Outcomes of a Seven Practice Pilot in a Pay-for-Performance (P4P)-Based Program in Pennsylvania

Rhonda M. Johnson · Twyla Johnson ·
Sarah D. Zimmerman · Gary M. Marsh ·
Oralia Garcia-Dominic

Received: 30 May 2014 / Revised: 29 July 2014 / Accepted: 29 August 2014 / Published online: 2 October 2014
© W. Montague Cobb-NMA Health Institute 2014

Abstract

Objective The objective of this study was to examine how targeted six-month interventions impacted best practice/patient outcomes for minority patients receiving primary care in physician practices participating in a pay-for-performance (P4P) program.

Methods P4P practices were invited to participate in a pilot intervention study designed to improve care for minority patients with hypertension, diabetes, or pediatric asthma. The following patient medical records were reviewed to assess how the interventions impacted ($n=7$ practices): body mass index, diet and exercise, smoking, compliance with visits as recommended, blood pressure, sodium intake and weight management counseling, medication reconciliation, HbA1c

testing, annual lipid profile, and anti-inflammatory medications.

Results Significant improvements in various clinical quality measures were observed in all seven practices. Of the 19 specified interventions, 13 were statistically significant at $\alpha=0.05$ level and 14 met the target proportion. This suggests that the best practice intervention had a significant impact on some of the health care processes in the physician practices.

Conclusions The most impactful interventions were those related to face-to-face educational discussions and patient medical chart documentations rather than those pertaining to medication adherence. Improvements in measuring reporting and recording of data at post-intervention were also observed.

R. M. Johnson
Health Equity & Quality Services at Highmark, 120 Fifth Avenue,
Suite FAP 733, Pittsburgh, PA 15222, USA
e-mail: Rhonda.moore.johnson@highmark.com

T. Johnson
Provider Engagement, Performance & Partnerships at Highmark, 120
Fifth Avenue, Suite FAP 893, Pittsburgh, PA 15222-3099, USA
e-mail: twyla.johnson@highmark.com

S. D. Zimmerman · G. M. Marsh
The Center for Occupational Biostatistics and Epidemiology,
Department of Biostatistics at University of Pittsburgh, A410
Crabtree Hall, A411 DeSoto Street, Pittsburgh, PA 15261, USA

S. D. Zimmerman
e-mail: scd27@pitt.edu

G. M. Marsh
e-mail: gmarsh@pitt.edu

G. M. Marsh
Biostatistics, Epidemiology and Clinical & Translational Sciences,
Graduate School of Public Health, University of Pittsburgh, A410
Crabtree Hall, 130 DeSoto Street, Pittsburgh, PA 15261, USA

G. M. Marsh
Highmark Health Equity Quality at Highmark, 1800 Center Street,
B1 L2 Suite 219, Camp Hill, PA 17011, USA

O. Garcia-Dominic (✉)
Departments of Biobehavioral Health at The Pennsylvania State
University, and Public Health Sciences at Milton S. Hershey Medical
Center, College of Medicine ASB 2200 Hershey,
Pennsylvania 17033, USA
e-mail: Oralia.dominic@highmark.com

O. Garcia-Dominic
e-mail: oag102@psu.edu

O. Garcia-Dominic

Allegheny-Singer Research Institute (ASRI), 30 Isabella Street, Suite
300, Pittsburgh, PA 15212-4772, USA

Keywords Pay-for-performance · Healthcare health disparities · BMI and diabetes · Primary care physicians · Smoking

Introduction

The 2011 National Healthcare Disparities Report documents that racial and ethnic disparities remain pervasive [1]. A decade prior, in 2001, the Institute of Medicine released the *Crossing the Quality Chasm* report that recommends incentives to encourage quality health care and better health outcomes [2]. To address these concerns, numerous private-sector initiatives to link portion of payments made to physicians and hospitals dependent on improved quality are underway. These initiatives are commonly referred to as pay-for-performance programs (P4P) [2]. P4P is broadly defined to include any type of performance-based provider payment arrangement, including those that target performance on cost measures [2]. The primary goal of P4P is rewarding or improving quality of care; a secondary goal is controlling cost by reducing errors and inappropriate utilizations [2]. P4P is only one of many proposed strategies that may improve quality and thus positively impact healthcare disparities [3–6]. A proposed strategy for addressing health disparities is engagement [7] as defined by general awareness, reflection/empowerment, and cues to action within their own practice and their communities in general.

