Association of National Hospital Quality Measure adherence with long-term mortality and readmissions

David M Shahian, Gregg S Meyer, Elizabeth Mort, Susan Atamian, Xiu Liu, Andrew S Karson, Lawrence D Ramunno, Hui Zheng

ABSTRACT

Background: In existing studies, the association between adherence with recommended hospital care processes and subsequent outcomes has been inconsistent. This has substantial implications because process measure scores are used for accountability, quality improvement and reimbursement. Our investigation addresses methodological concerns with previous studies to better clarify the process-outcomes association for three common conditions.

Methods: The study included all patients discharged from Massachusetts General Hospital between 1 July 2004 and 31 December 2007 with a principle diagnosis of acute myocardial infarction (AMI), heart failure (HF) or pneumonia (PN) who were eligible for at least one National Hospital Quality Measure. The number of patients analysed varied by measure (374 to 3020) depending on Centers for Medicare and Medicaid Services eligibility criteria. Hospital data were linked with state administrative data to determine mortality and readmissions. For patients with multiple admissions, the time-weighted impact of measure failures on mortality was estimated using exponential decay functions. All patients had follow-up for at least 1 year or until death or readmission. Cox models were used to estimate HRs adjusted for transfer status, age, gender, race, census block-group socioeconomic status, number of Elixhauser comorbidities, and do not resuscitate orders.

Results: Adjusted survival and freedom from readmission for AMI and PN showed superior results for 100% and 50–99% adherence compared with 0–49% adherence. For HF, the results were inconsistent and sometimes paradoxical, although several individual measures (eg, ACE inhibitor/angiotensin receptor blockade) were associated with improved outcomes.

Conclusion: Adherence with recommended AMI and PN care processes is associated with improved long-term outcomes, whereas the results for HF measures are inconsistent. The evidence base for all process measures must be critically evaluated, including the strength of association between these care processes and outcomes in real-world populations. Some currently recommended processes may not be suitable as accountability measures.

INTRODUCTION

Performance measurement and reporting are central features of healthcare reform, and hospitals devote considerable resources to optimising their scores on these measures. Because of their implications for accountability, referrals and reimbursement, performance measures should be based on the highest level of evidence. Important evaluation criteria include the quality, consistency and quantity of the aggregate evidence base, and high magnitude of net benefit.

Outcomes measures are the preferred modality for assessing performance because they integrate the net impact of measured and unmeasured processes and structures of care, an important principle noted by Donabedian nearly half a century ago. However, for many diagnoses, accurate estimation of outcomes is challenging because of small sample sizes, infrequent adverse outcomes, unavailable or reliable data, and inadequate risk models to account for patient severity. Because of these limitations, process and structure measures have been used to assess performance for certain diagnoses, either alone or in combination with outcomes. Some of these measures are used for public reporting, including the Centers for Medicare and Medicaid Services (CMS) Hospital Compare National Hospital Quality Measures (NHQM) for acute myocardial infarction (AMI), heart failure (HF) and pneumonia (PN).

Process measures used to assess provider performance should have a demonstrable,
proximate association with important outcomes such as mortality and readmission. Unfortunately, studies of these associations have had inconsistent and inconclusive results, in some instances raising serious concerns as to the suitability of such measures for provider profiling. For example, Werner and Bradlow studied AMI, HF and PN and found limited association of hospital-level, individual process measure adherence and hospital mortality at 30 days and 1 year.

Studies of AMI or acute coronary syndromes have provided differing results depending on timing of the endpoint and hospital versus patient-level outcomes. Overall, the evidence base for AMI process performance measures is relatively strong, particularly for some measures such as timeliness of revascularisation (eg, ‘door to balloon time’). However, many such studies have focused on short-term mortality and have not investigated the potential association of measure adherence with longer-term mortality or readmissions.

Studies of PN, the other acute NHQM, have shown consistently positive associations between measures such as antibiotic selection and short-term (eg, hospital or 30-day) mortality. Recent evidence also suggests that, like AMI, the full impact of PN and its treatment may not be apparent until patients are observed for a longer time period, and this has not been done in most process–outcomes studies. It is also difficult to assess the efficacy of inpatient measures such as smoking cessation counselling that may be confounded by post-discharge patient compliance.

