Research Methods

Measuring Practice Systems for Chronic Illness Care: Accuracy of Self-Reports from Clinical Personnel

Sarah Hudson Scholle, M.P.H., Dr.P.H.; L. Gregory Pawlson, M.D., M.P.H.; Leif I. Solberg, M.D.; Sarah C. Shih, M.P.H.; Stephen E. Asche, M.A.; Ann F. Chou, Ph.D.; Merry Jo Thoele, M.P.H.

Gaps in health care quality are widespread and costly. The McGlynn et al. landmark study showed that health care is consistent with widely recommended guidelines only about half of the time.[1] A widely cited reason for poor performance in ambulatory care is that clinical practice systems and processes are not optimally designed to address the care of persons with chronic diseases or preventive needs.[2]

According to the Chronic Care Model (CCM), an evidence–based conceptual framework for effective care of chronic conditions,[3] adopting a variety of practice systems is key to improving quality. These systems create the microenvironment needed in office practice to provide consistent and comprehensive care and to foster more productive interactions between patients and clinicians. By *practice systems*, we mean organized processes designed to ensure that certain information is collected and information or services are provided when needed to patients or health care personnel (for example, counseling, reminders, test results, or patient education).

Interventions using a systematic approach have improved quality and outcomes in diabetes, depression, and other chronic conditions[4–9] in a variety of settings.[10,11] Furthermore, implementation of the CCM was also positively associated with provision of recommended preventive services to address health risk behaviors.[12] All four clinically related elements of the CCM—clinical information systems, decision support, self-management support, and delivery system redesign that includes care managers and team-based care—appear important in relationship to quality. For most practices, implementation starts with a registry or organized list of patients with a particular condition that will allow tracking their need for services over time.

Even outside a comprehensive model, the research literature supports the benefits of specific aspects of the CCM. For example, a review of more than 70 studies showed that clinical decision support systems, such as clinical reminders and decision support tools, improved quality of care.[13] Electronic prescribing and decision support may also reduce costs.[14]

Although medical practices can implement a variety of sys-

Article-at-a-Glance

Background: Purchasers, plans, and clinical practices involved in quality improvement initiatives are increasingly interested in measuring practice systems, particularly in relation to clinical quality and as part of pay-for-quality initiatives. The validity of self-reports of the use of practice systems was examined.

Methods: In 11 medical groups in Minnesota, the Physician Practice Connections Readiness Survey, which was developed on the basis of the concepts and evidence base of the Chronic Care Model, was used to survey office practice personnel about practice systems. Participation rates by medical group ranged from 61% to 94%, with a mean of 76%, yielding surveys from 32 lead physicians and 241 other personnel. Survey results were compared with an on-site audit by trained surveyors.

Results: Overall agreement with the on-site audit ranged from 40.9% to 96.7% among lead physicians and from 33.9% to 81.9% among other personnel. Mean agreement was high for quality improvement (96.7% for lead physicians and 81.9% for other personnel), moderate for clinical information systems (71.2% for lead physicians and 66.0% for others), and low for the use of care management (less than 50% for both groups). Mean positive predictive value ranged from 55.2% to 100% among lead physicians and from 49.6% to 100% among other personnel. Both the presence of systems and the accuracy of reporting varied across medical groups.

Discussion: The accuracy of self-reports of practice systems varies by type of system being assessed and by type of respondent. Although self-assessment may be useful for quality improvement purposes, self-reported information on clinical practices systems should not be used for accountability purposes, including pay-for-quality efforts or public reporting unless additional documentation is required to ensure fair comparisons.

tems identified by the CCM, including decision support without the presence of electronic health records,[4,15,16] research has suggested that electronic information systems may facilitate effective implementation of practice systems. However, we reported that the use of information technology was only weakly related to overall use of systems.[15] Initiatives to monitor quality, pay providers on the basis of performance, and publicly report results are growing, often with rewards for adoption of practice systems.[17,18] A survey found a 40% increase between 2004 and 2005 in the use of information technology performance measures focused on improving care coordination and safety.[19]

Still, there is relatively little information about the extent to which practice systems are present in medical practices. Adoption of evidence-based practice systems appears to be low,[12,20,21] although this is primarily based on self-report survey data from larger medical groups. Valid and reliable tools for assessing practice systems are needed, both for understanding the extent of adoption and the use of systems and particularly for efforts to use the degree of critical systems use as the basis for performance rewards or other incentives.

Given the lack of valid and reliable tools, in 2003 the National Committee for Quality Assurance (NCQA) developed a self-report survey tool, now known as the Physician Practice Connections-Readiness Survey (PPC-RS), to assess the presence and function of clinical practice systems. The study reported in this article examines the tool's validity. Because of the interest in using systems surveys both to direct quality improvement (QI) initiatives and to serve as the basis of pay-for-performance awards, the research focused on identifying the accuracy of self-reports and the best respondent. Specific research questions were as follows:

- How well do self-reports of clinical practice systems from lead physicians and other staff agree with an on-site audit?
- Are there variations across medical groups in agreement with the audit?
- Does the accuracy of self-reports vary by the type of respondent?

