTIFICATION

1. **Need for Information**

The Agency for Healthcare Research and Quality (AHRQ) requests that the Office of Management and Budget (OMB) approve under the Paperwork Reduction Act of 1995 AHRQ's intention to collect information from two safety net health care organizations that have implemented quality-based payment programs. This data collection is responsive to AHRQ's request for field-based research released under its *Accelerating Change and Transformation in Organizations and Networks* (ACTION) initiative.[1] ACTION's mission is to promote innovation in health care delivery by accelerating the development, implementation, diffusion, and uptake of demand-driven and evidence-based products, tools, strategies and findings. Boston University is one of fifteen partnerships participating in ACTION. Boston University was competitively awarded a contract to collect data on the implementation and impact of pay-for-quality strategies in safety net settings in fulfillment of ACTION's mission. Boston University undertakes this data collection through its partnership with Boston Medical Center (Boston, MA) and Montefiore Medical Center (Bronx, NY).

AHRQ is the lead agency charged with supporting research designed to improve the quality of healthcare, reduce its cost, improve patient safety, decrease medical errors, and broaden access to essential services for all Americans. In recent years, quality-based payment strategies—that is, strategies that seek to better align payment with the provision of high quality healthcare—have been widely implemented in both private and public sectors to improve quality and reduce costs. Over 100 quality-based payment programs currently exist in the United States, covering both private and public sectors (Baker and Carter 2005). Recent federal legislation requires that the Secretary of Health and Human Services develop a plan for implementing a quality-based payment mechanism (referred to as a value-based purchasing program) within the Medicare program by fiscal year 2009 (Deficit Reduction Act of 2005). Thus, collecting and analyzing information regarding the impact and implementation of these strategies has the potential to inform policy changes at a national level.

Health care purchasers are increasingly looking to AHRQ to provide guidance on the design and implementation of quality-based payment strategies, especially with the rapid growth of these strategies in the health care sector. With the potential for widespread diffusion of quality-based payment strategies in Medicare, AHRQ is particularly concerned that healthcare purchasers and healthcare delivery system leaders in the safety net system have appropriate and timely tools to assist their decision-making. However, the majority of current evidence on the impact and implementation of quality-based payment programs informs decision-making in healthcare

---

[1] For more information on the ACTION initiative, see *Accelerating Change and Transformation in Organizations and Networks (ACTION): Field Partnerships for Applied Research*. Fact Sheet, AHRQ Publication No. 06-P011. Agency for Health Care Research and Quality, Rockville, MD. http://www.ahrq.gov/research/action.htm.

delivery systems that primarily serve well-insured patients. Because there is still very little evidence regarding the effectiveness of these strategies in safety net settings, AHRQ has contracted with Boston University to conduct case studies on two natural experiments in quality-based payment implemented in health care systems serving vulnerable populations. These populations include the uninsured, low-income groups, minorities, women, children, the elderly, and individuals with special heath care needs.

The application of quality-based payment strategies on safety-net setting warrants special study for at least two reasons. First, the effects of quality-based payment strategies on providers working in safety-net settings may be quite different than it is on providers working in commercial settings. Research on quality-based payment programs in work settings suggests that financial incentives may prove to be counterproductive when employees already have a strong intrinsic motivation to perform their jobs. It is likely that physicians who choose to work in safety-net setting have a high degree of intrinsic motivation for their jobs as they often are sacrificing some level of remuneration to care for underserved individuals. Accordingly, the imposition of financial incentives in such settings may theoretically crowd out intrinsic motivation leading ultimately to lower rather higher levels of performance. Second, providers in safety-net settings may be severely disadvantaged in their ability to achieve quality targets because their patients are typically much sicker and less likely to comply with screening or treatment recommendations due to socio-economic barriers. This disadvantage may become a source of unintended consequences from the application of quality-based payment strategies to safety-net settings. That is, providers in these settings may seek to avoid the most severely ill patients, restricting access to care for these patients. Such provider behavior was demonstrated in a study by Shen (2003), in which Medicaid patients with more severe substance abuse-related illness were less likely to be selected for treatment after the introduction of performance-based contracting. In such circumstances, providers may engage in activities that are both unintended and undesirable for quality of care. For example, providers may not be engaged in these programs in a meaningful way. Alternatively, they may attempt to game the programs in ways that can ultimately produce unintended consequences such as avoiding high-risk patients or cutting back on important but not explicitly rewarded clinical activities.

Boston University's collaborating partners are testing a diverse set of payment schemes that afford AHRQ a unique opportunity to assess quality-based payment strategies implemented in both safety net and non-safety net settings. The case study sites include safety net settings at Montefiore Medical Center in New York, New York, and Boston Medical Center in Boston, Massachusetts. Both programs share a common unit of accountability—individual physicians. However, it is the diversity in program experience, incentive structure, populations served, and implementation strategies among these programs that offer an exciting opportunity to learn whether and how quality-based payment strategies actually improve quality of care among vulnerable populations. On the basis of the findings of these case studies, AHRQ will develop resources to assist safety net providers in effectively utilizing quality-based payment strategies to improve quality in their setting.

Based on an extensive review of the literature and discussions with experts in the field, Boston University developed a conceptual model of the factors that may affect the impact of P4P programs (Young et al. 2005). The literature review focused on research addressing the

conditions under which providers have the motivation and capability to change their practice behavior (e.g., Dudley et al. 1998; Conrad and Christianson 2004), specifically studies on provider adoption of clinical/administrative innovation and adherence to clinical guidelines and best-evidence practice. These studies point to the importance of providers' practice setting, including culture, resource capabilities, and market characteristics (e.g., Shortell et al. 1995; Banaszak-Holl, Zinn and Mor 1996) as well as experience and demographics (e.g., Tambyn 2003). Further, the literature on guideline adherence highlights the importance of assessing what physicians perceive as barriers to their ability to adopt guidelines or innovative practices (e.g., Pathman et al. 1996; Cabana et al. 1999). Indeed, P4P programs essentially are initiatives to address what has been cited in this literature as a key perceived barrier to guideline adherence, namely lack of reimbursement or financial incentives (Institute of Medicine 2001).

