

The Challenges of Measuring, Improving, and Reporting Quality in Primary Care

Richard A. Young, MD¹

Richard G. Roberts, MD, JD²

Richard J. Holden, PhD³

¹JPS Hospital Family Medicine Residency Program, Fort Worth, Texas

²University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin

³Indiana University School of Informatics and Computing, Bloomington, Indiana

ABSTRACT

We propose a new set of priorities for quality management in primary care, acknowledging that payers and regulators likely will continue to insist on reporting numerical quality metrics. Primary care practices have been described as complex adaptive systems. Traditional quality improvement processes applied to linear mechanical systems, such as isolated single-disease care, are inappropriate for nonlinear, complex adaptive systems, such as primary care, because of differences in care processes, outcome goals, and the validity of summative quality scorecards. Our priorities for primary care quality management include patient-centered reporting; quality goals not based on rigid targets; metrics that capture avoidance of excessive testing or treatment; attributes of primary care associated with better outcomes and lower costs; less emphasis on patient satisfaction scores; patient-centered outcomes, such as days of avoidable disability; and peer-led qualitative reviews of patterns of care, practice infrastructure, and intrapractice relationships.

Ann Fam Med 2017;15:175-182. doi: <https://doi.org/10.1370/afm.2014>.

INTRODUCTION

The US National Quality Strategy has 3 overarching aims: improve the quality of care, improve the health of the population, and reduce the cost of care.¹ The achievement of these aims depends, in part, on the collection and reporting of quality measures, more than 400 of which are endorsed currently by the US National Quality Forum.² Supporters of quality metrics and physician scorecards, such as those required for patient-centered medical home (PCMH) certification, assume that better health can be achieved by following guidelines developed for single diseases, and that a summation of single-disease guidelines accurately describe the quality of work delivered by a primary care practice. These assumptions are aligned with traditional strategies for process and quality improvement (QI), such as Six Sigma and lean thinking, that have been powerful tools in mechanical systems and disease-specific care processes.^{3,4}

Many people think that systems are improved by deconstructing the overall system performance and management into component elements.⁵ In contrast, primary care is better conceptualized as a complex adaptive system—where learning people and institutions (“agents” in the complex adaptive system vernacular) interact with the environment in nonlinear patterns and self-organize, resulting in unpredictable, emerging creative behaviors rather than rigidly adhering to a standardized set of linear processes for diagnosing and treating single diseases.⁶⁻⁸ Failure to appreciate these complexities leads some to erroneously conclude that practices have failed by not implementing standardized interventions.⁶

Well-aligned quality measures for primary care should promote accountable performance and boost clinicians’ motivation by rewarding them for managing complexity, solving problems, and thinking creatively when addressing the unique circumstances of each patient.^{9,10} Instead, misaligned

Conflicts of interest: Richard Holden is a consultant for the nonprofit organization Peterson Center on Healthcare on a project to implement an evidence-based change package for improving primary care delivery in the United States. Young and Roberts report no conflicts of interest.

CORRESPONDING AUTHOR

Richard A. Young, MD
JPS Hospital Family Medicine Residency Program
1500 S Main
Fort Worth, TX 76104
ryoung01@jpshealth.org

QI metrics and other mandates as electronic health records (EHRs)¹¹ have contributed to burnout among physicians, especially those in primary care,¹² causing some to advocate for the Quadruple Aim by adding the goal of enhancing professional satisfaction and well-being to the Triple Aim.¹³ Most importantly, many primary care physicians believe the existing metrics may paradoxically encourage poor quality of care.^{14,15}

Given primary care's central role in health care, we believe that the inappropriate application of traditional QI strategies and misaligned metrics undermines primary care, and in turn, all patient care. We challenge the notion that care process strategies, outcome goals, and reporting devices that may work in mechanical areas of health care are valid in primary care (Table 1).⁴ We offer alternative approaches that we believe will better support primary care's important responsibilities in helping us achieve our national quality goals.

CARE PROCESSES IN LINEAR AND COMPLEX NONLINEAR ENVIRONMENTS

Process Complexity

Industrial QI approaches have improved health care delivery in mechanical linear domains, such as elective spine surgery,¹⁶ ventilator-associated pneumonia bundles,^{17,18} and central line bundles,^{19,20} where processes do not change appreciably for different patients. These processes have relatively few variables—5 for ventilator bundles and 8 for central lines—and it is reasonable to assume that measurable process steps are causally connected to patient-oriented outcomes. Such is not the case for primary care, which is more complex than other specialties (eg, cardiology, psychiatry, and obstetrics/gynecology), because so many more inputs and outputs are managed during each visit.^{21,22}

For example, a study of primary care physicians' workflow calculated a mean of 37 tasks performed per

Table 1. Differences in Processes and Outcomes Between Care Isolated to a Single Disease and Primary Care