Developing the PPC-RS*

The PPC-RS uses the CCM as its conceptual foundation along with elements from Six Sigma processes.[22] The developmental process included an extensive review of research,[10,18,23–26] and pretesting efforts with physician practices. An expert advisory panel, reviewing the evidence on the effectiveness of clinical practice systems, recommended specific elements of the CCM for assessing practice systems, with a particular focus on both preventive and chronic illness care. Some subelements had considerable empirical evidence (for example, use of patient registries improved glycosolated hemoglobin [HA1C] levels in patients with diabetes[27,28]). However, the evidence was less rigorous for other subelements; in the case of community involvement, given the lack of face validity at the individual office practice level, corresponding measures were not included in the initial measurement set. When possible, we used items consistent with the National Survey of Physician Organizations, whose range of topics includes but is not limited to practice systems.[20]

The concepts of the subelements were translated into survey items that would quantify a process or an element in the physician practice.

The test set of physician office measures was presented to several NCQA advisory councils. NCQA's work with the Bridges to Excellence program, a pay-for-performance initiative,[18] also informed the development of the PPC-RS. A panel of clinicians, purchaser medical directors, systems experts, and others identified areas in ambulatory clinical practice with deficiencies that it felt led to failures in clinical care. Input from these discussions on reducing defects in care processes was also considered in developing the systems survey.

The survey was revised on the basis of alpha testing of the survey with physician groups in New Mexico, Minnesota, and Washington, and interviews with several small practices in the Washington, DC, area. Further pretesting was conducted with clinicians in three practices in Minnesota before fielding the survey for the work reported here.

Examining the Validity of the Self-Report Surveys

Research Setting

In late 2004, we conducted this study in Minnesota in collaboration with Institute for Clinical Systems Improvement (ICSI), a QI collaborative that includes most of the medical groups and hospitals in the state. Those member organizations at the time included about 75% of the physicians in Minnesota.[29] To meet a recruitment goal of 12 medical groups, ICSI purposively selected 19 of its 38 medical group members as potential participants. These 19 medical groups all provided primary care to adults, had experience with quality and QI methods, and together represented a diversity of locations and panel sizes. The study was approved and monitored by the HealthPartners Institutional Review Board.

Sampling and Survey Process

Recruitment from the sample of 19 groups was conducted

* The PPC-RS is available from the authors by e-mail request.

by first sending a letter to the medical director (or equivalent) of each group, followed by phone calls from a local physician study investigator [L.S.]. Only 3 medical groups, each citing competing demands, declined participation, and 2 groups agreed too late to be included. Three more medical groups were excluded because they participated in pretesting the survey and audit, yielding a final sample of 11 groups. Each medical group's medical director was asked to complete the survey, along with the lead physician and up to eight randomly sampled physicians and four registered nurses (R.N.s) at each practice site. A list of potential respondents from each site was enumerated and sampled. The survey was mailed with a cover letter from ICSI. Postcard reminders were sent to nonrespondents two weeks after the initial mailings, followed by a second mailing of the survey two weeks later. Telephone follow-up with the practice was used to encourage participation.[30] Each person completing the survey received a $15 gift certificate.

On-Site Audit

Two trained and experienced nurse auditors conducted the on-site audits in early 2005. The auditors met with each participating medical group's QI leader, supplemented by meetings with other staff who had critical knowledge about any of these systems. The audit covered details and reviewed examples about each of seven separate but related systems, as noted in the "Measures" section.

For each practice system, the auditors required visual evidence that the system and each of its potential components was present and usable. An investigator or data collection supervisor accompanied the auditors to most of the site visits to monitor them, and several debriefing sessions were conducted with the entire investigator group to clarify and verify the information and its collection process.

Measures

The prototype PPC-RS addresses the following seven domains that are consistent with four of the CCM's elements:

- *Clinical Information Systems:* Three domains in this area assessed (1) information systems (including paper-based tools and electronic systems); (2) the presence of a registry or organized database allowing practices to group patients by diagnoses or other parameters, and (3) systematic monitoring of the patient population (using a database to monitor key indicators of chronically ill patients' medical conditions for information that may require attention).
- *Decision Support:* A single domain focused on clinician reminders, including specific communications intended to inform or alert clinicians of laboratory tests, visits, or guidelines (best practices) related to the care of individual patients.
- *Delivery System Redesign:* A single domain assessed defined services for managing patients with chronic illness involving multiple clinicians and care between office visits.
- *Health Care Organization:* Two domains in this area included performance tracking and feedback and clinical quality improvement. Performance tracking and feedback items assessed the process of using clinical information systems to aggregate and report on key indicators culled from a patient registry or other data source for the purposes of benchmarking performance and informing QI activities.
- *Clinical Quality Improvement:* Items evaluated the presence of formal processes to assess care, develop interventions, and use data to monitor the effects.

Each domain was assessed using one to four components, and each component included 1 to 13 survey items. For the analysis, a score was calculated for each component using the simple average of item scores. The summary score for the domain was the simple average of the component scores.