The framework, presented in Figure 1, consists of three broad domains: characteristics of the incentive program, characteristics of the practice environment, and provider-level characteristics. Within the provider characteristics domain, we distinguish between provider demographics and provider attitudes. More specifically, our framework identifies seven critical dimensions of provider attitudes related to quality targets and incentives: (1) *awareness and understanding* of the incentive program; (2) *salience* of the financial incentives; (3) *clinical relevance* of the quality targets; (4) *control* over the resources needed to achieve the quality targets; (5) *fairness* in the administration of the incentive program; (6) frequency and nature of *performance feedback* provided; and *(7)* possible *unintended consequences* associated with the pursuit of the quality targets.

Regarding *awareness and understanding,* providers must have some degree of familiarity with the definitions of the quality targets used to evaluate their performance if they are going to actively participate in the incentive program. Providers' motivation to pursue quality targets may be affected by their degree of understanding of the criteria and rules for distributing incentive money among participating providers. Assuming awareness and some understanding of the P4P program, we further hypothesize that providers' responses to incentives will be affected by *financial salience*; that is, the amount of the financial award compared to the costs in time and effort necessary to achieve the quality targets. Regardless of the amount of the incentive involved, we also hypothesize that providers' behavior will depend on their perceptions of several clinical issues related to the quality targets. These include providers' judgments of the *clinical relevance* of the quality targets, including consideration of such issues as whether or not the targets are based on sound medical science, and whether reaching the targets will truly improve the health of their patients. Additionally, providers' estimates regarding the potential for negative *unintended consequences* are likely to be important; that is, whether they believe that their efforts to achieve the quality targets will detract in any way from attending to other important aspects of care.

The framework also hypothesizes that providers' behavior relative to an incentive program will depend in part on whether they believe that they have adequate *control* over the activities and/or resources necessary to achieve the quality targets. If, for example, providers believe that achieving the quality targets depends more on patient behavior than their own efforts, or that they will not be able to secure the cooperation of other physicians or providers involved in the provision of program-required tests or services, then they may be less likely to be fully engaged

in the pursuit of the incentives. Additionally, the framework posits that providers' perceptions of the *fairness* of the incentive program affect their motivation to pursue P4P quality targets. Fairness in this context refers to the appropriateness of the proposed quality measure, including relevant case mix adjustment considerations. If providers believe that the characteristics of their patients – for example, age, educational attainment, health status, or co-morbidities – make it especially difficult to achieve the quality targets, then they might be less inclined to pursue those targets. We also propose that providers' perceptions of the helpfulness of the *feedback* they receive regarding their progress toward achieving program quality targets are important. For example, a program in which providers only received performance feedback once a year, and then only a short time before the annual incentive checks were distributed, might engender a different level of participation than an incentive program that involved monthly or quarterly performance progress reports.

The proposed data collection will seek to modify or enhance this framework based on the results of the investigation of quality-based payment in safety net settings. In particular, Boston University will closely examine whether the framework appears to be valid in safety net settings and examine the importance of various program characteristics – experience, populations served, incentive arrangements, implementation approach – as mediators between financial incentives and quality performance.

Figure 1. Conceptual Framework for Proposed Collection

4
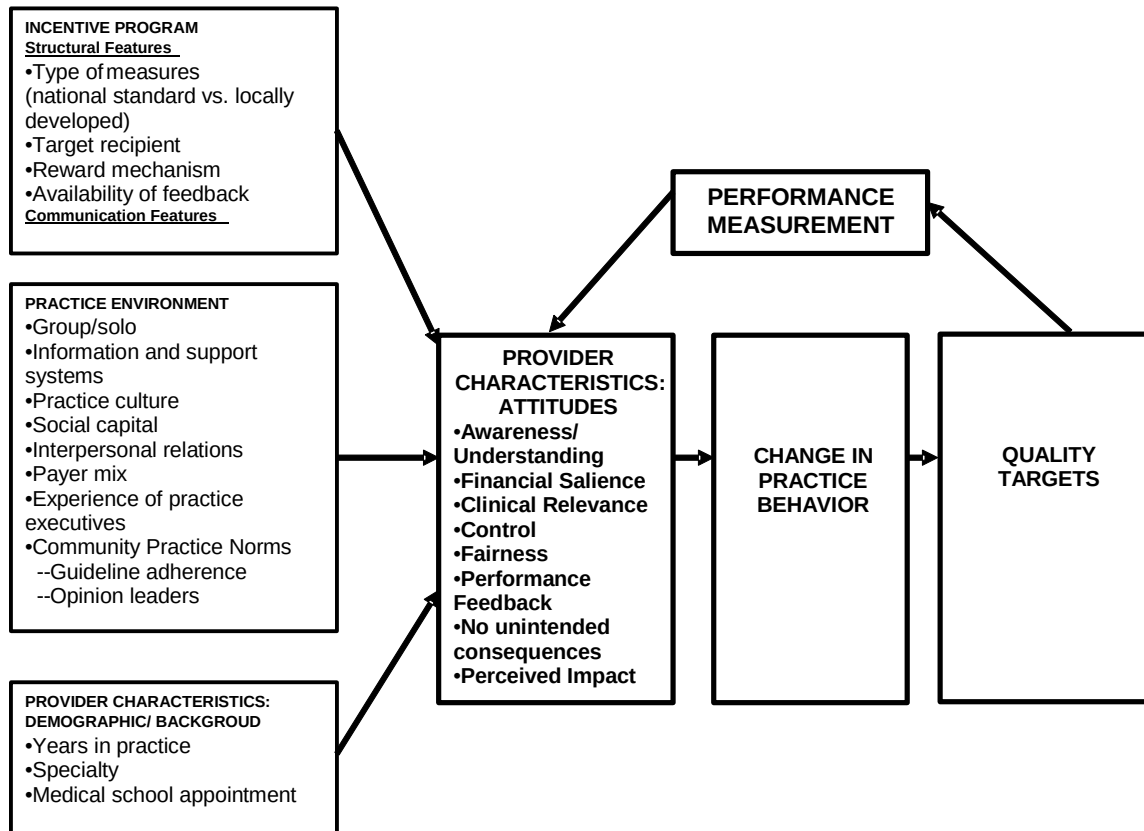
2. **How, by Whom, and for What Purpose Information Will Be Used**

Information collected from this project will be collected by Boston University. Below we briefly describe the settings from which information will be obtained, the types of information that will be used, and the purpose of collecting that information in fulfilling the purpose of the collection.

Intervention sites were selected from Boston University's collaborative ACTION partnership. These sites were chosen based on (1) whether they had implemented a quality-based payment program affecting safety net populations, and (2) the willingness of their leadership to actively participate in data collection. Two sites met these criteria: Montefiore Medical Center in the Bronx, New York, and Boston Medical Center in Boston, Massachusetts. We briefly describe these sites below as well as the quality targets to which financial incentives are tied.