Construct	Isolated Single Disease with Linear Mechanical Processes	Primary Care Nonlinear Adaptive Processes
Care process		
Process complexity	Few variables to be measured and controlled Example: central line bundles	Numerous variables that make accurate measurement problematic Example: patients with multiple medications, comorbidities, and socioeconomic challenges
Process standardization	Standard processes use consistent raw materials Example: antibiotics administered just before the incision is made in elective surgeries	Variable processes with variable raw materials Example: a wide range of disease severities and treatment options for the same diagnosis: eg, migraine, chronic low back pain, and fibromyalgia
Process controls	Machines and unconscious patients are largely controlled by their human operators Example: procedure not started until the pre-surgical checklist is completed and chlorhexidine antiseptic has been applied	The patient "machine" is controlled by a milieu of forces, including caregiver biases, unique patient beliefs, socioeconomic status, and the external environment Example: medication nonadherence associated with poverty, which is not controllable by the physician or the health care team
Outcome goals		
Goal clarity: multimorbidity	All team members and machines work toward one clear goal Example: titanium artificial hip placed in the appropriate position	There is no one right answer or goal, only an individualized understanding of risks and benefits where ideally the patients chooses the best answer for him or her Example: another round of chemotherapy for a patient with metastatic cancer vs hospice care
Goal clarity: unique patient priorities	Patients and caregivers agree on clear outcome Example: minimum days intubated on mechanical ventilation	Patients have different goals or priorities from their caregivers' recommendations Example: a diabetic patient who does not want to start taking insulin to reduce her blood glucose because of concerns about the affordability of the medicine and a belief that insulin killed her aunt
Goal timing	Standard processes have fixed expectations of the timing of interventions Example: daily trials of endotracheal tube extubation	The timing and order of addressing patient concerns are highly variable Example: the primary care physician and patient may negotiate and agree that a vague symptom be given more time to evolve, with no testing or treatment ordered the first time the concern is mentioned
Inadequate summative quality scorecards		
Poor risk-adjustment tools	Coexisting patient complexities rarely affect process metrics Example: postoperative thrombosis prophylaxis	Coexisting patient complexities often affect patient outcomes Example: any number of social determinant factors affect disease- and patient-oriented outcomes
Goal target number	Six Sigma-level outcomes Example: 0% infection rate or 100% vaccination uptake	Outcomes are dependent on a multitude of social and behavioral cofactors Example: much less than 100% of a population wants colon cancer screening no matter how strongly it is recommended and incentivized
Scorecard comprehensiveness	List of metrics for a physician represents most of the work performed Example: an overall rating for an orthopedist who only replaces hips and knees	Few simplistic quality measures capture only a tiny fraction of the work performed by a primary care physician. The alternative is a long, cumbersome list that is costly and burdensome to maintain and of questionable validity

visit in no discernable pattern within visit or physician, concluding "the number and sequence of tasks is varied and unpredictable."²³ This complexity is problematic for certain QI techniques, which have a practical limit on the number of variables that can be reliably measured and controlled. Typically, if more than 20 variables are included in a process control algorithm, the impact of measurement inaccuracies, for instance, results in cumulative noise that makes distinguishing individual variable effects very difficult (Thomas Edgar, PhD, Chair, Department of Chemical Engineering, University of Texas at Austin, personal communication).

Process Standardization

Traditional QI assumes that the highest quality is achieved when a linear process occurs the same way each time.⁴ Primary care, however, depends on meeting a wide variety of patient needs. More than one-third of health problems initially encountered at a primary care center do not lend themselves to a diagnosis, and about one-half are unlikely to result in a definitive diagnosis that would trigger a standard care pathway.²⁴ Care of patients in the primary care setting must account for each patient's comorbidities, disease severity, medication tolerance, beliefs, desires, and socioeconomic factors. Given the paucity of evidence, primary care physicians must often rely on creativity, problem solving, and adaptability to develop a custom care plan.

Process Controls

Input variables in processes, such as hospital central line bundles, are largely controlled by clinicians (washing hands, using chlorhexidine antiseptics, etc). In contrast, complex adaptive systems have no single point of control.⁵ Input variables for primary care processes and health outcomes are often beyond the control of the physician and patient,²⁵ such as socioeconomic factors²⁶ and poverty.²⁷ Some successes in hot spots of improving care for disadvantaged populations²⁸ may be generalizable, but the improved outcomes of these programs were more likely achieved by creating care plans adapted to the unique circumstances of high-risk patients. Additionally, for these approaches to spread more rapidly and consistently to other vulnerable populations, physician payment models must change.²⁹

DIFFERING OUTCOME GOALS IN LINEAR AND COMPLEX ENVIRONMENTS

Outcomes may be prioritized differently by independent agents in a complex adaptive system, may vary depending on the external environment of the system, and likely change with time,^{3,4} creating challenges for

policy makers attempting to measure the outcomes of primary care centers in different settings.

