Arthritis Care & Research
Vol. 68, No. 10, October 2016, pp 1402–1409
DOI 10.1002/acr.22985

SPECIAL ARTICLE

Methods for Developing the American College of Rheumatology's Electronic Clinical Quality Measures

JINOOS YAZDANY,[1] RACHEL MYSLINSKI,[2] AMY MILLER,[2] MELISSA FRANCISCO,[2] SONALI DESAI,[3] GABRIELA SCHMAJUK,[1] DIANE LACAILLE,[4] CLAIRE E. BARBER,[5] CATALINA OROZCO,[6] MATTHEW BUNYARD,[7] MARTIN J. BERGMAN,[8] MURRAY PASSO,[9] ERIC L. MATTESON,[10] RICHARD OLSON,[11] STUART SILVERMAN,[12] ROBERT WARREN,[13] KAMALA NOLA,[14] AND MARK ROBBINS[15]

Introduction

The widespread use of electronic medical records (EMRs) is creating new opportunities to reimagine quality measurement and improvement in rheumatology. Previous quality measurement efforts have relied on analyses of administrative billing claims, such as the National Committee on Quality Assurance disease-modifying antirheumatic drug measure (1), or on manual chart review, such as the American Medical Association's Physician Consortium for Performance Improvement's (PCPI) Physician Quality Reporting System measures (2). While there is some evidence that these programs have resulted in quality improvement, many have argued that current performance assessments have limited utility and do not support building a continuously learning health care system, as envisioned by the Institute of Medicine (1–3).

The limitations of current measurement approaches have led to calls for new systems to make quality measurement more efficient and meaningful. Rather than relying on retrospective assessments of care, newer measures should be fully integrated into clinical workflows and results available in real-time to clinicians (4). Performance information should be tailored to individual practices and health systems to allow flexibility and innovation in local quality improvement efforts. At the same time, aggregate data across practices should also be available to allow for benchmarking. The implementation of EMRs has created new infrastructure to begin to advance these goals. Electronic clinical quality measures (eCQMs) are a new approach to measurement that automatically extracts information from EMRs, potentially allowing timely generation of performance data and increasing the efficiency of data collection for quality improvement (5,6).

In this paper, we discuss the methodologic approach recommended by the American College of Rheumatology (ACR)

Supported by the American College of Rheumatology. Dr. Yazdany's work was supported by the Robert L. Kroc Chair in Rheumatic and Connective Tissue Diseases (I) (AHRQ-R01-HS024412) and the Russell/Engleman Medical Research Center for Arthritis.

[1]Jinoos Yazdany, MD, MPH, Gabriela Schmajuk, MD, MS: University of California, San Francisco; [2]Rachel Myslinski, MBA, Amy Miller, Melissa Francisco, MBA: American College of Rheumatology, Atlanta, Georgia; [3]Sonali Desai, MD, MPH: Brigham & Women's Hospital, Boston, Massachusetts; [4]Diane Lacaille, MD, FRCPC, MHSc: Arthritis Research Centre of Canada, Vancouver, British Columbia, Canada; [5]Claire E. Barber, MD, PhD, FRCPC: University of Calgary, Calgary, Alberta, Canada; [6]Catalina Orozco, MD: Rheumatology Associates, Dallas, Texas; [7]Matthew Bunyard, MD: Cleveland Clinic, Cleveland, Ohio; [8]Martin J. Bergman, MD: Drexel University College of Medicine, Philadelphia, Pennsylvania; [9]Murray Passo, MD: Children's Hospital, Medical University of South Carolina, Charleston; [10]Eric L. Matteson, MD, MPH: Mayo Clinic College of Medicine, Rochester, Minnesota; [11]Richard Olson, MD: Rockford Orthopedic Associates, Rockford, Illinois; [12]Stuart Silverman, MD: Osteoporosis Medical Center, Beverly Hills, California; [13]Robert Warren, MD, PhD, MPH: Medical University of South Carolina, Charleston; [14]Kamala Nola, PharmD, MS: Lipscomb University College of Pharmacy, Nashville, Tennessee; [15]Mark Robbins, MD, MPH: Harvard Vanguard Medical Associates, Atrius Health, Somerville, Massachusetts.

Dr. Bergman has received consultant fees, speaking fees, and/or honoraria from Amgen, Iroko, BMS, Horizon, Johnson & Johnson, and Novartis (less than $10,000 each), and AbbVie, Celgene, Genentech, and Pfizer (more than $10,000 each), and holds stock or stock options in Merck, Johnson & Johnson, and Pfizer. Dr. Passo has received consultant fees, speaking fees, and/or honoraria from Pfizer and Cincinnati Children's Hospital (less than $10,000 each). Dr. Matteson has received advisory fees from GlaxoSmithKline (less than $10,000).