Montefiore Medical Center has been engaged in an internally-funded quality-based payment initiative since 1996 and is the most experienced partner involved in our evaluation. A total of four medical groups owned by Montefiore are involved in the program, encompassing approximately 250 primary care physicians. Montefiore provides its salaried primary care physicians with annual lump-sum bonuses for reaching pre-defined clinical quality targets. The amount of bonus, which ranges from $800 to $3,000 per physician, is dependent on a physician's percent compliance with those targets in comparison to thresholds set annually by Montefiore. These compliance "scores" are assessed through an intensive chart abstraction process completed annually by abstractors external to the medical group using a strict abstraction protocol. Montefiore pays a different level of bonus to physicians for attaining different levels of performance. For example, physicians achieving 97 percent compliance might receive $2,500, whereas physicians achieving 94 percent compliance might receive $2,000. Each target, however, has an absolute minimum performance threshold. Montefiore uses different clinical targets to assess performance among the physicians participating in its quality-based payment program based on their primary care specialty (family practice, pediatrics, and internal medicine). For example, pediatricians are held accountable for meeting Early and Periodic Screening, Diagnostic, and Treatment schedule requirements such as immunizations, whereas internists are held accountable for adult targets such as comprehensive diabetes care. Montefiore also stratifies its clinical target thresholds by targets attainable by physician practice behavior versus those targets dependent on patient behavior. For example, physicians are held accountable to annual Hemoglobin A1c testing at a compliance level of 97 percent, whereas the threshold compliance for obtaining Hemoglobin A1c levels of less than 8 is 50 percent because it depends on patient behavior.

Including Montefiore in this data collection presents several advantages. First, it is unusual for a quality-based program to risk-stratify its quality targets by physician- and patient-dependent characteristics. This design feature could potentially correct for any adverse selection of patients on the basis of risk or chronicity of illness. Second, we can evaluate the question of how a quality-based payment program can be responsive to changes in quality over time, a question that has been raised but not adequately answered in the literature. Montefiore has taken steps to account for changing quality by increasing the thresholds required for physicians to receive a bonus over the course of its program. For example, in 2003, pediatricians with 91 to 94 percent scores received a bonus of $2,500, whereas scores of 95 to 97 were needed to achieve the same

bonus amount in 2004. Finally, Montefiore's four medical groups each serve very different populations. Two of the medical groups serve a population primarily covered by Medicaid, the third group serves mostly uninsured patients, and the fourth group serves a population primarily covered by commercial insurance. These differences afford our evaluation an opportunity to compare quality-based payment on a safety net population with a non-safety net population.

Boston Medical Center implemented a quality-based payment program in January 2006 targeting its primary care physicians practicing at affiliated community health centers and medical groups. This quality-based payment program is being implemented by Boston Medical Center Health Plan (BMCHP), a Medicaid managed care organization operated by Boston Medical Center. This network of community health centers and medical groups provide care to an underserved population in the Greater Boston metropolitan area. The incentive program involves approximately 250 primary care physicians. BMCHP's quality-based payment program rewards primary care physician performance in three areas: clinical quality, administrative quality, and enrollment. The incentive rewards performance at the group level—that is, each health center and primary care group practice will be measured against the BMCHP's average for each measure. The payout is awarded in the form of additional capitated payment for each of the six measures being assessed by BMCHP. For example, a health center or group practice will receive an additional $0.83 per member per month if performance targets are met. Payment is awarded as a lump sum, paid annually to the health center or group practice.

Boston University believes BMCHP's contribution to this data collection is three-fold. First, it is uncommon for quality-based payment programs to be evaluated concurrently with implementation. Boston University would be fortunate to be able to capitalize on the strong relationship it has formed with BMC and BMCHP through an evaluation and provide formative feedback to BMCHP on its implementation. Second, rewards are made at the group-level for collective performance among providers, a unique feature among the participating evaluation sites. This key difference will allow for important opportunities to compare performance across sites. Finally, there is a paucity of peer-reviewed literature evaluating Medicaid-sponsored quality-based payment programs despite the plethora of quality-based payment programs targeting Medicaid recipients across the United States. This case study would make a contribution to the peer-reviewed literature in this regard.

Each site differs on the number and type of clinical quality targets that incentive rewards are based. ***These targets may have changed as the programs matured. A list of all quality targets incentivized by each program in 2006 are presented in the Table 1 on the following page.***

The types of information that will be collected by Boston University in this study include the following:

(1) Administrative data measuring clinical performance of physicians participating in quality-based payment programs;
(2) Responses from a one-time mailed, self-administered survey of physicians participating in quality-based payment programs; and
(3) Responses from one-time, in-person, semi-structured interviews with quality-based payment program stakeholders.

Table 1. 2006 Quality Targets

Montefiore Medical Group	
Problem List (Medical Record)	**Nursing**
Up-to-date Problem List	VAR documented
Up-to-date Medication List	Nurses Note Documents pt Education Provided
Drug allergies noted on problem list	At end of neb rx peak flow or pulse ox doc
Health Maintenance	**Medical Record Quality**
Mammogram suggested	Age>21, weight recorded
Mammogram Report in chart	Legible documentation
PAP suggested	Pages in record secured
Flex Sig/Colonoscopy Suggested	All pages contain pt name and/or ID
Flex Sig/Colonoscopy Done	All reports signed by provider
3 Stools for blood recommended	Consult reports are signed by provider
3 Stools results in chart	Documentation addresses Advanced Dir
Cholesterol done in last 5 years	Self Assessment
Child Health	Barriers To Learning Documented
Age appropriate safety discussed	Smoking Status Assessed
Lead Screening	**Diabetes**
Hemoglobin level documented	Ophthalmology Exam
Vaccinations up-to-date	Diabetic Protein Algorithm Followed
Growth Chart up-to-date	2 or more Hb A1c
Developmental Milestones/ School Performance	Hb A1c < 8.0
Valid documentation of immunization	LDL checked in last 12 months
Hearing and Vision in last 5 years	LDL Cholesterol <100
Dentist referral > age 3	LDL Cholesterol <130
Asthma < 16 years	Foot exam documented in last 12 months
Severity Classified	Monofilament Foot Exam
Peak flow documented once a year	BP < 130/85 (140/90 in 2000)
Peak flow or pulse ox on sick visits	Fluvax documented in Diabetics
Environmental History Documented	Pneumovax documented
Medication reviewed at Asthma visits	**Pain**
Patients on Inhaled Corticosteroids	Patients are screened for pain
Asthma Action Plan in Chart	Pain scale used to qualify
Asthma Education	Pain re-assessment at next visit
Fluvax documented in last 12 months	
Boston Medical Center Health Plan P4P Program	
Clinical Quality Measures	**Administrative/Enrollment Measures**
Diabetes - Retinal exam	Member retention (PCP transfer rate)
Asthma – Appropriate medication all ages	Access/available (3rd next available appointment)
Well-child visits - % of members with 6+ visits in first 15 months of life	Encounter data completeness