Outcome Goal Clarity: Multimorbidity

A body of literature concludes that specialists are more likely than generalists to follow disease-specific guidelines, but this literature assumes one correct diagnostic or treatment option based on the organ of interest.³⁰ For example, if a patient with a history of a heart attack develops erectile dysfunction after starting a β -blocker, what is the correct measure of quality? The cardiac disease-specific answer would be to continue the β -blocker. The primary care physician addresses the human factors; a patient-centered discussion of risks and benefits might cause the patient to accept a small increased risk of sudden death in return for a viable sex life and therefore choose to discontinue the β -blocker. When primary care physicians deviated from single-disease guidelines, reasons were deemed medically appropriate in most cases.^{31,32}

Outcome Goal Clarity: Unique Patient Priorities

Traditional QI assumes that patients want and will do what experts say is best. In reality, however, human beings have a wide range of priorities that may not be aligned with those of their caregivers. Dr Iona Heath, former president of the Royal College of General Practitioners, commented on a reform proposal for the National Health Service by a British think tank in 2011, "... patients are presented as units of health need, indistinguishable but for their medical histories. They are assumed to be uniformly rational and committed to making decisions designed to promote their health"³³ She concluded, "I find no evidence of any understanding [by the think tank authors] that patients can also be abusive, manipulative, and self-destructive." Although most patients are well meaning and cooperative, their personal goals, attitudes, and beliefs often confound guideline writers and physicians. Primary care physicians provide the most patient-centered care when they attempt to understand the meaning behind varying patient priorities and adjust diagnostic and treatment plans accordingly.

Goal Timing

In traditional QI metrics, there is little variance of the timing of intervention goals. For example, the Center for Medicare and Medicaid Services (CMS) core quality indicator set includes a hospital metric that aspirin will be prescribed to every patient with an acute myocardial infarction at hospital discharge.³⁴ In primary care, the best time frame to attempt to achieve an idealized goal (if ever) is more complex. A well-supported patient with new-onset diabetes might achieve and

tolerate a hemoglobin A_{1c} of <7.0% within a matter of months. For another patient who recently experienced the loss of a spouse with a resulting change in diet and activity, a hemoglobin of A_{1c} <10% might represent a more reasonable goal until the social stressors improve.

SUMMATIVE QUALITY SCORECARDS AND QUALITY OF A PRIMARY CARE PRACTICE

A rigid list of quality metrics purporting to summarize the quality of a primary care practice disregards the reality of a complex adaptive system, where adaptability and learning cause the practice to self-organize and prioritize practice aspects most worthy of improvement efforts while recognizing these priorities evolve with time.⁵ Complex interactions and interdependencies emerge within such a system that cannot be understood or predicted simply by measuring individual elements of the system. One cannot assume that the whole is merely a sum of the measurable parts.^{8,35} These realities imply that measures based on linear models, such as Ambulatory Diagnostic Groups and the Ambulatory Severity Index, are inadequate summaries of primary care quality.³⁶

Poor Risk-Adjustment Tools

Risk adjustment of summative quality scorecards is crucial to insure that primary care physicians and practices are evaluated fairly, because the characteristics of the associated patient population have a large effect on the reported results of standard QI metrics. Studies of these quality scorecards conclude that caring for complex patients in a safety-net setting strongly predicts failure to meet common quality goals in hospitals^{37,38} and primary care settings.³⁹⁻⁴¹ The National Quality Forum has recognized that socioeconomic factors are important contributors to patient outcomes, that current measurements do not account for these factors, and that adequate risk adjustments for quality outcomes do not currently exist.⁴² QI leaders have discussed attempting risk adjustment for patient panels,⁴³ but rigorous methods are lacking, and existing models give different results.⁴⁴ As a result, current summative quality scorecards disadvantage physicians who care for the most vulnerable patient populations. Physicians who are more attentive to the social and cultural context, and who adjust care plans accordingly, are more likely to deliver positive patient outcomes,⁴⁵ though measuring such contextual care takes considerable effort.⁴⁶

Goal Target Number

There is a clash of expectations between the reality of primary care and traditional notions of industrial QI, such as Six Sigma targets, which imply adherence

to process measures 99.99966% of the time.⁴⁷ Uptake of recommended tests or treatments is never 100% in the outpatient setting, even in populations with little socioeconomic deprivation. In vulnerable populations, patient preferences and daily practicalities often disrupt recommended ambulatory care. For example, in primary care, even though many patients say they are willing to undergo colon cancer screening when asked by their physicians,⁴⁸ uptake of recommended screening is low, often measured at less than 50% of eligible patients.^{49,50} Even in primary care centers of excellence, chronic disease targets are met and sustained less than 50% of the time despite extra resources, such as health coaches.⁵¹ The impossibility of achieving 100% uptake makes it much more difficult to draw a summative conclusion about which primary care practices are providing high-quality care when contrasted against elective surgeries where nearly 100% compliance with preoperative antibiotic guidelines could reasonably be achieved.

