By Elaine C. Khoong, Roy Cherian, Natalie A. Rivadeneira, Gato Gourley, Jinoos Yazdany, Ashrith Amarnath, Dean Schillinger, and Urmimala Sarkar

DOI: 10.1377/hlthaff.2018.0709
HEALTH AFFAIRS 37,
NO. 11 (2018): 1760–1769

Accurate Measurement In California's Safety-Net Health Systems Has Gaps And Barriers

Elaine C. Khoong (elaine.khoong@ucsf.edu) is a primary care research fellow in the Division of General Internal Medicine, University of California San Francisco (UCSF), and the Zuckerberg San Francisco General Hospital and Trauma Center.

Roy Cherian is a research data analyst at the Center for Vulnerable Populations, UCSF, and the Zuckerberg San Francisco General Hospital and Trauma Center.

Natalie A. Rivadeneira is a research data analyst at the Center for Vulnerable Populations, UCSF, and the Zuckerberg San Francisco General Hospital and Trauma Center.

Gato Gourley is a project coordinator at the Center for Vulnerable Populations, UCSF, and the Zuckerberg San Francisco General Hospital and Trauma Center.

Jinoos Yazdany is an associate professor of medicine in the Division of Rheumatology, UCSF.

Ashrith Amarnath is a patient safety officer at the Sutter Medical Foundation and a former patient safety officer in the Office of the Medical Director, Department of Health Care Services, both in Sacramento, California.

Dean Schillinger is a professor of medicine at UCSF and chief of the Division of General Internal Medicine at Zuckerberg San Francisco General Hospital and Trauma Center.

ABSTRACT **Patient safety in ambulatory care has not been routinely measured. California implemented a pay-for-performance program in safety-net hospitals that incentivized measurement and improvement in key areas of ambulatory safety: referral completion, medication safety, and test follow-up. We present two years of program data (collected during July 2015–June 2017) and show both suboptimal performance in aspects of ambulatory safety and questionable reliability in data reporting. Performance was better in areas that required limited coordination or patient engagement—for example, annual medication monitoring versus follow-up after high-risk mammograms. Health care systems that lack seamlessly integrated electronic health records and patient registries encountered barriers to reporting reliable ambulatory safety data, particularly for measures that integrated multiple data elements. These data challenges precluded accurate performance measurement in many areas. Policy makers and safety advocates need to support the development of information systems and measures that facilitate the accurate ascertainment of the health systems, patients, and clinical tasks at greatest risk for ambulatory safety failures.**

The Institute of Medicine report *To Err Is Human* catalyzed efforts to improve patient safety in hospitals,[1] but less attention has been paid to outpatient settings.[2] In comparison to inpatient care, ambulatory care more often involves multiple health systems and is dependent on patient actions (for example, scheduling follow-up). Adverse outcomes more frequently do not require medical care and are therefore known only to the patient. Consequently, there are many opportunities for safety lapses in ambulatory care processes.

Furthermore, the diversity of work flows associated with the numerous dimensions of ambulatory safety (such as medication monitoring, diagnostic timeliness and accuracy, referral coordination, and test result management)[3,4] makes the measurement of ambulatory patient safety challenging.[5] The scarcity of validated measures for these areas inhibits quantification of the impact of safety lapses, and this results in their exclusion from pay-for-performance programs. Thus, despite the importance of ambulatory patient safety, population-level data are lacking, and the feasibility of wide-scale quality measurement remains unknown.

In recognition of these gaps, the Public Hospital Redesign and Incentives in Medi-Cal (PRIME) Program in California created incentives for safety-net health care systems—defined by the Institute of Medicine to include systems that "deliver a significant level of healthcare…to uninsured, Medicaid, and other vulnerable patients"[6]—to measure ambulatory patient safety. The PRIME Program, California's Medicaid waiver for safety-net hospitals, is a pay-for-performance program that will distribute ap-

proximately $7 billion over five years, starting in 2016.[7]

Through the PRIME Program, safety-net systems are rewarded for reporting performance in year 1; in subsequent years funding is distributed to systems that achieve a predetermined level of improvement. Available funding is determined based on a proportional allotment factor that reflects each system's number of Medicaid beneficiaries and costs incurred for those patients. Part of the program focuses on outpatient care and includes the option to measure ambulatory patient safety measures.

To our knowledge, PRIME is the first wide-scale pay-for-performance program that includes ambulatory patient safety measures. It provides a unique opportunity to assess the feasibility of wide-scale measurement and acquire population-level data from safety-net health systems. The objectives of this study were twofold: to report the performance of the seventeen California safety-net health care systems that participated in the first two years of this program, and to describe challenges encountered during implementation of this novel ambulatory safety measurement effort. We also suggest next steps for health care systems and policy makers to continue advancing ambulatory patient safety.

Study Data And Methods

STUDY DESIGN AND SETTING This observational study used data reported to the California Department of Health Care Services for the PRIME Program by seventeen safety-net public health care systems classified as Designated Public Hospitals. To maintain confidentiality, we do not name systems—five were University of California systems, and twelve were government (nonstate)-operated systems—but instead assign letters to systems. Collectively, over half of the patients who received care from the systems that participate in the program were uninsured or received Medicaid.[8]

MEASURES STUDIED The PRIME Program was designed with required and optional "projects" that contain related measures (for example, six behavioral health measures are grouped into a behavioral health project). Designated Public Hospitals were required to report measures associated with three mandatory outpatient-related projects. The hospitals were also required to select one of four optional projects.

We present data on seven PRIME measures that address four distinct aspects of ambulatory patient safety: referrals from one provider to another, medication safety, timely follow-up of test results, and timely diagnosis. These measures were chosen because they represent a variety of areas in which ambulatory patient safety gaps are likely to occur. Exhibit 1 briefly describes each measure and the available funding associated with it.[9] Of the seven measures we analyzed, only one (closing the referral loop) was part of a required project; the remaining six measures came from optional projects. Detailed descriptions of measures are in online appendix exhibit A1.[10] No participating health system had previously measured or reported these measures.

The first three measures in exhibit 1 had been validated in external settings, but only the third (annual medication monitoring) had been widely reported through the Healthcare Effectiveness Data and Information Set (HEDIS). If established measures did not exist, health care system leaders developed new ones, designated as PRIME innovative measures, through consensus.[11]

DATA COLLECTION Each health care system independently reported its performance to the California Department of Health Care Services, which provided measure performance data to the authors (after suppressing data for confidentiality per the department's data deidentification guidelines). Systems reported both the denominator (the number of eligible patients) and the numerator (the number of patients who had the desired outcome) for each measure. We present data from the first two years of the program: July 1, 2015–June 30, 2016 and July 1, 2016–June 30, 2017.

We collected descriptive data about health care systems from publicly available information and the California Health Care Safety Net Institute. Descriptive data included the number of patients eligible for PRIME measures as of July 2017, the presence of an academic medical program, and rurality (as designated by the county Rural-Urban Continuum Code).[12] Given the importance of electronic health record (EHR) usability in quality measurement,[13] we also determined whether each system used a comprehensive EHR system. We defined a *comprehensive EHR* as a system in which the same software was used for inpatient and outpatient care as well as population health patient registries, which are critical tools for data collection and reporting.

OUTCOMES AND DATA ANALYSIS We conducted descriptive analyses of data reported for the seven measures and described the systems that were unable to report data for certain measures, because of challenges in validating data.