It is well recognized [3, 4, 7–11] that there are both advantages (increased payment, improved efficiency, and improved care quality, especially among disparate populations) and disadvantages (cost of acquiring information technology, multiple programs and guidelines, data collection) to P4P programs currently designed and implemented as reported by the American Academy of Family Physicians (AAFP) [8]. One challenge associated with P4P is its potential to increase disparities if some groups (like medically underserved groups) are less able to respond to specific interventions when compared to others [12]. Another reported challenge is to implement P4P in the small practice setting [2, 4, 8], which experiences substantial barriers like limited staff and fewer resources than larger group practices, thus increasing the difficulty completing quality improvement projects [2, 4, 8]. According to AAFP, there are seven P4P guiding principles that programs should adhere to [8]: (1) focus on improved quality of care; (2) support the patient relationship; (3) utilize performance measures based on evidence-based clinical guidelines; (4) involve practicing physicians in program design; (5) use reliable, accurate, and scientifically valid data; (6) provide positive incentives; and (7) offer voluntary participation. Highmark's Quality Blue Physician P4P program incorporates all of these principles. Specifically,

the program recognizes and rewards primary care physicians (PCP) for collaborating with Highmark to improve the quality and effectiveness of health care services provided to our members.

Highmark has been engaged in both focused disparity-related activities as well as P4P for several years. These efforts have been published elsewhere [13–15]. Notably, Highmark's health disparities work has been recognized in the National Health Plan Collaborative (NHPC) Toolkit. NHPC was a novel public-private partnership between eight health plans covering approximately 95 million lives. Sponsors included the Agency for Health Care Research and Quality and the Robert Wood Johnson Foundation.

The Quality Blue Physician P4P program is offered in 49 counties of western and central Pennsylvania with approximately 6300 primary care physicians (PCP) in more than 1500 practices, providing services to more than 1,731,000 unique Highmark members. The primary care physician network is composed of 2448 practices; 1604 or 66 % of all primary care practices participate in the Quality Blue program.

The Best Practice component of Highmark's Quality Blue Physician P4P program rewards practices for conducting a population-based clinical initiative in their office setting or for completing a professional organization certification/recognition activity. The primary goal of the Best Practice component is to improve patient care offered by requiring practices to apply knowledge and experience gained via Deming's Plan-Do-Study-Act methodology. The practices explore opportunities for clinical-based activities, thereby empowering them to integrate improvement models into daily office activities and promote proactive rather than reactive care. The Best Practice studies carried out by practices have varied over the years and have included some of the following improvements:

- Improving diabetic LDL and HbA1c control;
- Reviewing medications with patients to address medication adherence;
- Conducting spirometry testing to assess for asthma and COPD;
- Counseling and other interventions to address childhood obesity; and
- Improving rates of colorectal cancer screening.

Although P4P strategies are increasingly common, little is known about the impact of applying this payment method to reducing healthcare disparities (HCD) in general or by geography (urban and rural areas), including Pennsylvania [16]. In this study, we hypothesize that existing P4P programs have the potential to decrease disparities in minority groups. This study examined the impact of various interventions on the Best Practice measures: as follows body mass index (BMI),

diet and exercise, smoking, compliance with visits as recommended, measuring blood pressure, sodium intake and weight management counseling, medication reconciliation, HbA1c, lipid profile, and anti-inflammatory medications among participating network practices utilizing a P4P approach in the Pennsylvania counties in Highmark's service region.

Methods

Participating Practices

Although all PCP practices that participate in the Quality Blue Physician P4P program are encouraged to conduct a Best Practice study, practices identified as serving sizable minority populations (e.g., African American and Latino/Hispanic) were recruited to participate in the pilot study. Highmark's clinical quality consultants (CQC) and provider relations staff met with each of the practices to inform them of the pilot and to gauge their interest in participating. Efforts were made to recruit practices where African American and Hispanic Highmark members with known cardiovascular and diabetes disparities received their primary care. Since previous data analysis had identified disparities in pediatric asthma clinical indicators, efforts were also made to recruit pediatric practices that served minority children.

