© Health Research and Educational Trust
DOI: 10.1111/j.1475-6773.2006.00604.x

Does Ambulatory Process of Care Predict Health-Related Quality of Life Outcomes for Patients with Chronic Disease?

Katherine L. Kahn, Diana M. Tisnado, John L. Adams, Honghu Liu, Wen-Pin Chen, Fang Ashlee Hu, Carol M. Mangione, Ronald D. Hays, and Cheryl L. Damberg

Objective. The validity of quality of care measurement has important implications for practicing clinicians, their patients, and all involved with health care delivery. We used empirical data from managed care patients enrolled in west coast physician organizations to test the hypothesis that observed changes in health-related quality of life across a 2.5-year window reflecting process of care.

Data Sources/Study Setting. Patient self-report data as well as clinically detailed medical record review regarding 963 patients with chronic disease associated with managed care from three west coast states.

Study Design. Prospective cohort study of change in health-related quality of life scores across 30 months as measured by change in SF-12 physical component scores.

Data Collection/Extraction Methods. Patient self-report and medical record abstraction.

Principal Findings. We found a positive relationship between better process scores and higher burden of illness ($p<.05$). After adjustment for burden of illness, using an instrumental variables approach revealed better process is associated with smaller declines in SF-12 scores across a 30-month observation window ($p=.014$). The application of the best quartile of process of care to patients currently receiving poor process is associated with a 4.24 increment in δ SF-12-physical component summary scores.

Conclusions. The use of instrumental variables allowed us to demonstrate a significant relationship between better ambulatory process of care and better health-related quality of life. This finding underscores the importance of efforts to improve the process of care.

Key Words. Quality of care, chronic disease, process–outcome link

The validity of quality of care measurement has important implications for practicing clinicians, their patients, and all involved with health care delivery.

The classic strategy for assessing the internal validity of quality of care measures is to examine whether patient outcomes are mediated by process (Donabedian 1982; Field et al. 2001). This can be accomplished either with randomized trials (Antiplatelet Trialists' Collaboration 1994; ISIS-2 [Second International Study of Infarct Survival Collaborative Group] 1988) or, in the case of observational data, with cohort studies. In either case, the fact that sicker patients receive more process, presents challenges in revealing a relationship between process and outcomes. A decade ago, a significant relationship between processes of care and outcomes for hospitalized patients was noted, but even then a paradoxical relationship between process and outcomes for the sickest patients was noted (Kahn, Keeler et al. 1990; Kahn, Rogers et al. 1990). We hypothesize that this paradoxical relationship is a consequence of the association between unmeasured burden of illness and both more process and also worse outcomes. Angrist, Newhouse, McClellan, and Brooks have documented a methodology and a clinical context for using instrumental variables to disentangle the endogeneity of unmeasured burden of illness and both processes and outcomes (McClellan, McNeil, and Newhouse 1994; Angrist, Imbens, and Rubin 1996; Brook, McGlynn, and Cleary 1996; Brooks, McClellan, and Wong 2000; McClellan and Newhouse 2000; Brooks et al. 2003).

We applied this methodology to the study of the quality of processes of care to evaluate whether better process was associated with better outcomes. We used empirical data from managed care patients enrolled in west coast physician organizations to test the hypothesis that after adjustment for burden of illness, the changes in health-related quality of life across a 2-year window reflecting process of care. To date, few studies have shown a relationship between explicit process measures and outcomes in the ambulatory setting (Nobrega et al. 1977; Romm and Hulka 1980; Berlowitz et al. 1998; Safran

---

Address correspondence to Katherine L. Kahn, M.D., Department of Medicine, University of California at Los Angeles, Division of General Internal Medicine and Health Services Research, 911 Broxton Plaza, Box 951736, Los Angeles, CA 90095-1736. Katherine L. Kahn, M.D., John L. Adams, Ph.D., Ronald D. Hays, Ph.D., and Cheryl L. Damberg, Ph.D., are with the RAND Corporation, Santa Monica, CA. Additionally, Dr. Kahn and Diana M. Tisnado, Ph.D., Honghu Liu, Ph.D., Fang Ashlee Hu, M.D., Carol M. Mangione, M.D., and Ronald D. Hays, Ph.D., are with the Department of Medicine, University of California at Los Angeles, Division of General Internal Medicine and Health Services Research, Los Angeles, CA. Wen-Pin Chen, M.S., is with the University of California at Irvine, Cancer Center, Irvine, CA. Cheryl L. Damberg, Ph.D., is also with the Pacific Business Group on Health, San Francisco, CA.

et al. 1998; Asch et al. 2005; Haynes et al. 1982; Higashi et al. 2005), and none have done so for insured managed care patients, or for a health status measure.