Scorecard Comprehensiveness

The comprehensiveness of services offered by primary care physicians creates a tension between (1) the desire of funders of health care systems—such as Medicare and insurance companies—to be reassured with numerical data that participating primary care practices provide high-quality care, and (2) the practical limits, practice burdens, and dubious validity of data collection and reporting. On the one hand, a few well-chosen measures will represent only a fraction of the work delivered at primary care centers. CMS has just begun to evaluate patient-oriented outcome measures, and only for hip and knee replacement surgery.⁵² For a practice that mostly performs such surgeries, this metric may allow policy makers to determine whether the practice is performing well. A 2-disease measure would be inappropriate and inadequate for primary care physicians.²²

On the other hand, hundreds of measures become cumbersome and expensive to collect and report. Even worse is that these myriad measures will still be unlikely to capture the richness of the interactions between the primary care practice professionals and patients. Direct observational research of family physicians has shown that even when diagnoses are clustered into groups (eg, low back pain syndromes and headaches), 25 of these clusters account for only 60% of the diagnoses.⁵³ Family physicians have been observed to manage nearly 500 different diagnoses in 3,344 recorded patient encounters in a national study.²² Such a complex interaction of variables makes it unrealistic to represent a complex adaptive system with a finite number of dependent and independent variables.⁸ The cost to a practice is also burdensome. Just to main-

tain a PCMH certification has been estimated to cost \$120,000 per physician per year.⁵⁴ Thus, given the current state of inadequate measurement in primary care, summative scorecards, such as a PCMH certification, which purport to distinguish “good” and “bad” primary care practices, are woefully inadequate at best and misleading at worst. Scorecards constructed using EHR data represent an ideal case as imagined by system designers that often include incomplete or inaccurate information and differ from the reality experienced by patients and physicians.⁵⁵

MOVING FORWARD

It is challenging to evaluate complex adaptive systems that often do not comprise easily observable processes that lead to easily measurable outcomes.^{5,8,35,56} Instead, these systems, such as in primary care, often have a plurality of aims and interests that sometimes conflict; they are influenced by their external environment, whose effects on the system could be minimal or profound; and the impact of the external forces often change with time.^{5,8} In traditional QI of mechanical processes, measurements are often at the level of an individual step. Ideally, it is more appropriate to measure complex adaptive systems at the level of the final outcome of interest.⁵

Which is not to suggest that all aspects of a complex adaptive system are completely unmeasurable. Parts of a primary care practice may be fairly linear and simple, others might best be imagined as a complicated set of linear processes, and other parts reflect an adaptive system.³⁶ At the macro level, decades of research show that places with more family physicians or general practitioners are associated with better health outcomes at a lower cost.⁵⁷ At the practice level, it may be meaningful to measure and improve such processes as registration times or managing preventive services, though complexity science theory concludes that one cannot also assume that those process changes necessarily lead to better population health.

The British Quality and Outcome Framework (QOF) pay-for-performance system is instructive in many ways for some of the changes and metrics proposed for primary care in America, and it raises the question of whether the pursuit of numeric goals has come at the cost of loss of therapeutic relationships, trust, and less effective primary care. The QOF may be a good example of the disconnect between a series of disease-specific measures and overall population outcomes.⁵⁸ The QOF system, which has been in place more than 10 years, has resulted in only modest improvements in quality often that do not endure,⁵⁹ decreased the person-centeredness and continuity of

care,⁵⁸ decreased quality of care for diseases not part of their pay-for-performance system,⁵⁸ cost many billions of pounds to implement,⁶⁰ decreased physician morale,⁶¹ and did not decrease premature mortality.⁶² The Scottish National Health Service announced it will abandon the QOF system in 2017.⁶³ This disappointing result is likely explained by the top-down nature of the QOF evaluation, which removed some of the freedoms of the practices to self-organize, an important feature of a complex adaptive system.

Ideally in complex adaptive systems, the role of the evaluator changes from the top-down mandates of QOF.⁸ The evaluator takes on the task of designing and implementing transforming feedback loops across systems, not merely compiling lists of single-disease measurements. This role falls into 2 major categories: absorbing uncertainty and making learning the primary outcome. Health care payers and regulators are not typically comfortable with uncertainty and will likely continue to insist on numerical metrics and quality scorecards, but more relevant and feasible measures for primary care are needed.⁶⁴ There have been several attempts to mandate a complex set of rules to govern health care.⁴ When these have not yielded desired results, the instincts of payers and regulators has been to create even more rules. Complexity science asserts that these instincts take us in the wrong direction.⁴ Acknowledging these external regulatory pressures, however, we propose the following principles for the next generation of quality metrics and scorecards in US primary care (summarized in Table 2).

Shared-Decision Reporting

Future quantitative quality scorecards should borrow a concept from the British National Health Service QOF pay-for-performance system currently called exception reporting.⁶⁵ British general practitioners are given a list of reasons that allow them to remove individual patients from quality reports. Examples include patients new to their panel, patients who decline recommended tests or treatments, and patients with comorbidities that render a disease-specific quality metric inappropriate. Exception reporting may not be the best label; we concur with others who suggested alternative descriptions, such as patient choice or shared decisions.⁶⁶

Target Ranges Without Absolute Goals

Quality scorecards should never include metrics where 0% or 100% are the targets. Intermountain Health expects variance of 5% to 15% in some of its quality measures and scrutinizes physicians for complying with a protocol too much.⁶⁷ Targets should also be adjusted for socioeconomic status, though existing identifiable covariates have not proven to be very useful.

