

PERSPECTIVE

Challenges in Assessing the Process–Outcome Link in Practice

Layla Parast, PhD¹, Brian Doyle, MD², Cheryl L. Damberg, PhD, MPH¹, Kanaka Shetty, MD, MS¹, David A. Ganz, MD, PhD^{1,2,3}, Neil S. Wenger, MD, MPH^{1,2}, and Paul G. Shekelle, MD, PhD^{1,2,3}

¹RAND Corporation, Santa Monica, CA, USA; ²David Geffen School of Medicine at UCLA, Los Angeles, CA, USA; ³VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA.

The expanded use of clinical process-of-care measures to assess the quality of health care in the context of public reporting and pay-for-performance applications has led to a desire to demonstrate the value of such efforts in terms of improved patient outcomes. The inability to observe associations between improved delivery of clinical processes and improved clinical outcomes in practice has raised concerns about the value of holding providers accountable for delivery of clinical processes of care. Analyses that attempt to investigate this relationship are fraught with many challenges, including selection of an appropriate outcome, the proximity of the outcome to the receipt of the clinical process, limited power to detect an effect, small expected effect sizes in practice, potential bias due to unmeasured confounding factors, and difficulties due to changes in measure specification over time. To avoid potentially misleading conclusions about an observed or lack of observed association between a clinical process of care and an outcome in the context of observational studies, individuals conducting and interpreting such studies should carefully consider, evaluate, and acknowledge these types of challenges.

KEY WORDS: quality improvement; program evaluation; medicare; preventive care; randomized trials.

J Gen Intern Med 30(3):359–64

DOI: 10.1007/s11606-014-3150-0

© Society of General Internal Medicine 2015

The use of clinical process-of-care measures to assess the quality of health care has grown rapidly in the past 20 years. Process measures are commonly used for internal quality assessment and improvement activities, for external accountability, for pay-for-performance (P4P) and value-based purchasing, and for regulatory purposes.^{1, 2} This use is motivated by the following chain of events: 1) scientific studies—often randomized controlled trials (RCTs)—find health benefits for a specific process of care for a specific

population; 2) professional consensus develops that delivering this process of care for a particular population represents good-quality care; and 3) the process of care is specified as a performance measure and used to assess, improve, and report quality of care. For example, RCTs found as early as 1981 that beta blocker use decreased mortality rates after acute myocardial infarction (AMI).^{3–6} The regular use of beta blockers in post-AMI patients was advocated in authoritative clinical practice guidelines a decade later,⁷ and the use of beta blockers was incorporated into the National Committee for Quality Assurance Health Plan Employer Data and Information Set in 1996⁸ and into multiple other performance measurements systems thereafter.^{9, 10} Similar stories exist for other processes of care, such as the use of influenza vaccinations, colorectal cancer screening, and angiotensin-converting enzyme inhibitors for systolic heart failure. As providers have increased their delivery of these processes of care, there was an expectation that this would lead to improvements in patient outcomes. Yet many studies have found minimal or even no differences in clinical outcomes between high- and low-performing providers on certain process-of-care measures (where “high-performing” indicates that the provider has delivered the process of care to a high percentage of eligible patients).^{11–15} This has left providers and health policymakers wondering whether the focus on processes of care is misplaced.

In this paper, we discuss important challenges with regard to analyses that attempt to relate delivery of processes of care with changes in patient outcomes. We highlight analytical issues that should be considered in attempts to assess the process–outcome link in practice, such that results of the analyses can be appropriately interpreted by clinicians and health policymakers. These challenges can be grouped into four general domains: the choice of outcome, the power to detect differences in outcome, the ability to explain or control for confounding, and the stability of measure specification over time. Acknowledging these challenges is important for both the individuals conducting the analysis as well as those interpreting and disseminating the results.

Electronic supplementary material The online version of this article (doi:10.1007/s11606-014-3150-0) contains supplementary material, which is available to authorized users.

Received June 6, 2014

Revised September 10, 2014

Accepted November 26, 2014

Published online January 7, 2015

WHAT IS THE EXPECTED OUTCOME OF THE PROCESS OF CARE?