Among the three diseases, studies of HF have demonstrated the most problematic link between process measure adherence and outcomes. As a chronic disease with frequent readmissions, longer-term follow-up is important, and the results of current studies raise questions. Using the OPTIMIZE-HF registry, Fonarow and colleagues found no association of American College of Cardiology/American Heart Association endorsed HF performance measures with inhospital mortality. Patterson and colleagues found similar results in a study of hospital-level measure adherence and 1-year mortality and readmissions among Medicare recipients, although both studies were limited by the lack of socioeconomic information, a potentially important confounder. In the study by Fonarow and colleagues, discharge ACE inhibitor/angiotensin receptor blockade (ACEI/ARB) use was associated with decreased 60–90-day mortality and readmission, similar to the findings in previous studies of long-term ACEI use. Discharge β-blockade, not a performance measure because of concerns regarding its use in potentially unstable hospitalised patients, was also associated with reduced mortality; similar findings have been observed for other emerging HF care processes. A few studies do show a process adherence–outcomes association. Kfoury and colleagues found a dose–response association between the number of HF processes used and 1-year mortality, and several studies suggest that adherence with recommended care processes correlated with lower readmission rates.

Validating the directionality and strength of the process–outcomes association is critical. If process adherence does not improve patient outcomes, then scarce resources have been misallocated, and stakeholders may be misinformed about quality of care. The motivation for this study was to expand our understanding of this issue and to address some of the methodological concerns with prior investigations. Many of these studies are based on older data; there was substantial heterogeneity among institutions from which data were aggregated; follow-up duration was inadequate; populations were not inclusive; aggregate hospital results rather than patient-level data were analysed; and potentially important confounders (eg, socioeconomic status) were unavailable.

We explored the process–outcomes association using contemporary data from a large urban tertiary medical center that serves both as a referral center and as the community hospital for a diverse local population. A robust institutional database provided numerous variables unavailable in most administrative and clinical registries. It was thus possible to adjust for these factors and to better isolate the association between process adherence and outcomes.

METHODS

Study population

This observational retrospective cohort study analysed patients discharged from the Massachusetts General Hospital (MGH) between 1 July 2004 and 31 December 2007 with a diagnosis of AMI, HF or PN who were eligible for at least one NHQM. Not all patients who qualified for study entry were eligible for each measure. Therefore, denominators for various measures vary according to CMS eligibility criteria.

Principle outcomes

The main outcomes are all-cause mortality and all-cause readmissions at 90 days and 1 year after hospital discharge. We obtained death and readmission information on each patient for at least 1-year follow-up and until 31 December 2008. For the mortality cohort, survival time was defined as the interval between the last eligible discharge that included a particular measure and either patient death or achievement of 90-day or 1-year survival, whichever was shorter. We chose discharge rather than admission date because we were
focusing on hospital survivors and long-term outcomes, and because many of the measures studied were provided at discharge. Sensitivity analyses were also performed which calculated survival time beginning at hospital admission for measures provided on arrival. For the readmission cohort, freedom from readmission was the period between the first discharge and subsequent readmissions or the 90-day/1-year endpoints, whichever was shorter. Sensitivity analysis was performed with outcome assessed from date of admission rather than discharge for arrival measures.

Patients in the mortality study cohort could have multiple discharges, each of which was regarded as having a time-dependent impact on outcomes (see below). Only one primary discharge per patient was included for the readmission cohort. Subsequent admissions were considered as readmissions, and only the first readmission was included in the study.

Exclusions
Patients who died before hospital discharge were excluded from the study cohort (49 patients with AMI, 25 with PN and none with HF), as were out-of-state patients because their death and readmissions could not be reliably ascertained. For the mortality endpoint, sensitivity analysis was performed in which we included patients who died in hospital.

Further exclusions were applied to the readmission cohort based on CMS readmission model specifications. We excluded patients who were discharged to another short-term general hospital or left against medical advice. For the AMI readmission cohort, patients were excluded if they stayed in hospital for only 1 day (CMS regards such patients as unlikely to have had AMI) and if they also had not been transferred from another acute hospital. The latter criterion is our addition to the CMS rule, reflecting our institutional experience. MGH and other referral centers often receive transfer patients for tertiary evaluation or treatment who have a length of stay of 1 day, but who are nonetheless eligible for at least one AMI measure.