Table 2. Proposals for Improved Quality Metrics and Reporting in Primary Care

Shared-decision reporting
Target ranges without absolute goals
Measure when physicians do not order tests or treatments
Measure other aspects of primary care capacity associated with better outcomes
Comprehensiveness of services offered at the primary care center
Physician-patient continuity
Smaller practice size
Rate of generic prescription writing
Increased office visit time for complex patients
Access to local clinic professionals 24/7
Careful selection of referral specialists
De-emphasize measures of patient satisfaction
Measure outcomes more important to patients
Peer-led qualitative reviews of patterns of care

Measure When Physicians Do Not Order Tests or Treatments

Physicians should be measured and acknowledged when they do not subject their patients to interventions of dubious value. In traditional QI, process measures almost always dictate what should be done, not what should be avoided. The true value of primary care often occurs when physicians choose to not order a test, refer a patient to another physician, or start a new treatment.⁶⁸ The Choosing Wisely campaign provides a thoughtful framework to develop new metrics to reduce unnecessary tests and treatments.⁶⁹

Measure Other Aspects of Primary Care Capacity Associated With Better Outcomes

Other aspects of primary care are associated with better outcomes or lower costs and should be measured. Examples include the comprehensiveness of services provided by family physicians that is associated with lower Medicare costs and hospitalizations⁷⁰; increased physician-patient continuity associated with lower costs and fewer complications of common chronic conditions,⁷¹ fewer hospitalizations,⁷² and lower overall mortality⁷²; smaller primary care practice size associated with reduced hospitalizations from preventable conditions⁷⁴; and the rate of generic prescription writing associated with lower costs.⁷⁵ Other measures associated with lower total costs for larger primary care systems could include increased time for office visits for complex patients, 24/7 access to local clinic professionals, and careful selection of referral specialists.⁷⁶

A better measurement of a practice may be that it regularly undertakes self-reflection and measures its patient care, with regulators worrying less about exactly what each practice is attempting to measure and improve. Doing so would encourage the positive features

of a complex adaptive system: self-organization, emergence, and coevolution with the practice's environment.⁵⁵

De-emphasize Measures of Patient Satisfaction

Improved patient satisfaction scores have been linked to inappropriate services, such as prescribing unnecessary antibiotics.⁷⁷ A full review of the advantages and disadvantages of patient satisfaction scores is beyond the scope of this article, other than to note that patients' goals and treatment preferences are not the same as patient satisfaction, and there is evidence that greater patient satisfaction scores are associated with worse outcomes, including higher mortality rates.⁷⁸⁻⁸⁰ The enormous gaps in knowledge about the relevance of patient satisfaction scores to quality demand caution and further research in this area.

LOFTIER PROPOSALS

Measure Outcomes More Important to Patients

A lofty goal would be to measure more patient-centric outcomes at a population level, such as premature death rates (adjusted for age, comorbidities, and socioeconomic factors), disability (absence from work or school, limitation of activities, etc), ease of access, and some measures of patient experience of care (eg, respect, feeling heard, sensitivity, and treatment burden). Such measures may be more difficult to develop and implement, and they may often be statistically insignificant, even for a relatively large primary care practice, but they are likely to be more meaningful for achieving societal goals.^{81,82}

Peer-Led Reviews of Patterns of Care

Perhaps a better approach to reassure payers, regulators, and patients of high-quality primary care would be to replace most quantitative measures with adequately funded peer-reviewed assessments of patterns of care, as proposed recently in Scotland.⁶³ Peer-reviewed assessments would also be more consistent with evaluation principles of a complex adaptive system, where qualitative assessments of relationships between agents provides a more meaningful understanding of the quality of care provided than externally mandated numeric targets.^{35,56}

Final Thoughts

Family physicians' comfort with complexity, ambiguity, and uncertainty—and their ongoing relationships with their patients—allow them to negotiate flexible diagnostic and treatment plans with their patients based on patient-specific risk factors and probabilities of serious disease.^{68,84} At its heart, traditional QI assumes there is a definitive and measurable right answer in a given situation. In contrast, primary care physicians often deliver high-value care by doing the best they

can with the patient care cards they are dealt., knowing that perfection will never be achieved. Adaptability rather than standardization should be the cornerstone of complex primary care and chronic disease care.⁸⁵ The national trends of rigid metrics and simplistic noncomprehensive scorecards must be reversed for primary care to do an even better job of delivering better patient care at a lower cost. At a minimum, measures that better respect the complexity and value of primary care would help promote the sustainable primary care workforce that is desperately needed.

To read or post commentaries in response to this article, see it online at <http://www.annfam.org/content/15/2/175>.

Key words: quality improvement; primary health care; health policy; complex adaptive systems

Submitted March 4, 2016; submitted, revised, September 19, 2006; accepted October 12, 2016.