The first challenge in analyses that investigate the effect of a process of care on a clinical outcome is in determining exactly

what outcome should be used.¹⁶ While the outcome evaluated in a clinical trial or other evidence that serves as the basis for a recommended process of care is known, this outcome is frequently not available from data that is accessible to the analyst, or it may be measured with more error than in the clinical trial. For example, RCT evidence has demonstrated that aspirin reduces the combined outcome of any serious vascular event by 25 % among patients with acute or previous vascular disease.¹⁷ However, in assessing, for example, whether receipt of aspirin upon hospital arrival after an AMI has led to improved outcomes in the Medicare population, it would be very difficult to obtain information about confirmed and adjudicated serious vascular events if one were restricted to Medicare claims or other administrative data. Instead, analyses attempting to gauge the impact of receipt of aspirin have often approximated the outcome with available data, using vascular event-related readmission or death.^{18, 19} While these are outcomes of importance to both patients and providers, their use in process–outcome analyses requires an assumption that the expected improvement attributable to the process of care is approximately equal to that for the outcome used in the clinical trial. The validity of the analysis rests on the strength of this assumption.

WHAT IS THE PROXIMITY OF THE OUTCOME TO THE PROCESS OF CARE?

Another issue in selecting an outcome is the proximity of the observed outcome to the delivered process of care. This is of particular concern with processes of care wherein multiple steps or long periods of time are required between the delivery of the process of care and the outcome. Preventive processes of care typically have both of these limitations. The more steps that are required, the more difficult it is to be confident that any difference in outcome is due to the delivery of the specific process of care. The longer the period of time that is needed to produce an observable effect on the outcome—thus involving a longer follow-up period—the more difficult it is to attribute differences in outcomes to care received in the distant past. This difficulty in determining an outcome that is proximate and appropriate for the process of care being evaluated means that many preventive processes of care may not be amenable to analyses aimed at evaluating associations between process measure performance and patient outcome.

WHAT IS THE POTENTIAL OBSERVED DIFFERENCE IN OUTCOME, GIVEN THE OBSERVED VARIATION IN PROVIDER PERFORMANCE?

In a provider-level analysis (i.e., where the unit of analysis is the hospital or long-term care facility), the lack of power is often attributable to small differences in performance among providers. When there is little variation in provider performance (e.g., hospital performance on a measure clustered

nationally around 80 %), the potential observable difference in health outcomes that is attainable is very small.²⁴ As an example, we consider the use of beta blockers following an AMI. RCT evidence in more than 20,000 patients has demonstrated a 1.2-percentage-point absolute reduction in mortality rates in AMI patients following the use of beta blockers (Fig. 1a).²⁵ In other words, by comparing a group where 0 % received therapy to a group where 100 % received therapy, it was estimated that receipt of beta blocker therapy decreased the probability of death by 1.2 %. Now, suppose that one wanted to compare a hospital with “low” performance on the process measure assessing beta blocker therapy at discharge (e.g., a hospital whose performance rate was at the 25th percentile) vs. a hospital with “high performance” (e.g., a hospital at the 75th percentile). According to Hospital Compare, hospital performance in 2004 at the 25th percentile was 87 % (meaning 87 % of patients received beta blocker therapy at discharge), while performance at the 75th percentile was 97 %.²⁶ Given this 10-percentage-point difference in performance and the 1.2 % expected reduction in mortality, we would expect to observe only a 0.12-percentage-point difference in mortality when comparing these two hospitals (Fig. 1b).¹² Hence, the “failure” to demonstrate large differences in mortality between high- and low-performing hospitals may not be due to any lack of effect of the process of care on the outcome, but rather to insufficient variation in performance.

ARE THOSE NOT RECEIVING THE PROCESS OF CARE A SMALL PROPORTION OF THE POPULATION?