Address correspondence to Jinoos Yazdany, MD, MPH, Division of Rheumatology, University of California, San Francisco, Box 0920, San Francisco, California 94143-0920. E-mail: jinoos.yazdany@ucsf.edu.

Submitted for publication November 20, 2015; accepted in revised form June 29, 2016.

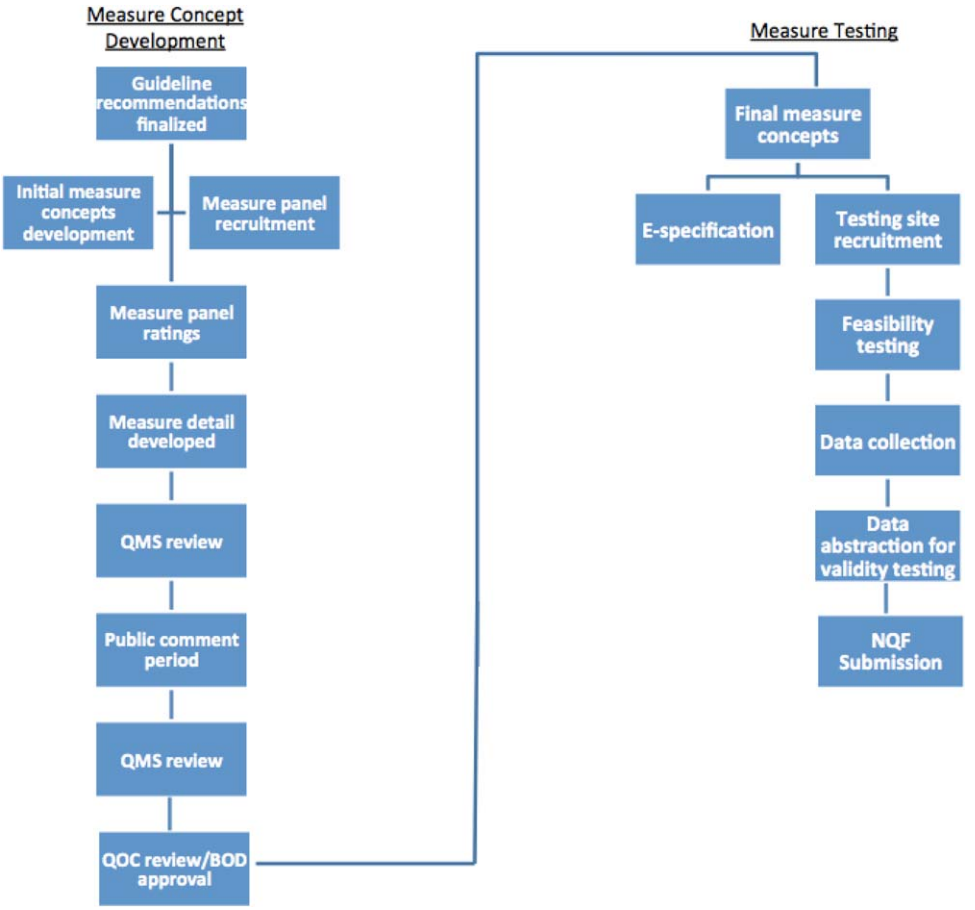

Figure 1. Figure depicting phases of the American College of Rheumatology's electronic clinical quality measure development program. QMS = Quality Measures Subcommittee; NQF = National Quality Forum; QOC = quality of care; BOD = board of directors.

to develop new eCQMs in priority areas to fulfill these goals. The ACR's eCQM development program involves seven phases, including prioritization, measure conceptualization, interdisciplinary consensus ratings, public comment, electronic specification, eCQM field testing, and submission for national endorsement.

Methods for electronic clinical quality measure development

The ACR's Quality Measures Subcommittee worked for several years to develop a detailed process for eCQM development (Figure 1). Below we outline how the Subcommittee's recommendations were applied to advance the first ACR eCQM development projects. The specific data for each project will be presented in separate publications.

In devising methods for eCQMs, the Subcommittee sought to align the proposed ACR process with national standards outlined by the National Quality Forum (NQF) and the American Medical Association's PCPI, but also to tailor methods to align with ACR policies and priorities (7). We recommend seven phases for measure development, described below.