Readmissions within 30 days for percutaneous coronary interventions (PCIs) or coronary artery bypass graft were not counted as AMI readmissions because they are typically planned. For the HF discharge instruction measure, we limited the study cohort to patients who were discharged to home with or without services (having discharge destination code of 01 and 06), thus mitigating any potential confounding by the care rendered at extended care facilities (13 patients excluded from the study cohort).

Data sources and linkages
Study patients were identified from an MGH registry that included patient demographic and clinical characteristics and measure adherence results. This registry included patient identifiers that allowed us to link measure adherence with Massachusetts state mortality and readmission data. Mortality data from 2004 to 2008 were obtained from Massachusetts Vital Records and Statistics. We merged MGH adherence data with state death records by patient name and date of birth. If a match was not found in the state mortality database, we assumed the patient was still alive at the end of the study.

For the readmission cohort, we obtained Massachusetts all-hospital inpatient data from the Massachusetts Division of Health Care Finance and Policy. These data included a unique state-level patient identifier, the admitting hospital and medical record number, admission and discharge dates, and principle and secondary diagnoses. We first merged MGH internal data with the state database to extract the unique patient identifiers, matching by MGH medical record number and admission and discharge dates. Using this approach, 99% of MGH study admissions were matched to the state inpatient database. We then used the unique patient identifier to identify subsequent readmissions from the state inpatient database. Following this initial linking, we stripped unique patient identifiers and all other patient identifiers from the analysis dataset.

Measure aggregation and extent of adherence
At the individual measure level, the term adherence in this study signifies that a patient was reported to CMS as having met the NHQM requirements for a given measure that were applicable at the time of their hospitalisation. As the intent of the study was to estimate the association between measure non-adherence and various outcomes, measure failure or non-adherence was coded 1, and adherence was coded 0. The extent of measure adherence was estimated for each discharge using a variety of techniques.

Mortality method 1—impact of any previous measure failure
For mortality, we included all discharges for a given patient for each specific diagnosis. We estimated the association of measure non-adherence with mortality using two different approaches. In the first method, for each individual measure, we estimated the mortality HRs associated with failure on any hospitalisation. Similarly, we estimated the HRs associated with failure to achieve 100% adherence with all components of a group of related measures (eg, arrival measures, discharge measures), as defined in online table 1, on any admission. Finally, we also estimated an all-or-none score, the most stringent approach. To receive credit, there could be no failures on any measure for which the patient was eligible, on any admission. All-or-none scoring is commonly used, although it may be criticised on
statistical grounds because the joint probability of failure increases with the number of eligible measures, which varies among patients.

Mortality method 2—time-weighted impact of previous measure failures

For the mortality cohort, we also used a second method to account for the fact that some patients had multiple admissions within the study period for the same diagnosis. We hypothesised that the impact on survival of each individual failure decreased exponentially over time. Patients with failure on all admissions would have a score of exactly 1, and adherence on a distant admission with process measure failure on all subsequent admissions would result in a score that approached 1. Conversely, a single remote admission with a measure failure followed by numerous subsequent admissions with perfect adherence would have a score close to 0.

We estimated continuous, time-weighted average adherence scores for multiple admission patients using exponential decay functions (online appendix 1). Online appendix 2 demonstrates how we used this approach to estimate the time-weighted average scores. Average time-weighted adherence scores were then used as independent variables in Cox models. The resulting adjusted HRs estimate the association between each 10% increase in average time-weighted adherence score (where failure=1 and adherence=0) and subsequent mortality. We further categorised all patients into three groups based on their average time-weighted adherence scores to estimate dose-response relationships.

Readmission method

For the readmission cohort, only the first readmission was studied. Subsequent readmissions were disregarded, and it was therefore unnecessary to use time-weighted decay. We studied the association of readmission with non-adherence to individual, group and all-or-none measure (1=measure non-adherence, 0=adherence). We also analysed three categories of adherence, based on the number of adherent measures divided by the number of eligible measures.