Many processes of care have mean performance rates now exceeding 90 %, or even 95 %. When a very small proportion of the population specified by the measure is not receiving the process of care, there are two problems, one statistical and the other inferential. The statistical problem is that with a fixed total sample size, the ability to detect differences in outcomes is greatly diminished as the number of patients not receiving the process of care decreases. Table 1 illustrates this point, showing the minimum detectable percentage point change in outcome with 80 % power, assuming a total population of 10,000 and an average outcome rate of 10 % without the process of care (e.g., 10 % mortality among those who do not receive a beta blocker prescription at hospital discharge after AMI). Table 1 shows that for a fixed total sample size, when the number of those receiving versus not receiving the process of care is more evenly split, smaller changes in the outcome can be detected, i.e., one could detect a 10.0 % versus 8.4 % difference with 5,000 patients in each group. However, when this split is largely unbalanced, only large changes in outcome are detectable, i.e., one could detect only a 10.0 % vs. 3.2 % difference in mortality with 9,900 in one group and 100 in the other. This makes it difficult to detect small effect sizes

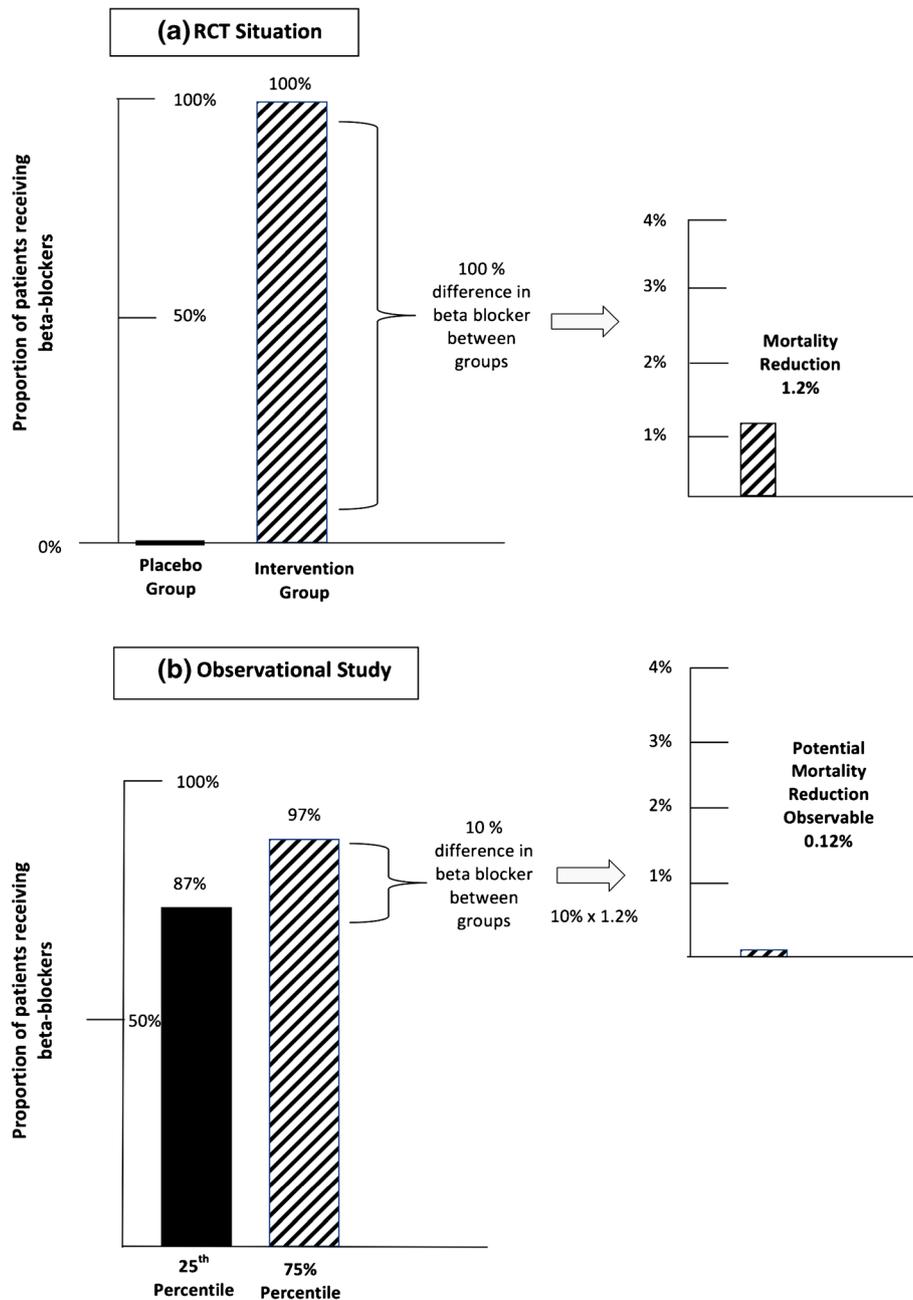


Fig. 1 (a) RCT evidence supporting beta blockers after an AMI demonstrated a 1.2-percentage-point absolute reduction in mortality; (b) a 10-percentage-point difference in performance between two hospitals would lead us to expect only a 0.12-percentage-point reduction in mortality.