In addition, we analyzed performance changes from year 1 to year 2 and determined the median change across sites for each measure. We described systems that reported divergent performance changes or changes of a greater magni-

Urmimala Sarkar is an associate professor of medicine in the Division of General Internal Medicine, UCSF, and a primary care physician at Zuckerberg San Francisco General Hospital's Richard H. Fine People's Clinic.

EXHIBIT 1

Selected ambulatory patient safety measures and maximum funding associated with each measure

Measure	Measures percent of:	Measure steward[a]	Funds available for distribution in: Year 1	Year 2	Years 1–5	No. of systems reporting measure[b]
REFERRALS						
Closing the referral loop	Patients with referrals for which referring provider received report from provider to whom patient was referred	Centers for Medicare and Medicaid Services	$16.0 million	$21.3 million	$94.1 million	17
MEDICATION SAFETY						
Laboratory monitoring for patients on warfarin	Patients whose INR is checked for every eight weeks of warfarin therapy	National Quality Forum	$7.5 million	$10.0 million	$44.1 million	5
Annual laboratory monitoring for patients on persistent high-risk medications	Patients who receive appropriate annual laboratory monitoring if they receive at least 180 days of certain high-risk medications	National Committee for Quality Assurance	$7.5 million	$10.0 million	$44.1 million	5
TIMELY FOLLOW-UP OF HIGH-ACUITY ABNORMAL TEST RESULTS						
Timely follow-up of abnormal INR	Times that INR outside of treatment range was rechecked in a timely manner	PRIME innovative measure	$2.5 million	$3.3 million	$14.7 million	5
Timely follow-up of abnormal potassium	Times that abnormal potassium was rechecked in a timely manner	PRIME innovative measure	$2.5 million	$3.3 million	$14.7 million	5
TIMELY DIAGNOSIS						
Follow-up of abnormal fecal immunochemical test (FIT)	Patients with abnormal FIT who received colonoscopy within 6 months	PRIME innovative measure	$6.8 million	$9.0 million	$39.9 million	5
Timely biopsy after high-risk abnormal mammogram	Patients who had mammogram interpreted as high risk who received biopsy within 14 days	PRIME innovative measure	$6.8 million	$9.0 million	$39.9 million	5

SOURCE Authors' analysis of data provided by the California Department of Health Care Services. **NOTES** Funding refers to federal and state incentive payments made to health care systems based on their performance on these measures. The funding available for all seven measures combined was $49.5 million in year 1 (July 1, 2015–June 30, 2016), $66.0 million in year 2 (July 1, 2016–June 30, 2017), and $291.5 million in all five program years. International normalized ratio (INR) is explained in the text. High-risk mammograms are defined in the text. PRIME is Public Hospital Redesign and Incentives in Medi-Cal Program. [a]Stewards are applicable only to measures that were validated in external settings by the organization listed as steward; PRIME innovative measures were new measures developed through consensus. [b]Of the measures reported in this study, only the first was required. Thus, all seventeen safety-net public health care systems in the study reported that measure. The other measures were optional projects, and only five systems reported each of them.

tude than expected, based on observed changes across participating systems. Because large year-to-year changes are infrequent across health care performance measurement in general, divergent results suggest that the data should be viewed with caution. Initially, we defined *divergent changes* as >2.5 median absolute deviations from the median.[14] Because of our sample size, some of these deviations were small, and we did not want to identify results as divergent purely based on statistical considerations. We therefore designated a performance change as divergent if it was both >2.5 median absolute deviations from the median and a change of > 20 percent—a value chosen because this magnitude of change is rare and suggests data-reporting challenges rather than true performance change. For example, California 2017 HEDIS data present thirty ambulatory measures from twenty-six health plans, and only two plans reported a change of >20 percent from the prior year for one measure.[15] Given our small sample size, we included divergent results in analyses and calculations of median performance.

LIMITATIONS There were limitations to this analysis. First, since we wanted to report on the same systems in both years, we do not present data from district and municipal public hos-

pitals (DMPHs). Smaller and more rural than Designated Public Hospitals, these hospitals did not start reporting data until year 2, to allow time to develop data infrastructure.

Second, consistent with most pay-for-performance programs, including HEDIS,[16] all PRIME Program data were independently collected by health systems, so data collection methods varied.

Third, only some systems chose to report data for the optional measures we analyzed. As a result, data from only five health care systems are presented for each optional measure. Despite this, for every measure, the systems that reported data collectively represented at least 300,000 outpatients each year.

Study Results

Appendix exhibit A2[10] shows characteristics of the seventeen health care systems. All PRIME Program health systems are urban and affiliated with training programs. Nine used comprehensive EHR systems, but two of the comprehensive systems were locally developed (versus commercial) EHRs. The other eight systems used nonintegrated, noncomprehensive EHR systems, meaning that different software systems were deployed for patient registries, outpatient care, and inpatient care.

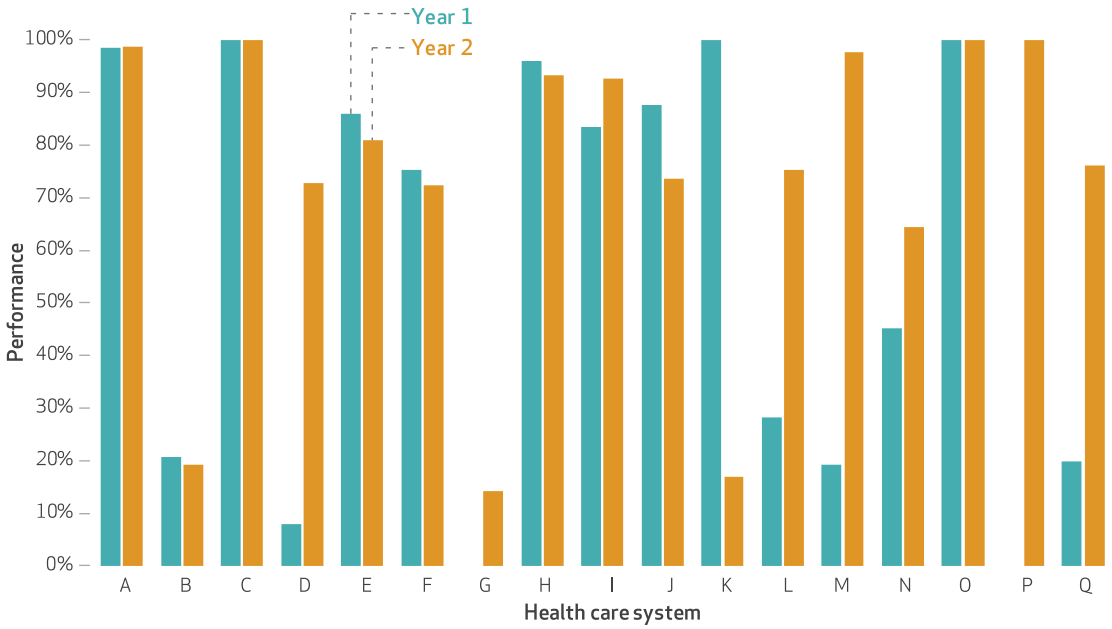
Depending on their participation in optional projects, different systems reported each measure. Therefore, we report results for each measure independently.