Use of Clinical Performance Data: The first part of the data collection involves the collection of clinical performance data for secondary analysis. Boston University will use this data to conduct pre/post-test analysis assessing the impact of quality-based payment on the quality measures that are the focus of two participating programs. The design of the data collection requires the sites to provide this information directly to Boston University.   Both sites' participation in the project was predicated on their willingness and capacity to provide this information to Boston University. For the program at Boston Medical Center Health Plan, which began in January

2006, Boston University has access to one year of baseline data and one year of post-intervention data for all quality measures that have been the focus of the quality-based payment program. This data is claims-based. Boston University has access to considerably more pre- and post-intervention data from Montefiore, which began its quality-based payment program in 1996. For Montefiore, the data are abstracted from medical chart reviews, which provide the source of information for its quality-based payment program. The programs in both settings use quality measures that adhere to or modify the specifications of measures included in the Health Plan Employer Data and Information Set (HEDIS). Boston University will use the clinical data to identify possible reductions in the frequency of important but not explicitly rewarded (within program) clinical activities. Data analyzed in this part of the project also will be used to provide formative feedback to both Boston Medical Center Health Plan and Montefiore Medical Center to enhance their planning and future implementation of their quality-based payment programs. While these programs attach incentives to different quality targets, to the extent possible Boston University will attempt to compare the results obtained in this project to previous analyses done on quality-based payment programs in non-safety net settings.

<u>Use of a Physician Survey</u>: The second part of the data collection is a survey of physicians participating in the quality-based payment programs. Data collected by the survey will be used to assess physician attitudes toward a wide range of issues related to pay-for-quality. To achieve this goal, Boston University will use a reliable, valid, and field-tested instrument that it developed as part of its evaluation of the Rewarding Results Demonstration Program and is based on the conceptual framework in item 1 above. Briefly, the instrument (hereafter referred to as the PAI-26) is a 26-item self-report measure that yields scores on seven key dimensions related to quality incentives. These dimensions are: (1) awareness and understanding of the incentive program; (2) salience of the financial incentives; (3) clinical relevance of the quality targets; (4) control over the resources need to achieve the quality targets; (5) fairness in the way the program is administered (6) possible unintended consequences associated with the pursuit of the quality targets; and (7) the perceived impact of the quality targets and incentives on the respondents clinical behavior. The PAI-26 instrument also measures impact by asking physicians to judge the extent to which their practice behavior has changed in response to the quality targets and associated financial rewards. The psychometric properties of the PAI-26 questionnaire were examined and found to be adequate in a large-scale study involving exploratory factor analysis and multi-trait analysis applied to data from over 600 physicians.

For the survey, we intend to recruit physicians that practice at medical groups and community health centers that contract with Boston Medical Center Health Plan and all medical groups at Montefiore Medical Center. Practices must serve a predominantly safety net population to be included. To be eligible to receive a survey, physicians must be a current or former participant in the quality-based payment program at their institution. Information collected through the survey will be compared to responses obtained through a previous administration of the survey conducted in one non safety net setting. Such a comparison will yield information that can be used to customize quality-based payment programs based on the patient population served and the type of providers who serve them. ***Our comparison site is located in Rochester, New York. The site consists of a partnership between Excellus, Inc. and the Rochester Individual Practice Association (RIPA). Excellus, a single health maintenance organization (HMO), has enrolled more than 70 percent of the commercial (non-Medicare and non-Medicaid) population in the***

nine county region surrounding Rochester, New York. RIPA is a contracting entity comprised of more than 800 primary care physicians who collectively care for more than 420,000 commercial patients, approximately 80 percent of whom are members of Excellus. Excellus and RIPA have collaborated to share administrative data for the purpose of their P4P program. The sample for the survey consisted of physicians who qualified to receive a financial reward from RIPA for achieving quality targets related to care pathways for at least one of four conditions: asthma, diabetes, otitis media, or sinusitis. A total of 597 physicians (152 family practitioners, 290 internists, and 155 pediatricians) met this criterion; all were included in the survey sample. Survey methodology was similar to the method proposed in this collection.

Use of Semi-Structured Interviews: The final part of the data collection involves semi-structured interviews with stakeholders representing each participating quality-based payment program. Boston University will conduct semi-structured interviews with up to 8 individuals at each site (a total of 24 interviews) in the study, including such informants as physician organization executives, practice leaders, physicians, and clinical staff. ***All sites will be eligible for interviews and we will attempt to interview at least one key informant from each practice site. None of the sites Boston University will interview are non-participants.*** Data collected from these interviews will be used to investigate the implementation of the quality-based program incentives and to assess the attitudes and perceptions toward these programs. Boston University will interview selected informants from the practices in each study organization. ***Interviewees will be selected from short lists of key informants developed by the project liaisons at each research site based on their availability for interview during times when the project staff can travel to them. Boston University will attempt to accommodate as many informants as our travel schedule allows.*** At the management level, Boston University will interview the clinical and administrative leaders of both physician organizations and the practices within them that are participating in the incentives. At the front-line, Boston University will use a comparable sampling approach to identify physicians and clinical staff. To identify these informants, Boston University will ask practice executives or leaders at each site to provide us a list of personnel at the management, clinical, and administrative levels identified above who are participating in or affiliated with the program.

