

Primary Care Physician Resource Use Changes Associated With Feedback Reports

Eva Chang, PhD, MPH; Diana S.M. Buist, PhD, MPH; Matt Handley, MD; Eric Johnson, MS; Sharon Fuller, BA; Roy Pardee, JD, MA; Gabrielle Gundersen, MPH; and Robert J. Reid, MD, PhD

Fueled by concerns of rising healthcare costs and the national dialogue on overuse and misuse, healthcare systems are increasingly interested in developing better tools for measuring and reporting healthcare quality and efficiency.¹ Reducing variability in physician practice patterns is an opportunity for healthcare systems to improve quality and reduce costs.² Healthcare organizations are addressing variability in care with increased internal transparency in organizations to create a culture of high-value care.³⁻⁵ Numerous organizations, from Medicare and Medicaid to small practices, use audit and feedback reports to support physician behavior change and performance improvement.^{6,7} Although feedback reports have been shown to effectively increase accountability,⁸⁻¹⁰ the literature on implementing feedback reports and their impact on physician practice patterns in healthcare systems is inconclusive.^{8,11-16}

In 2012, Kaiser Permanente Washington (KPWA; formerly Group Health Cooperative) implemented a Resource Stewardship quality improvement (QI) initiative to reduce low-value care, or care that does not improve patient outcomes and can harm patients. The initiative focused on helping KPWA physicians become better stewards of healthcare resources. Recognizing that daily decisions drive quality and cost, an internally transparent peer-comparison feedback report was developed to bring personalized information to physicians that informed their clinical decision making and facilitated internal conversations about medical resource use.

The annual resource stewardship report showed individual physician-, clinic-, and system-level ordering rates for specialty referrals, high-end imaging, laboratory tests, and 30-day prescriptions. It was designed to show variations in global service use, because improving a few specific areas was deemed unlikely to be sufficient for reducing low-value care at the healthcare-system level. Feedback was combined across multiple services to help identify areas of greater overall resource use rather than targeting single areas of utilization that may result in unintended shifts in resource utilization (eg, reducing specialty referrals may cause increases in high-end imaging).

We evaluated whether distribution of internally transparent peer-comparison feedback reports within an integrated delivery system

ABSTRACT

OBJECTIVES: To measure changes in primary care physician (PCP) ordering rates for 4 global resource use measures before and after dissemination of physician feedback reports that provided peer-comparison resource use rates. We also explored whether physician practice characteristics (panel size, clinic size, and years of experience) were associated with resource use changes.

STUDY DESIGN: Pre-post implementation study measuring physician resource use in an integrated healthcare system (2011-2014).

METHODS: Kaiser Permanente Washington PCPs (N = 210) were provided annual feedback reports showing their personal ordering rates compared with those of their peers. Monthly physician ordering was measured from November 2011 to September 2014 (including prereport and postreport periods). We examined 4 physician ordering rates (specialty referrals, high-end imaging, laboratory tests, and 30-day prescriptions) per 1000 patients, adjusted for patient age, gender, and clinical complexity.

RESULTS: After accounting for physician practice characteristics, monthly PCP ordering rates for high-end imaging significantly decreased by 0.8 images per 1000 patients ($P < .01$). In contrast, orders for laboratory tests and 30-day prescriptions significantly increased by 15.0 tests and 84.7 prescriptions per 1000 patients (both $P < .01$). We observed greater changes following feedback in physicians with fewer years of experience (≤ 10 years), who had 4.2 fewer specialty referrals ($P = .01$) and 101.3 more 30-day prescriptions ($P < .01$) compared with those with more experience (> 20 years).

CONCLUSIONS: Physician feedback reports may be associated with changes in physician resource use, and physicians with fewer years of experience may be more responsive to feedback reports. Better understanding of factors associated with changes in resource use is necessary for future targeted development of physician interventions.

Am J Manag Care. 2018;24(10):455-461

TAKEAWAY POINTS

Healthcare systems use feedback reports to document physician practice patterns and increase physician accountability for resource use, but there are few reports of whether personalized resource use feedback affects provider performance. We evaluated the association between distribution of feedback reports to primary care physicians in an integrated delivery system and changes in 4 global measures of physician resource use.

- ▶ Feedback reports were associated with varied changes in resource use.
- ▶ Physicians with fewer years of experience may be more responsive to feedback reports.
- ▶ Better understanding of factors associated with changes in resource use is necessary for future targeted development of physician interventions.

was associated with changes in individual physician resource use. We also explored whether physician practice characteristics (panel size, clinic size, and years of experience) were associated with changes in resource use. Although evidence for the link between practice characteristics and resource use is mixed,¹⁷⁻²² better understanding of these characteristics could be used to target feedback reports to those who would benefit most.

METHODS

Setting and Data

KPWA is a mixed-model delivery system providing insurance and healthcare to approximately 710,000 patients in Washington state. Approximately 370,000 KPWA patients are cared for within the integrated delivery system (group practice); the remainder receive care through a network of contracted physicians and other healthcare providers across the state. Within the group practice, approximately 1000 salaried multispecialty clinicians practice in 25 medical clinics; approximately 300 are primary care physicians (PCPs). We included KPWA PCPs specializing in family medicine or internal medicine with a panel size of 250 or more patients during each month of the study period. This QI initiative and subsequent analyses were determined to be exempt from human subjects review in compliance with the Office for Human Research Protections (45 CFR Part 46). We followed healthcare QI reporting guidelines.²³

We extracted data on physician orders and characteristics from KPWA's automated clinical and administrative data systems. We used physician orders to accurately attribute and assess physicians' intended resource use. We created monthly measures of physician ordering from November 2011 to September 2014, including the 11 months before distribution of the reports (prereport period, November 2011-September 2012) and the 11 months after distribution (postreport period, November 2013-September 2014). We included a 13-month implementation period (October 2012-October 2013) during which 2 reports were distributed to each PCP.

Feedback (Resource Stewardship) Reports

Resource stewardship reports were designed to identify practice pattern variations and make resource use visible to KPWA's PCPs. Physicians were given clinic-specific reports that included ordering

rates and physician names for all PCPs in their clinic. Bar graphs of annual, physician-level, comorbidity-adjusted ordering rates in the prior calendar year for specialty referrals, high-end imaging, laboratory tests, and 30-day prescriptions per 1000 patients were displayed for all PCPs within the clinic (see [eAppendix A](#) [eAppendices available at [ajmc.com](#)] for a sample 2015 report; some reported ordering rates changed since the 2013 report). The graphs included lines to indicate the system or clinic average and 1 SD above and below the average. Bar graphs in the

second-year report (September 2013) also showed physicians' ordering rates in the prior year to allow physicians to see personal changes over time. Reports included physician names to promote within-system transparency and give physicians the opportunity to learn from the practice patterns of other physicians within their clinic.