Analysis

All analyses were conducted using SAS version 9.1 (SAS Institute, Carey, North Carolina). To assess whether self-report responses agreed with the reports of the on-site auditor, the percent agreement and positive predictive value (PPV) of self-reported responses were selected as a simple and intuitive approach for summarizing the large amount of detailed data. We calculated the average percent agreement (the percentage of respondents who agreed with the auditor's decision about whether the system was present or not present) across all components for each of the seven system domains. We also assessed the accuracy of self-report by examining the predictive value of self-reported responses (that is, among respondents who reported that a system was present, the percentage whose reports were validated by the auditor). To determine whether lead physicians are more accurate reporters than other clinical staff, we compared the validity of self-reports for the lead physicians in the medical group, including the medical director and the responsible physician at each practice site, to that of reports by other physicians and nurses. We did not compute statistical significance tests because our review of the data suggested that the large number of comparisons could identify chance findings and the differences between reporters did not appear to be meaningful. Finally, we present descriptive data on reporting across each of the 11 medical groups participating in the study.

Table 1. Description of the Participating Medical Groups

Location	
■ Metropolitan area	3
■ Cities (50,000–100,000)	3
■ Towns (10,000–20,000)	5
Ownership	
■ Hospital system	6
■ Physicians	5
Number of Sites	1–6
Staffing	
■ No. of adult primary care physicians	6–129
■ Nurse practitioners	0–14
■ Physician assistants	0–17
Clinical information systems	
■ Paper only	1
■ Paper records supplemented with electronic systems	6
■ Partial electronic health record	3
■ Full electronic health record	1
All sites use same systems?	
■ Yes	3
■ No	5
■ Not applicable (single site)	3

Results

Description of the Medical Groups

The 11 participating medical groups, which varied in size, ownership, and geographic location (Table 1, above), completed a total of 273 surveys. The number of responses per group ranged from 10 to 43, with 32 responses by the lead physician (medical group director or lead physician at a particular practice site). Response rates with various medical groups ranged from 61% to 94% (mean, 76%).

Self-Report/Audit Agreement and PPV

Table 2 (page 411) shows the average agreement for the audit and average PPV scores for each domain. The remaining tables provide detailed results for each domain and item, except for the single-item registry. Agreement with the on-site audit was highest for the QI domain (for example, 96.7% for lead physicians, 81.8% for other clinicians), moderate for registry and clinical information systems, and lower for the remaining domains. In general, there was little difference in the level of agreement between lead physicians and other clinical staff, although the direction of the difference always favored the lead physicians. The PPVs of reports were usually higher than average agreement, which is expected for systems that were present in all or nearly all medical groups. For example, overall agreement with the presence of care management was low (40.9%), but lead physicians who reported the presence of care management were nearly always validated by the audit (PPV, 88.9%) because all medical groups had this system, according to the audit (Table 2).The mean PPV ranged from 55.3% to 100% among lead physicians and from 49.6% to 100% among other personnel.

As shown in Table 3 (page 411), some items in the clinical information systems domain had high agreement and PPVs. For example, among tools for organizing clinical data, staff reports of problem lists and medication lists were highly accurate, whereas reports of flow sheets, checklists, and tools for assessing patient motivation were less so. Among electronic functions, the accuracy of medication order entry reports was high, as were alerts on drug-drug interactions. Lead physicians tended to have more accurate reports than others among items concerning functions of electronic systems.

Table 4 (page 412) shows that self-report of systems for monitoring the needs of the patient populations had moderate to low agreement with the on-site audit. The highest agreement was found for an item that was not present in any medical group. Specifically, survey respondents accurately reported that practices did not have systematic monitoring to identify when patients filled prescriptions. Percent agreement for items concerning clinician reminder systems was low (Table 5, page 413), whereas some PPVs were high, suggesting that respondents were not aware of some elements of these systems that were frequently present in medical groups. Accuracy of reports for performance tracking systems tended to be moderate, on average, for lead physicians but low for other positions (Table 6, page 413). The specific elements of these systems were widely present among medical groups, leading to high PPVs.

As shown in Table 7 (page 414), all 11 medical groups had access to care management services on site or through a health plan or other organizations. The elements in this table had the lowest average percent agreement among all items examined. The combination of poor percent agreement and some high PPVs suggests that respondents were not reporting elements of this system that were generally present.

Table 8 (page 415) illustrates variation across medical groups in (1) the average agreement between survey responses and the on-site audit on the presence of practice systems and (2) the proportion of elements present within a practice system according to the on-site audit. For example, the average agreement between respondents and the audit of clinician reminder elements varied across medical groups from 37% to 80%. The proportion of nine possible clinician reminder elements found by the audit at each medical group ranged from 0% to 100%.