The types of information Boston University will question informants about include: (1) knowledge of the quality-based payment program in their organization and more generally, financial rewards for quality; (2) perception of the impact of the quality-based payment program on physician behavior in their practice setting, including the relevance of the quality measures, adequacy of financial incentives, and the fairness of payout formulas; (3) perception of unintended consequences of quality-based payment including impact on other clinical or support staff; description of the implementation of the quality-based payment program, including challenges encountered and strategies for addressing them; and (4) perception of factors within the physician organization, physician practice, and larger hospital setting that affected the structure and implementation of the quality-based payment program. In addition, physician organization executives will be asked about the ***decision-making process that shaped the structuring of the quality-based payment program and the arrangements with the health plan or insurers providing the incentives. We will also ask about the selection of quality targets and, if relevant, why some targets deviate from national standards (i.e., HEDIS measures).*** This

data will be used to generate hypotheses about the potential for quality-based payment programs to improve quality in the safety net setting and the unintended consequences of implementing such a program. *Semi-structured interviews of senior managers in each study site will also be used to identify whether other types of incentive arrangements have been established with respect to physician performance. Should other incentive arrangements exist, such information will be considered as a possible confounder relative to any significant changes that we observed in quality measures following the introduction of the quality-related incentives. Boston University will also inquire about any decisions that were made to exclude certain practices/ clinics or individual physicians from the incentive programs.*

The interviews will be conducted concurrently with survey fielding. Boston University will ask each interviewee if we can follow-up with them by phone should a question come up following the survey analysis.

3. **Use of Improved Information Technology**

Because this is a small project, investing in improved electronic technology (e.g., web-based materials) would not be cost-effective nor reduce the burden of data collection on the public.

4. **Efforts to Avoid Duplication**

Data collection instruments have been designed to reflect the specifics of the targeted population. The voluntary data collection instruments being used for this study have already been validated in a previous study conducted by AHRQ and Boston University staff worked with AHRQ staff to make sure this effort does not duplicate previous efforts or any other efforts to assess financial incentive programs in safety net settings.

5. **Small Businesses**

The physician survey and key informant interview guides were designed to minimize burden on all respondents and will not have a significant effect on small businesses or other small entities. None of the physician group practices at which the survey respondents practice are small businesses. Responses are entirely voluntary.

6. **Consequences of Less Frequent Collection**

Collection of data only occurs once during the project for both the survey and interviews. The data collection instruments are appropriate vehicles to assess quality-based payment strategies implemented in safety net settings. A previous project evaluating the Rewarding Results Demonstration Program used these same instruments, used only once, and the results have been published in several peer-reviewed journals.

7. **Consistency with the Guidelines in 5 CFR 1320.5(d)(2)**

The data collection efforts will be consistent with the guidelines at 5 CFR 1320.5(d)(2).

8. **Applicability to 5 CFR 1320.8(d)**

5 CFR 1320.8(d) is not applicable to this request.

9. **Remuneration of Respondents**

Physician Survey Respondents: It is well established that response incentives are an effective technique for raising survey response rates, especially among busy professionals such as physicians. Physicians in private practice are, in fact, losing money when they take time to participate in research activities such as surveys, making financial incentives all the more welcome and effective in improving response rate. However, no remuneration will be provided to survey respondents in this data collection for two reasons. First, using a cash response incentive in connection with a survey about quality-based payment could be rightly criticized as a methodology that would produce a biased sample—namely, one that contained a disproportionate number of individuals who had favorable views regarding financial incentives. Second, the data collected from survey respondents in this project will be compared with data collected from a previous administration of the survey instrument in which no response incentives were used. In order to make sure study populations are comparable, the current data collection must use identical data collection procedures.

Key Informant Interview Respondents: Interview participants will be given a one-time honorarium for participating in an hour-long interview. This honorarium is meant to reimburse the participants for their time. We will pay a $50 honorarium per individual. Remuneration for key informant interview participation is a recognized standard industry practice, without which, it would be difficult to achieve appropriate and adequate participation.

10. **Assurance of Confidentiality**

Individuals and organizations contacted will be assured of the confidentiality of their replies under Section 924(c) of the Healthcare Research and Quality Act of 1999. Respondents will be advised that surveys and interviews are entirely voluntary and that any information they provide will be combined and summarized with information provided by others and no individually identifiable information will be released. In instances where respondent identity is needed to facilitate data collection (e.g., follow-up surveys), the information collection will fully comply with all respects of the Privacy Act. No waiver is necessary for this project.

11. **Questions of a Sensitive Nature**

No questions of a sensitive nature are anticipated under this clearance.

12. **Estimates of Hour Burden Including Annualized Hourly Costs**

Response burden estimates are shown in table 12. The survey questionnaire will require 15 minutes to complete and the key informant interviews will require one hour. Only one response is required from each respondent. The total annual burden hours is estimated at 114 hours, as shown in table 12. The total annualized cost to respondents is estimated at $7,614.06, as shown in table 12.

Table 12

Type of Respondent	Estimated Number of Respondents	Average Burden Hours per Response	Total Burden Hours	Average Wage Rate	Total Respondent Cost Burden
Physicians	360	0.25	90	$66.79*	$6,011.10
Key Informants**	24	1.00	24	$66.79*	$1,602.96
Totals	240	--	114		$7,614.06

*Based upon the mean of the average wages for physicians and medical service managers, National Compensation Survey: Occupational wages in the United States May 2005, US Department of Labor, Bureau of Labor Statistics. Available at: http://www.bls.gov/oes/current/oes_nat.htm.
** Including such informants as physician organization executives, practice leaders, physicians, and clinical staff. Since the exact distribution of respondents across these different occupations is not known the highest average wage rate is used so as to not underestimate respondent cost burden.

13. **Estimates of Other Total Annual Cost Burden to Respondents**

The only cost to the respondent will be that associated with their time to respond to the information collection, as shown in table 12. There will not be any costs for capital equipment or operational expenses.

14. **Estimates of Annualized Cost to the Government**

The maximum cost to the Federal Government is $193,941.00.

15. **Change in Burden**

Not applicable. This is a new clearance.

16. **Plans for Analyses**

The purposes of these information collections are to evaluate quality-based payment programs in two safety net settings. The analyses will be descriptive and, to the extent that they can, inferential and generalizable. The results of these findings are primarily for internal use but may be shared with key government policy and management officials, AHRQ staff, public and private health providers, and members of the general public.