Feedback reports were disseminated in November 2012 and September 2013 through monthly within-clinic physician meetings and individual meetings between physicians and their local clinic chiefs. The Medical Director for Quality (M.H.) and clinic chiefs explained the Resource Stewardship initiative and the report's purpose, presented the data, and gave physicians opportunities to discuss intraclinic variability as a group. Clinic chiefs were trained to present the reports as a chance for physicians to learn from each other rather than as a punitive tool.

Measures

We constructed 4 physician ordering measures reported in the feedback reports: specialty referrals, high-end imaging, laboratory tests, and 30-day prescriptions. Specialty referrals included all referral orders made by the physician. High-end imaging orders included computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography. Laboratory tests and prescriptions included all laboratory tests and prescriptions ordered; prescriptions were standardized to a 30-day supply. Rates were calculated as per 1000 paneled patients. Nearly all KPWA patients within the group practice select a PCP and are empaneled to the physician's care team; care teams are accountable for the care coordination of all empaneled patients.²⁴ Given the focus on physician stewardship and accountability, patients who were not paneled to a PCP were excluded. We adjusted ordering rates by the panel's average Adjusted Clinical Group (ACG) score to account for diagnosis mix and comorbidity burden. The ACG system is a widely used patient case-mix-adjustment system for populations receiving outpatient services.²⁵

For each physician, we extracted 3 physician practice characteristics hypothesized to be associated with physician resource utilization: panel size, clinic size (number of physicians in the clinic), and years of experience (years since graduation from medical school).¹⁷⁻²¹ Median panel size was calculated using the postreport period and adjusted for full-time equivalency. We also collected physician characteristics (gender, race [white, nonwhite]) for control variables.

Statistical Analysis

We calculated descriptive statistics for physicians and assessed variability across physician orders using the coefficient of variation (CV; SD divided by mean). All models to test differences in orders were performed at the physician-month level with a repeated measures random effects model, using an identity link and Gaussian error term with robust standard errors. Dummy variables for each calendar month were included to control for variations in average utilization between months. We controlled for correlation due to repeated measures within physicians and clustered physicians within clinics. The model specifications allowed us to estimate the average resource utilization rate using marginal standardization over both the prereport and postreport time periods. We tested for differences in average resource utilization rates between the prereport and postreport time periods. We then separately tested interactions between each of the physician practice characteristics with the time periods to identify physician practice characteristics associated with changes in resource use. All models controlled for physician gender and race; age was highly correlated with years of experience and thus was not included. We did not control for multiple comparisons; instead, we chose to present results from all analyses that were carried out. All analyses were conducted with Stata 13.0 statistical software (StataCorp; College Station, Texas). All hypothesis tests were performed using 2-sided $\alpha = .05$.

RESULTS

Sample Population

Our study included 210 PCPs (202 family physicians and 8 general internists) (eAppendix B). Physicians were balanced on gender (54% female) and were predominantly white (72%) (Table 1). Forty-four percent had panels with 1400 to 1700 patients, 46% had practiced for more than 20 years, and 51% worked in a large multispecialty clinic with more than 20 other physicians. Most worked less than full time; mean physician full-time equivalency was 0.73 (SD = 0.18; data not shown).

Trends in Observed Orders

Medical resource use over time is illustrated in the Figure. Over the entire study period, the CV of case-mix-adjusted mean monthly rates was highest for imaging (CV, 65%; mean = 5.0 images per 1000 patients; range, 4.2-6.1) and lowest for 30-day prescriptions (CV, 19%; mean = 1049.4 prescriptions per 1000 patients; range, 942.0-1128.4). Specialty referrals and laboratory tests had similar CVs of 38% (mean = 47.5 referrals per 1000 patients; range, 38.7-54.2) and 35% (mean = 289.1 tests per 1000 patients; range, 241.3-423.5), respectively. Prereport and postreport period CVs were consistent with the CV over the entire study period and with each other (Table 2).

Differences in Physician Orders

Regression results for case-mix-adjusted orders of specialty referrals, high-end imaging, laboratory tests, and 30-day prescriptions

TABLE 1. Baseline Characteristics of the Study Population (N = 210), Kaiser Permanente Washington

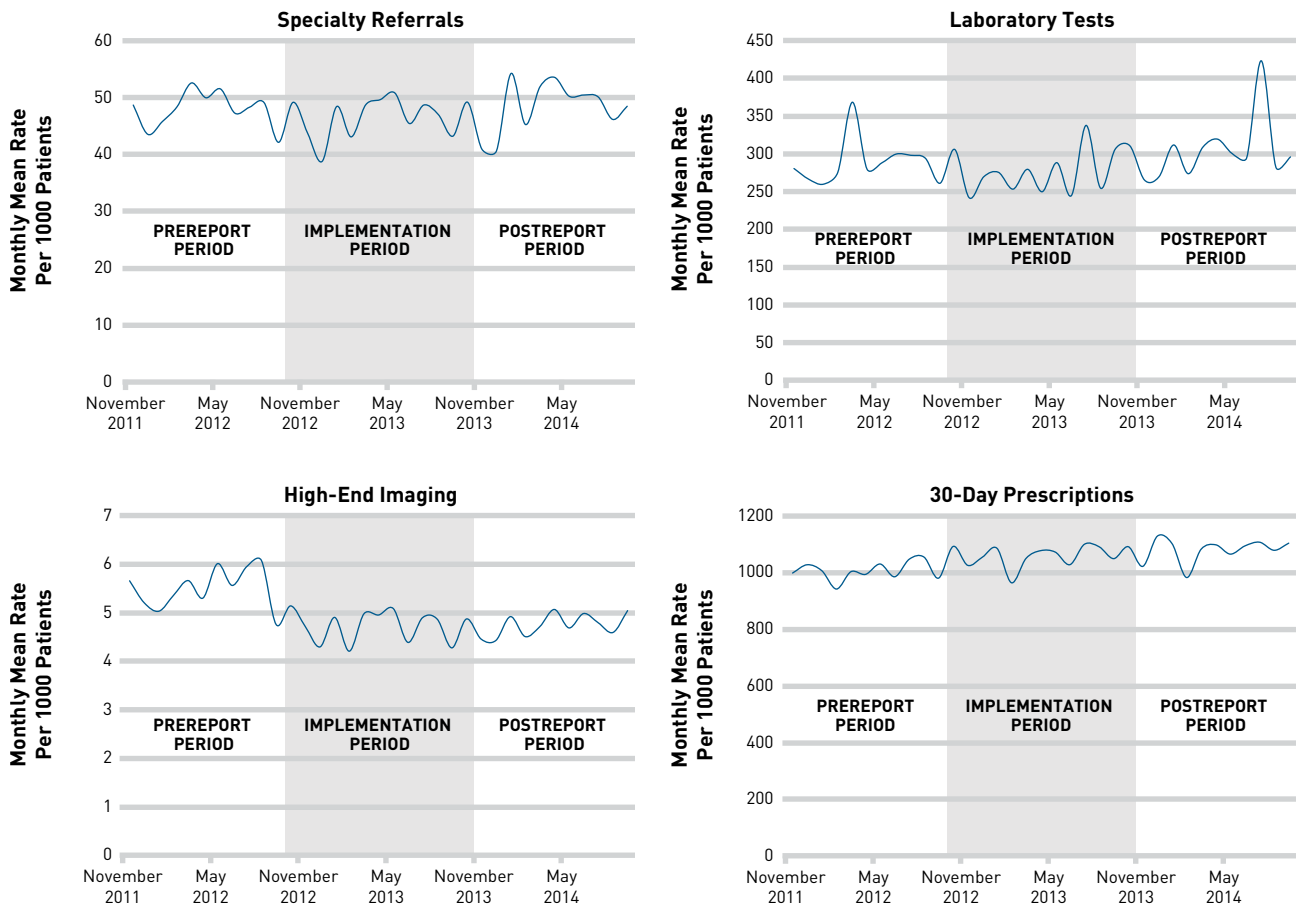
Characteristics	Physicians n (%)
Physician Characteristics	
Gender	
Female	113 (54)
Race	
White	151 (72)
Physician Practice Characteristics	
Panel size (no. of patients)	
≤1400	57 (27)
1401-1700	93 (44)
>1700	59 (28)
Years of experience	
≤10	48 (23)
11-20	60 (29)
>20	97 (46)
Clinic size (no. of physicians)	
≤10	22 (10)
11-20	78 (37)
>20	108 (51)