Discussion

Physicians and practice staff accurately reported on the presence of some clinical practice systems, including QI activities

Table 2. Validity of Self-Report of Practice Systems Compared with On-Site Audit, by Type of Respondent*

	Percent Agreement with On-Site Audit Mean (S.D.)		Positive Predictive Value of Self-Report (Compared with On-Site Audit) Mean (S.D.)	
	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)
Clinical Information Systems (2 components, 19 items)	71.2 (20.5)	66.0 (18.0)	74.8 (23.7)	72.2 (24.5)
Registry* (1 item)	78.1	69.3	85.0	79.2
Systematic Monitoring of Patient Population (2 components, 14 items)	65.6 (10.2)	64.0 (11.7)	55.3 (31.3)	49.6 (30.9)
Clinician Reminders (2 components, 9 items)	55.1 (13.2)	53.0 (12.0)	69.4 (19.9)	78.7 (19.8)
Performance Tracking and Feedback (4 components, 10 items)	69.3 (13.4)	53.7 (9.5)	91.5 (13.0)	90.9 (14.7)
Quality Improvement (1 item)	96.7	81.9	100.0	100.0
Care Management (2 components, 4 items)	40.9 (16.8)	33.9 (14.9)	88.9 (12.8)	84.5 (27.2)

* S.D., standard deviation.

Table 3. Validity of Self-Report of Clinical Information Systems Compared with On-Site Audit, by Type of Respondent*

		Percent Agreement with On-Site Audit		Positive Predictive Value (PPV) of Self-Report (Compared with On-Site Audit)	
	Number of Medical Groups with System, Based on On-Site Audit	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)
Tools for Organizing Clinical Data					
Problem lists	11	100.0	95.9	100.0	100.0
Medication lists	11	100.0	98.8	100.0	100.0
Checklists of tests or interventions for prevention or monitoring of chronic illness	8	56.3	72.5	58.1	86.5
Flow sheets completed during visits	3	22.6	28.0	20.0	21.8
Written records or printouts for patient symptom or treatment tracking	10	93.8	83.2	96.8	97.0
Assessment tool to assess patient interest in changing behavior	4	43.8	45.2	35.0	33.1
Clinical guidelines in patient care area	11	81.3	63.8	100.0	100.0
Component Mean (S.D.)		**71.1 (30.6)**	**69.6 (26.1)**	**72.8 (34.7)**	**76.9 (34.3)**
Functions of Electronic Systems					
Scheduling	10	68.8	83.7	83.3	89.4
Laboratory order entry	5	56.3	59.8	60.0	55.6
Medication order entry	6	84.4	81.9	88.9	92.2
Decision support to help M.D. choose meds based on effectiveness	5	71.0	54.7	75.0	53.6
Decision support to help M.D. choose meds based on efficiency	6	91.3	52.9	90.0	84.7
Alerts on drug-drug interactions	7	93.8	83.7	100.0	95.2
Alerts on disease-drug interactions	4	77.4	52.3	71.4	43.8
Alerts on abnormal test results that are clinically important	8	58.1	60.1	100.0	80.0
Search clinical info in visit notes such as BP	5	62.5	59.4	57.1	52.5
Identify patients or patient info based on a specific chronic condition	7	53.1	56.3	58.3	62.0
Identifying patients on a specific med	4	68.8	63.6	66.7	60.5
Searching for clinical guidelines	5	68.8	58.6	60.0	64.5
Component Mean (S.D.)		**71.2 (13.3)**	**63.9 (12.0)**	**75.9 (16.1)**	**69.5 (17.8)**

* S.D., standard deviation; M.D., physician; BP, blood pressure. Percent agreement and PPVs > 80.0 are highlighted.

Table 4. Validity of Self-Report of Systems for Systematic Monitoring of Patient Population and Patient Reminders Compared with On-Site Audit, by Type of Respondent*

		Percent Agreement with On-Site Audit		Positive Predictive Value (PPV) of Self-Report (Compared with On-Site Audit)	
	Number of Medical Groups with System, Based on On-Site Audit	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)
Practice has systematic monitoring for the following purposes:					
To identify when patients are due for preventive services	6	62.5	65.7	68.8	68.0
To identify when tests are needed for monitoring a chronic condition	7	62.5	61.3	68.2	77.0
To identify patients needing follow-up due to missed appointments	4	68.8	64.0	63.6	44.0
To identify when patients are due for follow-up visits	6	65.6	66.0	81.8	73.1
To identify when patients are due for medication renewals	0	71.0	81.7	0.0	0.0
To identify when patients fill Rx	0	95.7	88.7	0.0	0.0
To determine if treatment is consistent with guidelines	6	50.0	52.3	66.7	25.0
To track lab test report until reports return to clinic	9	68.8	53.0	86.4	85.5
To track radiology reports until reports return to clinic	9	59.4	40.8	84.2	77.8
To track consultation reports until reports return to clinic	2	65.6	70.0	0.0	7.8
Component Mean (S.D.)		**67.0 (11.7)**	**64.4 (14.0)**	**52.0 (36.7)**	**45.8 (34.8)**
Practice has systematic monitoring for the following conditions:					
Asthma	4	61.3	64.5	54.5	37.1
Cardiovascular disease	7	58.1	61.7	71.4	75.7
Depression	5	61.3	62.7	60.0	49.3
Diabetes	8	67.7	64.2	69.2	73.5
Component Mean (S.D.)		**62.1 (4.0)**	**63.3 (1.3)**	**63.8 (7.9)**	**58.9 (18.8)**

* S.D., standard deviation; Rx, prescription. Percent agreement and PPVs > 80.0 are highlighted.

and clinical information systems. However, the accuracy was low for other systems, such as care management. The variability in the accuracy of reporting based on domain, staff position, and medical group signals the need for additional documentation when information is used for accountability purposes (such as public reporting and pay-for-performance rewards).