For the types of information data collections described in item 2, the following analyses will be employed:

Clinical Performance Data Analysis: A central question of the proposed evaluation is whether it appears that the quality-based payment concept can lead to improvement in selected quality measures in safety net settings. To address this question, Boston University will conduct analyses for the previously noted study settings for which data on clinical quality measures are available before and after the introduction of financial incentives. Boston University will conduct for each setting a separate pre-test/post-test analysis to assess whether and to what degree there was improvement in relevant quality measures following the introduction of financial incentives. For the analysis itself, Boston University will conduct a two-way repeated measures analysis of variance (ANOVA) for a particular quality measure that considers changes in both performance levels (pre-intervention vs. post-intervention) and trends (time point 1, 2, 3, within pre- and post-intervention periods). In this analysis, statistically significant interactions between changes in performance levels and changes in trends indicate an actual change in performance beyond an extension of the pre-intervention pattern. Additionally, Boston University will conduct Neuman-Keuls t-tests for post-hoc multiple comparisons to assess on a year-to-year basis whether and where changes in scores for each measure were statistically significant throughout the period for which we have data at any particular setting. To help determine whether observed improvements are attributable to particular quality-based payment programs versus secular trends, we will where possible obtain data (e.g., HEDIS) for comparing a particular settings' performance for a given quality measure to national or regional trends for that measure.

Physician Survey Analysis: Analysis of the physician surveys will be used to address the previously noted questions about unintended consequences and important implementation factors. Basic descriptive statistics and frequencies will be generated on all survey items to check for out-of-range values, high percentages of missing data, and other anomalies. If observed, appropriate corrective actions will be taken. Once the data have been cleaned, scale scores will be computed. The reliability of the scales for the present population will be checked by computing internal consistency reliability coefficients (Cronbach's alpha) for all scales.

Boston University will compare mean scale scores for Montefiore physicians to Boston Medical Center Health Plan's physicians to examine whether the quality-based payment programs at these two institutions have had a differential impact on physicians' perceptions of quality-based payment, and if so, on what dimensions. This will be done by using independent-sample t-tests. To control for the increased likelihood of Type 1 error associated with multiple analyses, a Bonferroni adjustment will be applied to the criterion alpha level used to determine statistical significance. Boston University will also make comparisons of the mean scale scores obtained in this project with the mean scale scores of physicians who previously responded to the survey who participated in quality-based payment programs in non-safety net settings.

Cohen's *d* statistic (the difference between means as a proportion of the combined groups' standard deviation) is the appropriate effects size measure for the comparison of independent means by t-test (Cohen 1977). Assuming a modest (d=.30) to moderate (d=.50) effect size (ES), samples of 64 (moderate ES) to 175 (modest ES) per group would be sufficient to achieve power of .80 to detect a two-tailed difference a p=.05. Given our initial population sizes of 300 at each site and an anticipated overall response rate of 60%, we anticipate having more than adequate power for the analyses described above.

Key Informant Interview Analysis: Analyses of the key informant interviews will be used to understand the effects of context on the effectiveness quality-based payment programs; to describe the implementation of quality-based payment and the factors that affect it in different settings; and to identify and understand unintended consequences of quality-based payment to physicians. Boston University will use an explanation-building analytic strategy. Consistent with Miles and Huberman's guidelines for comparative case studies, Boston University will work to understand each case before proceeding to cross-site explanations, and then will cycle back and forth between analytic strategies aimed at understanding case dynamics and understanding the effect of key variables (Miles and Huberman 1994). The site summaries will be the starting point for developing our understanding of each case. Based on them, Boston University will develop analytic memos that describe the key factors that appear to influence quality-based payment program effectiveness within that site. For cross-case analysis, Boston University will begin by coding and sorting the interview notes into descriptive meta-matrices organized by components of the initial conceptual framework. Each team member will be assigned one or more components for further analysis of themes and interactions, and the team will re-visit the individual case analyses in order to make sure that the cross-site explanations of important components are consistent with the explanations developed within each case site. Through extensive team discussion, we will identify key factors in addressing the evaluation questions and the relationships among them. Boston University will test and further refine the findings through discussion with and feedback from the study sites.

17. **Exemption for Display of Expiration Date**

No exemption is being requested.

18. **Certifications**

These activities will comply with the requirements of 5 CFR 1320.9.

STATISTICAL METHODS

1. **Potential Respondent Universe and Sample Selection Method**

Survey: Data collections will be designed to minimize burden on respondents while obtaining essential information. The respondent universe includes all physicians who practice at (a) Montefiore Medical Center within the previously noted practices of the Montefiore Medical Group, and (b) the health centers and group practices within the network of the Boston Medical Center Health Plan. However, in order to appropriately assess the effect of the quality-based payment program, Boston University's potential respondent universe is limited to physicians who have formerly or are currently participating in their institution's quality-based payment program. Our partners at these institutions have indicated that there are approximately 300 physicians each in these settings (a total of 600 physicians) who meet these criteria. Boston University intends to survey all 600 physicians in the study universe and expect to achieve a least a 60% overall response rate based on site liaisons' prior experience surveying this same population for other research projects. ***Boston University also expects a very high level of commitment from the liaisons in terms of promoting the survey and motivating respondents to complete and return the questionnaires.*** Because Boston University anticipates yielding responses from 360 physicians which will provide sufficient power for the analysis, no sampling selection method will be used for the survey. Boston University expects that the physician survey will take approximately 15 minutes to complete per physician.

Boston University will investigate possible response bias by comparing characteristics of respondents with non-respondents. From each of the study sites, Boston University will be able to obtain basic demographic characteristics for all physicians in the sampling frame. These demographic characteristics include age, sex, number of years in practice, and number of years since joining the current medical center. On the basis of these characteristics, Boston University will compare respondents to non-respondents. Boston University has performed similar analyses in our previous surveys of physicians participating in quality-based payment programs within commercial settings and have not found any substantial evidence of response bias.

Interviews: There are a finite number of practice executives and senior managers and practice executives at the participating clinics and community health centers (up to 45 potential interview respondents). Boston University only plans to interview 24 respondents overall. The interviews with senior managers and practice executives will be completed in approximately 1 hour.

For both the interviews and surveys, Boston University will obtain readily available respondent lists from project liaisons at each site. Identification is simplified because the participants are employees of the institutions implementing the quality-based payment programs included in the evaluation.

2. **Information Collection Procedures**

In this section, we describe how Boston University will collect data as specified in item 2 above. All information collections will be conducted in a manner that is consistent with the following guidelines:

- Participation will be fully voluntary, and non-participation will have no effect on eligibility for, or receipt of, future AHRQ-sponsored health services research.

- Information collection will be limited to that needed to assess the perceptions and attitudes toward financial incentives at the site being studies and will only be conducted once.

- Given the voluntary nature of the information collections, efforts will be made to obtain the highest possible response rates. Efforts will also be made to assess non-response bias, to the extent feasible.