account for physician characteristics, physician practice characteristics, and calendar month (Table 2). Compared with the prereport period, monthly physician orders for high-end imaging decreased by 0.8 images per 1000 patients, which corresponds to a -14.5% (95% CI, -20.0% to -9.1%) change. Orders of laboratory tests and 30-day prescriptions increased by 15.0 tests (5.2% change; 95% CI, 2.6%-7.8%) and 84.7 prescriptions (8.4% change; 95% CI, 5.7%-11.1%), respectively, per 1000 patients. No statistically significant change in specialty referral orders was observed.

Physician Practice Characteristics Associated With Change in Physician Orders

Although specialty referral orders showed no overall change, stratification by years of experience revealed statistically significant changes among physicians; physicians with 10 or fewer years of experience were associated with a relatively greater decrease in referral orders (-4.2 referrals per 1000 patients; -7.7% change; 95% CI, -13.4% to -1.8%) compared with physicians with more than 20 years of experience (Table 3). Having fewer years of experience was associated with a relatively greater increase in prescription orders (86.0 orders for physicians with 11-20 years [8.8% change; 95% CI, 2.6%-15.0%] and 101.3 for physicians with ≤10 years [11.7% change; 95% CI, 4.2%-19.2%] compared with physicians with >20 years). Practicing in a clinic with 11 to 20 physicians was also associated

FIGURE. Kaiser Permanente Washington Physician Orders of Specialty Referrals, Laboratory Tests, High-End Imaging, and 30-Day Prescriptions (per 1000 paneled patients) Over Time, November 2011–September 2014



with a relatively greater increase in prescription orders (46.7 orders; 2.6% change; 95% CI, 0.1%–9.1%) compared with clinics with more than 20 physicians. No practice characteristics were associated with statistically significant changes in imaging or laboratory tests.

DISCUSSION

Distribution of internally transparent peer-comparison resource use feedback reports at KPWA was associated with changes in physician resource use, specifically reductions in high-end imaging orders and increases in laboratory test and 30-day prescription orders. Years of experience were associated with changes in practice patterns, with greater changes in some resource use among less experienced physicians.

This study is an important example of using physician ordering data to evaluate changes over time following physician interventions. Compared with the claims and survey data normally used to assess utilization, physician order data may more accurately reflect intended

use rather than realized use (as measured in claims) or perceived use (as measured in surveys). Claims data may be confounded by patient behaviors, whereas survey data may be confounded by response biases. We previously reported physicians' recollection and experience receiving resource stewardship reports.²⁶ Although we could not directly measure who received reports and associated discussions, nearly four-fifths recalled receiving the report from their clinic chief. Three-quarters of these respondents participated in a discussion about their individual and group reports, and the majority found them useful for understanding their own practice patterns and provoking discussion among their peers. Notably, although more than three-fifths of all respondents reported changing their practice in each of the 4 measures, we observed important changes in 3 of the 4 measures in this study. The ability to examine physician orders through electronic health record systems allows healthcare organizations to measure their physicians' intended resource use.

Based on current best practices for effective feedback reporting systems and our survey findings,^{9,26–28} we believe that clearer specific

performance targets and more timely feedback with greater frequency could improve the effectiveness of our feedback reports. Like previous interventions with physician feedback reports and resource use,^{8,11,12} the direction of the change in resource use after receipt of the report varied by the specific measure. How feedback was presented to physicians may help explain the inconsistent direction of change in orders. Although KPWA's and clinics' average rates and peer physician rates were intended to be viewed as comparators, the feedback reports did not include explicit target goals or benchmarks to guide how physicians should change specific behaviors. Verstappen et al found that feedback reports reduced the use of some low-value services among PCPs.¹⁵ However, given the challenges of using utilization data to identify appropriate and inappropriate care and systemwide interest in addressing variation in global service use,²⁹ these reports only provided overall resource use. It is likely that higher resource users have greater appropriate and inappropriate use. Physicians also could have been influenced to reduce resource utilization by other sources. For example, KPWA's promotion of safe reduction of CT and MRI scans since 2009 and external campaigns to reduce imaging overuse may have reinforced the message to reduce overall utilization of high-end imaging.³⁰ Similarly, KPWA physicians were sent weekly "clinical pearls" emails that occasionally highlighted low-value care themes that targeted specific conditions and therapies. Conversely, low resource users may have increased their ordering patterns after seeing how they compared with their peers. Use of global ordering rates limited our ability to ascertain how feedback reports impacted inappropriate care compared with appropriate care, and further research is needed to decompose these effects.

The frequency and timeliness of feedback data have been found to be associated with more effective feedback reports.^{4,8,9} Our feedback reports used annual data from the previous calendar year; the lag in the data (11 and 9 months for the first and second reports, respectively) may have reduced the reports' effectiveness because physicians were unable to link their current performance with the data reported. Despite known best practices

TABLE 2. Adjusted Changes in Mean Physician Ordering Rate Per 1000 Panelled Patients Before and After Feedback Reporting

Physician Ordering Measures	Unadjusted Rates				Statistical Comparison of Postreport Period vs Prereport Period ^a	
	Prereport Period ^b		Postreport Period ^b		Change (95% CI)	P
	Mean (SD)	CV	Mean (SD)	CV		
Specialty referrals	47.9 (20.2)	42%	48.3 (17.3)	36%	0.0 (-1.8 to 1.7)	.98
High-end imaging	5.5 (3.6)	65%	4.7 (3.1)	66%	-0.8 (-1.1 to -0.5)	<.01
Laboratory tests	288.2 (111.8)	39%	303.5 (96.0)	32%	15.0 (7.5-22.6)	<.01
30-day prescriptions	1006.3 (214.3)	21%	1078.0 (192.8)	18%	84.7 (57.8-111.5)	<.01

CV indicates coefficient of variation.