REFERRALS The median performance for closing the referral loop (a required measure that assesses whether referring providers receive information from consulting providers) was 83 percent in year 1 and 76 percent in year 2. Each system's performance is shown in exhibit 2. Of the fifteen sites that reported data in both years, five sites reported changes in performance that were divergent, as described above. Appendix exhibit A3 provides details on each system's performance on this measure.[10]

MEDICATION SAFETY AND HIGH-ACUITY ABNORMAL TEST FOLLOW-UP Of the five systems that reported the four optional measures discussed in this section, median performance was >80 percent in both years for annual monitoring of persistent medications, follow-up of abnormal international normalized ratio (INR), and follow-up of abnormal potassium. (The INR measures how well a blood thinner is working; blood

EXHIBIT 2

Seventeen safety-net public health care systems' performance on closing the referral loop in years 1 and 2 of their participation in the California Public Hospital Redesign and Incentives in Medi-Cal (PRIME) Program



SOURCE Authors' analysis of data provided by the California Department of Health Care Services. **NOTES** Year 1 was July 1, 2015–June 30, 2016, and year 2 was July 1, 2016–June 30, 2017. All seventeen systems were Designated Public Hospitals. Systems G and P were unable to capture validated data in year 1. Systems A–I used comprehensive electronic health record (EHR) systems. Systems G and H used locally developed EHRs. Percentages were determined using the number of eligible patients as the denominator and the number of patients who had the desired outcome as the numerator. Systems D, K, L, M, and Q reported divergent changes from year 1 to year 2; see the text for details.

that is too thin or not thin enough can be dangerous. Similarly, potassium levels that are low or high can be immediately life-threatening by affecting heart and nerve function.) Warfarin is a blood thinner medication, and guidelines advise assessing its efficacy by measuring the patient's INR every eight weeks. Performance on the measure of warfarin monitoring was lower than that on the other three measures: 51 percent in year 1 and 66 percent in year 2. Exhibit 3 shows performance in warfarin monitoring and follow-up of abnormal INR. (Appendix exhibit A4 graphs performance in annual monitoring of persistent medications and follow-up of abnormal potassium.)[10] Validated data were captured by all systems in both years. Only system P reported a performance change that was divergent, and it did so for both annual monitoring of persistent medications and follow-up of abnormal INR. Details about performance on these four measures are shown in appendix exhibit A5.[10]

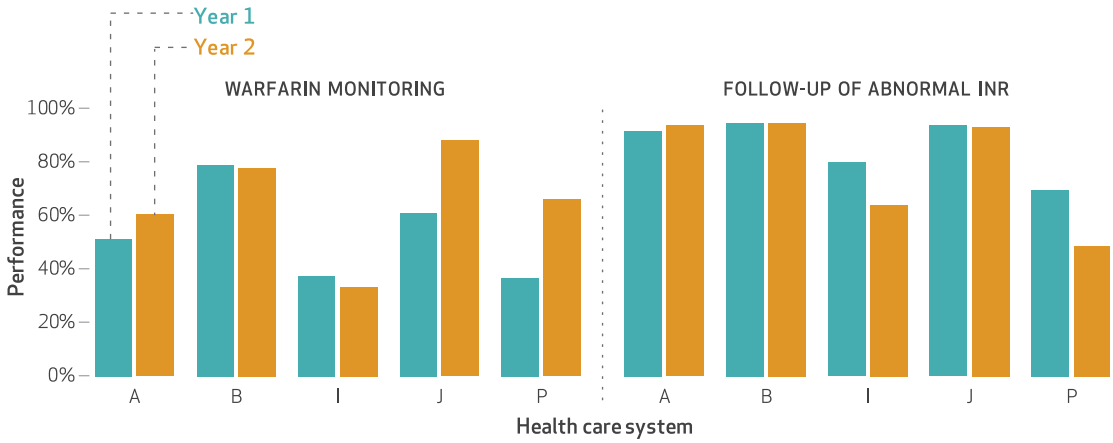
TIMELY DIAGNOSIS The five health care systems that measured the two optional measures discussed in this section reported poor performance. (Exhibit 4 graphs the systems' performance for the measures, with more details shown in appendix exhibit A6.)[10] For follow-up of an abnormal fecal immunochemical test (FIT, a stool-based test used to screen for colon cancer), the median performance was 49 percent in year 1 and 36 percent in year 2 (data not shown). Similarly, the median performance for a timely biopsy after a high-risk mammogram was 52 percent in year 1 and 48 percent in year 2. System O reported divergent performance changes for both measures.

DATA VALIDATION DIFFICULTIES AND RELIABILITY CONCERNS The majority of systems that struggled to acquire reliable data—whether they could not report validated data or reported divergent levels of performance change—did not have a comprehensive EHR. Of the three systems unable to capture validated data in year 1 (systems G and P for closing the referral loop and system Q for timely follow-up for FIT and biopsy for high-risk mammogram), two did not have a comprehensive EHR, and one (system G) used a locally developed comprehensive EHR. As appendix exhibit A2 shows, all three systems served 20,000–40,000 patients eligible for the PRIME Program.[10] Of the seven systems that reported divergent changes in performance, six did not have a comprehensive EHR.

FUNDING DISTRIBUTED Over five years, the federal and state government will distribute nearly $300 million in incentive payments to health care systems based on performance on these measures. In year 1, all participating systems received incentive payments (collectively, $49.5 million), even if they could not provide validated data. Up to $4.0 million (of $66.0 million) in year 2 was eligible for distribution based on divergent levels of performance improvement (exhibit 1).

EXHIBIT 3

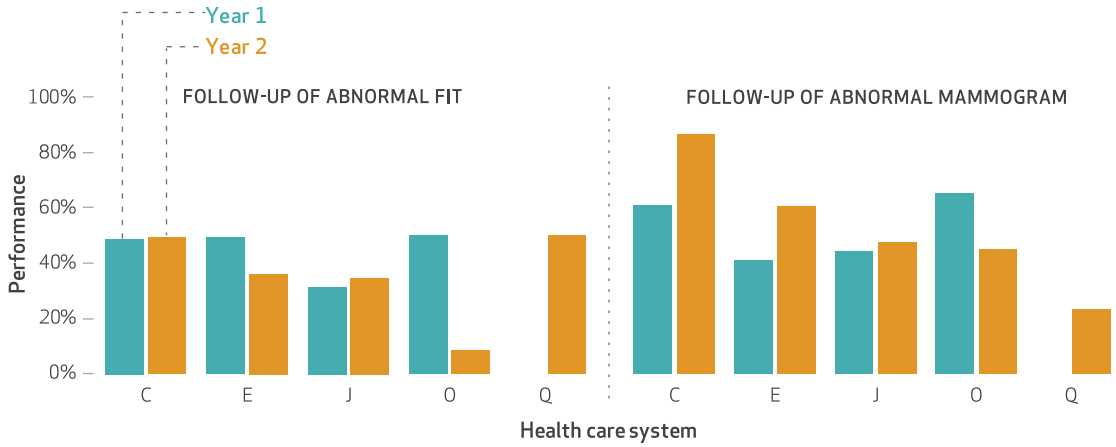
Five safety-net public health care systems' performance on warfarin monitoring and timely follow-up of abnormal international normalized ratio (INR) in years 1 and 2 of their participation in the California Public Hospital Redesign and Incentives in Medi-Cal (PRIME) Program



SOURCE Authors' analysis of data provided by the California Department of Health Care Services. **NOTES** Year 1 was July 1, 2015–June 30, 2016, and year 2 was July 1, 2016–June 30, 2017. These optional measures were reported by only these five systems. Systems A, B, and I used comprehensive electronic health record systems. Percentages were determined as described in the notes to exhibit 2. System P reported a divergent change from year 1 to year 2 for follow-up of abnormal INR; see the text for details.