All PCPs serving minority (African American and Latino/Hispanic) patients previously identified with disparities were identified by Highmark trained staff in April 2010. These PCPs were identified through a query report matching all minority patients with known healthcare disparities in diabetes and cardiovascular disease HEDIS indicators with their primary care physician practices who were participating in the Quality Blue Physician P4P program. From these PCPs, practices were identified that included a total of at least five eligible minority patients. The five-patient criterion was used to allow sufficient numbers of minority patients for the statistical evaluation of the pilot and to help identify practices that may be more willing to participate. Figure 1 illustrates how the practices were selected for the pilot study. Of the 515 practices identified from the eligible PCPs, only 29 met the minimum five-patient criterion. In addition, 30 practices were identified by Highmark staff between April and August 2010 that were not included in the query report but were recommended as having a larger volume of minority patients with hypertension or pediatric asthma. Thus a total of 59 practices were available for the pilot. Of those, 50 practices met the eligibility criteria of having previously participated in the best practice option. Of the 50 practices, 25 chose to participate in the pilot, and among those, a total of seven completed the pilot study. The Allegheny Singer Research Institute's Institutional Review Board approved this study.

Fig. 1 Data flow for pay-for-performance best practices/patient outcomes for minority patients

Study Design

Best practice interventions were decided within each of the seven participating practices. Therefore, each practice participated in different interventions. The clinical Best Practices were measured at the end of each month for six consecutive months. A random sample of medical charts of minority patients with the identifiable healthcare disparity was selected to audit for evidence of the intervention-specific measurement. The baseline measurement was taken before the start of the intervention and was compared with results at month six to evaluate the effectiveness of the various interventions within each practice. Interventions are described in the left most column of Table 1.

Statistical Analyses

A two-tailed, two-sample test of proportions was performed, which compares the baseline results (proportion of patients satisfying the BPPO measure) for each BPPO with the final results at month six. The baseline and final samples were treated as independent samples. A statistically significant result is indicated if the p value is less than or equal to 0.05. Due to the exploratory nature of this pilot, no adjustments were made to account for multiple comparisons. It is possible that some of the subjects were measured at both baseline and month six; in this scenario, the independent two-sample test of proportions will likely yield a conservative p value. Additionally, most of the outcomes are measures of documentation rate of the practice which would be unaffected by the particular patient being documented. As such, even if patients did overlap in the samples, the resulting pre- and post-documentation rates of the practice would be independent of each other. Practices were identified if it met or exceeded the target proportion. Targets were specified separately by each practice.

Results

Practice Demographics

Of the 50 eligible intervention practices that met the requirements for this study, only 25 chose to participate with seven completing the P4P program. As shown in Table 1, two of the seven practices were hospital owned and multi-specialty, and the remaining five were independently owned. Four practices were located in urban Pennsylvania while two were in rural areas; one was located in a suburban area. Four of the seven practices had physician(s) reporting that he/she spoke more than one language. Most practices offered only weekday hours. Only one practice reported not having an electronic medical record (EMRs) system.

Patient Demographics

Table 2 shows a description of the patient population of the participating practices. Three geographic regions were represented across practices: Southwestern, Northwestern, and Central Pennsylvania. In general, there were equal representation of both male (48 %) and female (52 %) patients across practices, but there was slightly more females. Patients were heterogeneous relative to race/ethnicity across all practices. Approximately one in five patients in practices 1, 3, and 4 were black patients. For practices 6 and 7, one in three patients was Hispanic. The average income level was \$33,000 (range \$28,197 to \$39,529).

Best Practice/Patient Outcomes

As shown in Table 3, most of the differences in the proportion of patients satisfying the BPPO measure between the baseline, which was measured before the start of the intervention, and month six were statistically significant. Of the 19 specified interventions, 13 were statistically significant at $\alpha=0.05$ level, and 14 met the target proportion. This suggests that the best practice intervention had a significant impact on some of the health care processes in the physician practices.