Prioritization of topics. To identify priority areas for quality measure development, the ACR previously convened a working group led by Saag et al, consisting of experts in measure development and clinical rheumatology (8). Using a combination of formal consensus processes and literature reviews, the group recommended areas for future ACR quality measure development. Adult conditions with the highest rankings were rheumatoid arthritis (RA), osteoporosis, and gout. The ACR's Quality Measures Subcommittee considered these recommendations along with priorities identified by the NQF's Measure Prioritization Advisory Committee (9). This NQF committee identified priorities for measure development as well as gaps in the national quality measurement portfolio. Among the high-impact Medicare conditions examined, both RA and osteoporosis were ranked among the top 20 national priorities for measure development (9). Although gout was not identified as a Medicare priority, the subcommittee recommended that gout be considered a priority condition given demonstrated gaps in quality of care (10–12).

The Quality Measures Subcommittee also recommended that quality measure development efforts should align, when possible, with ACR clinical practice guidelines. This will allow measure developers to leverage the rigorous scientific evidence reviews and consensus processes in guideline development, which are often crucial for national endorsement.

With these considerations and priorities in mind, the first three areas chosen for ACR quality measure development were RA, glucocorticoid-induced osteoporosis, and

gout. These areas had been the focus of major ACR guideline efforts (13–15), were priority areas identified in the ACR white paper on quality measurement (8), and two of the areas (RA, osteoporosis) were also national priorities (9). Below we outline the different phases of eCQM development recommended by the Subcommittee, and frame these recommendations by discussing the methods used in the first ACR eCQM projects.

Measure conceptualization. Once priorities are identified, the next step in eCQM development is conceptualization of the measure. For example, in the initial priority areas, the ACR convened separate working groups for each of the prioritized topic areas to draft preliminary measure concepts. Each working group included a member who was chair of the Quality Measures Subcommittee, at least one rheumatologist who was involved in relevant ACR guideline development, and methodologic experts in measure development. The work groups drafted and refined potential measures in an iterative manner over a period of months, including reviewing existing measures and developing new measures derived from ACR guidelines. Consideration was given to whether potential quality measures 1) reflected strong scientific evidence based on review of ACR guidelines, 2) had clear and auditable actions, 3) were under the control of rheumatologists or rheumatologic practices, 4) were feasible and meaningful in clinical practice, 5) had documented care gaps or expert agreement on opportunities for improvement, 6) did not overlap with existing measures, and 7) would advance the field based on lessons learned from previous applications of related measures in reporting programs.

For the initial priority areas, both process measures (i.e., what clinicians do in providing care) and outcome measures (i.e., health outcomes that result from care) were considered.

Interdisciplinary consensus ratings. The purpose of interdisciplinary consensus ratings is to formally assess the face validity of eCQM concepts with a panel of experts. The ACR Subcommittee recommended that a formal group process is used; an overview of the group processes used in the first eCQM projects is provided as an example below.

Nomination process. For the initial eCQM priority areas, measure validity was assessed using a validated consensus process (16). Nominations for an interdisciplinary expert panel were sought from the ACR, Association of Rheumatology Health Professionals, national quality organizations, payers, and members of other relevant professional societies (e.g., American College of Physicians, American Society for Bone and Mineral Research, American Academy of Orthopedic Surgery). Care was taken to include rheumatologists in clinical practice and those with content expertise in the topic area.

Conflicts of interest. The ACR has established policies to manage conflicts of interest. For example, for the initial eCQM projects, both the chair of the expert panel and the majority of its members (≥50%) had no financial conflicts of interest with any drug, device, or other product made for the management of the condition; remaining members were required to fully disclose all relationships, including potential conflicts of interest, to ensure transparency.

Expert panel meetings and rating methods. For the initial eCQM projects, expert panel members participated in a webinar introducing the project in which they were involved. Members were provided a summary of relevant quality measures currently used in public and private reporting programs, as well as drafted measures derived from recently published ACR guidelines. For each quality measure, materials were distributed that summarized the measure concept, referenced the relevant evidence in the ACR guideline, and outlined the measure numerator (specific clinical action required by the measure), denominator (eligible cases for a measure), and potential exclusions. In addition, a summary of existing analogous measures in national and other reporting programs (e.g., Physician Quality Reporting System, Healthcare Effectiveness Data and Information Set) and current gaps in care based on these reporting programs were provided.