Statistical analysis

In the primary analysis, we calculated patient demographic and clinical characteristics and performance for all measures. We used the log-rank test to identify potential confounders and a multivariable Cox proportional hazards model to adjust for confounding effects. The final model covariates included the patient’s age, gender, race (white, other), census block-group socioeconomic status (SES, lowest 25th vs other) using the multi-indicator approach of Diez-Roux and colleagues, the number of Elixhauser comorbidities, do not resuscitate order on admission for arrival measure models, do not resuscitate order anytime during hospitalisation for other measure models, and transfer to MGH for non-arrival measures (CMS rules regard transfer patients as ineligible for arrival measures). Adjusted HRs and 95% CIs were computed for various time points. Cox models were fit separately for individual, group and composite measure outcomes at 90 days and 1 year for all-cause mortality and readmissions.

All statistical analyses were performed with SAS software, V.9.2.

Human subject protection

This research was reviewed and approved by the MGH/Partners Institutional Review Board. Use of level VI Massachusetts state data was conducted in conformity with all applicable state regulations.

RESULTS

Online table 1 lists the percentage adherence for each individual, group and overall measure for the three conditions. Individual measure adherence ranged from 54.8% (influenza vaccine assessment) to 100% (O2 assessment for PN). For each condition, online table 2 lists the bivariate demographic characteristics, clinical features, outcomes (mortality and readmissions), and p values for patients with 100% adherence to all measures versus <100% adherence.

Table 1 shows the 90-day and 365-day adjusted mortality HRs for non-adherence to individual, group and all-or-none measures.

For AMI, failure on the arrival aspirin or β-blocker measures were independently associated with significantly reduced 1-year survival. Arrival measures, discharge measures and the all-or-none measures were all strongly associated with increased risk of death at 90 days and 365 days. After excluding in-hospital deaths, there were insufficient patients to analyse the association of PCI <90 min and late outcomes. Sensitivity analyses (using in-hospital deaths and time from admission rather than discharge) changed the absolute values of most HRs but not their directionality or statistical significance (results available upon request). However, with in-hospital deaths included, failure on the PCI <90 min measure was a highly significant predictor of 90-day (adjusted HR 5.48, 95% CI 1.22 to 24.65) and 365-day mortality (HR 3.62, 95% CI 1.07 to 12.28). HRs for the time-weighted approach (table 1) are directionally similar for most predictors, although in some cases they change from statistically significant to non-significant; the time-weighted categorical variable shows a dose-response relationship. Sensitivity analyses, as described previously, also showed that failure to achieve PCI <90 min was a significant predictor of 90-day and 365-day mortality.
Table 1  Association of measure non-adherence with subsequent all-cause mortality