Several factors may explain this variation as well as the rather limited agreement in some areas. First, the pattern of survey results and discussions with the auditors suggests that some systems were not implemented consistently across all practice sites within a group or among all clinicians within a site. For example, the auditors identified cases where electronic functions were demonstrated to be present in some sites but were not reported as being available by the medical group director or other staff. In other cases, individual physicians had developed their own clinical reminder systems that they used with their nursing staff, but these were not supported by the medical group or implemented across multiple sites. Second, the pattern of results suggests that some staff may have been more familiar with some systems than others, depending on the nature of their work. For example, lead physicians more accurately reported decision support systems for medication prescribing, whereas other staff more accurately reported on scheduling systems. Our observation here is consistent with other reports describing variations in the use of clinical systems. For example, Agrawal and Mayo-Smith noted that adherence to reminders in Department of Veterans Affairs (VA) sites was high (86.2%) overall but varied by clinic, clinician, and the content of the clinical reminder.[31]

Third, questionnaire wording and lack of understanding of

Table 5. Validity of Self-Report of Clinician Reminder Systems Compared with On-Site Audit, by Type of Respondent*

		Percent Agreement with On-Site Audit		Positive Predictive Value (PPV) of Self-Report (Compared with On-Site Audit)	
	Number of Medical Groups with System, Based on On-Site Audit	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)
Practice Has Clinician Reminders					
Alerts on preventive services needed	8	62.9	56.9	76.5	88.9
Alerts on tests needed for monitoring chronic condition	8	46.9	45.8	55.6	80.0
Alerts on drug-drug interactions	6	59.3	60.9	85.7	96.6
Alerts on disease-drug interactions	2	74.2	79.7	25.0	31.6
Alerts on abnormal test results of clinical importance	8	32.3	43.5	75.0	86.9
Component Mean (S.D.)		**55.1 (16.0)**	**57.4 (14.5)**	**63.6 (24.2)**	**76.8 (26.0)**
Asthma patients	7	56.7	53.0	70.0	73.1
Cardiovascular disease patients	9	40.0	40.4	71.4	84.5
Depression patients	7	56.7	51.5	70.0	70.8
Diabetes patients	10	66.7	45.6	95.2	96.3
Component Mean (S.D.)		**55.0 (11.1)**	**47.6 (5.8)**	**76.7 (12.4)**	**81.2 (11.7)**

* S.D., standard deviation. Percent agreement and PPVs > 80.0 are highlighted.

Table 6. Validity of Self-Report of Performance Tracking Systems Compared with On-Site Audit, by Type of Respondent*

		Percent Agreement with On-Site Audit		Positive Predictive Value (PPV) of Self-Report (Compared with On-Site Audit)	
	Number of Medical Groups with System, Based on On-Site Audit	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)
Practice collects or uses data to track clinic performance to compare with guidelines or indicators.	**11**	**87.5**	**71.2**	**100.0**	**100.0**
Practice reports results of performance tracking to clinicians.	**11**	**86.7**	**64.6**	**100.0**	**100.0**
Performance reports give information on the practice's performance on:					
Clinical processes	11	65.6	39.5	100.0	100.0
Clinical outcomes	11	81.3	59.6	100.0	100.0
Service data	9	53.1	45.2	87.5	84.3
Patient-reported experience of care	10	71.9	55.2	95.8	99.2
Component Mean (S.D.)		**68.0 (11.8)**	**49.9 (9.2)**	**95.8 (5.9)**	**95.9 (7.7)**
Performance reports give information on the performance of individual clinicians on:					
Clinical processes	8	59.4	48.5	94.1	91.8
Clinical outcomes	10	75.0	54.8	96.0	97.6
Service data	6	50.0	48.1	58.3	53.5
Patient-reported experience of care	7	62.5	50.6	83.3	83.0
Component Mean (S.D.)		**61.7 (10.3)**	**50.5 (3.1)**	**82.9 (17.3)**	**81.5 (19.6)**

* S.D., standard deviation. Percent agreement and PPVs > 80.0 are highlighted.

systems also appeared to have contributed to the limited validity of the self-report. Despite cognitive pretesting during survey development and use of some questions that had been previously tested in other settings, the lack of agreement with the audit suggests that some items were clearly not understood by those doing the surveys and was subsequently reflected in poor item

Table 7. Validity of Self-Report of Care Management Systems Compared with On-Site Audit, by Type of Respondent