A modification of the RAND/UCLA Appropriateness Method was employed to rate the preliminary set of possible quality measures from the work group (16,17). Briefly, this approach entailed two rounds of anonymous ratings on a standardized scale, with the voting conducted via an online survey mechanism. Between these rounds, there was a moderated discussion. For each round, panel members rated quality measures on two scales, one for validity and one for feasibility (range 1–9 for both, where 1 = not valid or not feasible and 9 = definitely valid or definitely feasible). For validity, panelists were instructed to consider the following questions: 1) Is there adequate scientific evidence or professional consensus to support the measure? 2) Are there identifiable health benefits to patients who receive care specified by the measure? 3) Based on your professional experience, would you consider physicians with significantly higher rates of adherence to the measure higher quality providers? and 4) Are the majority of factors that determine performance on the measure under the control of the physician? For feasibility, panelists were asked to consider the following questions: 1) Can the measure be interpreted for use in the typical clinical setting? 2) Can the measure be integrated into existing workflows and health information systems to collect, manage, and manipulate the required data elements? and 3) Can this aspect of care be measured with reasonable cost and level of effort?

Round 1 ratings were compiled and sent to panelists approximately one week before a second webinar, allowing them to compare their responses with the ratings of their colleagues. During the second webinar, the moderator used the ratings to guide discussion, focusing on areas with greatest disagreement. No attempt was made to force the panel to consensus; instead, the discussion sought to determine whether divergent ratings resulted from real clinical disagreement, or simply reflected different interpretations of the measures. After incorporating suggested revisions, expert panel members again anonymously rated the validity and feasibility of each revised measure.

Statistical analysis. For the initial projects, the statistical analysis plan for measure ratings adhered to published methods; measures with a median validity rating of ≥7 and without disagreement were considered valid. Disagreement was defined as a specified number of panelists rating the indicator in the highest tertile range (7–9), while others rated it in the lowest tertile (16). Measures with a median feasibility rating of ≥4 were considered potentially feasible. Only measures rated as both valid and potentially feasible progressed to the next phase of each project.

Public comment and ACR Committee review. The Subcommittee recommended that measure concepts that are found to have face validity through an expert consensus process are distributed to the ACR membership for public comment and also reviewed by appropriate ACR committees. For example, in the initial eCQM projects, measures that were rated as valid and potentially feasible by the expert panel were forwarded to the Quality of Care Committee for final review before a public comment period. The public comment period entailed posting the selected measures on the ACR web site, solicitation of ACR member feedback through e-mail announcements, and letters to relevant professional societies and stakeholder groups.

After the public comment period, the work group reviewed each comment and considered revisions to the measures to improve clarity or modify content. Measures were then sent to the ACR Board of Directors for final approval.

Electronic specification. For quality measures using EMR-derived data, additional steps to construct a feasible and reliable measure are required. EMR-derived measures, or eCQMs, are specified in the Health Quality Measures Format (HQMF). HQMF is a standard for electronic documentation, and ensures consistency in the measure's structure, metadata, definitions, and logic (18). Below we outline ACR recommendations regarding the specification of ACR eCQMs, using the methods employed in the initial measure projects as examples.

Measure-authoring tool (MAT) and quality data model (QDM). The MAT is software maintained by the Centers for Medicaid and Medicare Services (CMS) to develop quality measures in electronic format. For the initial eCQM projects, measure concepts were written in a format consistent with the QDM in the MAT. The QDM is an information model that includes electronic data standards to allow consistent interpretation of quality measures for EMR data abstraction. QDM elements have a category, the datatype in which that category is used, and then a value set of specific codes that are used to identify the element. In addition, QDM elements also have related attributes (known as metadata), which provide additional information about the data element (Figure 2). For example, a category may refer to the fact that the data element is a medication or a condition. The value set then defines the specific instance of the category by assigning a set of values, e.g., standard drug codes. All QDM elements also have attributes, such as timing (e.g., time of occurrence and stop or start times) (19).

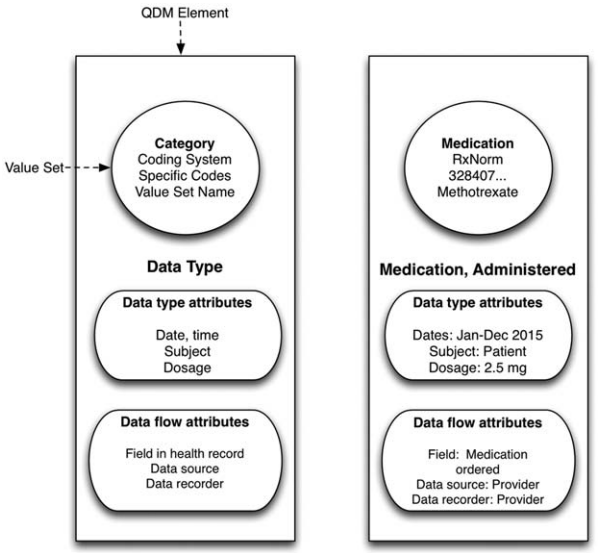

Figure 2. Example of a quality data model (QDM) element. Each QDM element is comprised of the components depicted on the left. A specific example of a QDM element is provided on the right. QDM elements have a category, the data type in which that category is used, and then a value set of specific codes that are used to identify the element. In addition, QDM elements have specific attributes (metadata), such as the data type, and they also indicate the flow of data into the electronic health record. This figure was modified from one included in public guidance from the National Quality Forum (with permission) (26).