Table 1 Minimum detectable effect sizes, assuming an average outcome rate of 10 % without the process of care and a total sample size of 10,000

Hypothetical proportion (N) of care	Hypothetical proportion (N) not receiving process of care	Minimum detectable percentage (absolute) change in outcome with 80 % power
50 % (5,000)	50 % (5,000)	1.6
75 % (7,500)	25 % (2,500)	1.9
95 % (9,500)	5 % (500)	3.5
99 % (9,900)	1 % (100)	6.8

equivalent in magnitude to those generally seen in RCTs, which often report absolute risk reduction of less than 5%.^{17, 20–22}

The inferential problem arises if the small fraction of individuals who do not receive the process of care differ in important ways from those who do. With any process measure, it is improbable that the measure will be applicable to the entire target population. There will always be some subset of individuals with characteristics that make the process of care inappropriate for them, such as rare contraindications or prior rare side effects, but which lie outside the exclusion criteria.

Supporting this conclusion is the experience of the UK National Health Service Quality and Outcomes Framework (QOF), a P4P program that allows general practitioners (GPs) to exclude from measurement any patient for whom the GP believes the process of care should not be applied (“exception reporting”). A recent UK study examining exception reporting in the QOF program found that a median of 5.3 % of patients were excluded from quality calculations.²³ If such patients exist equally in the United States (where exception reporting is not in place), they likely comprise the majority of patients not receiving “topped out” processes of care, those where 90–95 % of patients nationally are receiving the process of care. When this occurs, there is an equal or greater likelihood that any differences in outcomes between patients who do and do not receive the process of care are as a result of differences in the patients being compared rather than the receipt of the process of care.

IS THE ANALYSIS ROBUST TO POTENTIAL UNMEASURED CONFOUNDING?

Another potential problem inherent in all observational studies is the possibility of unmeasured confounding. When this occurs, the observed relationship between the process of care and the outcome may be biased. For example, in an observational study examining the effect of beta blockers in a population-based cohort of elderly patients with heart failure, patients were less likely to be prescribed a beta blocker if they were older or had several comorbid conditions.²⁷ Such patients are also at a higher risk of death and hospital readmission. Thus, if information concerning comorbid conditions is not available, the estimated effect of receiving the beta blockers may be overly optimistic, as it does not account for the fact that patients with these conditions are less likely to receive the process of care and more likely to experience negative outcomes.

While it is impossible to ensure that there are no unmeasured factors influencing the relationship between the process of care and outcome, it is possible to quantitatively assess the robustness of analysis to such unmeasured confounders.^{21, 27–30} To illustrate, we simulated a study investigating the association between beta blocker prescription at discharge after an AMI and a negative outcome (the details of which are in the [online appendix](#)). As shown in Table 2, confounders with relatively modest associations between receipt of the process of care and outcome (odds ratios of 1.3–1.5) can increase the probability of finding an association when none exists to as high as 1 in 10 or even 1 in 8, or they can increase the chance of not finding an association when one does exist to nearly 50 %. Thus, an important component in conducting these analyses should be to consider potential unmeasured confounders not simply by mentioning them as possible limitations, but by quantifying this possibility. While clinical expertise is needed to identify potential confounding factors,

Table 2 Simulation assessing the percentage of patients prescribed beta blockers at discharge after an AMI: sensitivity of type I and type II (power) error rates to an unmeasured confounder

Odds ratio reflecting the hypothetical relationship between the receipt of the process of care and the unmeasured confounder	Odds ratio reflecting the hypothetical relationship between the outcome and the unmeasured confounder	Probability that a significant association will be detected when there is truly no association (type I error)	Probability that a significant association will be detected when there is a true effect (power)
1	1	0.05	0.86
1.2	1.2	0.05	0.81
1.4	1.2	0.06	0.76
1.3	1.4	0.08	0.70
1.4	1.4	0.09	0.65
1.5	1.4	0.11	0.59
1.5	1.5	0.13	0.53

statistical simulations such as this can provide estimates of the sensitivity of the results to the specified unmeasured confounders.^{28–30}

WHAT ARE THE POTENTIAL ANALYTICAL IMPACTS OF MEASURE SPECIFICATION CHANGES?