^aValues are adjusted for sex, race, panel size, years of experience, and clinic size.

^bPrereport period estimates are from November 2011 to September 2012, and postreport period estimates are from November 2013 to September 2014. The CVs (SD divided by mean) are shown for the monthly case-mix-adjusted resource use.

TABLE 3. Adjusted Changes in Mean Physician Ordering Rate Per 1000 Panelled Patients Before and After Feedback Reporting, by Physician Characteristics

Physician Characteristic	Unadjusted Rates Mean (SD)		Statistical Comparison of Postreport Period vs Prereport Period ^a	
	Prereport Period ^b	Postreport Period ^b	Change (95% CI)	P
Specialty Referrals				
Panel size (no. of patients)				
≤1400	47.0 (19.4)	47.0 (17.2)	1.5 (-2.0 to 5.0)	.39
1401-1700	49.2 (21.9)	50.7 (17.9)	2.0 (-1.2 to 5.2)	.22
>1700	46.9 (18.2)	45.8 (15.8)	Reference	
Years of experience				
≤10	54.8 (18.9)	52.7 (18.1)	-4.2 (-7.4 to -1.0)	.01
11-20	51.5 (23.3)	51.0 (16.3)	-2.2 (-5.4 to 0.9)	.17
>20	42.7 (17.4)	44.2 (16.3)	Reference	
Clinic size (no. of physicians)				
≤10	51.8 (21.5)	51.3 (19.8)	-1.0 (-6.5 to 4.6)	.73
11-20	48.2 (22.1)	47.6 (17.2)	-2.6 (-5.9 to 0.7)	.12
>20	46.8 (18.5)	48.2 (16.8)	Reference	
High-End Imaging				
Panel size (no. of patients)				
≤1400	5.9 (3.7)	5.2 (3.2)	-0.1 (-0.7 to 0.5)	.84
1401-1700	5.6 (3.8)	4.8 (3.3)	-0.1 (-0.8 to 0.7)	.86
>1700	4.9 (3.0)	4.2 (2.8)	Reference	
Years of experience				
≤10	5.5 (3.3)	4.9 (3.2)	0.0 (-0.7 to 0.6)	.90
11-20	5.9 (4.1)	4.8 (3.5)	-0.4 (-1.0 to 0.1)	.10
>20	5.3 (3.3)	4.6 (2.9)	Reference	
Clinic size (no. of physicians)				
≤10	5.9 (4.5)	5.4 (3.9)	-0.3 (-1.1 to 0.5)	.53
11-20	5.7 (3.5)	4.6 (3.0)	-0.6 (-1.1 to 0.0)	.06
>20	5.3 (3.4)	4.7 (3.1)	Reference	

(continued)

CLINICAL

TABLE 3. (Continued) Adjusted Changes in Mean Physician Ordering Rate Per 1000 Panelled Patients Before and After Feedback Reporting, by Physician Characteristics

Physician Characteristic	Unadjusted Rates Mean (SD)		Statistical Comparison of Postreport Period vs Prereport Period ^a	
	Prereport Period ^b	Postreport Period ^b	Change (95% CI)	P
Laboratory Tests				
Panel size (no. of patients)				
≤1400	301.5 (122.8)	309.2 (105.1)	-7.9 [-26.5 to 10.6]	.40
1401-1700	291.4 (109.8)	310.0 (94.6)	0.8 [-13.3 to 14.9]	.91
>1700	271.0 (102.8)	288.2 (87.7)	Reference	
Years of experience				
≤10	290.2 (102.5)	298.6 (84.8)	-14.3 [-32.5 to 3.9]	.12
11-20	306.7 (121.2)	318.0 (94.3)	-10.5 [-22.9 to 1.9]	.10
>20	277.0 (109.5)	297.5 (102.1)	Reference	
Clinic size (no. of physicians)				
≤10	365.4 (157.5)	365.1 (135.2)	-11.5 [-38.8 to 15.9]	.41
11-20	297.9 (108.6)	313.1 (94.6)	-2.0 [-17.8 to 13.9]	.81
>20	267.1 (94.4)	284.9 (80.3)	Reference	
30-Day Prescriptions				
Panel size (no. of patients)				
≤1400	1055.6 (246.8)	1109.4 (204.2)	-7.1 [-57.1 to 43.0]	.78
1401-1700	983.4 (194.8)	1072.0 (197.5)	25.2 [-7.2 to 57.7]	.13
>1700	987.7 (194.5)	1059.0 (170.3)	Reference	
Years of experience				
≤10	865.5 (182.0)	998.1 (183.2)	101.3 [36.2-166.4]	<.01
11-20	978.3 (223.5)	1089.4 (153.5)	86.0 [25.1-146.8]	<.01
>20	1088.3 (183.5)	1117.3 (208.1)	Reference	
Clinic size (no. of physicians)				
≤10	1079.7 (243.3)	1146.0 (192.5)	12.5 [-118.5 to 143.5]	.85
11-20	1013.6 (223.9)	1093.3 (209.9)	46.7 [1.5-91.9]	.04
>20	993.0 (193.9)	1054.3 (166.1)	Reference	

^aValues are adjusted for sex, race, panel size, years of experience, and clinic size.

^bPrereport period estimates are from November 2011 to September 2012, and postreport period estimates are from November 2013 to September 2014.

suggesting greater frequency and timeliness of producing feedback reports,^{27,28,31} we were restricted by the lack of dedicated resources. Reporting intervals likely need to be shorter because more than 20% of surveyed KPWA physicians did not remember having the report shared with them.²⁶ Current recommendations suggest providing feedback in shorter intervals (eg, monthly) so that physicians have sufficient time to change their behavior.^{27,28,31}

Our exploratory examination of physician practice characteristics and changes in resource use found significantly greater changes in utilization among physicians with fewer years in practice. Having fewer years of experience has been associated with greater resource use (eg, higher cost, greater prescribing and imaging), which is hypothesized to be due to greater exposure to newer and costlier care, lack of experience, or lack of established practice patterns.¹⁷⁻¹⁹ Similarly, physicians with less experience may be more flexible in

their practice patterns and may have changed their practice patterns to be more comparable with patterns observed in more experienced physicians. This finding, coupled with a high percentage of physicians reporting conversations about their practice patterns,²⁶ may suggest that internally transparent reports may be prompting physicians with less experience to learn from their colleagues with more experience. These conversations among physicians may be a key element of creating a culture and environment of continuous learning that promotes high-value care.³⁻⁵ Conversely, feedback reports appeared to be less effective in modifying ordering rates among physicians with more experience. More active and multifaceted interventions, such as outreach programs and workshops combined with feedback reports, may be needed to change their ordering practices.³²

Limitations

We evaluated healthcare resource utilization patterns following a systemwide feedback initiative designed to bring better information to physicians by informing their clinical decision making on the global use of resources. Most importantly, we did not have a control group to evaluate whether observed changes were reflective of historical trends or important unmeasured confounders. For example, high-end imaging rates have been slowing down nationally during this time.^{33,34} KPWA physicians are salaried in an integrated delivery system, which may limit the generalizability of our findings because working in a capitated payment system may influence physician practice patterns.