For example, in practice 1, only three of the 19 patients had their BMI measured and documented before the start of the intervention. At the end of the intervention, all 19 of the patients selected in month six had their BMI measured and documented. The most impactful interventions were those related to education, discussions, and patient medical documentations rather than those pertaining to medication adherence.

Discussion

We examined the change in selected Best Practice measures among participating practices utilizing a P4P approach in the Pennsylvania counties in Highmark's service region. We found significant improvements in blood pressure measurement and documentation (from 15.8 to 100 %), exercise and eating plan documentation (from 15.8 to 100 %) and patient adherence to diet and exercise (from 15.8 to 84.2 %), smoking assessment and counseling provided (from 73.7 to 100 % for each parameter), sodium and weight discussion (from 0 to 100 % for each parameter), medication reconciliation (from 2.5 to 87.5 %), annual lipid profile testing (from 60 to 90 %), and asthma education (from 22.2 to 100 %) (see Table 3). While there was no formal comparison group, the magnitude of these findings suggests that P4P had a significant impact on some of the physician practices. This study supports the findings of others reporting the use of P4P as a proposed

Table 1 Practice demographics

Practice	Type of practice	Number of practice sites	Region	Location description	Number of physicians	Physician specialty	Patient age	Electronic medical records (EMR) and/or electronic prescribing (erx)	Evening and weekend hours	Languages spoken by physicians
1	Hospital-owned multi-specialty	2	Southwestern	Urban located inside two teaching hospitals	Six physicians	Emergency Medicine Internal Medicine Pediatrics	6–100+ years	eRx	No	English Spanish Polish Castilian
2	Independent Solo Practitioner	2	Southwestern	Suburban	One physicians	Family Practice	6–100+ Years	NNo	No (E) Yes (W)	English Spanish Castilian
3	Hospital-owned multi-specialty	1	Northwestern	Urban	Five physicians, one nurse practitioner	Family practice orthopedic surgery neonatology podiatry	0–100+ Years	EEMR eRx	Yes	English Hindi Telugu
4	Independent	1	Northwestern	Urban	Two physicians	Family practice ob-gyn	0–100+ Years	EEMR eRx	No	English
5	Independent	1	Northwestern	Rural	Two physicians, one physician assistant	Family practice pediatrics	6–100+ Years	EEMR eRx	No	English
6	Independent	1	Central	Rural	Six physicians, one physician assistant	Family practice general practice	0–100+ Years	EEMR eRx	No	English
7	Independent	1	Central	Urban	Three physicians, one physician assistant	Family practice general practice	0–100+ years	EEMR eRx	No	English Philippine Spanish Castilian

Table 2 Patient demographics for each of the seven practices

Practice	Geographic region	Gender distribution	Race	Age distribution	Median household income	Education level for ≥25 years old
1	Southwestern Pennsylvania	52 % Female 48 % Male	66 % White 26 % Black 4 % Asian 2 % Hispanic 3 % Other	0–19=21 % 20–64=65 % 65+=14 %	\$28,588	88 %=High school graduates 34 %=Bachelors degree
2	Southwestern Pennsylvania	54 % Female 46 % Male	91 % White 5 % Black 1 % Asian 2 % Hispanic 1 % Other	0–19=23 % 20–64=59 % 65+=18 %	\$39,529	95 %=High school graduates 33 %=Bachelors degree
3	Northwestern Pennsylvania	52 % Female 48 % Male	72 % White 17 % Black 2 % Asian 7 % Hispanic 2 % Other	0–19=28 % 20–64=60 % 65+=12 %	\$32,218	85 %=High school graduates 20 %=Bachelors degree
4	Northwestern Pennsylvania	52 % Female 48 % Male	72 % White 17 % Black 2 % Asian 7 % Hispanic 2 % Other	0–19=28 % 20–64=60 % 65+=12 %	\$32,218	85 %=High school graduates 20 %=Bachelors degree
5	Northwestern Pennsylvania	54 % Female 46 % Male	93 % White 3 % Black 1 % Asian 1 % Hispanic 4 % Other	0–19=24 % 20–64=57 % 65+=19 %	\$34,732	91 %=High school graduates 21 %=Bachelors degree
6	Central Pennsylvania	52 % Female 48 % Male	62 % White 4 % Black 1 % Asian 30 % Hispanic 3 % Other	0–19=29 % 20–64=57 % 65+=14 %	\$37,438	76 %=High school graduates 9 %=Bachelors degree
7	Central Pennsylvania	52 % Female 48 % Male	48 % White 13 % Black 1 % Asian 37 % Hispanic 1 % Other	0–19=36 % 20–64=54 % 65+=10 %	\$28,197	65 %=High school graduates 10 %=Bachelor's degree

strategy that may improve quality and thus positively impact healthcare disparities [1–6].