Clinical Performance Data Collection: Sites will directly supply Boston University with clinical performance data, either in electronic or paper-based format.   Both sites' participation in the project was predicated on their willingness and capacity to provide this information to Boston University. Information will be provided in a secure format and stored on a password-protected computer. *For each research site, Boston University intends to examine whether improvement occurred on the targeted quality measures following the introduction of the programs. Boston University will compare baseline scores with post-intervention scores by applying the statistical procedure of repeated measures analysis of variance. This procedure is useful for assessing whether the post-intervention scores are significantly different from the pre-intervention pattern in terms of both levels and trends. Boston University applied this statistical procedure to assess the impact of an incentive program in the Rewarding Results demonstration (Young et al., Journal of General Internal Medicine, in press). Boston University may also consider using other procedures that take into accounting nesting of observation as we recognize that physicians working together in the same practice or clinic are not necessarily truly independent observations from a statistical point of view. Thus, Boston University may also examine data using repeated measures analysis depending on whether preliminary tests indicate moderate correlations among physician performance scores within the same practices or clinics.*

At each site, Boston University in collaboration with managers and clinicians will identify a set of performance measures that have not been explicitly rewarded but still serve as important indicators of quality of care.   One concern about linking financial incentives with quality measures is that it will encourage providers to adopt a "teach to the test" mentality" so that they may pay less attention to clinical activities that are not financially rewarded even though they are though are clinically important.

Boston University will also have data on patient caseload and will use this information in the analyses. In particular, Boston University will include in the impact analyses only those physicians who had a sufficient number of relevant cases for a quality measure so that performance scores are reliable.

Physician Survey Data Collection: Boston University will identify physicians through senior managers participating in the key informant interview who will provide Boston University with lists of names of all physicians participating in his or her organization's quality-based payment program. The questionnaire will be administered as a confidential mail survey using a modified version of the methodology proposed by Dillman (2000). Boston University will send surveys to a site liaison who will distribute them individually to each physician participating in their organization's program. The survey will be administered as a confidential mail survey which will take between 10 and 15 minutes to complete. Questionnaire booklets will be sent by first class mail address to a ***project site liaison*** accompanied by a cover letter from the CEO of the relevant medical group and a pre-paid business reply envelope addressed to a third-party data entry vendor. ***The project site liaison will distribute the surveys to individual physicians.*** Both the questionnaire and cover letter are included in the Appendix of this package. The first page of the survey booklet will contain the informed consent. Physicians will convey their consent by completing the survey and returning it in the envelope provided. No signature is required. The physician can decide whether or not to move forward to complete the survey. The survey will not collect any information about provider performance specific to an individual. Rather, Boston University is collecting information about how providers feel about the financial incentive program in which they are participating and how providers collectively perform in meeting institutional performance targets.

Each questionnaire will bear a unique identification number that will be used to log in respondents and thereby drop them from follow-up mailing lists. Not only does this approach spare those who have responded the annoyance and potential confusion associated with receiving unnecessary follow-up contacts, but it is also much more resource efficient than the blanket follow-up that is necessary if an individual identifier is not used. Two weeks after the initial questionnaire is mailed, a second questionnaire will be mailed to a site liaison to distribute to non-respondents. ***Boston University will not be using thank you or reminder postcards after the first round of surveys are distributed, as was the case in prior administrations of the survey for the aforementioned sites in question.*** Data collection will be closed four weeks after the second questionnaire mailing. The entire duration of the survey fielding will take place over a six-week period.

Key Informant Interview Data Collection: To identify these informants, Boston University will ask practice executives or leaders at each site to provide us a list of personnel at the management, clinical, and administrative levels identified above who are participating in or affiliated with the program. Boston University will contact each individual by letter to solicit participation and follow-up by phone to request his or her participation in the interview. The advance letter is contained in the Appendix of this package. If the individual agrees verbally by phone to an interview, Boston University will fax or email him or her ***an abbreviated copy of the interview protocol*** (see Appendix for full-length protocol). Boston University will then schedule an in-person interview with that individual. ***At that time, the research team member scheduling the interview will inform the interviewee that there are questions during the interview that may require preparation (e.g., demographic questions). The scheduler will highlight those questions during the call and on the protocol copy.*** Boston University will obtain written informed consent on-site, including consent to audio-tape the interview. If the interviewee does not grant us permission to tape the interview, Boston University will conduct the interview

without taping it. Boston University will also request that we can contact the individual with follow-up question if clarification is needed, but will not do so without consent.

Informants will be asked to participate in an hour-long interview with two members of the evaluation team; one member of the team will be responsible for guiding the interview, while the second will take detailed notes. Informed consent will be obtained *in writing* before the interview. Interviews will semi-structured. ***The interview protocol is a modification of a protocol that was piloted and used extensively in the aforementioned Rewarding Results evaluation. Two articles presenting results from data collection using the interview protocol have been published (Bokhour et al., 2006, and Sautter et al., 2007).*** The interview protocol will suggest questions and areas of interest to pursue; however, the interviewer will allow participants to discuss issues of greatest concern to them and follow up on themes generated by the participants. After each interview, both members of the interview team will review the notes for completeness and will add any additional observations about the interview. For each site visit, the evaluation team will generate a site summary that summarizes the major research questions and themes observed based on all the interviewees' responses at that site. Boston University anticipates this part of the data collection will take place over a 6-month period.

3. **Methods to Maximize Response Rate**

The design of each information collection will include approaches to maximize response rates, while retaining the voluntary nature of the effort, consistent with appropriate survey methodology. Boston University will employ several methods to maximize response rates for both the survey and key informant interviews. For the survey, Boston University will capitalize on its close relationships with site liaisons to distribute the survey to physicians in groups (e.g., at staff meetings or in-services). The site liaisons will personally distribute the survey to physicians and stress the importance participation could have in contributing to quality improvement efforts. The site liaisons have also committed to doing individual follow-up with physicians who do not respond to the first round of surveys.

For the key informant interviews, Boston University will pay respondents an honorarium of $50 per individual, which will be an hour in length. Project leaders in each site will make personal calls to potential interviewees to encourage their participation.

4. **Tests of Procedures**

The individuals who will be involved in the statistical design and analysis of this project are Mark Meterko, PhD, and Hai Lin, MD, MPH.

Because Boston University is using a survey instrument previously tested and validated, no pilot testing of the survey is necessary. A paper documenting the instrument was published in Health Services Research (Meterko et al. 2006).