Acknowledgment: John Epling, MD helped us understand complexity science and how to evaluate complex adaptive systems.

References

- Agency for Healthcare Research and Quality. About the National Quality Strategy (NQS). <http://www.ahrq.gov/workingforquality/about.htm#aims>. Accessed Jun 10, 2016.
- Agency for Healthcare Research and Quality. NQF-Endorsed Measures. <https://www.qualitymeasures.ahrq.gov/browse/nqf-endorsed.aspx>. Accessed Jun 10, 2016.
- de Brantes F, Galvin R, Lee T. Bridges to excellence: building a business case for quality care. *J Clin Outcomes Manag*. 2003;10(8):439-446.
- Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, D.C.: National Academies Press; 2001.
- Rouse W. Health care as a complex adaptive system: implications for design and management. *The Bridge*. 2008;38(1):17-25.
- Sturmberg JP, Martin CM, Katerndahl DA. Systems and complexity thinking in the general practice literature: an integrative, historical narrative review. *Ann Fam Med*. 2014;12(1):66-74.
- McDaniel RR Jr, Lanham HJ, Anderson RA. Implications of complex adaptive systems theory for the design of research on health care organizations. *Health Care Manage Rev*. 2009;34(2):191-199.
- Eoyang G, Berkas T. Evaluating performance in a CAS. Human Systems Dynamics Institute. <http://www.hdinstitute.org/learn-more/library/articles/Evaluating-Performance-in-a-CAS.pdf>. Accessed Aug 19, 2016.
- Pink D. *Drive: The Surprising Truth About What Motivates Us*. New York, NY: Riverhead Books; 2011.
- Cassel CK, Jain SH. Assessing individual physician performance: does measurement suppress motivation? *JAMA*. 2012;307(24):2595-2596.
- Rabin RC. A growing number of primary-care doctors are burning out. How does this affect patients? *Washington Post*. Mar 31, 2014.
- Linzer M, Levine R, Meltzer D, Poplous S, Warde C, West CP. 10 bold steps to prevent burnout in general internal medicine. *J Gen Intern Med*. 2014;29(1):18-20.
- Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med*. 2014;12(6):573-576.
- Greenhalgh T, Howick J, Maskrey N; Evidence Based Medicine Renaissance Group. Evidence based medicine: a movement in crisis? *BMJ*. 2014;348:g3725.
- Waddimba AC, Burgess JF Jr, Young GJ, Beckman HB, Meterko M. Motivators and hygiene factors among physicians responding to explicit incentives to improve the value of care. *Qual Manag Health Care*. 2013;22(4):276-292.
- Meyer H. At UPMC, improving care processes to serve patients better and cut costs. *Health Aff (Millwood)*. 2011;30(3):400-403.
- Morris AC, Hay AW, Swann DG, et al. Reducing ventilator-associated pneumonia in intensive care: impact of implementing a care bundle. *Crit Care Med*. 2011;39(10):2218-2224.
- Bird D, Zambuto A, O'Donnell C, et al. Adherence to ventilator-associated pneumonia bundle and incidence of ventilator-associated pneumonia in the surgical intensive care unit. *Arch Surg*. 2010;145(5):465-470.
- Galpern D, Guerrero A, Tu A, Fahoum B, Wise L. Effectiveness of a central line bundle campaign on line-associated infections in the intensive care unit. *Surgery*. 2008;144(4):492-495, discussion 495.
- Guerin K, Wagner J, Rains K, Bessesen M. Reduction in central line-associated bloodstream infections by implementation of a postinsertion care bundle. *Am J Infect Control*. 2010;38(6):430-433.
- Katerndahl DA, Wood R, Jaén CR. A method for estimating relative complexity of ambulatory care. *Ann Fam Med*. 2010;8(4):341-347.
- Katerndahl D, Wood R, Jaén CR. Complexity of ambulatory care across disciplines. *Healthc (Amst)*. 2015;3(2):89-96.
- Holman GT, Beasley JW, Karsh BT, Stone JA, Smith PD, Wetterneck TB. The myth of standardized workflow in primary care. *J Am Med Inform Assoc*. 2016;23(1):29-37.
- Rosendal M, Carlsen AH, Rask MT, Moth G. Symptoms as the main problem in primary care: A cross-sectional study of frequency and characteristics. *Scand J Prim Health Care*. 2015;33(2):91-99.
- Hershberger PJ, Bricker DA. Who determines physician effectiveness? *JAMA*. 2014;312(24):2613-2614.
- Stringhini S, Sabia S, Shipley M, et al. Association of socioeconomic position with health behaviors and mortality. *JAMA*. 2010;303(12):1159-1166.
- Foegen WH. Social determinants of health and health-care solutions. *Public Health Rep*. 2010;125(Suppl 4):8-10.
- Gawande AA. The hot spotters: can we lower medical costs by giving the neediest patients better care? *The New Yorker*. January 24, 2011.
- Phillips RL Jr, Pugno PA, Saultz JW, et al. Health is primary: family medicine for America's health. *Ann Fam Med*. 2014;12(Suppl 1):S1-S12.
- Harrold LR, Field TS, Gurwitz JH. Knowledge, patterns of care, and outcomes of care for generalists and specialists. *J Gen Intern Med*. 1999;14(8):499-511.
- Persell SD, Dolan NC, Friesema EM, Thompson JA, Kaiser D, Baker DW. Frequency of inappropriate medical exceptions to quality measures. *Ann Intern Med*. 2010;152(4):225-231.
- Steinman MA, Dimaano L, Peterson CA, et al. Reasons for not prescribing guideline-recommended medications to adults with heart failure. *Med Care*. 2013;51(10):901-907.
- Heath I. Who's complacent now? The King's Fund on general practice. *BMJ*. 2011;342:d2254.
- Center for Medicare & Medicaid Services. Clinical quality measures finalized for eligible hospitals and critical access hospitals beginning with FY 2014. https://www.cms.gov/regulations-and-guidance/legislation/ehrincentiveprograms/downloads/2014_cqm_eh_finalrule.pdf. Accessed Dec 15, 2016.
- Kordowicz M. Do systems and complexity theories aid our understanding of poor performing general practices. *J Contem Dev Manag Syst*. 2012;1(2):1-11.
- Green LA. The implications of measuring complexity. *Ann Fam Med*. 2010;8(4):291-292.
- Rajaram R, Chung JW, Kinnier CV, et al. Hospital characteristics associated with penalties in the centers for medicare & medicaid services hospital-acquired condition reduction program. *JAMA*. 2015;314(4):375-383.