<table>
<thead>
<tr>
<th>Measure non-adherence on any admission</th>
<th>90-day adjusted* HR (95% CI)</th>
<th>365-day adjusted* HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute myocardial infarction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin at arrival</td>
<td>6.90 (0.85 to 56.13)</td>
<td>6.71 (1.54 to 29.30)</td>
</tr>
<tr>
<td>β blocker at arrival</td>
<td>3.82 (0.76 to 19.08)</td>
<td>4.25 (1.28 to 14.07)</td>
</tr>
<tr>
<td>PCI &lt;90 min</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Aspirin (D/C)</td>
<td>6.23 (0.83 to 46.78)</td>
<td>2.51 (0.35 to 18.27)</td>
</tr>
<tr>
<td>Beta blocker (D/C)</td>
<td>NA</td>
<td>1.65 (0.23 to 11.96)</td>
</tr>
<tr>
<td>ACEI/ARB</td>
<td>1.58 (0.58 to 4.27)</td>
<td>1.24 (0.62 to 2.51)</td>
</tr>
<tr>
<td>Smoking counselling</td>
<td>2.51 (0.71 to 8.93)</td>
<td>1.56 (0.54 to 4.52)</td>
</tr>
<tr>
<td>Arrival group measure</td>
<td>3.46 (1.31 to 9.15)</td>
<td>2.71 (1.28 to 5.74)</td>
</tr>
<tr>
<td>Discharge group measure</td>
<td>2.85 (1.47 to 5.56)</td>
<td>2.18 (1.32 to 3.59)</td>
</tr>
<tr>
<td>Failure on any measure, any admission</td>
<td>2.67 (1.51 to 4.73)</td>
<td>2.08 (1.36 to 3.19)</td>
</tr>
<tr>
<td><strong>Heart failure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACEI/ARB for LVSD</td>
<td>1.98 (1.10 to 3.55)</td>
<td>1.54 (1.03 to 2.29)</td>
</tr>
<tr>
<td>LV assessment</td>
<td>1.88 (0.83 to 4.28)</td>
<td>1.74 (0.92 to 3.29)</td>
</tr>
<tr>
<td>Discharge instruction</td>
<td>1.22 (0.84 to 1.78)</td>
<td>1.23 (0.97 to 1.55)</td>
</tr>
<tr>
<td>Smoking counselling</td>
<td>2.22 (0.83 to 5.95)</td>
<td>2.28 (1.24 to 4.21)</td>
</tr>
<tr>
<td>Patient information measures</td>
<td>1.19 (0.82 to 1.72)</td>
<td>1.25 (0.99 to 1.57)</td>
</tr>
<tr>
<td>Failure on any measure, any admission</td>
<td>0.97 (0.74 to 1.28)</td>
<td>1.03 (0.85 to 1.24)</td>
</tr>
<tr>
<td><strong>Pneumonia</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotic &lt;4 h</td>
<td>1.13 (0.68 to 1.85)</td>
<td>1.05 (0.74 to 1.49)</td>
</tr>
<tr>
<td>Blood culture timing</td>
<td>1.86 (1.07 to 3.21)</td>
<td>1.59 (1.10 to 2.30)</td>
</tr>
<tr>
<td>O2 assessment</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Antibiotic selection</td>
<td>1.42 (0.55 to 3.70)</td>
<td>1.28 (0.69 to 2.36)</td>
</tr>
<tr>
<td>Influenza vaccine assessment</td>
<td>0.97 (0.57 to 1.68)</td>
<td>1.04 (0.71 to 1.52)</td>
</tr>
<tr>
<td>Pneumonia vaccine assessment</td>
<td>1.06 (0.71 to 1.59)</td>
<td>1.18 (0.90 to 1.56)</td>
</tr>
<tr>
<td>Smoking counselling</td>
<td>1.19 (0.53 to 2.68)</td>
<td>1.59 (0.93 to 2.70)</td>
</tr>
<tr>
<td>Arrival group measures</td>
<td>1.46 (0.98 to 2.18)</td>
<td>1.20 (0.91 to 1.57)</td>
</tr>
<tr>
<td>Health maintenance measures</td>
<td>1.07 (0.73 to 1.56)</td>
<td>1.10 (0.86 to 1.44)</td>
</tr>
<tr>
<td>Failure on any measure, any admission</td>
<td>1.23 (0.87 to 1.75)</td>
<td>1.24 (0.98 to 1.58)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time-weighted measure non-adherence</th>
<th>90-day adjusted* HR† (95% CI)</th>
<th>365-day adjusted* HR† (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute myocardial infarction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin at arrival</td>
<td>1.05 (0.69 to 1.59)</td>
<td>1.16 (0.96 to 1.39)</td>
</tr>
<tr>
<td>Beta blocker at arrival</td>
<td>1.16 (0.98 to 1.37)</td>
<td>1.20 (1.06 to 1.36)</td>
</tr>
<tr>
<td>PCI &lt;90 min</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Aspirin (D/C)</td>
<td>1.20 (0.98 to 1.47)</td>
<td>1.10 (0.90 to 1.34)</td>
</tr>
<tr>
<td>Beta blocker (D/C)</td>
<td>NA</td>
<td>1.05 (0.86 to 1.28)</td>
</tr>
<tr>
<td>ACEI/ARB</td>
<td>1.05 (0.95 to 1.16)</td>
<td>1.02 (0.95 to 1.09)</td>
</tr>
<tr>
<td>Smoking counselling</td>
<td>1.10 (0.97 to 1.25)</td>
<td>1.05 (0.94 to 1.17)</td>
</tr>
<tr>
<td>Arrival group measures</td>
<td>1.14 (0.98 to 1.33)</td>
<td>1.15 (1.03 to 1.27)</td>
</tr>
<tr>
<td>Discharge group measures</td>
<td>1.22 (1.09 to 1.38)</td>
<td>1.18 (1.08 to 1.29)</td>
</tr>
<tr>
<td>Average non-adherence score</td>
<td>1.16 (1.04 to 1.29)</td>
<td>1.17 (1.08 to 1.26)</td>
</tr>
<tr>
<td>0–49% adherence (vs all)</td>
<td>3.06 (1.30 to 7.16)</td>
<td>3.06 (1.72 to 5.43)</td>
</tr>
<tr>
<td>50–99% adherence (vs all)</td>
<td>2.45 (1.19 to 5.06)</td>
<td>1.53 (0.84 to 2.78)</td>
</tr>
<tr>
<td><strong>Heart failure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACEI/ARB for LVSD</td>
<td>1.06 (0.99 to 1.13)</td>
<td>1.04 (0.99 to 1.08)</td>
</tr>
<tr>
<td>LV assessment</td>
<td>1.03 (0.93 to 1.14)</td>
<td>1.03 (0.95 to 1.11)</td>
</tr>
<tr>
<td>Discharge instruction</td>
<td>1.01 (0.97 to 1.05)</td>
<td>1.01 (0.99 to 1.04)</td>
</tr>
<tr>
<td>Smoking counselling</td>
<td>1.07 (0.96 to 1.19)</td>
<td>1.06 (0.99 to 1.14)</td>
</tr>
<tr>
<td>Patient information measures</td>
<td>0.99 (0.95 to 1.03)</td>
<td>1.01 (0.98 to 1.03)</td>
</tr>
<tr>
<td>Average non-adherence score</td>
<td>0.97 (0.91 to 1.03)</td>
<td>0.99 (0.95 to 1.03)</td>
</tr>
<tr>
<td>0–49% adherence (vs all)</td>
<td>0.68 (0.46 to 1.01)</td>
<td>0.80 (0.62 to 1.03)</td>
</tr>
<tr>
<td>50–99% adherence (vs all)</td>
<td>1.31 (0.94 to 1.81)</td>
<td>1.30 (1.03 to 1.63)</td>
</tr>
</tbody>
</table>