Notably, KPWA had concurrent components of this initiative and other QI interventions that may have been potential confounders for the ordering measures. The Resource Stewardship initiative included 2 additional components to promote reducing low-value care. The first component was an organization-wide effort to promote shared decision making and the second focused on evidence-based clinical improvement, in the spirit of the Choosing Wisely campaign.³⁵ In the first, physicians were offered training in multiple shared decision-making skills, including discussing low-value care with patients. In the second, as noted previously, KPWA sent clinical emails that highlighted Choosing Wisely themes and implemented a prostate-specific antigen (PSA)/Pap smear trigger tool that electronically identified inappropriate Pap tests in women and PSA tests in elderly men from June 2013 to May 2014. During the study period, 3 emails (1 each for Pap tests and PSA tests and

1 introducing Choosing Wisely lists) targeted PCPs, whereas other Choosing Wisely-themed emails targeted specialists. Given the focus on reducing Pap tests and PSA tests, the third component of the initiative had the greatest likelihood of impacting changes in laboratory tests. This suggests that our finding of an increase in laboratory tests may be conservative. Additionally, high-end imaging had the most robust concurrent QI initiative (described previously), and additional data from the feedback reports may have accelerated ongoing improvements. Other concurrent initiatives (numerous pharmacy initiatives to promote use of preferred agents and KPWA's primary care medical home transformation) were less likely to change global utilization rates.³⁶ Despite these limitations, we used a robust pre-post design to assess the impact of implementing feedback reports in a large primary care setting.

CONCLUSIONS

Implementation of systemwide dissemination of feedback reports across multiple medical resources was associated with modest changes in ordering behavior among PCPs. Changing physician ordering behavior through greater transparency of physician behavior may help healthcare systems move toward creating a culture of high-value care. ■

Author Affiliations: RTI International (EC), Waltham, MA; Kaiser Permanente Washington Research Institute (DSMB, EJ, SF, RP, GG), Seattle, WA; Washington Permanente Medical Group (MH), Seattle, WA; Trillium Health Partners, Institute for Better Health (RJR), Mississauga, ON, Canada.

Source of Funding: This work was funded by a Partnership for Innovation grant from the Group Health Foundation.

Prior Presentation: The findings in this article were presented at the 2015 HMO Research Network Conference, Long Beach, CA, and the 2015 AcademyHealth Annual Research Meeting, Boston, MA.

Author Disclosures: Dr Buist has received Robert Wood Johnson grant funding on reducing low-value care and has attended the V-BID Health Task Force on Low-Value Care. The remaining authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (EC, DSMB, MH, RP, RJR); acquisition of data (DSMB, SF, RP); analysis and interpretation of data (EC, DSMB, EJ, SF, RP, RJR); drafting of the manuscript (EC, DSMB, MH, EJ); critical revision of the manuscript for important intellectual content (EC, DSMB, MH, GG, RJR); statistical analysis (EC, DSMB, EJ); provision of patients or study materials (DSMB, GG); obtaining funding (DSMB, MH); administrative, technical, or logistic support (DSMB, SF, GG); and supervision (DSMB, MH).

Address Correspondence to: Eva Chang, PhD, MPH, RTI International, 307 Waverly Oaks Rd, Ste 101, Waltham, MA 02452. Email: echang@rti.org.

REFERENCES

- Lee VS, Kawamoto K, Hess R, et al. Implementation of a value-driven outcomes program to identify high variability in clinical costs and outcomes and association with reduced cost and improved quality. *JAMA*. 2016;316(10):1061-1072. doi: 10.1001/jama.2016.12226.
- Institute of Medicine. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington, DC: National Academies Press; 2013.
- Cosgrove D, Fisher M, Gabow P, et al. A CEO checklist for high-value health care. National Academy of Medicine website. nam.edu/wp-content/uploads/2015/06/CEOHIGHValueChecklist.pdf. Published June 5, 2012. Accessed October 4, 2016.
- Parchman ML, Henrikson NB, Blasi PR, et al. Taking action on overuse: creating the culture for change. *Healthc (Amst)*. 2017;5(4):199-203. doi: 10.1016/j.hidsi.2016.10.005.
- Taking Action on Overuse website. takingactiononoveruse.org. Published 2017. Accessed September 30, 2017.