		Percent Agreement with On-Site Audit		Positive Predictive Value (PPV) of Self-Report (Compared with On-Site Audit)	
	Number of Medical Groups with System, Based on On-Site Audit	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)	Lead Physicians at Medical Group or Practice Sites (*n* = 32)	Other Physicians and Nurses (*n* = 241)
Clinic offers or arranges clinical care management for patients with chronic illness.	**11**	**35.5**	**29.9**	**100.0**	**100.0**
Care management details					
Previsit planning offered to patients with chronic illness	11	28.1	17.5	100.0	100.0
After-visit follow-up offered to patients with chronic illness	6	65.6	53.3	77.8	44.0
Protocol-based referrals for special services offered to patients with chronic illness	8	34.4	35.0	77.8	94.1
Component Mean (S.D.)		**42.7 (20.1)**	**35.3 (17.9)**	**85.2 (12.8)**	**79.4 (30.8)**

* S.D., standard deviation. Percent agreement and PPVs > 80.0 are highlighted.

performance. For example, reports of laboratory order entry had only moderate agreement with the audit and similar PPVs, perhaps because this item could be interpreted to mean either that physicians entered orders directly or that staff entered orders electronically. "Ability to search on clinical information in visit notes" was intended to capture whether clinical information such as blood pressure was captured in a searchable field, but self-reported data on this item had poor accuracy.

Fourth, participating medical group staff members may have been unfamiliar or uncertain about the scope and meaning of some of the systems assessed. In particular, the concept of "systematic monitoring" appeared to be unfamiliar to many respondents, so the accuracy of reporting on these items was low. Indeed, most physicians and staff do not tend to think of managing their patients as a population but rather focus on managing care one-to-one within visits. We revised the language and explanations used in the subsequent versions of the PPC-RS to explore alternative phrasing and other wording changes.

One of the subgoals of this study was to determine which staff person could most accurately report about practice systems. The results demonstrate that data collection from any single source will have limitations. Relying on responses from the lead physician (medical director or lead physician at a practice site) appears to be the best approach because fewer surveys need to be completed, and their reports were somewhat more valid than other sources for most questions. However, explicit documentation of systems capability, functioning, and implementation appears to be needed.

Limitations

Although this study contributes to our understanding of the accuracy and utility of self-reports, a number of limitations should be noted. The findings suggest that health care personnel are more likely to underreport rather than overreport systems; however, the extent of over- and underreporting may differ in other settings. For example, overreporting might occur when financial incentives are attached to the demonstration of increased "systemness." Furthermore, the medical groups in this study are engaged in a community QI collaborative, making it more likely that they may have adopted a higher level of systems (for example, one third had electronic health records). Their leaders and other personnel may be more aware of systems than in other regions. This may have some influence on the accuracy of reporting, but it is not clear if the effect would be positive or negative.

Implications for Policy

Overall, these results raise a number of issues regarding the use of self-reports of clinical practice systems. The agreement of self-reports with an on-site audit is generally good for some systems but not others. For this reason, self-reported information on clinical practices systems *should not be* used for accountability purposes, including pay-for-quality efforts or public reporting unless additional documentation is required to ensure fair comparisons. As these reports are used in ways that can affect financial or public standing, there will be increasing incentives to report systems that are not present or not consistently used.

In this research study, on-site audits were used to verify self-

Table 8. Validity of Self-Report of Practice Systems Compared with On-Site Audit, by Medical Group

	Medical Groups										
	A	B	C	D	E	F	G	H	I	J	K
Number of respondents	43	24	10	11	30	18	27	38	28	34	10
Clinical Information Systems											
Mean % agreement	65.6	73.1	67.3	64.1	60.9	72.3	60.2	61.6	75.9	69.6	75.7
% items found in on-site audit	72.2	38.9	52.6	75.0	68.4	89.5	73.7	73.7	100.0	27.8	64.3
Registry											
Mean % agreement	65.1	70.8	100.0	72.7	56.7	100.0	55.6	81.6	85.7	55.9	50.0
% items found in on-site audit	100.0	0.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	0.0	100.0
Systematic Monitoring of Patient Population and Patient Reminders											
Mean % agreement	77.4	54.6	62.3	68.6	58.9	69.8	38.9	68.2	70.8	69.3	52.9
% items found in on-site audit	0.0	64.3	64.3	15.4	64.3	71.4	84.6	50.0	53.8	14.3	45.5
Clinician Reminders											
Mean % agreement	56.5	57.4	37.3	85.3	6.7	80.0	37.9	47.8	71.3	68.9	58.3
% items found in on-site audit	50.0	55.6	100.0	0.0	100.0	100.0	66.7	88.9	88.9	22.2	85.7
Performance Tracking and Feedback											
Mean % agreement	62.9	50.9	81.0	44.6	53.7	77.4	23.1	54.9	52.3	59.1	75.9
% items found in on-site audit	80.0	90.0	100.0	70.0	70.0	80.0	100.0	100.0	100.0	100.0	62.5
Quality Improvement											
Mean % agreement	88.1	100.0	100.0	90.9	46.4	100.0	84.0	86.5	69.2	100.0	44.4
% items found in on-site audit	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Care Management											
Mean % agreement	39.7	53.1	30.0	50.0	11.7	78.0	15.5	40.8	10.7	35.5	36.4
% items found in on-site audit	75.0	50.0	75.0	0.0	100.0	100.0	100.0	75.0	100.0	75.0	100.0