For the initial eCQM projects, a physician with professional expertise in health informatics then used specified QDM elements to elaborate all possible value sets available in current electronic health records, including International Classification of Diseases, Ninth and Tenth Revision (ICD-10), Systematized Nomenclature of Medicine-Clinical Terms, Logical Observation Identifiers Names and Codes, Current Procedural Terminology, Healthcare Common Procedure Coding System, and RxNorm. With the full list of possible value sets, two physicians in each working group independently reviewed the items associated with each QDM element, assessing their clinical appropriateness for inclusion. After this independent review, differences were adjudicated through discussion between the two reviewers and final value sets were agreed upon.

Based on the measure specifications in the QDM as well as the adjudicated value sets, ACR staff used the MAT to further specify each eCQM. For each project, investigators defined population criteria, including defining the initial patient population (e.g., based on encounters, patient characteristics, risk category/assessment), the denominator, denominator exceptions, numerators, and exclusions. These criteria can be turned into corresponding Health Level 7 Continuity of Care Document modules for application in EMRs. Completed eCQMs then progressed to field testing.

eCQM field-testing. Field testing of eCQMs is critical to establish the feasibility of accurate data capture and validity of measure specifications. The ACR Subcommittee

recommended that each eCQM is tested in at least three different EMRs to assess these parameters given significant heterogeneity in data systems used in current rheumatology practices. This is also the minimum number currently required for testing eCQMs by the NQF.

Clinical sites. For the initial eCQM projects, at least three clinical sites, each using distinct EMR or data systems, were selected for each project. In addition, clinical sites needed to have some experience in establishing workflows to electronically capture required information and the necessary health information technology support staff to build and test the eCQM locally.

Feasibility. Because the application of eCQMs in the health care system is a new endeavor, feasibility assessments occur on a continuum. For example, some practices have long-established workflows to capture specific data elements (e.g., functional status score) in a structured field that is easily queried. These practices often have mature EMR systems and have already implemented eCQMs locally for the purposes of quality improvement. Other practices may have newly implemented an EMR, or may have instituted more recent changes to clinical workflow. These practices may capture required information less consistently. The purpose of feasibility assessment is to capture this continuum when possible and to allow adjustments to eCQMs prior to formal testing of reliability and validity (20).

Feasibility in the initial eCQM projects was assessed by asking clinical sites to complete a structured survey, using methods developed by the American Medical Association's PCPI and used in previous national electronic measure development work. Sites were asked to assess the following: 1) availability of specific data elements in a structured format, including data accuracy (i.e., correctness) and completeness, 2) data standards (i.e., the extent to which data elements are coded using nationally accepted health information technology standards, including definitions and value sets, 3) workflows currently in place to capture the measure, 4) measure logic (i.e., Did the measure specifications make sense clinically?), and 5) measure aggregation and reporting (i.e., whether performance can be aggregated and reported in the practice). For specific data elements, a rating scale created by the NQF was used for each of these elements (3 = data element exists in structured format; 2 = data element is required for certified EMR, but is not available in structured format in the practice's EMR; and 1 = data element is not currently required for certified EMRs) (20). Summary scores reflecting means on the rating scales across the domains assessed were generated, along with aggregation of free text comments provided by clinical sites to allow determination of whether measures met the a priori criteria for feasibility. The intent of the feasibility assessment is to allow adjustments to eCQMs prior to formal testing of reliability and validity (20).

Reliability. Unlike data abstraction from manual chart review, data extracted from EMRs using automated algorithms is inherently reliable (i.e., accurately repeatable). Therefore, operationalizing an eCQM that includes only data elements from the QDM (or new data elements that are submitted for inclusion in the model) ensures reliability. However, because specific data elements for eCQMs can be inaccurate (i.e., do not accurately reflect care provided), a major focus of testing is on validity (7,21).

Concurrent validity. The focus of validity testing was concurrent validity, or whether the information from the computer-generated EMR data pull was similar to the information that a physician abstractor obtains by reading the front-end of the EMR. Examples of how these parameters were assessed in the initial eCQM projects are provided below.