The ability to examine the association between changes in measure performance over time and outcomes can be hindered by changes in measure specifications. Changes to both the numerator and denominator of a measure often occur once a measure is in use. While these changes generally lead to better-specified measures, analyzing changes in performance and associated changes in outcomes across time is very difficult. For example, one of the Hospital Inpatient Quality Reporting Program process-of-care measures has had eight specification changes since January 2006.³¹ Following one of these changes, which added further detail on patients that should be excluded from eligibility, the number of eligible patients decreased dramatically, by 4,000 patients per month, and mortality rates increased among those receiving the process of care (26 % to 30 % mortality rate). This would indicate that the change likely removed from eligibility those patients at a lower risk of mortality. As such, any comparison between processes and outcomes over time for this measure would need to consider whether this change in specification may have influenced any observed relationship.

IMPLICATIONS FOR RESEARCHERS AND POLICYMAKERS

Despite a strong desire to test for links between processes of care and outcomes in practice, such analyses need to be carefully considered. The clinical and analytic issues involved in conducting such studies are substantial, and lack of attention to the issues described in this paper will increase the chance of producing a result that may mislead providers and

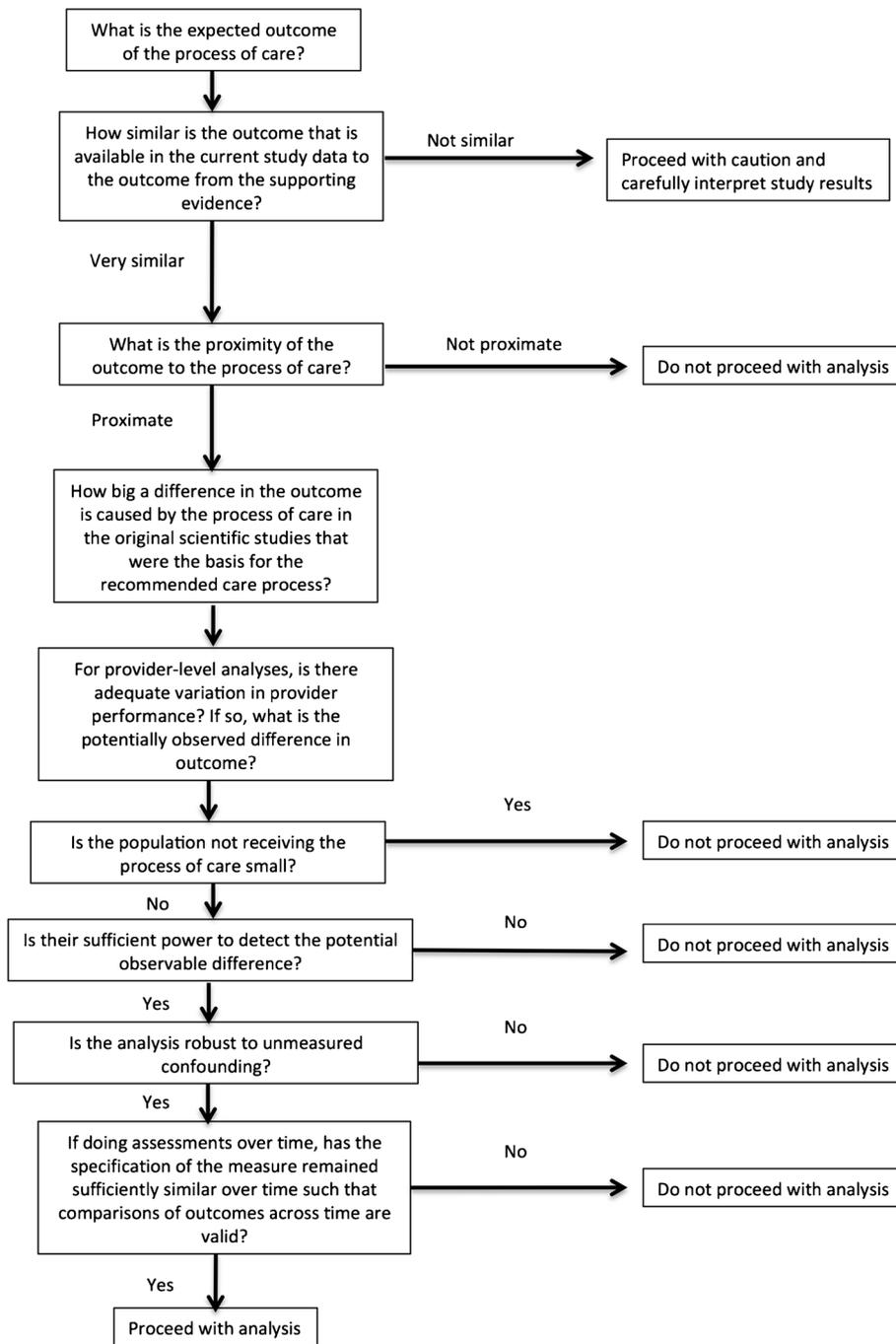


Fig. 2 Questions to address before undertaking or when interpreting results from a process–outcome analysis

policymakers regarding associations, or lack thereof, between recommended processes of care and outcomes.