- Kahi CJ, Ballard D, Shah AS, Mears R, Johnson CS. Impact of a quarterly report card on colonoscopy quality measures. *Gastrointest Endosc*. 2013;77(6):925-931. doi: 10.1016/j.gie.2013.01.012.
- VanLare JM, Blum JD, Conway PH. Linking performance with payment: implementing the Physician Value-Based Payment Modifier. *JAMA*. 2012;308(20):2089-2090. doi: 10.1001/jama.2012.14834.
- Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 2012;(6):CD000259. doi: 10.1002/14651858.CD000259.pub3.
- Hysong SJ. Meta-analysis: audit and feedback features impact effectiveness on quality. *Med Care*. 2009;47(3):356-363. doi: 10.1097/MLR.0b013e3181893f6b.
- Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Does telling people what they have been doing change what they do? a systematic review of the effects of audit and feedback. *Qual Saf Health Care*. 2006;15(6):433-436. doi: 10.1136/qshc.2006.018549.
- Jain S, Frank G, McCormick K, Wu B, Johnson BA. Impact of physician scorecards on emergency department resource use, quality, and efficiency. *Pediatrics*. 2015;136(3):e670-e679. doi: 10.1542/peds.2014-2363.
- Tavarez MM, Ayers B, Jeong JH, Coombs CM, Thompson A, Hickey RW. Practice variation and effects of e-mail-only performance feedback on resource use in the emergency department. *Acad Emerg Med*. 2017;24(8):948-956. doi: 10.1111/acem.13211.
- Myakis S, Karamanof G, Lontos M, Mountokalakis TD. Factors contributing to inappropriate ordering of tests in an academic medical department and the effect of an educational feedback strategy. *Postgrad Med J*. 2006;82(974):823-829. doi: 10.1136/pgmj.2006.049551.
- Chinnaiyan KM, Peyser P, Goraya T, et al. Impact of a continuous quality improvement initiative on appropriate use of coronary computed tomography angiography: results from a multicenter, statewide registry, the Advanced Cardiovascular Imaging Consortium. *J Am Coll Cardiol*. 2012;60(13):1185-1191. doi: 10.1016/j.jacc.2012.06.008.
- Verstappen WH, van der Weijden T, Sijbrandij J, et al. Effect of a practice-based strategy on test ordering performance of primary care physicians: a randomized trial. *JAMA*. 2003;289(18):2407-2412. doi: 10.1001/jama.289.18.2407.
- Wong JH, Lubkey TB, Suarez-Almazor ME, Findlay JM. Improving the appropriateness of carotid endarterectomy: results of a prospective city-wide study. *Stroke*. 1999;30(1):12-15.
- Mehrotra A, Reid RD, Adams JL, Friedberg MW, McGlynn EA, Hussey PS. Physicians with the least experience have higher cost profiles than do physicians with the most experience. *Health Aff (Millwood)*. 2012;31(11):2453-2463. doi: 10.1377/hlthaff.2011.0252.
- Hartley RM, Charlton JR, Harris CM, Jarman B. Patterns of physicians' use of medical resources in ambulatory settings. *Am J Public Health*. 1987;77(5):565-567.
- Tamblyn R, McLeod P, Hanley JA, Girard N, Hurlay J. Physician and practice characteristics associated with the early utilization of new prescription drugs [erratum in *Med Care*. 2003;41(10):1177]. *Med Care*. 2003;41(8):895-908. doi: 10.1097/01.MLR.0000078145.41828.3E.
- Eisenberg JM, Nicklin D. Use of diagnostic services by physicians in community practice. *Med Care*. 1981;19(3):297-309.
- Sistrom C, McKay NL, Weilburg JB, Atlas SJ, Ferris TG. Determinants of diagnostic imaging utilization in primary care. *Am J Manag Care*. 2012;18(4):e135-e144.
- Ping NK, Wei Ling NC. The effects of practice size on quality of care in primary care settings: a systematic review. *JBI Libr Syst Rev*. 2012;10(27):1549-1633.
- Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for Quality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. *Perm J*. 2015;19(4):65-70. doi: 10.7812/TPP/15-141.
- Neuwirth EEB, Schmittiel DJ, Tallman K, Bellows J. Understanding panel management: a comparative study of an emerging approach to population care. *Perm J*. 2007;11(3):12-20.
- Weiner JP, Starfield BH, Steinwachs DM, Mumford LM. Development and application of a population-oriented measure of ambulatory care case-mix. *Med Care*. 1991;29(5):452-472.
- Buist DS, Chang E, Handley M, et al. Primary care clinicians' perspectives on reducing low-value care in an integrated delivery system. *Perm J*. 2016;20(1):41-46. doi: 10.7812/TPP/15-086.
- McNamara P, Shaller D, De La Mare J, Ivers NM. *Confidential Physician Feedback Reports: Designing for Optimal Impact on Performance*. Rockville, MD: Agency for Healthcare Research and Quality; 2016. ahrq.gov/sites/default/files/publications/files/confidreportguide_0.pdf. Accessed June 3, 2016.
- Ivers NM, Grimshaw JM, Jamtvedt G, et al. Growing literature, stagnant science? systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. *J Gen Intern Med*. 2014;29(11):1534-1541. doi: 10.1007/s11606-014-2913-y.
- Buist D, Collado M. Promoting the appropriate use of health care services: research and policy priorities. AcademyHealth website. academyhealth.org/sites/default/files/publications/files/HealthCareResourceUse/ResourceUseIssueBrief2014.pdf. Published July 2014. Accessed October 19, 2016.
- Rao VM, Levin DC. The overuse of diagnostic imaging and the Choosing Wisely initiative. *Ann Intern Med*. 2012;157(8):574-576. doi: 10.7326/0003-4819-157-8-201210160-00535.
- Hysong SJ, Best RG, Pugh JA. Audit and feedback and clinical practice guideline adherence: making feedback actionable. *Implement Sci*. 2006;1:9. doi: 10.1186/1748-5908-1-9.
- Mostofian F, Ruban C, Simunovic N, Bhandari M. Changing physician behavior: what works? *Am J Manag Care*. 2015;21(1):75-84.
- Levin DC, Rao VM, Parker L, Frangos AJ, Sunshine JH. Bending the curve: the recent marked slowdown in growth of noninvasive diagnostic imaging. *Am J Roentgenol*. 2011;196(1):W25-W29. doi: 10.2214/AJR.10.4835.
- Lang K, Huang H, Lee DW, Federico V, Menzin J. National trends in advanced outpatient diagnostic imaging utilization: an analysis of the Medical Expenditure Panel Survey, 2000-2009. *BMC Med Imaging*. 2013;13:40. doi: 10.1186/1471-2342-13-40.
- Good Stewardship Working Group. The "top 5" lists in primary care: meeting the responsibility of professionalism. *Arch Intern Med*. 2011;171(15):1385-1390. doi: 10.1001/archinternmed.2011.231.
- Reid RJ, Johnson EA, Hsu C, et al. Spreading a medical home redesign: effects on emergency department use and hospital admissions. *Ann Fam Med*. 2013;11(suppl 1):S19-S26. doi: 10.1370/afm.1476.

Visit ajmc.com/link/3217 to download PDF and eAppendix

Resource Stewardship – Primary Care

Introduction to 2015 Dashboard



The Golden Mean

Aristotle's concept of the Golden Mean is "the ideal balance between two extremes, one of excess and the other of deficiency". The Golden Mean isn't the average between the two extremes, but the desirable middle that depends on the situation. For example, Aristotle wrote that courage is a virtue, but if taken to excess it would manifest as recklessness, and, in deficiency, cowardice.

In health care, we are constantly striving to find the golden mean. Every day we make tough choices about what tests and treatments our patients need. Finding the ideal balance of providing the right amount of care—not too little and not too much—is an ongoing challenge for all clinicians.

The purpose of this report is to help you understand how your use of new specialty visits, referrals, chem panels, and imaging compare with other providers in your clinic. The variation we see in use of these services can only be explained by clinician preferences for providing more or less care, or the "propensity to act". That variation drives unnecessary utilization and higher costs. The closer we get to achieving the right balance, the more our patients and Group Health benefit.

How to use the dashboard

We ask you to review these data and discuss them with your colleagues. Most of us are likely to see comparison data that make us uncomfortable. Our hope is that we will support and learn from each other to understand why we make the choices we do. In many ways, the conversation is more important than the data. Facing tough questions and wrestling with doing what's right is what makes up the fabric of our culture. We are each other's greatest resource.

What's new in this edition?

Chem panels replace lab tests in previous reports. Ordering only the needed individual tests is better for patient care than routinely ordering panels. Panels will often show abnormal results that are clinically insignificant, prompting unnecessary follow up.

Glargine prescribing patterns. The vast majority of patients with Type 2 diabetes who need basal insulin do just as well on NPH as on glargine. Glargine insulin is more than 10 times more expensive than NPH insulin, yet randomized trials comparing NPH and glargine show they have almost identical outcomes. Glargine has only been shown to be more effective in a few uncommon situations, such as patients with Type 1 diabetes or insulin deficient Type 2 diabetes who need both basal insulin and pre-meal rapid-acting insulin. In most cases, use of Glargine does not result in better outcomes, only a higher financial burden on our patients.