Assessment of parameters. *Local measure construction.* To assess validity, clinical sites first worked with local health information technology staff to build the eCQM extraction algorithms from the EMR. This required review of the HQMF files containing eCQM specifications, including measure background information, required data elements, measure logic and measure calculation instructions, human-readable formats of the measure in html, as well as a Microsoft Excel spreadsheet with value sets (i.e., code sets) for each measure.

Sample size calculation, sample data set construction, and statistical analysis. Once the automated extraction algorithm was identified, a sample size calculation for data element validity was performed to identify a subset of charts for manual EMR review (i.e., reviewing data through the EMR's user interface). A calculation was made to determine what sample size would be necessary to calculate interrater reliability. A simple random sample of patients was then chosen for analysis. The statistical measure of agreement in these analyses was the kappa coefficient, a measure of interrater agreement. Cut points that are commonly accepted were used to assess the degree of agreement (22).

Data extraction. Using a structured data collection form, a manual EMR chart abstraction was performed by reviewers at each site to compare data elements electronically extracted and those manually abstracted from the EMR. Agreement was also assessed between performance scores obtained through both electronic and manual abstraction.

Submission for national endorsement. Because an explicit goal of ACR eCQM development is to contribute toward a coherent and uniform performance measurement strategy for US rheumatologists, an important priority is to submit selected measures for national endorsement. The NQF is currently the legislatively authorized consensus-based endorsement entity of quality measures in the US. Comprising many health care stakeholders, the NQF endorses consensus standards for performance measurement. As part of this effort, the NQF convenes working groups comprising representatives from both public and private sectors to evaluate quality measures. Harmonization of measures across the health care system and across public and private sectors, often through work done by the Measures Application Partnership (a multi-stakeholder partnership that guides the US Department of Health and Human Services on the selection of performance measures for federal health programs), is also a key NQF priority.

Measures submitted to the NQF undergo rigorous review in multiple phases (7). Measures can be submitted for full endorsement, or more recently, eCQMs are also eligible for trial endorsement until complete testing is performed. Below, we review the processes followed to date to achieve national endorsement for ACR-developed eCQMs.

NQF standing committee review. For some measures (i.e., RA and gout), eCQMs were presented to the appropriate NQF standing committees comprised of diverse health care stakeholders (e.g., patients, clinicians, health system leaders, payers, industry leaders) for review against NQF consensus standards. Criteria for evaluation (for both standard measures and eCQMs) include the following: 1) importance to measure and report, 2) scientific acceptability of measure properties, 3) feasibility, 4) usability and use, and 5) related and competing measures. Each of these criteria had specific subcomponents that required committee voting for assessment of consensus. For example, for the criteria regarding importance to measure and report, committee members considered the following: 1) scientific evidence to support the measure, 2) evidence of a performance gap or disparities in care, and 3) content in a high priority area (e.g., in an area that is a specific national health goal/priority, or likely to be high impact on health) (7).

NQF national public comment period. Measures approved by the NQF standing committees for full or trial endorsement were posted for national public comment. The ACR was asked to formally respond to public comments and clarify or revise measures as appropriate.

Consensus Standards Approval Committee and NQF Board of Directors approval. NQF staff members summarized the measures, the standing committee reviews, and the public comments in a publically available report. The report was then presented to NQF leadership for final endorsement.

Discussion

More than a decade has passed since the first rheumatology quality measures were applied in the US health care system (1). The science of quality measurement, including both technical methods and measurement infrastructure, has evolved considerably since then. While the focus of early efforts was often to use readily accessible data such as administrative claims to assess performance on what were often minimal standards of care, current efforts are shifting to focus on clinical outcomes and improving both provider and health system performance (3). Increasing the clinical relevance and usefulness of quality measures, decreasing measurement burden, and promoting local innovation are important goals for the next decade of quality measurement. To support this new paradigm, the ACR has developed methods to create and maintain valid and feasible quality measures in rheumatology.

ACR measure development methods outlined here leverage the growing infrastructure of EMRs, facilitated by the Health Information Technology for Economic and Clinical Health Act and enacted as part of the American Recovery and Reinvestment Act of 2009. Recent estimates suggest that more than 80% of US rheumatologists now use EMRs. This enables a new type of quality measure, an eCQM, that uses the full breadth of information available in EMRs. Although eCQMs require a significant upfront investment in development, testing, and deployment, they have potential to decrease measurement burden over time for practicing clinicians (5). Moreover, generating reports at both local and national levels through the use of aggregated patient registries has started to facilitate rapid cycle quality improvement efforts in some specialties (23,24).