EXHIBIT 4

Five safety-net public health care systems' performance on timely follow-up after an abnormal fecal immunochemical test (FIT) and timely biopsy after a high-risk abnormal mammogram in years 1 and 2 of their participation in the California Public Hospital Redesign and Incentives in Medi-Cal (PRIME) Program



SOURCE Authors' analysis of data provided by the California Department of Health Care Services. **NOTES** Year 1 was July 1, 2015–June 30, 2016, and year 2 was July 1, 2016–June 30, 2017. These optional measures were reported by only these five systems. System Q was unable to capture validated data in year 1. Systems C and E used comprehensive electronic health record systems. Percentages were determined as described in the notes to exhibit 2. System O reported changes from year 1 to year 2 that were divergent for both measures; see the text for details.

Discussion

These results establish baseline performance in ambulatory patient safety among California safety-net health systems. Although there was variation in performance, the median performance on each measure was stable from year 1 to year 2. Since significant performance changes in one year are unlikely, this stability suggests that these data are reasonable estimates of baseline performance, particularly for innovative measures that have not been previously widely measured (abnormal INR or potassium follow-up, and timely diagnostic tests after abnormal FIT or high-risk mammogram). However, these results also suggest significant barriers to the widescale measurement of ambulatory patient safety measures. In particular, systems without robust health data infrastructure, such as comprehensive EHR systems, might not be able to access data to accurately ascertain quality in multiple areas of ambulatory patient safety.

AREAS OF STRONGER PERFORMANCE Health systems performed better in follow-up of tests that required action within twenty-four to seventy-two hours (abnormal INR or potassium) than those that required action within weeks to months (abnormal FITs or mammograms). This suggests that work flows for more immediately life-threatening results are more robust, while those for abnormalities that do not require immediate action are underdeveloped—consistent with the results of prior studies.[17,18]

Performance was better on measures that required a single contact with a patient (annual monitoring of persistent medications or follow-up of abnormal INR or potassium) rather than repeated contact (warfarin monitoring). Systems also struggled with measures that required substantial patient engagement, such as follow-up of abnormal FIT. This supports assertions that the achievement of optimal ambulatory patient safety requires patient engagement.[19–21] Similarly, some systems struggled to achieve high performance when coordination with other providers (and health care systems) was required, such as closing the referral loop.

COMPARISON TO PRIOR STUDIES AND OTHER HEALTH SYSTEMS Although closing the referral loop and warfarin monitoring are established measures, they were not previously widely measured in the hospitals we studied. Prior literature on closing the referral loop showed a wide range of performance estimates (32–77 percent),[22–23] which was also reflected in the PRIME Program data. Similarly, our data on warfarin monitoring are consistent with those in a previous single-site study that found that approximately 60 percent of patients received adequate monitoring.[24]

Of the three established measures, only annual monitoring of persistent medications has been widely measured. HEDIS data show performance of 81–84 percent for patients with commercial insurance, 87–88 percent for Medicaid, and 91–93 percent for Medicare.[25] The PRIME Program

data show a similar level of performance (median: 92 percent in year 1 and 94 percent in year 2), which suggests that these data are reasonable estimates for a broad range of health care systems.

PRIME systems performed better than previously reported for two innovative measures: follow-up of abnormal INR and potassium. Earlier studies showed that over half of abnormal INR results received delayed or no follow-up.[26] Data from the PRIME Program suggest higher rates (approximately 85 percent) for follow-up of abnormal INR. However, the PRIME measure included patients with therapeutic INR levels (2.0–3.5) (appendix exhibit A1);[10] therefore, the measure only partially assessed abnormal INR follow-up. Similarly, while studies on follow-up of abnormal potassium suggest that 55–67 percent of patients receive timely follow-up,[27–31] PRIME system performance was >85 percent. However, once again, normal potassium levels were included in the measure, thereby inflating performance since both normal and abnormal test results were measured.

Unlike the other two innovative measures, performance on abnormal FIT and high-risk mammogram follow-up were consistent with that reported in prior literature. Previous studies documented that 40–60 percent of patients receive timely follow-up of abnormal FIT and 50–70 percent of patients receive timely biopsies of abnormal mammograms.[28,32–36] PRIME systems had a median performance of approximately 50 percent for both measures.

PREDICTORS OF DATA QUALITY CONCERNS Three health care systems were unable to validate data in year 1 for closing the referral loop and timely follow-up of abnormal FIT. Unlike measures that rely entirely on laboratory data, these require that data be captured from elements of the EHR, which may be stored in separate electronic systems. This supports assertions that measures that integrate disparate data elements (for example, pathology, imaging, and procedure notes) may be difficult to accurately measure in systems with less robust health data integration.[13,37]

Although only eight of the seventeen health care systems had a noncomprehensive EHR, they were disproportionately represented in systems that could not capture validated data (two of three) or reported divergent changes in performance (six of seven). While inaccuracies in EHR data capture for quality measures are well documented,[38–40] we further assert that inaccuracies are more likely to occur in systems with underdeveloped data infrastructure (including a noncomprehensive EHR), a more likely occurrence in underresourced settings.

Recommendations For Policy Makers And Ambulatory Patient Safety Advocates

ENSURE ACCURATE QUALITY MEASUREMENT Our findings support concerns that accurate performance measurement is difficult without a fully integrated data infrastructure.[41] Advocates of ambulatory patient safety need to consider how to support all health systems in acquiring the data and information system tools and personnel needed to support accurate performance reporting. Use of a certified EHR alone does not ensure ease of extracting accurate, complex data. Policy makers can regulate EHR vendors to create low-cost products that enable easy data collection for performance reporting, instead of requiring highly trained expensive analysts and local customization to support reporting. Performance reporting agencies must guarantee that less well resourced health systems have access to technical support on how to capture accurate data.

ENCOURAGE MEANINGFUL MEASUREMENT These data show that systems perform better on measures that have been more widely used (for example, annual medication monitoring) and worse in areas with newer measures (such as timely follow-up after abnormal tests). Ambulatory patient safety can improve only if performance is first measured, especially in areas where fewer data exist: diagnostic errors or delays, the management of test results, referrals between providers, transitions and care coordination, and administrative errors.[3,4,16] For areas where validated measures exist, use and adoption of the measures must be encouraged so that health care system leaders and policy makers have epidemiological data on the prevalence of safety concerns. In areas without measures, patient safety experts must develop and validate new ones.

However, given the burden that measurement places on health systems, particularly in low-resourced settings, measures must be meaningful (unlike the PRIME innovative measures for timely follow-up of abnormal INR and potassium). Organizations that create and validate new performance measures must continue efforts to develop measures across all areas of ambulatory safety that meet standards of acceptability, feasibility, reliability, sensitivity to change, and validity.[42]

Similarly, measure developers need to ensure that measures recognize that ambulatory safety requires patient engagement. Among the seven measures we studied, six required patients' cooperation and engagement to varying degrees—from presenting to care for a repeat laboratory blood draw to preparing for and attending a co-

Since all patients deserve safe care, state and nongovernment funding agencies should provide support for improving safety.

lonoscopy. Vulnerable patients, who are disproportionately cared for by safety-net systems, may encounter barriers to completing these actions. We know that current pay-for-performance programs penalize safety-net systems.[43] Measure developers should ensure that measures not only accurately assess the quality of care but also do not disproportionately penalize systems that care for vulnerable patients because of barriers to patient engagement.