"Early" pap tests for cervical cancer screening, and antibiotic use for URI. We have achieved significant reductions in overuse in these areas and continue to improve. Under our Choosing Wisely® grant, we are collaborating with Swedish, WSMA and the WHA to help others in the state make similar improvements.

We welcome your feedback and suggestions. Please send them to:

Medical Director, Group Practice Division

Medical Director, Quality

Resource Stewardship

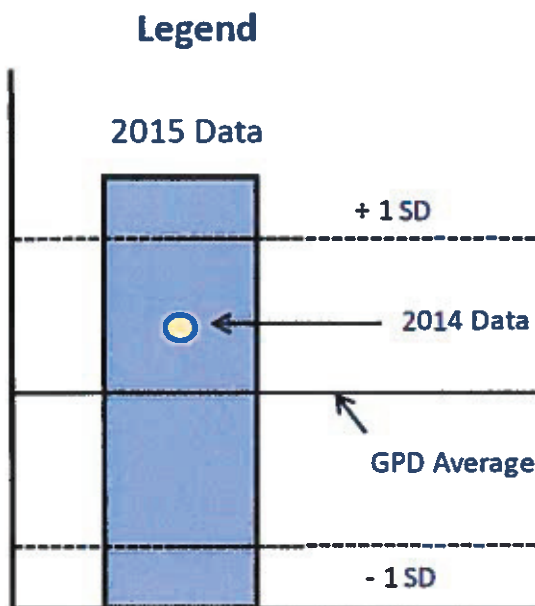
2015 Dashboard – Primary Care



GroupHealth.

Utilization and cost measures are adjusted imperfectly - utilization rates and costs use a denominator of DxCG adjusted panel size (per 1000 DxCG adjusted patients). We recognize that this does not capture all of the complexity between patients, but does present a good faith adjustment. We can never risk adjust perfectly.

In most clinics and measures, the variation in utilization and cost is great enough that it cannot be accounted for by differences in clinical circumstances or patients preferences. A significant factor in the variation is our differing practice patterns/styles. These data are presented to help prompt reflection and conversations on cost effective practice - we are each other's greatest resource.

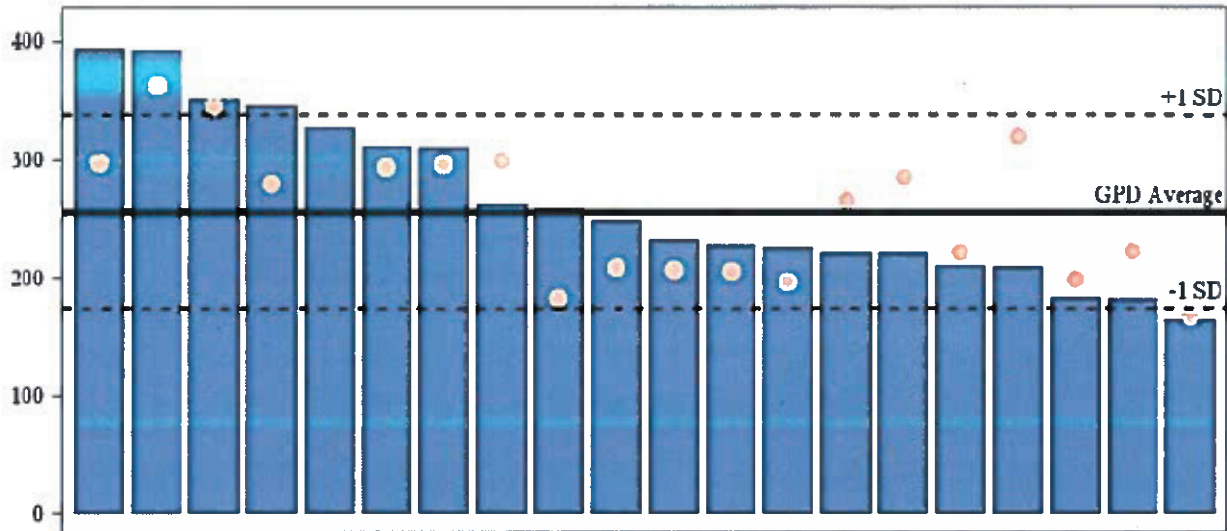


- DxCG is a tool for risk adjustment. For Resource Stewardship, we use a different version of DxCG than you generally see. The DxCG score used for panel size adjustment is a prospective score, which predicts future usage. The version of DxCG we use for Resource Stewardship is a concurrent score, which maps better to current utilization. As a result, the value you see for DxCG will likely be different than DxCG values you may have seen associated with your panel in other contexts.
- DxCG scores have a “floor” value of 1, so low scores (young healthy people) do not distort the data (we do not divide by a very low number, amplifying utilization). We applied this methodology to the 2014 and 2015 data.
- Note that variation in DxCG only accounts for about 40% of the variation in utilization and cost in a population.
- Data are not presented for clinicians with panel sizes smaller than 500.
- The clinic average is used for the referral comparison, not GPD average.
- If you don't have 2014 data reported, the following are the most likely reasons:
 - You are a new provider to GH as of 2015
 - Your 2014 panel size was less than 500

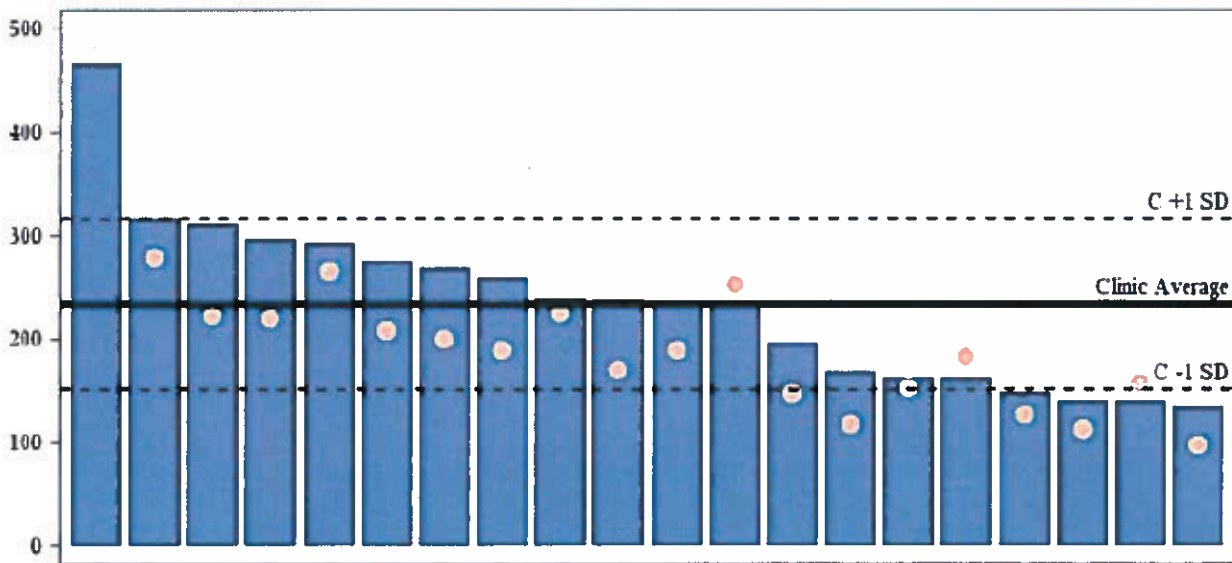
Resource Stewardship Dashboard PRIMARY CARE FP-IM