reports; however, it would be intrusive and prohibitively expensive to conduct such audits on a routine basis. Requiring detailed documentation, combined with the potential for a small proportion of random on-site audits, represents a more feasible approach for ensuring the accuracy of self-reported information on clinical practice systems than routine audits. NCQA created a Web-based tool for data collection that allows the respondent to attach specific types of documentation to support each response.[32] A growing number of initiatives are using the PPC-RS and the 2008 Physician Practice Connections Patient-Centered Medical Home Version (PPC-PCMH), which was adapted with input from four primary care physician specialties societies for qualifying practices as patient-centered medical homes in demonstration projects.

Although assessment may be a necessary step, practices that are trying to introduce new systems have a major and largely unmet need for educational programs, tools, and other remote and on-site assistance. Efforts to improve awareness of practice systems and their application in QI are being implemented in a variety of areas. For example, for maintenance of certification for internists, the American Board of Internal Medicine (ABIM) requires completion of a practice systems survey, which is virtually identical to the PPC-RS and was developed in conjunction with NCQA. The ABIM Web-based survey includes links to explanatory information and evidence-based literature, which may enhance the accuracy of its responses.[33] The American Academy of Family Physicians' Metric Program also assesses practice systems and provides continuing education for physicians related to enhancing those systems.[34]

Implications for Measurement and Research

More research is needed to improve the measurement of clinical practice systems. First, additional measures that capture all aspects of the CCM, including community involvement, are needed. Second, efforts should be made to understand which practices systems are most strongly related to improvements in quality and outcomes. This information could be used to reduce the burden of survey and documentation. Third, the methods of data collection warrant more attention. It remains a challenge to gather information about complex topics such as practice systems through a relatively brief self-report survey. The pretest version of our survey was much longer, with information about the intent of systems and explanation of terms. To reduce burden on respondents, we eliminated much of this explanatory information and tried to make the individual questions as detailed and practical as possible. Judging from our validity results and discussions with participants, more explanatory information may be needed to ensure accurate responses (as is present in the web-based PPC-RS tool). However, gaining participation from busy clinicians is difficult, even when a modest honorarium is included. Physicians appear to be most cooperative when they are personally approached by a respect-

ed member of their community or when they find the survey content to be directly informative to ongoing pay-for-performance activities. In cases where these types of incentives are not present, increasing the amount of the honorarium or offering multiple modes for data collection may improve response rates, but the main challenge appears to be getting physicians to believe that the information is useful. J

The study reported in this article was supported by Grant #048908 from the Robert Wood Johnson Foundation.

Sarah Hudson Scholle, M.P.H., Dr.P.H., is Assistant Vice President, Research, National Committee for Quality Assurance (NCQA), Washington, D.C., and **L. Gregory Pawlson, M.D., M.P.H.,** is Executive Vice President, NCQA. **Leif I. Solberg, M.D.,** is Associate Medical Director for Care Improvement Research, HealthPartners Research Foundation (HPRF), Minneapolis. **Sarah C. Shih, M.P.H.,** formerly Research Scientist, NCQA, is City Research Scientist, New York City Department of Health and Mental Hygiene. **Stephen E. Asche, M.A.,** is Statistician, HPRF. **Ann F. Chou, Ph.D.,** formerly Research Scientist, NCQA, is Assistant Professor, Health Administration & Policy, College of Public Health & College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City. **Merry Jo Thoele, M.P.H.,** is Research Project Manager, HPRF. Please address correspondence to Sarah Hudson Scholle, Scholle@ncqa.org.