Given the relative novelty of many ambulatory patient safety measures and the role of patients in ambulatory safety, we agree with others who have expressed skepticism about pay-for-performance programs.[44] We further assert that pay-for-performance currently has limited potential to advance ambulatory safety. As noted above, many health care systems do not have the data systems necessary to ensure accurate measurement. Moreover, outpatient safety requires patient engagement and shared decision making, which are difficult to measure.[45,46] This results in process-focused measures that incentivize actions that are not always tied to patient outcomes.[44,47] Pay-for-performance in its current form is not the right approach to improving outpatient safety; instead, initial investments in robust data infrastructure and the development of meaningful, valid measures are needed to ensure accurate data capture. Measurement is crucial, but its results should not be tied to reimbursement. Moreover, it is only one aspect of a multipronged approach to improve ambulatory patient safety that should also include leadership commitment, front-line engagement, a strong safety culture, and team-based work flows.

IDENTIFY SAFETY GAPS AND SHARE BEST PRACTICES The wide range of performance we observed suggests that there is substantial room for improvement. These data support the need for efforts to identify both the system-level characteristics that result in poor performance and the patients who are at greatest risk of falling into safety gaps (for example, people with low incomes or limited English proficiency). After high-risk populations and low-performing systems are identified, on-the-ground safety investigations are needed to understand the reasons for poor performance.[48]

By comparing high and low performers, patient safety advocates can begin to identify and develop approaches used by the former that successfully improve safety. Researchers and health care providers should encourage the sharing of innovative best practices that overcome barriers to patient safety, particularly for vulnerable systems and patients. Since all patients deserve safe care, instead of penalizing health systems that have suboptimal safety performance, state and nongovernment funding agencies should provide support for improving safety.

Conclusion

We found that wide-scale measurement of ambulatory patient safety faces challenges, particularly for complex measures that require the integration of different types of data. We also showed that there continues to be wide variation in performance on a broad range of ambulatory patient safety measures. In general, health systems perform better in areas that require only a single contact with a patient and limited patient engagement or coordination with other providers. To prevent harm to patients in ambulatory care settings, hospital systems need research and policies that incentivize the adoption of robust data infrastructure as well as the development of measures and measurement in all areas of ambulatory patient safety (especially test follow-up, diagnostic error, and care coordination). These data from the PRIME Program in California hold lessons for future measurement efforts and should inform local improvement initiatives in ambulatory patient safety. ■

Preliminary data and analyses were presented at the Annual Meeting of the Society of General Internal Medicine in Denver, Colorado, on April 11, 2018. Elaine Khoong is supported by the National Institutes of Health (NIH) through a National Research Service Award (Grant No. T32HP19025). Jinoos Yazdany received support for this work from the Agency for Healthcare Research and Quality (AHRQ) (Grant No. R01HS024412). Dean Schillinger was supported for work in this publication by the National Institute of Diabetes and Digestive and Kidney Diseases, NIH (Award No. P30DK092924). Urmimala Sarkar received support for this work from AHRQ (Grant No. R01HS024426) and the National Cancer Institute, NIH (Grant No. K24CA212294). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, AHRQ, or the US government. The authors acknowledge the data and insight provided by David Lown at the California Health Care Safety Net Institute, their SPARKNet collaborators, and other colleagues. The authors thank the California Department of Health Care Services for sharing and providing clarification on these data.

NOTES

1 Corrigan JM, Kohn LT, Donaldson MS, editors. To err is human: building a safer health system. Washington (DC): National Academies Press; 1999.

2 Wachter RM. Is ambulatory patient safety just like hospital safety, only without the "stat"? Ann Intern Med. 2006;145(7):547–9.

3 Shekelle PG, Sarkar U, Shojania K, Wachter RM, McDonald K, Motala A, et al. Patient safety in ambulatory settings [Internet]. Rockville (MD): Agency for Healthcare Research and Quality; 2016 Oct [cited 2018 Oct 3]. (Technical Brief No. 27). Available from: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/ambulatory-safety_technical-brief.pdf.

4 Sokol PE, Neerukonda KV. Safety risks in the ambulatory setting. J Healthc Risk Manag. 2013;32(3):21–5.

5 McGlynn EA, McDonald KM, Cassel CK. Measurement is essential for improving diagnosis and reducing diagnostic error: a report from the Institute of Medicine. JAMA. 2015;314(23):2501–2.

6 Institute of Medicine. America's health care safety net: intact but endangered. Washington (DC): National Academies Press; 2000. p. 21.

7 California Association of Public Hospitals and Health Systems. Priorities: Medi-Cal 2020 Waiver: Public Hospital Redesign and Incentives in Medi-Cal (PRIME) [Internet]. Oakland (CA): CAPH; c 2018 [cited 2018 Oct 9]. Available from: https://caph.org/priorities/medi-cal-2020-waiver/prime/

8 California Department of Health Care Services. Public Hospital Redesign and Incentives in Medi-Cal Program [Internet]. Sacramento (CA): DHCS; c 2018 [cited 2018 Oct 9]. Available from: http://www.dhcs.ca.gov/provgovpart/Pages/PRIME.aspx

9 California Department of Health Care Services. California Medi-Cal 2020 Demonstration [Internet]. Sacramento (CA): DHCS; 2017 Dec 22 [cited 2018 Oct 16]. Available from: https://www.dhcs.ca.gov/provgovpart/Documents/Medi-Cal2020STCsAmended060718.pdf

10 To access the appendix, click on the Details tab of the article online.

11 Ackerman SL, Gourley G, Le G, Williams P, Yazdany J, Sarkar U. Improving patient safety in public hospitals: developing standard measures to track medical errors and process breakdowns. J Patient Saf. 2018 Mar 14. [Epub ahead of print].

12 Department of Agriculture, Economic Research Service. Rural-Urban Continuum Codes: documentation [Internet]. Washington (DC): USDA; [last updated 2017 Nov 27; cited 2018 Oct 9]. Available from: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/

13 Chan KS, Fowles JB, Weiner JP. Review: electronic health records and the reliability and validity of quality measures: a review of the literature. Med Care Res Rev. 2010;67(5):503–27.

14 Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J Exp Soc Psychol. 2013;49(4):764–6.

15 California Department of Health Care Services. Medi-Cal managed care external quality review technical report: July 1, 2016–June 30, 2017 [Internet]. Sacramento (CA): DHCS; 2018 Apr [cited 2018 Oct 9]. Available from: https://www.dhcs.ca.gov/dataandstats/reports/Documents/MMCD_Qual_Rpts/TechRpt/CA2016-17_EQR_Technical_Report_F1.pdf

16 National Committee for Quality Assurance. HEDIS data submission [Internet]. Washington (DC): NCQA; [cited 2018 Oct 9]. Available from: http://www.ncqa.org/hedis-quality-measurement/hedis-data-submission

17 Blagev DP, Lloyd JF, Conner K, Dickerson J, Adams D, Stevens SM, et al. Follow-up of incidental pulmonary nodules and the radiology report. J Am Coll Radiol. 2014;11(4):378–83.

18 McDonald KM, Su G, Lisker S, Patterson ES, Sarkar U. Implementation science for ambulatory care safety: a novel method to develop context-sensitive interventions to reduce quality gaps in monitoring high-risk patients. Implement Sci. 2017;12(1):79–95.

19 Verstappen W, Gaal S, Bowie P, Parker D, Lainer M, Valderas JM, et al. A research agenda on patient safety in primary care. Recommendations by the LINNEAUS collaboration on patient safety in primary care. Eur J Gen Pract. 2015;21(Suppl):72–7.

20 Sarkar U, Wachter RM, Schroeder SA, Schillinger D. Refocusing the lens: patient safety in ambulatory chronic disease care. Jt Comm J Qual Patient Saf. 2009;35(7):377–83, 341.