Early quality measures in rheumatology relied on administrative billing claims (1,2). Such measures have the advantage of collecting information from all of the clinicians and entities that have submitted bills for clinical care, thus capturing a broad picture of health care received. They are also relatively easy for payers to access and analyze. However, the clinical information contained in billing claims has limited scope, and many relevant elements of the care that rheumatologists provide are not captured. Moreover, the use of codes in billing is often incomplete or inconsistent. Claims data are often not accessible to practicing rheumatologists in real-time due to the intrinsic lag in processing, and therefore does not lend itself to rapid cycle quality improvement. Other quality measures in rheumatology have relied on manual medical chart abstraction. Although this method can yield more clinically detailed data, it is labor intensive, incomplete, and difficult to implement for continuous quality improvement. Natural language processing (deriving information by performing structured queries of notes) may hold promise for streamlining chart abstraction, and contributing quality information beyond structured eCQM data; however, as yet its reliability for this purpose has not been widely demonstrated.

eCQMs have the potential to address some of these limitations by facilitating automated access to more clinically detailed information in the EMR. This obviates the need for duplicate data entry or retrieval. Generating performance reports both locally and nationally in real time to facilitate quality improvement is an important advantage of eCQMs. However, because methods for application of eCQMs are still evolving, there are significant challenges ahead as well. For example, documentation within EMRs may be inconsistent, and nonstructured information is difficult to systematically query. Coding systems continue to evolve, particularly with the implementation of ICD-10, requiring that eCQMs adapt to new coding systems. Moreover, achieving seamless interoperability of data systems remains a major obstacle, threatening efforts to obtain relevant clinical information that paints a more complete picture of clinical care. Finally, implementation of eCQMs at the local level requires health information technology support and successful quality reporting may require substantial changes to current workflows. Despite these obstacles, it is likely that eCQMs will take on an increasingly important role in efforts to continuously improve health care quality.

The ACR's measure development effort arose as a response to the growing number of quality measures affecting rheumatologists. Previous measures were often led by payers (e.g., Resolution Health, Ingenix), national quality organizations (e.g., National Committee on Quality Assurance), or other physician groups (e.g., the American Medical Association's PCPI) and implemented in both the private and public sectors. Measures were sometimes not aligned and rheumatologists

had little leverage in revising measures that were either difficult to implement or not consistent with current evidence (2). The need for rheumatologists to lead measure development, therefore, became more pressing as regulatory requirements began to shift from small incentive programs to larger penalty programs. One example of such a shift is the CMS Physician Quality Reporting System (PQRS). In 2016, this program penalizes physicians who do not participate (25). Key goals of the ACR's recent quality measure development projects are to develop rheumatology-specific measures for these types of programs, in which rheumatologists will be largely expected to participate, while promoting meaningful quality measures that advance rheumatologic practice.

An important mechanism for disseminating ACR quality measures will be a nationwide EMR-based registry. The ACR's Rheumatology Information System for Effectiveness (RISE) registry uses a scalable platform to facilitate automated data abstraction from EMR systems of participating practices. RISE has received certification as a Qualified Clinical Data Registry by the Centers for Medicare and Medicaid Services, an important milestone that streamlines quality reporting. Rheumatologists can use RISE for PQRS reporting. Participation in RISE also satisfies an objective in the Meaningful Use program: reporting to a special registry. The eCQMs developed by the ACR have been programmed into RISE and, as new measures become available, they can be rapidly implemented in RISE. Because data collected in RISE are available to practices in real-time, eCQMs can be used for local rapid cycle quality improvement projects and also to benchmark performance against other clinical sites nationally. In addition, revision of eCQMs to reflect updates in scientific evidence or to incorporate new or important concepts will be significantly easier moving forward.

The science of quality measurement as well as data infrastructure to support measurement is likely to continue to evolve. The ACR's eCQM development program should also evolve to address newly identified gaps in care that are amenable to quality improvement. While technology holds promise in making quality measurement more meaningful and more efficient, input from rheumatologists will continue to be a key ingredient for a measurement strategy that seeks not just to assess performance but to improve outcomes for patients.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Yazdany had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Yazdany, Myslinski, Miller, Desai, Lacaille, Orozco, Bunyard, Matteson, Silverman, Warren, Nola, Robbins.

Acquisition of data. Yazdany, Myslinski, Francisco, Desai, Bunyard, Passo.

Analysis and interpretation of data. Yazdany, Francisco, Desai, Schmajuk, Barber, Bunyard, Bergman, Matteson, Olson, Silverman, Robbins.