Similarly, the senior manager/practice executive interview protocol is an adaptation of a protocol used in a previous project. A paper using this interview protocol was published in Medical Care Research and Review (Bokhour et al., 2005). Apart from making slight modifications to the protocol reflecting the specific institution by which the informant is employed, we do not anticipate making changes to these instruments in view of their prior validation.

5. **Statistical Consultation and Independent Review**

Input from statisticians regarding the development, design, conduct, and analysis of information collections was sought by Boston University prior to submitting a proposal responding to AHRQ's request for research on pay-for-quality programs. Specifically, Boston University consulted Errol Baker, PhD, a researcher based at the Center for Organization, Leadership, and Management Research in the VA Boston Healthcare System's Research Institute, who consulted on the design of the validated survey instrument. Boston University will consult with other statisticians outside the project team during the course of the project, principally faculty based at the Health Policy and Management department.

6. **Study Limitations**

Apart from the previously discussed issue of response rate for the survey, the data collection presented herein is limited in the following ways. As noted, there is growing concern about the impact of quality-based payment programs on providers and patients in safety net settings. At the same time, evaluating the impact of quality-based payment programs in safety net setting has some very stringent challenges that Boston University is very well positioned to meet. One major challenge is being able to identify the possible unintended or negative effects that such programs produce. Possible negative effects include providers seeking to avoid sicker or less compliant patients, and providers ignoring important clinical activities that are not explicitly recognized as quality targets in a quality-based payment program. The challenge to researchers in identifying such negative effects is developing and operationalizing the measures as well as obtaining the necessary data. Boston University is well positioned to meet this challenge because during as part of its Rewarding Results Demonstration Program evaluation, Boston University developed the skills and measurement instruments for detecting unintended consequences of quality-based payment programs. In particular, Boston University's previously discussed provider survey instrument includes a scale that is designed to assess physician attitudes and experiences regarding unintended consequences from quality-based payment programs. The interview protocols for senior managers also include questions and related probes that we have used successfully to inquire about unintended consequences. We also have been developing a set of clinical tracers—*that is, quality measures which are not explicitly rewarded but are still important indicators of quality of care*—for detecting negative effects from quality-based payment programs that can be applied to administrative and clinical data sets. The selection of tracers depends on which type of quality targets are the focus of the program. Boston University will work with its colleagues at each study setting to identify tracers that are appropriate for their own programs to detect negative effects.

Another major challenge to studying quality-based payment in safety net settings is being able to identify and study early-stage implementation problems as they occur or soon thereafter. Studying implementation problems in a real-time fashion allows for more reliable and valid collection of data from those individuals who are engaged in the implementation process. However, very often, the researcher must study the implementation of such programs retrospectively long after early-stage problems have occurred. Fortunately, Boston University will be able to capitalize on the fact that one the study settings is currently in the early stages of their implementation effort, thus allowing us the opportunity to identify and examine important problems/events as they occur.

References

Baker, G., and B. Carter. 2005. *Provider Pay-for-Performance Incentive Programs: 2004 National Study Results*. San Francisco: Med-Vantage.

Banaszak-Holl, J., J. S. Zinn, and V. Mor. 1996. "The Impact of Market and Organizational Characteristics on Nursing Care Facility Service Innovation: A Resource Dependency Perspective." *Health Services Research* 31 (1): 97-117.

Bokhour, B. G., J. F. Burgess, J. M. Hook, B. White, D. Berlowitz, M. R. Guldin, M. Meterko, and G. J. Young. 2006. "Incentive implementation in physician practices: a qualitative study of practice executive perspectives on pay for performance." *Medical Care Research and Review* 63 (1): 73S-95S.

Cabana, M. D., C. S. Rand, N. R. Powe, A. W. Wu, M. H. Wilson, P. A. Abboud, and H. R. Rubin. 1999. "Why Don't Physicians Follow Clinical Practice Guidelines? A Framework for Improvement." *Journal of the American Medical Association* 282 (15): 1458-1465.

Cohen, J. 1977. *Statistical Power Analysis for the Behavioral Sciences*. Revised edition. New York: Academic Press.

Conrad, D. A. and J. B. Christianson. 2004. "Penetrating the "Black Box": Financial Incentives for Enhancing the Quality of Physician Services." *Medical Care Research and Review* 61 (3, Supplement): 37S-68S.

Dillman, D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons, Inc.

Dudley, R. A., R. H. Miller, T. Y. Korenbrot, and H. S. Luft. 1998. "The Impact of Financial Incentives on Quality of Health Care." *Millbank Quarterly* 76 (4): 649-686.

Institute of Medicine. Committee on Quality of Health Care in America. 2001. *Crossing the Quality Chasm: A New Health System for the Twenty-first Century*. Washington, DC: National Academies Press.

Meterko, M., G. Y. Young, B. White, B. G. Bokhour, J. F. Burgess, Jr., D. Berlowitz, M. R. Guildin, and M. N. Seibert. 2006. "Provider Attitudes toward Pay-for-performance Programs: Development and Validation of a Measurement Instrument." *Health Services Research* 41:5:

Miles, M. B., and A. M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Pathman, D. E., T. R. Konrad, G. L. Freed, V. A. Freeman, and G. G. Koch. 1996. "The Awareness-to-Adherence Model of the Steps to Clinical Guideline Compliance: the Case of Pediatric Vaccine Recommendations." *Medical Care* 34 (9): 873-889.

Shortell, S. M., J. L. O'Brien, J. M. Carman, R. W. Foster, E. F. Hughes, H. Boerstler, and E. J. O'Connor. 1995. "Assessing the Impact of Continuous Quality Improvement/Total Quality Management: Concept versus Implementation." *Health Services Research* 30 (2): 377-401.

Shen, Y. 2003. "Selection Incentives in a Performance-based Contracting System." *Health Services Research* 38: 535-52

Tamblyn, R., P. McLeod, J. A. Hanley, N. Girard, and J. Hurley. 2003. "Physician and Practice Characteristics Associated with the Early Utilization of New Prescription Drugs." *Medical Care* 41 (8): 895-908.

U. S. Congress. 2006. *Deficit Reduction Act of 2005*. 109th Congress, 1st session. Public Law 109–171, Sec. 1932.

Young, G. J., B. White, J. F. Burgess, Jr., D. Berlowitz, M. Meterko, M. R. Guldin, and B. G. Bokhour. 2005. "Conceptual Issues in the Design and Implementation of Pay-for-Quality Programs." *American Journal of Medical Quality* 20 (3):144-50.