21 Schwappach DL. Review: engaging patients as vigilant partners in safety: a systematic review. Med Care Res Rev. 2010;67(2):119–48.

22 Gandhi TK, Sittig DF, Franklin M, Sussman AJ, Fairchild DG, Bates DW. Communication breakdown in the outpatient referral process. J Gen Intern Med. 2000;15(9):626–31.

23 Patel MP, Schettini P, O'Leary CP, Bosworth HB, Anderson JB, Shah KP. Closing the referral loop: an analysis of primary care referrals to specialists in a large health system. J Gen Intern Med. 2018;33(5):715–21.

24 Mehrotra A, Forrest CB, Lin CY. Dropping the baton: specialty referrals in the United States. Milbank Q. 2011;89(1):39–68.

25 Stille C. Electronic referrals: not just more efficient but safer, too. Jt Comm J Qual Patient Saf. 2016;42(8):339–40.

26 Hurley JS, Roberts M, Solberg LI, Gunter MJ, Nelson WW, Young L, et al. Laboratory safety monitoring of chronic medications in ambulatory care settings. J Gen Intern Med. 2005;20(4):331–3.

27 National Committee for Quality Assurance. Annual monitoring for patients on persistent medications (MPM) [Internet]. Washington (DC): NCQA; c 2018 [cited 2018 Oct 9]. Available from: http://www.ncqa.org/report-cards/health-

plans/state-of-health-care-quality/2017-table-of-contents/persistent-medications
28 Chen ET, Eder M, Elder NC, Hickner J. Crossing the finish line: follow-up of abnormal test results in a multisite community health center. J Natl Med Assoc. 2010;102(8):720–5.
29 Lin JJ, Moore C. Impact of an electronic health record on follow-up time for markedly elevated serum potassium results. Am J Med Qual. 2011;26(4):308–14.
30 Moore C, Lin J, McGinn T, Halm E. Factors associated with time to follow-up of severe hyperkalemia in the ambulatory setting. Am J Med Qual. 2007;22(6):428–37.
31 Moore CR, Lin JJ, O'Connor N, Halm EA. Follow-up of markedly elevated serum potassium results in the ambulatory setting: implications for patient safety. Am J Med Qual. 2006;21(2):115–24.
32 Martin J, Halm EA, Tiro JA, Merchant Z, Balasubramanian BA, McCallister K, et al. Reasons for lack of diagnostic colonoscopy after positive result on fecal immunochemical test in a safety-net health system. Am J Med. 2017;130(1):93.e1–7.
33 Etzioni DA, Yano EM, Rubenstein LV, Lee ML, Ko CY, Brook RH, et al. Measuring the quality of colorectal cancer screening: the importance of follow-up. Dis Colon Rectum. 2006;49(7):1002–10.
34 Carlson CM, Kirby KA, Casadei MA, Partin MR, Kistler CE, Walter LC. Lack of follow-up after fecal occult blood testing in older adults: inappropriate screening or failure to follow up? Arch Intern Med. 2011;171(3):249–56.
35 Jones BA, Dailey A, Calvocoressi L, Reams K, Kasl SV, Lee C, et al. Inadequate follow-up of abnormal screening mammograms: findings from the Race Differences in Screening Mammography Process Study (United States). Cancer Causes Control. 2005;16(7):809–21.
36 Wernli KJ, Aiello Bowles EJ, Haneuse S, Elmore JG, Buist DS. Timing of follow-up after abnormal screening and diagnostic mammograms. Am J Manag Care. 2011;17(2):162–7.
37 Garrido T, Kumar S, Lekas J, Lindberg M, Kadiyala D, Whippy A, et al. e-Measures: insight into the challenges and opportunities of automating publicly reported quality measures. J Am Med Inform Assoc. 2014;21(1):181–4.
38 Roth CP, Lim YW, Pevnick JM, Asch SM, McGlynn EA. The challenge of measuring quality of care from the electronic health record. Am J Med Qual. 2009;24(5):385–94.
39 Danford CP, Navar-Boggan AM, Stafford J, McCarver C, Peterson ED, Wang TY. The feasibility and accuracy of evaluating lipid management performance metrics using an electronic health record. Am Heart J. 2013;166(4):701–8.
40 Shin EY, Ochuko P, Bhatt K, Howard B, McGorisk G, Delaney L, et al. Errors in electronic health record–based data query of statin prescriptions in patients with coronary artery disease in a large, academic, multispecialty clinic practice. J Am Heart Assoc. 2018;7(8):e007762.
41 Smith PC, Mossialos E, Papanicolas I. Performance measurement for health system improvement: experiences, challenges and prospects [Internet]. Copenhagen: World Health Organization Regional Office for Europe; 2008 [cited 2018 Oct 9]. (Background Document). Available from: https://www.who.int/management/district/performance/PerformanceMeasurementHealthSystemImprovement2.pdf
42 Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ. 2003;326(7393):816–9.
43 Damberg CL, Elliott MN, Ewing BA. Pay-for-performance schemes that use patient and provider categories would reduce payment disparities. Health Aff (Millwood). 2015;34(1):134–42.
44 Frakt AB, Jha AK. Face the facts: we need to change the way we do pay for performance. Ann Intern Med. 2018;168(4):291–2.
45 Barr PJ, Elwyn G. Measurement challenges in shared decision making: putting the "patient" in patient-reported measures. Health Expect. 2016;19(5):993–1001.
46 Sepucha KR, Scholl I. Measuring shared decision making: a review of constructs, measures, and opportunities for cardiovascular care. Circ Cardiovasc Qual Outcomes. 2014;7(4):620–6.
47 Pronovost PJ, Wachter RM. Progress in patient safety: a glass fuller than it seems. Am J Med Qual. 2014;29(2):165–9.
48 National Patient Safety Foundation. RCA^2: improving root cause analyses and actions to prevent harm. Boston (MA): The Foundation; 2015.

Appendix Exhibit A1: Measure Specifications

Measure 1: Closing the Referral Loop

Brief Description
Percentage of patients with referrals, regardless of age, for which the referring provider receives a report from the provider to whom the patient was referred.

Denominator
Number of patients, regardless of age, who were referred by any provider to a PRIME Entity specialty provider, and who had a visit with the PRIME Entity specialty care provider during the measurement period.

Denominator Exclusions
None

Numerator
Number of patients with a PRIME Entity specialty care referral, for which the referring provider received a report from the PRIME Entity specialty care provider to whom the patient was referred.

Numerator Exclusions
None

Measure 2: Laboratory monitoring for patients on warfarin

Brief Description
Percentage of individuals at least 18 years of age as of the beginning of the measurement period with at least 56 days of warfarin therapy who receive an International Normalized Ratio (INR) test during each 56-day interval with warfarin. Warfarin is a blood thinner, and INR is a lab value used to monitor how well the warfarin in thinning blood. The efficacy of warfarin is impacted by many factors and thus monitoring of its impact is warranted to ensure that the INR level is within the desirable range.

Denominator
Individuals at least 18 years of age as of the beginning of the measurement period with warfarin therapy for at least 56 days during the measurement period.

Denominator Exclusions
- Individuals who are monitoring INR at home
- Individuals who are in long-term care (LTC) during the measurement period

Numerator
The number of individuals in the denominator who have at least one INR monitoring test during each 56-day interval with active warfarin therapy

Numerator Exclusions
An interval with a hospitalization of more than 48 hours is considered an interval with an INR test.