Similar to findings reported by Coleman¹⁴, improvements were observed in diabetes for our patients that tested their HbA1c (from 21.7 to 30.4 %). However, these changes were not statistically significant. Coleman's study showed that the P4P program significantly increased the likelihood that patients received two HbA1c scores per year as recommended by the American Diabetes Association, but do not contribute to improved blood sugar control [17]. Though we only measured HbA1c testing rate and did not directly measure diabetes control parameters, we do agree with Coleman's reporting that it appears that P4P programs may help with improvement in compliance with HbA1c testing recommendations; however, a more comprehensive strategy [17] may be necessary to better understand the link between P4P and diabetes control. One proposed strategy presented by Coleman is to include a patient social support [17] component in the P4P approach. This may be a reasonable starting point for examining this link as

minorities report lack of social support as a barrier to their health care behaviors (cues to action), including compliance [17].

Because in our study the most impactful interventions were those related face-to-face educational discussions and patient medical chart documentations rather than those pertaining to medications, we propose that strategies for examining this particular link also includes an education, discussion, and patient medical documentation component. Whether these components should be tested alone or in combination needs further research.

Strengths

There are five notable strengths to this study. First, this study included feedback from a disparate population that could benefit from improved health care quality. Second, this study builds upon Highmark's existing P4P program capacity to focus on disparity-related activities and provides a catalyst for extending its Best Practice component to address

Table 3 Final reporting and re-measurement of minority patients of a seven practice pilot in a pay-for-performance (P4P)-based program in Pennsylvania

Best practice intervention	Variable assessed from patient medical chart	Total patients participated: n	Baseline measurement: x_1, p_1^a	Month six measurement (final): x_2, p_2^a	p value for $H_0: p_1 \neq p_2$	Target ^b : p (%)	Did final measurement meet or exceed target?
Practice 1							
BMI measured within parameters—follow up plan for healthy habits	Both documented	19	3, 15.8 %	19, 100 %	<0.00001*	85	Yes
BMI measured outside parameters—exercise and eating plan developed	Both documented	19	3, 15.8 %	19, 100 %	<0.00001*	85	Yes
Screen for tobacco use	Patient adherence to diet and exercise documented	19	3, 15.8 %	16, 84.2 %	<0.00001*	85	No
	Assessed and documented	19	14, 73.7 %	19, 100 %	0.017*	95	Yes
	Counseling documented	19	14, 73.7 %	19, 100 %	0.017*	95	Yes
Compliance with at least two visits or as recommended	Patients that quit Documented	14	6, 42.9 %	10, 71.4 %	0.128	75	No
	Following initiation of TX uncontrolled/seen every 3 months	19	11, 57.9 %	16, 84.2 %	0.074	85	No
BP control <140/90	Controlled BP—seen every 6 months	4	4, 100 %	4, 100 %	1.00	100	Yes
	Documented measurement	19	4, 21.1 %	17, 89.5 %	<0.00001*	85	Yes
Practice 2							
Hypertension dietary/weight management	Sodium restriction discussion documented	20	0, 0 %	20, 100 %	<0.00001*	80	Yes
	Weight reduction/management discussions documented	20	0, 0 %	20, 100 %	<0.00001*	80	Yes
	BMI documented	20	20, 100 %	20, 100 %	1.00	100	yes
Hypertension medication management	Medication reconciliation documented	20	0, 0 %	20, 100 %	<0.00001*	80	Yes
Practice 3							
Hypertension medication management	Medication reconciliation documented	40	1, 2.5 %	35, 87.5 %	<0.00001*	85	Yes
Practice 4							
Diabetes management HbA1c testing rate	HbA1c testing documented	23	5, 21.7 %	7, 30.4 %	0.502	40	No
Practice 5							
Hypertension management lipid profile testing	Annual lipid profile documented	20	12, 60.0 %	18, 90.0 %	0.019*	90	Yes
Practice 6							
Persistent pediatric asthma treatment	Documentation of anti-inflammatory meds prescribed	30	27, 90.0 %	28, 93.3 %	0.644	95	No
	Documentation of Spacer device prescribed	48	21, 43.8 %	35, 72.9 %	<0.00001*	50	Yes
Practice 7							
Management of pediatric asthma	Asthma education and parent(s) verbalized understanding documented	9	2, 22.2 %	9, 100 %	0.0007*	95	Yes