38. Skinner J, Chandra A, Staiger D, Lee J, McClellan M. Mortality after acute myocardial infarction in hospitals that disproportionately treat black patients. *Circulation*. 2005;112(17):2634-2641.
39. Reschovsky JD, O'Malley AS. Do primary care physicians treating minority patients report problems delivering high-quality care? *Health Aff (Millwood)*. 2008;27(3):w222-w231.
40. Bach PB, Pham HH, Schrag D, Tate RC, Hargraves JL. Primary care physicians who treat blacks and whites. *N Engl J Med*. 2004;351(6):575-584.
41. Hong CS, Atlas SJ, Chang Y, et al. Relationship between patient panel characteristics and primary care physician clinical performance rankings. *JAMA*. 2010;304(10):1107-1113.
42. National Quality Forum. Risk Adjustment for Sociodemographic Factors. 2014. http://www.qualityforum.org/risk_adjustment_sessions.aspx. Accessed Apr 29, 2014.
43. Fiscella K, Burstin HR, Nerenz DR. Quality measures and sociodemographic risk factors: to adjust or not to adjust. *JAMA*. 2014;312(24):2615-2616.
44. Kent DM, Shah ND. Risk models and patient-centered evidence: should physicians expect one right answer? *JAMA*. 2012;307(15):1585-1586.
45. Weiner SJ, Schwartz A, Sharma G, et al. Patient-centered decision making and health care outcomes: an observational study. *Ann Intern Med*. 2013;158(8):573-579.
46. Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: connecting patients to context. *J Am Med Inform Assoc*. 2015;22(1):2-10.
47. Henderson KM, Evans JR. Successful implementation of Six Sigma: benchmarking General Electric Company. *Benchmarking*. 2000;7(4):260-282.
48. Leard LE, Savides TJ, Ganiats TG. Patient preferences for colorectal cancer screening. *J Fam Pract*. 1997;45(3):211-218.
49. Anderson WF, Guyton KZ, Hiatt RA, Vernon SW, Levin B, Hawk E. Colorectal cancer screening for persons at average risk. *J Natl Cancer Inst*. 2002;94(15):1126-1133.
50. Green AR, Peters-Lewis A, Percac-Lima S, et al. Barriers to screening colonoscopy for low-income Latino and white patients in an urban community health center. *J Gen Intern Med*. 2008;23(6):834-840.
51. Sharma AE, Willard-Grace R, Hessler D, Bodenheimer T, Thom DH. What happens after health coaching? Observational study 1 year following a randomized controlled trial. *Ann Fam Med*. 2016;14(3):200-207.
52. Damburg CL, Baker DW. Improving the quality of quality measurement. *J Gen Intern Med*. 2016;14(3):200-207.
53. Stange KC, Zyzanski SJ, Jaén CR, et al. Illuminating the 'black box'. A description of 4454 patient visits to 138 family physicians. *J Fam Pract*. 1998;46(5):377-389.
54. Magill MK, Ehrenberger D, Scammon DL, et al. The cost of sustaining a patient-centered medical home: experience from 2 states. *Ann Fam Med*. 2015;13(5):429-435.
55. Hollnagel E. *The ETTO principle: Efficiency-Thoroughness Trade-Off*. Aldershot, UK: Ashgate; 2009.
56. Crabtree BF, Nutter PA, Miller WL, et al. Primary care practice transformation is hard work: insights from a 15-year developmental program of research. *Med Care*. 2011;49(Suppl):S2-S35.
57. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q*. 2005;83(3):457-502.
58. Gillam SJ, Siriwardena AN, Steel N. Pay-for-performance in the United Kingdom: impact of the quality and outcomes framework: a systematic review. *Ann Fam Med*. 2012;10(5):461-468.
59. Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay for performance on the quality of primary care in England. *N Engl J Med*. 2009;361(4):368-378.
60. Spence D. Kill the QOF. *BMJ*. 2013;346:f1498.
61. Berwick DM. Measuring physicians' quality and performance: adrift on Lake Wobegon. *JAMA*. 2009;302(22):2485-2486.