Continued
For PN, the other acute condition, only failure on the blood culture timing variable had a significant association with 90-day and 365-day survival. Time-weighted average non-adherence was associated with a marginally increased risk of 365-day mortality and the 0–49% adherence group had shorter survival time. Sensitivity analyses (including in-hospital deaths and calculating survival time from admission rather than discharge) did not reveal substantial changes in the associations of process adherence and survival for arrival measures.

For HF, non-adherence with the ACEI/ARB measure was associated with an increased hazard of long-term mortality, as was failure on the smoking counselling measure. The results for the time-weighted measures were generally insignificant and did not show a dose–response association.

Table 2 shows the association between measure non-adherence and 90-day and 1-year readmission.

For AMI, failure on the PCI <90 min measure was a strong predictor of readmission for both time-periods. Failure on the smoking counselling measure was associated with a greater risk of 90-day readmission. Failure on the discharge instruction measure was each associated with an increased risk of 90-day and 1-year readmission. Failure to receive all components of the Patient Information Measure group (online table 1) was associated with increased risk of 1-year readmission, as was 50–99% categorical measure adherence (but not 0–49% adherence) compared with 100% adherence. Finally, for PN, non-adherence with the antibiotic selection and smoking counselling measures was associated with increased risk of readmission, as was 0–49% (vs 100%) overall adherence. Associations between arrival measure adherence and readmission were not substantially changed in sensitivity analyses using time from admission (rather than discharge) to readmission.

For HF, failure on the ACEI/ARB measure and the discharge instruction measure were each associated with an increased risk of 90-day and 1-year readmission. Failure to receive all components of the Patient Information Measure group (online table 1) was associated with increased risk of 1-year readmission, as was 50–99% categorical measure adherence (but not 0–49% adherence) compared with 100% adherence. For PN, non-adherence with the antibiotic selection and smoking counselling measures was associated with increased risk of readmission, as was 0–49% (vs 100%) overall adherence. Associations between arrival measure adherence and readmission were not substantially changed in sensitivity analyses using time from admission (rather than discharge) to readmission.

Figure 1A,B present the adjusted survival and freedom from readmission curves for AMI. In both instances the curves for 100% and 50–99% adherence were virtually superimposable, whereas <50% adherence was associated with significantly lower survival and higher probability of readmission.