References

1. McGlynn E.A., et al.: The quality of health care delivered to adults in the United States. *N Engl J Med* 348:2635–2645, Jun. 26, 2003.
2. Institute of Medicine: *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington, DC: National Academy Press, 2001.
3. Von Korff M., et al.: Collaborative management of chronic illness. *Ann Inter Med* 127:1097–1102. Dec. 15, 1997.
4. Bodenheimer T., Wagner E.H., Grumbach K.: Improving primary care for patients with chronic illness. *JAMA* 288:1775–1779, Oct. 9, 2002.
5. Bodenheimer T., Wagner E.H., Grumbach K.: Improving primary care for patients with chronic illness: The Chronic Care Model, Part 2. *JAMA* 288:1909–1914, Oct. 16, 2002.
6. Wells K.B., et al.: Impact of disseminating quality improvement programs for depression in managed primary care: A randomized controlled trial. *JAMA* 283:212–220, Jan. 12, 2000. Erratum in: *JAMA* 283:3204, Jun. 28, 2000.
7. Piatt G.A., et al.: Translating the chronic care model into the community: Results from a randomized controlled trial of a multifaceted diabetes care intervention. *Diabetes Care* 29:811–817, Apr. 2006.
8. Katzelnick D.J., et al.: Applying depression-specific change concepts in a collaborative breakthrough series. *Jt Comm J Qual Patient Saf* 31:386–397, Jul. 2005.
9. Improving Chronic Illness Care: *Evidence of Better Care.* http://www.improvingchroniccare.org/index.php?p=Evidence_for_Better_Care&s=5 (last accessed May 28, 2008).
10. Renders C.M., et al.: Interventions to improve the management of diabetes in primary care, outpatient, and community settings: A systematic review. *Diabetes Care* 24:1821–1833, Oct. 2001.
11. Fleming B., et al.: The relationship between organizational systems and clinical quality in diabetes care. *Am J Manag Care* 10:934–944, Dec 2004.
12. Hung D.Y., et al.: Rethinking prevention in primary care: Applying the Chronic Care Model to address health risk behaviors. *Milbank Q* 85:69–91, Mar. 2007.
13. Kawamoto K., et al.: Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ* 330:765, Apr. 2, 2005.
14. McMullin S.T., Lonergan T.P., Rynearson C.S.: Twelve-month drug cost savings related to use of an electronic prescribing system with integrated decision support in primary care. *J Manag Care Pharm* 11:322–324, May 2005.
15. Solberg L.I., et al.: Practice systems for chronic care: Frequency and dependence on an electronic medical record. *Am J Manag Care* 11:789–796, Dec. 2005.
16. Berlin A., Sorani M., Sim I.: Characteristics of outpatient clinical decision support systems: A taxonomic description. *Medinfo* 11(pt. 1):578–581, 2004.
17. Integrated Healthcare Association: *Pay For Performance Overview.* http://www.iha.org/P4POVIEW.htm (last accessed May 22, 2008).
18. Bridges to Excellence: *Bridges to Excellence Programs.* http://www.bridgestoexcellence.org/Content/ContentDisplay.aspx?ContentID=23 (last accessed Jun. 2, 2008).
19. Baker G., Carter B.: *Provider Pay-for-Performance Incentive Programs: 2004 National Study Results.* 2005. http://www.bridgestoexcellence.org/Content/ContentDisplay.aspx?ContentID=4 (last accessed May 22, 2008).
20. Casalino L., et al.: External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases. *JAMA* 289:434–441, Jan. 22–29, 2003.
21. Burt C.W., Sisk J.E.: Which physicians and practices are using electronic medical records? Survey data show limited use of these information tools. *Health Aff (Millwood)* 24:1334–1343, Sep.–Oct., 2005.
22. Pande P.S., et al.: *The Six Sigma Way: How GE, Motorola, and Other Top Companies Are Honing Their Performance.* New York City: McGraw-Hill, 2000.
23. Schmittdiel J., et al.: The use of patient and physician reminders for preventive services: Results from a national study of physician organizations. *Prev Med* 39:1000–1006, Nov. 2004.
24. Bodenheimer T., et al.: What are the facilitators and barriers in physician organizations' use of care management processes? *Jt Comm J Qual Saf* 30:505–514, Sep. 2004.
25. Greenfield S., et al.: Profiling care provided by different groups of physicians: Effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med* 136:111, Jan. 15, 2002.
26. Greenfield S., Kaplan S.H.: Creating a culture of quality: The remarkable transformation of the Department of Veterans Affairs health care system. *Ann Intern Med* 141:316–388, Aug. 17, 2004.
27. Stroebel R.J., et al.: A randomized trial of three diabetes registry implementation strategies in a community internal medicine practice. *Jt Comm J Qual Improv* 28:441–450, Aug. 2002.
28. East J., et al.: Impact of a diabetes electronic management system on patient care in a community clinic. *Am J Med Qual* 18:150–154, Jul.–Aug. 2003.
29. Farley D.O., et al.: *Regional Health Quality Improvement Coalitions: Lessons Across the Life Cycle.* Oct. 2003. http://www.bridgestoexcellence.org/Content/ContentDisplay.aspx?ContentID=4 (last accessed May 22, 2008).
30. Dillman D.A.: *Mail and Internet Surveys: The Tailored Design Method.* New York City: John Wiley and Sons, Inc., 2000.
31. Agrawal A., Mayo-Smith M.F.: Adherence to computerized clinical reminders in a large healthcare delivery network. *Medinfo* 11(pt. 1):111–114, 2004.
32. National Committee for Quality Assurance: *Physician Practice Connections.* http://www.ncqa.org/ppc/ (last accessed May 22, 2008).
33. American Board of Internal Medicine: *Practice Performance Options.* http://www.abim.org/pims/details/sepp.aspx (last accessed May 22, 2008).
34. American Academy of Family Practitioners: *METRIC: Measuring, Evaluating and Translating Research Into Care.* http://www.aafp.org/online/en/home/cme/selfstudy/metric.html (last accessed May 22, 2008).