Just as researchers contemplating a clinical trial of a new therapy must consider beforehand whether their patient selection criteria are optimal, whether they are appropriately measuring the relevant outcome, and whether they have enough power to detect a clinically important effect, individuals assessing the relationship between a process of care and an outcome or interpreting the results of such studies should carefully consider the questions outlined in Fig. 2.

Specifically, researchers should ensure that study designs and analyses are robust to these potential issues and that analysis results are appropriately presented in the context of the study limitations. Similarly, policymakers interpreting these results should review the analyses to determine whether the challenges discussed here raise questions regarding the validity of the findings.

Observational studies of associations between processes of care and relevant short-term outcomes are an important component in the evaluation of policies that encourage better

health care outcomes. However, the interpretation of study results must consider the strengths and weaknesses of the analysis as well as the strength of the existing evidence base supporting the process–outcome relationship. One observational study that does not find a process–outcome relationship is unlikely to be stronger evidence than a series of RCTs that has.

Acknowledgments: This project was funded by the Centers for Medicare and Medicaid Services (CMS) under contract number HHSM-500-2011-AZOS.SIP.02.

Conflict of Interest: David A. Ganz is a co-investigator on a UCLA Contract from SCAN Health Plan and a co-investigator on a RAND Contract from Health Services Advisory Group. All other authors declare no conflicts of interest.

Financial Support: This project was funded by the Centers for Medicare and Medicaid Services under contract number HHSM-500-2011-AZOS.SIP.02.

Corresponding Author: Layla Parast, PhD; RAND Corporation, 1776 Main Street, PO Box 2138, Santa Monica, CA 90407-2138, USA (e-mail: parast@rand.org).