Figure 2A,B depict survival and freedom from readmission for HF, and the results are paradoxical. Adherence of 0–49% has the best survival, full adherence has intermediate results (not statistically different from 0–49%), and 50–99% adherence has the worst survival. Figure 2B shows that 100% adherence and 0–49% are nearly superimposable and associated with the highest freedom from readmission, whereas 50–99% adherence has significantly lower freedom from readmission.

The results for PN (figure 3A,B) are similar to those for AMI. Survival and readmission results for 100% and
50–99% adherence are similar and superior, whereas 0–49% adherence is associated with significantly inferior results for both.

**DISCUSSION**

Injudicious selection of performance measures may misallocate scarce performance improvement resources, divert attention from more efficacious processes of care, inappropriately penalise or reward providers and misinform consumers.\(^{25\text{--}40}\) If not associated with substantial positive effects, efforts to optimise scores on some publicly reported measures may even have a net negative impact because of adverse unintended consequences such as premature activation of cardiac catheterisation labs in AMI\(^{41\text{--}42}\) or administration of antibiotics to patients before a diagnosis of PN has been firmly established.\(^{38\text{--}39}\) Masoudi\(^ {40}\) suggested that the net positive and unintended negative consequences be assessed by ‘balancing measures’.

Numerous explanations have been hypothesised for the absent, weak or inconsistent association between process measures and outcomes. For example, even efficacious care processes may explain only a small proportion of variation in patient outcomes.\(^7\) Some recommended processes may have been selected based on limited observational studies, or conversely they may have been chosen based on randomised trials with rigid...
eligibility and exclusion criteria. Results from the latter may not generalise well to broader, real-world patient populations.52 52 84 34 4 The value of some process measures may be offset by their unintended adverse consequences.37 38 40 For this reason, some processes such as beta blockade for HF that favourably impact outcomes with chronic use are not included in current inpatient measure sets, as their use may be associated with higher risk in hospitalised patients.30 45 Some processes (eg, cardiac resynchronisation therapy or implantable defibrillators) may have substantial utility in selected subgroups but their inclusion/exclusion criteria and potential adverse consequences may preclude broad use as performance measures.46—49 Failure to individualise the application of guidelines could produce inferior results by encouraging inappropriate care of specific patients.39 50—52

Inconclusive results from existing process–outcomes studies may also result from flaws in study design, documentation or follow-up. These include the inability to adjust for important clinical, socioeconomic or hospital confounders8 25 26 28 53—63, ceiling effects for some ‘topped-out’ measures5 7; inadequate duration of follow-up to detect significant differences in outcomes attributable to process adherence (eg, smoking counselling, anticoagulation for atrial fibrillation)7 26 32; dissociation between care prescribed at discharge and subsequent patient adherence24 28 64 65; ‘check-the-box’ mentality, while failing to conform with the ‘spirit’ of the measure28; and inaccurate hospital documentation that does not accurately reflect care delivered.28

**Contributions of the current study**

Our study addresses important methodological concerns with previous investigations of the process–outcomes association. We use contemporary, all age and payer, patient-level data from a single large institution with a diverse patient population. This mitigates concerns regarding between-hospital heterogeneity. We have access to detailed, census block-group socioeconomic data,36 a potentially important confounder that is lacking in most studies and that may impact post-discharge patient behaviour. Patients of low SES may be less likely to consistently use prescribed medications or

**Figure 1** Adjusted 1-year survival (A) and freedom from readmissions (B) by measure adherence for patients with acute myocardial infarction (AMI).
attend outpatient clinics, thus adversely influencing survival and readmission rates, irrespective of the efficacy of inpatient care processes.

Our study also investigated process measure adherence at several levels of aggregation. Performance on one measure does not necessarily correlate with performance on other measures, and the degree to which an entire ‘bundle’ of measures is necessary to achieve the optimal effect is uncertain. Finally, our study employs a novel statistical approach, exponential decay, to address the problem of measure adherence or failure on repeated admissions for the same patient.