REFERENCES

1. Schmajuk G, Trivedi AN, Solomon DH, Yelin E, Trupin L, Chakravarty EF, et al. Receipt of disease-modifying antirheumatic drugs among patients with rheumatoid arthritis in Medicare managed care plans. JAMA 2011;305:480–6.
2. Desai SP, Yazdany J. Quality measurement and improvement in rheumatology: rheumatoid arthritis as a case study. Arthritis Rheum 2011;63:3649–60.
3. Institute of Medicine. Best care at lower cost: the path to continuously learning health care in America. Washington (DC): National Academies Press; 2013.
4. McGlynn EA, Schneider EC, Kerr EA. Reimagining quality measurement. N Engl J Med 2014;371:2150–3.
5. 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;21:181–4.
6. Gardner W, Morton S, Byron SC, Tinoco A, Canan BD, Leonhart K, et al. Using computer-extracted data from electronic health records to measure the quality of adolescent well-care. Health Serve Res 2014;49:1226–48.
7. National Quality Forum. Guidance for measure testing and evaluating scientific acceptability of measure properties. 2011. URL: http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx.
8. Saag KG, Yazdany J, Alexander C, Caplan L, Coblyn J, Desai SP, et al. Defining quality of care in rheumatology: the American College of Rheumatology white paper on quality measurement. Arthritis Care Res (Hoboken) 2011;63:2–9.
9. National Quality Forum. Measure Prioritization Advisory Committee report. Prioritization of high-impact Medicare conditions and measure gaps. 2010 URL: http://www.qualityforum.org/projects/prioritization.aspx.
10. Singh JA, Hodges JS, Toscano JP, Asch SM. Quality of care for gout in the US needs improvement. Arthritis Rheum 2007;57:822–9.
11. Singh JA. Quality of life and quality of care for patients with gout. Curr Rheum Report 2009;11:154–60.
12. Mikuls TR. Quality of care in gout: from measurement to improvement. Clin Exp Rheumatol 2007;25 Suppl 47:114–9.
13. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:625–39.
14. Khanna D, Khanna PP, Fitzgerald JD, Singh MK, Bae S, Neogi T, et al. 2012 American College of Rheumatology guidelines for management of gout. Part 2: therapy and antiinflammatory prophylaxis of acute gouty arthritis. Arthritis Care Res (Hoboken) 2012;64:1447–61.
15. Grossman JM, Gordon R, Ranganath VK, Deal C, Caplan L, Chen W, et al. American College of Rheumatology 2010 recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis. Arthritis Care Res (Hoboken) 2010;62:1515–26.
16. Bell BG, Spencer R, Avery AJ, Campbell SM. Tools for measuring patient safety in primary care settings using the RAND/UCLA appropriateness method. BMC Fam Pract 2014;15:110.
17. Lawson EH, Gibbons MM, Ko CY, Shekelle PG. The appropriateness method has acceptable reliability and validity for assessing overuse and underuse of surgical procedures. J Clin Epidemiol 2012;65:1133–43.
18. Chronaki C, Jaffe C, Dolin B. eMeasures: a standard format for health quality measures. Stud Health Technol Inform 2011;169:989–91.
19. Centers for Medicare and Medicaid Services. Measure authoring tool user guide. 2014. URL: http://www.emeasuretool.cms.gov/documents/10179/13108/MAT+User+Guide/c50cc2fb-01ee-4b97-9cd1-2e33feca8763.
20. National Quality Forum. eMeasure feasibility assessment: technical report. 2015. URL: http://www.qualityforum.org/Projects/e-g/eMeasures/Electronic_Quality_Measures.aspx.

21. Briggs JB, Weiner J, Chan K, Kind E, Awwad S. Performance measures using electronic health records: five case studies. New York: The Commonwealth Fund; 2009. URL: http://www.commonwealthfund.org/publications/fund-reports/2008/may/performance-measures-using-electronic-health-records-five-case-studies.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometr 1977;33:159–74.
23. Parker SL, McGirt MJ, Bekelis K, Holland CM, Davies J, Devin CJ, et al. The National Neurosurgery Quality and Outcomes Database Qualified Clinical Data Registry: 2015 measure specifications and rationale. Neurosurg Focus 2015;39:E4.
24. Bhatt DL, Drozda JP Jr, Shahian DM, Chan PS, Fonarow GC, Heidenreich PA, et al. ACC/AHA/STS statement on the future of registries and the performance measurement enterprise: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures and The Society of Thoracic Surgeons. Ann Throac Surg 2015;100:1926–41.
25. Burwell SM. Setting value-based payment goals: HHS efforts to improve U.S. health care. N Engl J Med 2015;372:897–9.
26. National Quality Forum. Quality data model. 2012. URL: https://ecqi.healthit.gov/system/files/qdm/qdm_122012.pdf.