Measure 3: Annual laboratory monitoring for patients on persistent medications

Brief Description
Percentage of individuals age 18 and older who received at least 180 treatment days of ambulatory medication therapy for select therapeutic agents during the measurement year and at least one therapeutic monitoring event for the therapeutic agent in the measurement year.

Denominator
Rate 1: Annual Monitoring for Individuals on ACE Inhibitors or ARBs
Individuals who received at least 180 treatment days of ACE inhibitors or ARBs, during the measurement year. Individuals may switch therapy with any included medication during the measurement year and have the day's supply for those medications count toward the total 180 treatment days (i.e., an individual who received 90 days of ACE inhibitors and 90 days of ARBs meets the denominator definition for rate 1).

Rate 2: Annual Monitoring for Individuals on Digoxin
Individuals who received at least 180 treatment days of digoxin during the measurement year.

Rate 3: Annual Monitoring for Individuals on Diuretics
Individuals who received at least 180 treatment days of a diuretic during the measurement year.

Denominator Exclusions
None

Numerator
Rate 1: Annual Monitoring for Individuals on ACE Inhibitors or ARBs
At least one serum potassium and a serum creatinine therapeutic monitoring test in the measurement year.

Rate 2: Annual Monitoring for Individuals on Digoxin
At least one serum potassium, at least one serum creatinine, and at least one serum digoxin therapeutic monitoring test in the measurement year.

Rate 3: Annual Monitoring for Individuals on Diuretics
At least one serum potassium and a serum creatinine therapeutic monitoring test in the measurement period. Any of the following during the measurement period meet criteria:

Numerator Exclusions
Exclude individuals from each eligible population rate who had an inpatient (acute or non-acute) claim/ encounter during the measurement period.

Measure 4: Timely follow-up of abnormal INR

Brief Description
For patients 18+ years who receive warfarin therapy for at least 56 days, at least one INR monitoring test during each 56-day interval with active warfarin therapy, and follow-up appropriate to the result. INR levels that are too high / low can be dangerous to patients by increasing bleeding or clotting risk.

Denominator
The number of INR's in the ambulatory setting for individuals at least 18 years old at the beginning of the measurement period with warfarin therapy for at least 56 days during the measurement period and who have at least one INR monitoring test during each 56-day interval with active warfarin therapy.

Denominator Exclusions
- Individuals monitoring INR at home or in long-term care during the measurement period.
- INR tests ordered from the emergency department or inpatient setting.
- INR tests ordered from urgent care settings that are located within an emergency department

Numerator
Number of denominator INR's that have had appropriate follow-up as per the following:

- INR < 2 → follow up with a new lab in 4 weeks
- INR 2-3.5→No lab follow-up
- INR > 3.5 and <4.9 →follow up with a new lab in 10 days
- INR >4.9 →Documented as a critical lab value and that the ordering clinician (or responsible delegate) has been informed within 24 hours

Numerator Exclusions
None

Measure 5: Timely follow-up of abnormal potassium

Brief Description
This measure assesses the percentage of ambulatory serum potassium tests performed on patients at least 18 years old who received at least 180 treatment days of ACE, ARB or Diuretic therapy during the measurement year, at least one potassium monitoring event and follow-up appropriate to the results of that potassium monitoring event. ACE, ARB, and diuretics are medications that can impact potassium levels. A potassium level that is too low or too high can be dangerous by affecting heart function.

Denominator
The number of serum potassium tests completed in the ambulatory setting during the measurement year in patients age 18 and older, as of the end of the measurement period, who are on selected persistent medications (ACE Inhibitors/ARB or Diuretics) for at least 180 days.

Denominator Exclusions
- Tests ordered from the emergency department or inpatient setting.
- Tests ordered from urgent care settings that are located within an emergency department

Numerator
Number of denominator serum Potassium tests that have had appropriate follow-up, as per the following:

- Potassium 3.5 – 5.1→ Normal range, no lab follow-up
- Potassium ≥3 and <3.5 → Follow up lab in 4 weeks
- Potassium ≥5.2 and < 6 → Follow up lab in 10 days
- Potassium <3 or ≥6 → Documented as a critical lab value and that the ordering clinician (or responsible delegate) has been informed within 24 hours

Numerator Exclusions None

Measure 6: Follow-up of abnormal FIT

Brief Description
This measure assesses if patients receive a follow-up colonoscopy after a positive stool-based colon cancer screening test (FIT or FOBT). Abnormal results for these tests warrant a colonoscopy to determine if colon cancer is present.

Denominator
Total number of individuals 51-75 years of age with a positive FIT/FOBT during the first six months of the measurement period.

Denominator Exclusions
None

Numerator
Number of individuals in the PRIME Project Target Population 51-75 years of age receiving a colonoscopy within 6 months of the date of the positive stool test. Measurement period for numerator is all 12 months of the measurement period.

Numerator Exclusions
None

Measure 7: Timely biopsy of high-risk abnormal mammogram

Brief Description
Timely follow-up after abnormal mammogram to ensure timely diagnosis of breast cancer

Denominator
Total number of individuals in the eligible population who received either a screening or diagnostic mammogram by the public health system during the measurement year that was assessed as a BIRADS 4 or 5.

Denominator Exclusions
None

Numerator
Number of individual in the denominator for whom a breast biopsy was performed within 14 business days of the result date of a screening or diagnostic mammogram being given a BIRADS 4 or 5; includes mammograms and biopsies ordered by the system that have been outsourced

Numerator Exclusions
- When the patient is offered, but declines to make an appointment in 14 business days (i.e. vacation, going out of the country, personal reasons, deceased, getting care at another facility, incarcerated). This must be documented in the mammogram report. This documentation may occur on the date of report. If the information is obtained later, the delay should be documented as an addendum to the report of a BIRADS 4 or 5.
- When the patient is refusing an imaging guided biopsy. This must be documented in the diagnostic mammogram report.
- When the doctor requests a delayed biopsy because other treatments take priority (i.e. chemotherapy or other medical treatments planned). This should be documented in addendum or report

Appendix Exhibit A2

Traits of Participating Healthcare System

Healthcare System	Comprehensive Electronic Health Record	Number of Outpatients [a]	State-Funded	Reported Measure 1	Reported Measures 2 - 5	Reported Measures 6 & 7
A	X	> 100,000	X	X	X	
B	X	> 100,000		X	X	
C	X	80 - 100,000		X		X
D	X	60 - 80,000	X	X		
E	X	40 - 60,000	X	X		X
F	X	20 - 40,000	X	X		
G	X [b]	20 - 40,000		X		
H	X [b]	< 20,000		X		
I	X	< 20,000	X	X	X	
J		> 100,000		X	X	X
K		80 - 100,000		X		
L		40 - 60,000		X		
M		40 - 60,000		X		
N		40 - 60,000		X		
O		40 - 60,000		X		X
P		20 - 40,000		X	X	
Q		20 - 40,000		X		X

[a] PRIME eligible population at end of second year (June 30 2017)
[b] Locally-developed system

Appendix Exhibit A3

Performance on Closing the Referral Loop (n = 17)