Our results suggest that even when in-hospital mortalities are excluded, process measure adherence (individual and group) for two common acute conditions, AMI and PN, is often positively associated with long-term outcomes. Furthermore, these outcomes improve with higher degrees of process adherence. Presumably, aspirin and β-blockade on arrival limit myocardial ischaemia and the subsequent extent of infarcted myocardium, thereby reducing early mortality. However, the extent of infarct also has an association with late outcomes. When appropriate medications are not given and the resulting infarct is larger, more extensive scar formation and post-infarct remodelling lead to a higher incidence of late deaths and readmissions from HF and arrhythmias. By a similar mechanism, failure to open the infarct artery in less than 90 min may be associated with increased in-hospital mortality and late outcomes.

The association between failure on the PN blood culture timing measure and long-term outcomes is interesting. Blood cultures drawn prior to antibiotic administration may allow more rapid determination of the causative organism in the most severely ill patients, and early institution of tailored, specific antibiotic therapy may favourably influence late outcomes. However, it is also possible that this process measure mainly reflects other unmeasured aspects of care, such as the use of standardised protocols.

For HF, only ACEI/ARB adherence and smoking counselling were associated with lower mortality and readmission. For the various composites, the results were generally inconsistent and insignificant, emphasising the importance of validating entire measure bundles and their individual constituent measures. Currently...
endorsed HF measures may lack a strong association with outcomes, or unmeasured outpatient factors may outweigh the importance of adherence with inpatient care processes in this chronic condition.

Limitations

We believe our single institutional study cohort is advantageous for the reasons noted previously, but it does raise the issue of generalisability, despite our diverse community and referral population.

We were unable to document compliance with and changes to prescribed care regimes following discharge. This potential unmeasured confounder may be most relevant for HF, a chronic illness requiring systematic outpatient follow-up.66 67

Because we limited the study of the HF discharge instruction measure to those patients discharged home, our results may not be applicable to patients discharged to extended care facilities.

The current study is observational, and there may be unmeasured confounders. These limitations should be considered in evaluating the strength of any causal inferences derived from our analyses.

CONCLUSIONS

In aggregate, these findings demonstrate the potential limitations of using process measures to assess provider performance. If appropriately selected, as in the case of AMI and PN, measure adherence may promote the fundamental goal of profiling—to improve short-term and long-term patient outcomes. Conversely, in our study only some of the current HF measures are favourably associated with long-term outcomes, and there is no overall dose-response association. This may reflect fundamental issues with the selected process measures, the impact of unmeasured processes of care or the confounding effect of subsequent outpatient care. Focusing on measures without a strong evidence base may divert scarce quality improvement resources, encourage marginally effective care practices and misclassify providers.
As performance measurement evolves, the emphasis should increasingly shift to direct outcomes metrics, including mortality and morbidity, patient-reported outcomes and satisfaction. When process measures are used, their association with outcomes should ideally be validated in randomised trials and real-world observational studies. Some recommended care processes may not be suitable as accountability measures.

Funding Internal.

Competing interests Dr Ramunno is Chief Quality Officer for Northeast Health Care Foundation, a Medicare QIO. He has been involved with the development, revision, and implementation of National Hospital Quality Measures as a contractor with the federal government.

Patient consent Retrospective review of previously collected administrative claims data and state all-payer administrative records. Not feasible to obtain permission, and risk considered minimal.

Ethics approval Partners IRB (2008P000003).

Contributors All authors satisfy the following requirements: substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; and final approval of the version to be published.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Online supplemental material includes detailed information on numbers of eligible patients for each measure; descriptive characteristics and bivariate associations; and examples of the exponential decay approach used in this study.

REFERENCES
Original research


Association of National Hospital Quality Measure adherence with long-term mortality and readmissions

David M Shahian, Gregg S Meyer, Elizabeth Mort, Susan Atamian, Xiu Liu, Andrew S Karson, Lawrence D Ramunno and Hui Zheng

BMJ Qual Saf 2012 21: 325-336 originally published online March 2, 2012
doi: 10.1136/bmjqs-2011-000615

Updated information and services can be found at:
http://qualitysafety.bmj.com/content/21/4/325

These include:

Supplementary Material
Supplementary material can be found at:
http://qualitysafety.bmj.com/content/suppl/2012/03/01/bmjqs-2011-000615.DC1

References
This article cites 65 articles, 12 of which you can access for free at:
http://qualitysafety.bmj.com/content/21/4/325#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Editor's choice (113)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/