*Significant at $\alpha=0.05$ level^a $p_1 = x_1/n$ and $p_2 = x_2/n$ ^b Target p was determined by each practice practitioner separately for each intervention

healthcare disparities among minority patients residing in urban and rural areas in Pennsylvania. Thirdly, this study provided an opportunity for one health insurer and several local practices serving minority patients to work together and identify potential strategies aimed at rewarding or improving quality of care. Another strength of the pilot is the absence of selection bias. Because the P4P incentive was based solely on completing the intervention and not on the results, there was no pre-selection of patients within each practice.

Lastly, all doctors and staff were required to complete a cultural competency course as a prerequisite for the participating in the pilot study. Completion of the course was optional but encouraged for nurses, physician assistants, and other staff at the practices. The course, Quality Interactions®, is a proprietary e-learning program that provides case-based instruction on cross-cultural health care. This interactive program focuses on common clinical and/or cross-cultural scenarios that build a framework of knowledge and skills for delivering quality care to diverse patient populations. A paired *t* test analysis was performed on the pre- and post- intervention test scores for the nurses, doctors, and physicians' assistants. The results were all statistically significant at $\alpha=0.001$ level, suggesting that the course improved their cultural awareness within and around the practice.

We attribute the statistically significant improvements observed in the Best Practices measures following intervention to our P4P program's study design, which was informed by the literature. For example, our P4P program focused on improving quality of care as defined by the selected biological and behavioral endpoints at baseline and monthly assessments over a six-month period. Our program supported the physician/patient relationship by providing a direct and open line of communications between our Highmark staff and each practice as needed. Our program utilized performance measures based on evidence-based clinical guidelines. It also included ongoing input during each phase of the study from Highmark's CQC and provider relations staff. One of the primary roles of Highmark's CQC staff is to act as practice coaches educating and guiding primary care physicians and their staff through quality improvement and practice transformation activities throughout the year. Staff from Highmark's Health Equity and Quality Services (HEQS) department collaborated on the pilot. The pilot physician and nurse participants were given access to a complimentary E-learning Quality Interactions Continuing Medical Education (CME) course and other cultural competency educational resources and the tools to implement the healthcare disparities activities. These resources remained available to network providers post-intervention.

Based on the opinion of and feedback from these intervention (pilot) practices, this P4P program can be easily adopted by other health plans provided they have an engagement strategy in place to support the spread of evidence-based guidelines, process, and quality

improvement activities and a commitment to eliminating healthcare disparities. If true, then this Quality Blue P4P Best Practice program can potentially be used as a template for practices aiming to reduce racial and ethnic disparities in health care among minority patients with diabetes, pediatric asthma and/or hypertension.

Results of this study may also help to inform next steps for public health initiatives aiming to improve patient medical chart documentations. Reported strategies include the following [18]: (1) to focus on how proper documentation can improve quality, having impact is important. This may help to encourage and provide added reasons for physicians to improve their medical documentation; (2) to provide training opportunities for the practices. Medical documentation is mostly the responsibility of a physician and medical staff, and thus across-the-board trainings become increasingly important. This may help provide consistency in documentations across staff members; and (3) to improve existing coding structures used by a practice. Review of coding structures should be revisited on an ongoing basis. This may help improve accuracy or consistency of claims data that is needed when codes change and/or do not provide the necessary level of detail. It is well documented that targeted interventions yield better outcomes when compared to non-targeted approaches. If true, then future studies examining effective strategies for improving patient medical documentation by physician specialty, practice size and geography, patient's race/ethnicity, gender, and staff trainings options to name a few are warranted.