Clinic: XXXXXXXXXX

Rate of NEW SPECIALTY VISITS for your panel - measure of total impact of your panel in specialty (includes self-referral and referrals you and your partners made for your patients)
DxCG Adjusted per 1000 patients



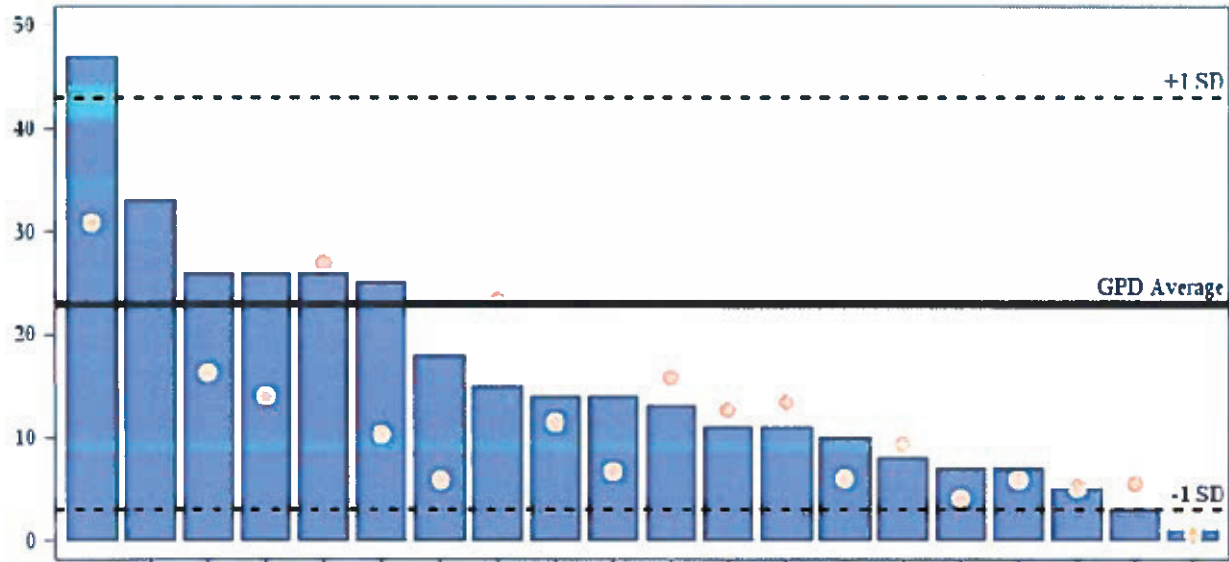
Rate of SPECIALTY REFERRALS ordered by you
DxCG Adjusted per 1000 patients



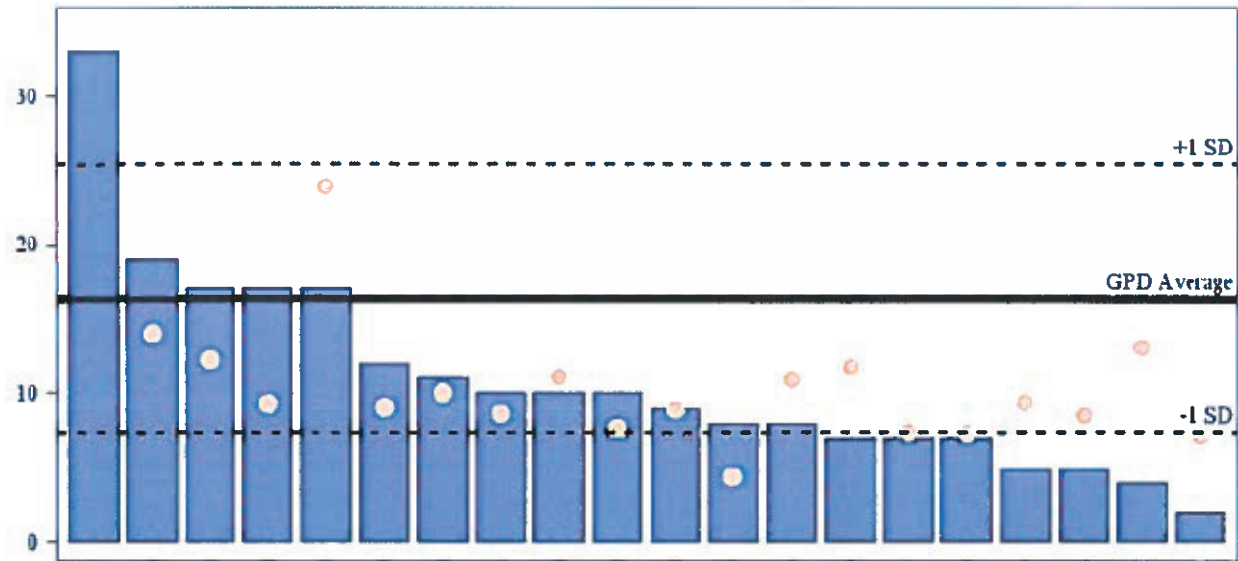
Resource Stewardship Dashboard PRIMARY CARE FP-IM

Clinic: [REDACTED]

Rate of CHEM PANELS ordered by you
DxCG Adjusted per 1000 patients



Rate of HIGH END IMAGING ordered by you
DxCG Adjusted per 1000 patients



Resource Stewardship Dashboard PRIMARY CARE FP-IM

Clinic: [REDACTED]

Explanatory Note:

The numbers in the graphs are expressed per DxCG and per 1000 patients, with a floor DxCG of 1.0. To translate values shown in this table, first increase any DxCG < 1.0 up to 1.0, then divide by DxCG, then multiply by 1000 divided by Avg Raw Panel Size.

For example:

- If HEI is 80, DxCG is 2.0, and Panel Size is 800, then $80/2 * 1000/800 = 50$
- If Pharmacy Rx is 12,000, DxCG is 0.7 (rounded up to 1.0 remember!), and Panel Size is 1200, then $12000/1.0 * 1000/1200 = 10,000$

Data sources:

New Specialty Visits: Cost Management Database (CMD), Claims

Specialty Referrals: Referrals System, [REDACTED]

High-End Imaging: Epic

Chem Panels: Lab Systems, [REDACTED]