REFERENCES

- Centers for Medicare and Medicaid Services. Hospital Quality Initiative. Available at: <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/index.html>. Accessed October 16, 2014.
- Centers for Medicare and Medicaid Services. CMA Quality Initiatives. Available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/index.html?redirect=/QualityInitiativesGenInfo/01_overview.asp. Accessed October 16, 2014.
- Beta-Blocker Heart Attack Study Group.** The beta-blocker heart attack trial. *JAMA*. 1981;246(18):2073–2074.
- Norwegian Multicenter Study Group.** Timolol-induced reduction in mortality and reinfarction in patients surviving acute myocardial infarction. *N Engl J Med*. 1981;304(14):801–807.
- Hjalmarson A, Elmfeldt D, Herlitz J, et al.** Effect on mortality of metoprolol in acute myocardial infarction. A double-blind randomised trial. *Lancet*. 1981;2(8251):823–827.
- Yusuf S, Peto R, Lewis J, Collins R, Sleight P.** Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis*. 1985;27(5):335–371.
- Gunnar RM, Bourdillon PD, Dixon DW, et al.** ACC/AHA guidelines for the early management of patients with acute myocardial infarction. A report of the American College of Cardiology/American Heart Association Task Force on Assessment of Diagnostic and Therapeutic Cardiovascular Procedures (subcommittee to develop guidelines for the early management of patients with acute myocardial infarction). *Circulation*. 1990;82(2):664–707.
- National Committee for Quality Assurance (NCQA). HEDIS and Performance Measurement.; Available at: <http://www.ncqa.org/HEDISQualityMeasurement.aspx>. Accessed on October 16, 2014.
- Kramer JM, Hammill B, Anstrom KJ, et al.** National evaluation of adherence to beta-blocker therapy for 1 year after acute myocardial infarction in patients with commercial health insurance. *Am Heart J*. 2006;152(3):454 e451–458.
- Mencke NM, Alley LG, Etchason J.** Application of HEDIS measures within a Veterans Affairs medical center. *Am J Manag Care*. 2000;6(6):661–668.
- Stulberg JJ, Delaney CP, Neuhauser DV, Aron DC, Fu P, Koroukian SM.** Adherence to surgical care improvement project measures and the association with postoperative infections. *JAMA*. 2010;303(24):2479–2485.
- Werner RM, Bradlow ET.** Relationship between Medicare’s hospital compare performance measures and mortality rates. *JAMA*. 2006;296(22):2694–2702.
- Jha AK, Joynt KE, Orav EJ, Epstein AM.** The long-term effect of premier pay for performance on patient outcomes. *N Engl J Med*. 2012;366(17):1606–1615.
- Ingraham AM, Cohen ME, Bilimoria KY, et al.** Association of surgical care improvement project infection-related process measure compliance with risk-adjusted outcomes: implications for quality measurement. *J Am Coll Surg*. 2010;211(6):705–714.
- Werner RM, Konetzka RT, Stuart EA, Norton EC, Polsky D, Park J.** Impact of public reporting on quality of postacute care. *Health Ser Res*. 2009;44(4):1169–1187.
- Lilford RJ, Chilton PJ, Hemming K, Girling AJ, Taylor CA, Barach P.** Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. *BMJ*. 2010;341:c4413.
- Antithrombotic Trialists C.** Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002;324(7329):71–86.
- Werner RM, Bradlow ET.** Public reporting on hospital process improvements is linked to better patient outcomes. *Health Aff*. 2010;29(7):1319–1324.
- Ryan AM, Nallamothu BK, Dimick JB.** Medicare’s public reporting initiative on hospital quality had modest or no impact on mortality from three key conditions. *Health Aff*. 2012;31(3):585–592.
- Moher D, Dulberg CS, Wells GA.** Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA*. 1994;272(2):122–124.
- Isis I, Collaborative Group.** Randomized trial of intravenous atenolol among 16027 cases of suspect acute myocardial infarction. *Lancet*. 1986;2:57–66.
- Moberley S, Holden J, Tatham DP, Andrews RM.** Vaccines for preventing pneumococcal infection in adults. *Cochrane Database Syst Rev*. 2013;1:CD000422.
- Doran T, Fullwood C, Reeves D, Gravelle H, Roland M.** Exclusion of patients from pay-for-performance targets by English physicians. *N Engl J Med*. 2008;359(3):274–284.
- Mant J, Hicks N.** Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ*. 1995;311(7008):793–796.
- Freemantle N, Cleland J, Young P, Mason J, Harrison J.** beta Blockade after myocardial infarction: systematic review and meta regression analysis. *BMJ*. 1999;318(7200):1730–1737.
- Official Hospital Compare Data. Centers for Medicare & Medicaid Services. Available at: <https://data.medicare.gov/data/hospital-compare>. Accessed on October 16, 2014.
- Sin DD, McAlister FA.** The effects of beta-blockers on morbidity and mortality in a population-based cohort of 11,942 elderly patients with heart failure. *Am J Med*. 2002;113(8):650–656.
- Greenland S.** Basic methods for sensitivity analysis of biases. *Int J Epidemiol*. 1996;25(6):1107–1116.
- Lin DY, Psaty BM, Kronmal RA.** Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics*. 1998;54(3):948–963.
- Vanderweele TJ, Arah OA.** Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*. 2011;22(1):42–52.
- Centers for Medicare and Medicaid Services. Hospital Inpatient Quality Reporting Program: Process of Care Measures. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalProcessOfCareMeasures.html>. Accessed October 16, 2014.