62. Kontopantelis E, Springate DA, Ashworth M, Webb RT, Buchan IE, Doran T. Investigating the relationship between quality of primary care and premature mortality in England: a spatial whole-population study. *BMJ*. 2015;350:h904.
63. 'Outdated' QOF GP payment system scrapped in Scotland. *BBC News*. Oct 1, 2015. <http://www.bbc.com/news/uk-scotland-scotland-politics-34419123>. Accessed Mar 3, 2016.
64. McGlynn EA, Adams JL. What makes a good quality measure? *JAMA*. 2014;312(15):1517-1518.
65. Doran T, Kontopantelis E, Fullwood C, Lester H, Valderas JM, Campbell S. Exempting dissenting patients from pay for performance schemes: retrospective analysis of exception reporting in the UK Quality and Outcomes Framework. *BMJ*. 2012;344:e2405.
66. McCartney M, Treadwell J, Maskrey N, Lehman R. Making evidence based medicine work for individual patients. *BMJ*. 2016;353:i2452.
67. Goitein L, James B. Standardized best practices and individual craft-based medicine: a conversation about quality. *JAMA Intern Med*. 2016;176(6):835-838.
68. Young RA, Bayles B, Benold TB, Hill JH, Kumar KA, Burge S. Family physicians' perceptions on how they deliver cost-effective care: a qualitative study from the Residency Research Network of Texas (RRNet). *Fam Med*. 2013;45(5):311-318.
69. ABIM Foundation. Choosing Wisely. <http://www.choosingwisely.org>. Accessed Sep 15, 2012.
70. Bazemore A, Petterson S, Peterson LE, Phillips RL Jr. More comprehensive care among family physicians is associated with lower costs and fewer hospitalizations. *Ann Fam Med*. 2015;13(3):206-213.
71. Hussey PS, Schneider EC, Rudin RS, Fox DS, Lai J, Pollack CE. Continuity and the costs of care for chronic disease. *JAMA Intern Med*. 2014;174(5):742-748.
72. Saultz JW, Lochner J. Interpersonal continuity of care and care outcomes: a critical review. *Ann Fam Med*. 2005;3(2):159-166.
73. Wolinsky FD, Bentler SE, Liu L, et al. Continuity of care with a primary care physician and mortality in older adults. *J Gerontol A Biol Sci Med Sci*. 2010;65(4):421-428.
74. Casalino LP, Pesko MF, Ryan AM, et al. Small primary care physician practices have low rates of preventable hospital admissions. *Health Aff (Millwood)*. 2014;33(9):1680-1688.
75. Federman AD, Halm EA, Siu AL. Use of generic cardiovascular medications by elderly Medicare beneficiaries receiving generalist or cardiologist care. *Med Care*. 2007;45(2):109-115.
76. Milstein A, Gilbertson E. American medical home runs. *Health Aff (Millwood)*. 2009;28(5):1317-1326.
77. Wexler R. Patient satisfaction as a quality metric promotes bad medicine. *JAMA Intern Med*. 2014;174(8):1418-1419.
78. Kupfer JM, Bond EU. Patient satisfaction and patient-centered care: necessary but not equal. *JAMA*. 2012;308(2):139-140.
79. Fenton JJ, Jerant AF, Bertakis KD, Franks P. The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch Intern Med*. 2012;172(5):405-411.
80. Mold JW, Lawler F, Schauf KJ, Aspy CB. Does patient assessment of the quality of the primary care they receive predict subsequent outcomes? An Oklahoma Physicians Resource/Research Network (OKPRN) study. *J Am Board Fam Med*. 2011;24(5):511-523. Retracted in *J Am Board Fam Med*. 2012;25(4):550.
81. Reuben DB, Tinetti ME. Goal-oriented patient care—an alternative health outcomes paradigm. *N Engl J Med*. 2012;366(9):777-779.
82. May C, Montori VM, Mair FS. We need minimally disruptive medicine. *BMJ*. 2009;339:b2803.
83. Martin JC, Avant RF, Bowman MA, et al.; Future of Family Medicine Project Leadership Committee. The future of family medicine: a collaborative project of the family medicine community. *Ann Fam Med*. 2004;2(Suppl 1):S3-S32.
84. Katerndahl D. Providing complex (rather than complicated) chronic care. *J Am Board Fam Med*. 2014;27(1):6-7.