Limitations

Three limitations must be noted. First, this was a one-time assessment of seven of 50 potential practices in Highmark service areas. Thus, the patients who participated in this study might not have been representative of all practices, including those serving outside the Highmark service area. However, many of these health professionals had experience working within and outside Highmark and stated that they had no reason to believe that the P4P approach used would not work in other practices serving minority patients. We did not assess whether their minority patients were any different compared to other minority patients. This was outside the scope of this study. Second, of 25 practices that committed to participation, only seven completed the pilot study. One explanation for the low completion rate by practices was due to internal changes/dynamics that were not foreseen. A major programming change in the Quality Blue P4P program requirements and scoring was implemented in 2011. The focus of the Quality Blue P4P component shifted to support meaningful use and patient-centered medical home activities. We do not

believe that this reason is associated with any of our measured outcomes. As such, the low completion rate does not pose a statistical concern of biased results.

A great deal of the CQC and provider relations staff focus was devoted to communicating the program changes to network PCP practices during the latter part of the year. This may have contributed to practices who had originally committed to the disparity pilot to withdraw; or they may have selected a different best practice study activity instead of ours. Also, some practices that we recruited for the disparity pilots were not due to start a new best practice until 2012 and they may have decided not to complete a best practice. This may also be due to structural features at practices that are associated with participation and may provide leverage points for (or for not) fostering participation as reported by others [19]. Future studies assessing barriers to self-enrolled participation completion rates among practices participating in P4P programs are warranted.

Lastly, this was a one-time assessment feedback about how to reduce health disparities among minority populations using a P4P approach. Whether health disparities can be further improved with continuous or repeated cyclical feedback from practices is unknown, but is worthy of further study.

Conclusion

This was a pilot study of participating networks utilizing a P4P approach in the Pennsylvania counties in Highmark's service region. The results of this study showed that there were significant improvements in the selected Best Practice measures. Of the 19 specified interventions, 13 were statistically significant at $\alpha=0.05$ level and 14 met the target proportion. The most impactful interventions were those related to education, discussions, and patient medical documentations rather than those pertaining to medication adherence. Improvement in measuring reporting and recording of data at post-intervention were also observed.

Acknowledgment The authors are grateful to our participating pay-for-performance (P4P) practices and former Highmark employee Christine Heasley that served our minority communities in Pennsylvania. This study was funded by Highmark Inc. Authors Rhonda Johnson, MD, MPH, Oralia Garcia Dominic, PhD, MS, MA, and Twyla Johnson, MPH, MBA, are employees of Health Equity and Quality Services at Highmark Inc. Oralia Garcia Dominic is a National Institutes of Health (NIH) National Cancer Institute (NCI) K01 grantee. Authors Gary Marsh, PhD, FACE, and Sarah D. Zimmerman, BS, MS, are Highmark Consultants. The authors thank Judah Abberbock, MS, for his technical support. Gary M. Marsh, PhD, FACE, declares that he has received

grants. He is a principal investigator of a sponsored research project at the University of Pittsburgh funded by Highmark, Inc. Oralia Garcia Dominic, PhD, MS, MA, declares that she has received grants from NIH, PA DOH, ACS. She is an NIH NCI K01 grantee and principal investigator of this sponsored research project at the Allegheny Singer Research Institute (ASRI) funded by NIH NCI CRCHD. She is an employee and the primary lead for health disparities and clinical interventions for Health Equity and Quality Services at Highmark IncCamp Hill, Pennsylvania.

Conflict of Interest Authors Rhonda M. Johnson, MD, MPH, Twyla Johnson, MPH, MBA, and Sarah D. Zimmerman, BS, MS, declare that they gave no conflict of interest.

Informed Consent The Allegheny Singer Research Institute's Institutional Review Board approved this study. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all participating practices/physicians participating in a pay-for-performance (P4P) program for being included in the study. No identifying information about patients is included in the article.

Animal Studies No animal or human studies were carried out by the authors for this article.