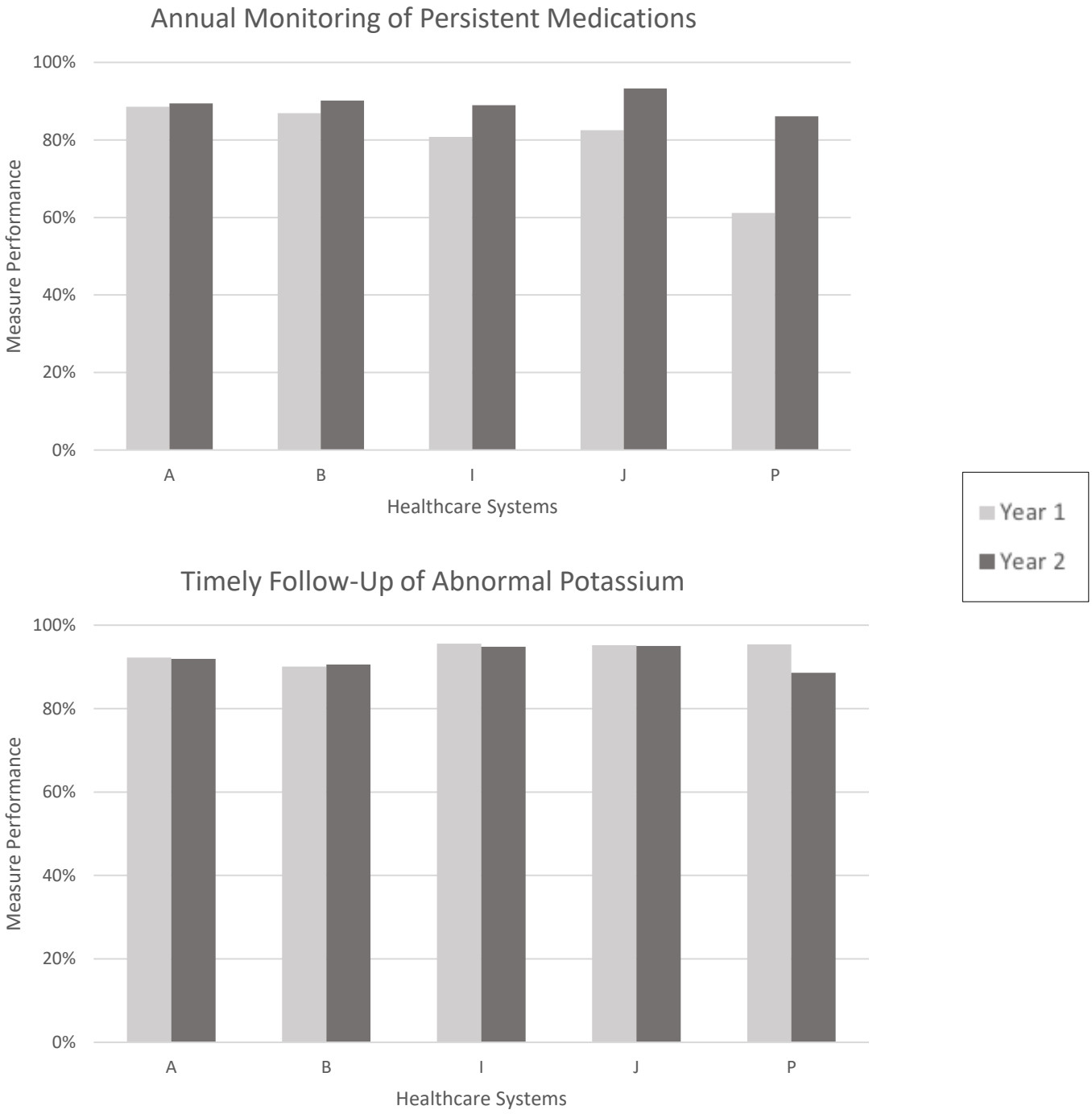
Healthcare System	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2
A	98%	99%	0%
B	21%	19%	- 1%
C	100%	100%	0%
D	8%	73%	+ 65% *
E	86%	81%	- 5%
F	75%	72%	- 3%
G [a]	--	14%	--
H [a]	96%	93%	- 3%
I	83%	93%	+ 9%
J [b]	88%	74%	- 14%
K [b]	100%	17%	- 83% *
L [b]	28%	75%	+ 47%*
M [b]	19%	98%	+ 79%*
N [b]	45%	64%	+ 19%
O [b]	100%	100%	0%
P [b]	--	100%	--
Q [b]	20%	76%	+ 56%*
All systems: median (MAD [c])	83% (17%)	76% (17%)	0% (10%)

Source: Authors' analysis of data provided by the California Department of Health Care Services

Notes:
[a] Systems used a locally-developed non-commercial electronic health record
[b] Systems that did not use a comprehensive electronic health record system
[c] Median absolute deviation (MAD): measure of spread around the median
* Change more than 2.5 median absolute deviations from median change and at least 20%

Appendix Exhibit A4

Performance on Annual Monitoring of Persistent Medications & Follow-Up of Abnormal Potassium (n = 5)



Source: Authors' analysis of data provided by the California Department of Health Care Services
Notes: These are optional measures reported by only these five systems. Systems A, B, and I used comprehensive EHR systems. System P reported change of at least 20% and more than 2.5 median absolute deviations from median change for annual monitoring of persistent medications.

Appendix Exhibit A5

Performance on Medication Safety & Follow-Up of High-Acuity Abnormal Test Follow-Up (n = 5)

Healthcare System	Measure 2: Warfarin Monitoring			Measure 3: Annual Monitoring of Persistent Medications		
	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2
A	51%	60%	+ 9%	89%	89%	+ 1%
B	79%	78%	- 1%	87%	90%	+ 3%
I	37%	33%	- 4%	81%	89%	+ 8%
J [a]	61%	88%	+ 27%	53%	93%	+ 11%
P [a]	37%	66%	+ 29%	61%	86%	+ 25% *
All systems: median (MAD [b])	51% (14%)	66% (12%)	+ 9% (13%)	83% (8%)	89% (1%)	+ 8% (5%)
	Measure 4: Timely Follow-Up of Abnormal INR			Measure 5: Timely Follow-Up of Abnormal Potassium		
	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2
A	92%	94%	+ 2%	92%	92%	0%
B	95%	95%	0%	90%	90%	0%
I	80%	64%	- 16%	96%	95%	- 1%
J [a]	94%	94%	0%	95%	95%	0%
P [a]	70%	48%	- 21% *	95%	89%	- 7%
All systems: median (MAD [b])	92% (3%)	94% (1%)	0% (2%)	95% (1%)	92% (3%)	0% (0%)

Source: Authors' analysis of data provided by the California Department of Health Care Services

Notes: These are optional measures reported by only these five systems.
[a] Systems that did not use a comprehensive electronic health record system
[b] Median absolute deviation (MAD): measure of spread around the median
* Change more than 2.5 median absolute deviations from median change and at least 20%

Appendix Exhibit A6

Performance on Timely Diagnosis (n = 5)

Healthcare System	Measure 6: Timely Follow-Up of Abnormal FIT			Measure 7: Timely Biopsy of High-Risk Abnormal Mammogram		
	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2	Performance in Year 1	Performance in Year 2	Change from Year 1 to Year 2
C	49%	49%	+ 1%	61%	86%	+ 25%
E	49%	36%	- 13%	41%	60%	+ 19%
J [a]	31%	34%	+ 3%	44%	48%	+ 4%
O [a]	50%	8%	- 42% *	65%	45%	- 20% *
Q [a]	--	50%	--	suppressed	23%	--
All systems: median (MAD [b])	49% (0.5%)	36% (13%)	- 6% (8%)	52% (10%)	47% (12%)	12% (11%)

Source: Authors' analysis of data provided by the California Department of Health Care Services

Notes: These are optional measures reported by only these five systems.
[a] Systems that did not use a comprehensive electronic health record system
[b] Median absolute deviation (MAD): measure of spread around the median
* Change more than 2.5 median absolute deviations from median change and at least 20%