METHODS

Study Cohort and Data Sources

We used patient self-report data from a cohort of 963 chronically ill patients in 1996 (Damberg and Bloomfield 1997) and 2.5 years later in 1998 (Pacific Business Group on Health 2004), as well as clinically detailed data abstracted from medical records of patients associated with 39 west coast physician organizations to assess quality of care for 30 consecutive months (Kahn et al. 1999, 2003). Registered nurses experienced with both clinical practice and medical record abstraction used an abstraction instrument designed specifically for this project to abstract records from 963 patients representing data from 5,095 unique patient–physician dyads.

We analyzed adherence (yes or no) to 120 explicit process criteria based upon data from the 1996 Core Survey (Damberg and Bloomfield 1997), the 1998 Chronic Condition Survey (Pacific Business Group on Health 2004), and medical record review of all visits that occurred between the two surveys. Explicit process criteria were based on clinical practice guidelines, literature, and clinical judgment. To assess interrater reliability we compared the performance of 11 pairs of abstractors who independently assessed adherence to explicit process measures from the medical records of 54 unique patients. Concordance between abstractors was excellent with no significant difference noted across abstractors in overall process scores. The aggregate kappa score across process measures was 0.87 (Landis and Koch 1977).

The cohort for this analysis of the relationship between process scores and health-related quality of life outcomes is defined as patients with at least one of three diseases where literature and trials suggest such a relationship between process and outcomes might be expected. Specifically, we studied 963 patients with at least one of: ischemic heart disease, asthma and/or emphysema, or diabetes diagnoses described as present with 1996 self-report and corroborated with evidence from the 1996 and 1998 surveys and medical record data.

Assessing the Process of Medical Care

Study patients were evaluated with disease-specific explicit process criteria pertinent to hypertension, diet/nutrition, obesity, exercise, smoking, hyper-

lipidemia, thyroid disease, menopause, depression, medication management, substance abuse, and follow-up or continuity for patients. The 120 process measures (58 generic and 62 disease specific) were specified as applicable to individual patients according to their age, gender, and clinical characteristics.

The explicit process measures were selected to collectively represent one of six domains of clinical care. *Domain 1, Cognitive Diagnostic Process,* uses 31 explicit process criteria to evaluate the extent to which the provider systematically collected patient historical data (e.g., presence or absence of symptoms, precipitating or relieving factors) (Hunt and Gerstein 1999) necessary for the clinician to make an adequate assessment of the patient's clinical needs. *Domain 2, Physical Examination,* uses 20 explicit criteria to evaluate the provider's use of pertinent components of the physical exam (e.g., lung exam for patients with asthma, foot exam for diabetic patients) (American Diabetes Association 1998). *Domain 3, Laboratory Studies,* uses 24 criteria to evaluate the use of laboratory studies for diagnostic or surveillance purposes (e.g., monitoring of creatinine in patients using angiotensin converting enzymes) (Knight and Avorn 2001). *Domain 4, Procedures,* uses 7 criteria to evaluate the provider's use of diagnostic procedures (ACC/AHA 1997). *Domain 5, Medications,* uses 26 criteria to evaluate provider recommendations for, and patient use of medications (e.g., use of β-blocker for patients with myocardial infarction and no contraindication) (Ryan et al. 1999). *Domain 6, Counseling,* uses 12 criteria to score the provider's counseling interventions (as an alternative or supplement to pharmacological or procedural interventions) (U.S. Preventive Services Task Force 1996).

For each patient, the proportion of applicable domain-specific process criteria passed was used to calculate domain-specific process scores. Process criteria were weighted equally within domains and domains were weighted equally regardless of the number of criteria comprising the domain. Aggregate observed process is specified as the mean of six domain-specific process scores, with each of those scores representing a proportion defined as the number of THEN criteria met, conditional on the IF being applicable to the patient. This variable defines the aggregate of adherence to the 120 explicit process measures in the ordinary least squares (OLS) regression model.

Burden of Illness

For each patient, we calculated three dimensions of burden of illness: a count of up to 39 patient comorbidities noted by either patient self-report or medical record review; the severity of cardiac, pulmonary, or diabetic disease; and

body mass index (BMI) (Iezzoni 1994;Field et al. 2001) (see Table 1). Items eligible for scoring a point in the comorbidity index include: cardiovascular problems; cerebrovascular disease; cancer; diabetes; chronic lung disease; common ambulatory problems; depression; measures of functional impairment; habits associated with medical problems; and patient report of worsening health status as documented in Table 1. Severity of the patient's coronary heart disease, lung disease (asthma or emphysema), or diabetes were calculated in a disease-specific manner defined to be independent of use of services (Table 1). To test the validity of the comorbidity and staging systems, we checked the relationships between the comorbidity and staging scores and the construct of burden of illness as measured by the number of drug categories the patient used.

Patient Demographics

Patient demographics were categorized as age, gender, race, Hispanic ethnicity, education, and income.