References

1. Reports: Agency for Healthcare Research and Quality 2011. National healthcare disparities report. Rockville, MD. <http://ahrq.gov/qual/nhdr11/key.htm>. Accessed 10 Apr 2014.
2. Reports: Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press, 2001.
3. Klein J, Schubiner L, Gadon M, Wynia M. Physicians' experience and opinions regarding strategies to improve care for minority patients. *JHRRP*. 2008;2(2):75–90.
4. Reports: Institute of Medicine. Unequal treatment: confronting racial & ethnic disparities in health care. National Academy of Sciences. <http://www.iom.edu/~media/Files/Report%20Files/2003/Unequal-Treatment-Confronting-Racial-and-Ethnic-Disparities-in-Health-Care/Disparitieshcproviders8pgFINAL.pdf>. Accessed 10 Apr 2014.
5. Reports: National Health Plan Collaborative. Toolkit to reduce racial and ethnic disparities in health care 2008. AHRQ Rockville, MD. <http://www.rwjf.org/files/research/nhpc toolkit.pdf>. Accessed 10 Apr 2014.
6. Ho K, Moy E, Clancy C. Can incentives to improve quality reduce disparities? *Health Serv Res*. 2010;45(1):1–5.
7. Alexander GC, Lin S, Sayla MA, Wynia MK. Development of measure of physician engagement in addressing racial and ethnic health care disparities. *Health Serv Res*. 2008;43(2):773–84.
8. American Academy of Family Physicians: pay-for-performance <http://www.aafp.org/online/en/home/policy/policies/p/payforperformance.html>. <http://www.aafp.org/about/policies/all/pay-performance.html> Accessed 10 Apr 2014.
9. Baker G, Carter B. The evolution of pay for performance models for rewarding providers. Case studies in health plan pay for performance (introduction). Atlantic Information Services: Washington, DC; 2004.
10. Epstein AM, Lee TH, Hamel MB. Paying physicians for high-quality care. *N Engl J Med*. 2004;350:406–10.
11. Christianson JB, Knutson DJ, Mazze RS. Physician pay-for-performance: implementation and research issues. *J Gen Intern Med*. 2006;21 Suppl 2:S9–13.

12. Casolino LP, Elster A. Will pay-for-performance and quality reporting affect health care disparities? *Health Aff.* 2007;26(3): w405–14.
13. Lurie N, Fremont A, Somers SA, Coltin K, Gelzer A, Johnson R, et al. The national health plan collaborative to reduce disparities and improve quality. *The Joint Commission Journal on Quality and Patient Safety.* 2008;34(5):256–65.
14. America's Health Insurance Plan (AHIP). Innovations in recognizing and rewarding quality. Recognizing Highmark Blue Cross Blue Shield–Quality BLUE Physician Pay for Performances Program. March 2009: 30–1.
15. Highmark's Quality Blue 2011 Achievement Compendium. Quality Blue Hospital Pay-for-Performance Program. Copyright Highmark Inc. 2011. Available at: www.highmark.com . Accessed 10 Apr 2014.
16. Pennsylvania Department of Health, Minority Health Disparities. <http://www.portal.state.pa.us/portal/server.pt?open=514&objID=596652&mode=2> . Accessed 10 Apr 2014.
17. Coleman K, Reiter KL, Fulwiler D. The impact of pay-for-performance on diabetes care in a large network of community health centers. *J Health Care Poor Underserved.* 2007;18(4):966–83.
18. Leon-Chisen N. Coding and quality reporting: resolving the discrepancies, finding opportunities. *Journal of AHIMA.* 2007;78(7):26–30.
19. Herald LR, Alexander JA, Shi Y, Casalino LP. Pay-for-Performance and public reporting program participation and administrative challenges among small- and medium-sized physician practices. *J Medical Care Research and Review.* 2014;71: 299–312.
20. Rees CA, Karter AJ, Young BA. Race/ethnicity, social support, and association with diabetes self-care and clinical outcomes in NHANES. *Diabetes Edu.* 2010:435–45.
21. Partner YHC. Press, quality blue. 2012. Available at: <https://www.highmark.com/hmk2/pdf/presskits/qualityblue2012.pdf> .; Accessed April 10, 2014.