SF-12 Physical Component Summary Scores (SF-12-PCS)

Health-related quality of life scores were computed for each patient in 1996 and again, 2 years later. We calculated change in SF-12 scores as the simple arithmetic difference between each patient's raw 1998 and earlier 1996 SF-12 SF-12-PCS (Ware, Kosinski, and Keller 1996). A positive change value is interpreted to mean the patient's health-related quality of life improved with time; a negative value indicates a decrement across the 2-year period.

Predicting Changes in SF-12-PCS from 1996 to 1998

We had an a priori concern that better process might be associated with both greater measured burden of illness, and also with greater unmeasured burden of illness (and/or provider challenges to implementing process and facilitating better patient outcomes), and that the greater burden of illness was an important predictor of both more process and worse outcomes. This concern led to the use of instrumental variables methods to address the potential endogeneity of process (McClellan et al. 1994). We postulated instrumental variables would be useful because unmeasured burden of illness is, by definition unobservable (i.e., not measurable); unmeasured burden of illness influences both processes (a key independent variable) and outcomes (the dependent variable); and unmeasured burden of illness (the omitted variable) is an important predictor of both processes and outcomes when using OLS regression.

Table 1: Patient Characteristics

Variable Description	*n*	*Mean*	*SD*	*Median*	*Min–Max*
Comorbidity*	963	7.90	3.49	7	1–26
Disease-specific severity proportion†	963	0.49	0.29	0.50	0–1
Body mass index	963	29.08	6.47	28.14	16.46–65.55
Heart cohort severity	239	2.16	1.05	2	1–4
Lung cohort severity	318	1.79	0.88	1	1–3
Diabetes cohort severity	387	2.03	0.90	2	1–4
Change SF-12-PCS: 1998–1996	963	−1.20	9.83	−0.69	−40.19–+29.42
1996 SF-12-PCS	963	41.97	11.60	44.33	14.36–63.77
1998 SF-12-PCS	963	40.77	12.03	42.61	14.21–63.34
Proportion all process criteria applicable to study cohort	963	0.33	0.09	0.33	0.08–0.60
Proportion generic process criteria applicable to study cohort	963	0.44	0.11	0.45	0.12–0.79
Proportion heart process criteria applicable to heart cohort	268	0.66	0.08	0.68	0.53–0.89
Proportion lung process criteria applicable to lung cohort	318	0.54	0.25	0.57	0.19–0.95
Proportion diabetes process criteria applicable to diabetes cohort	387	0.75	0.10	0.77	0.50–0.91
Frequency of clinical visits in the first 12 months	915	8.06	6.63	6.00	1–57
Observed overall process‡	963	−0.02	0.62	0.04	−2.34–+1.40
Predicted overall process‡	963	−0.03	0.62	0.00	−1.84–+1.36

*Items eligible for scoring a point in the comorbidity index include: cardiovascular problems (heart disease, coronary bypass surgery or angioplasty, myocardial infarction within the last year, angina, left ventricular dysfunction, family history coronary disease at an early age; peripheral vascular disease, history of deep venous thrombosis, hypercholesterolemia, or hypertension); cerebrovascular disease (stroke or carotid disease); cancer; diabetes; chronic lung disease (bronchitis, asthma, emphysema, sleep apnea); common ambulatory problems (arthritis; kidney problems; migraine headaches, chronic or seasonal allergies, sinus trouble, chronic back problems, osteoporosis, ulcers, hemorrhoids, dermatitis; hepatobiliary disease, epilepsy, thyroid problems, prostate problems [males only], urinary incontinence [women only]); depression; measures of functional impairment (blindness or blurred vision, deafness, limitations in the use of an arm or leg); habits associated with medical problems (remote smoking, concurrent smoking, drug abuse problem, alcohol abuse problems); patient report of worsening health status.

†Severity of the patient's coronary heart disease, lung disease (asthma or emphysema), or diabetes was calculated according to the stage of a patient's single disease (69% of patients) or as the mean of the patient's diseases for the (16%) with more than one of these conditions. Heart stages were assigned as follows: stage 1 (prior myocardial infarction, coronary surgery, or unstable angina ever prior or during the 30 month study window: 39% of heart patients); stage 2 (stable angina or congestive heart failure: 14%); stage 3 (new or worsening angina or congestive heart failure: 38%); stage 4 (hospitalized during the 30-month study period with acute myocardial infarction, bypass surgery or angioplasty: 9%). Lung stages were assigned according to the proportion of clinical visits during the first 12-month abstraction window that were associated with acute shortness of breath as a current problem: stage 1 (fewer than 10% of clinical visits: 51% of lung patients); stage 2 (10% to $<25\%$ of visits: 19%); stage 3 ($\geq$ 25% of visits: 30%). Diabetes stages were assigned according to longer diabetes duration, lipid disorder, hypertension, known coronary or cerebrovascular disease, diabetic nephropathy, diabetic retinopathy risk factors for diabetic foot ulcers; obesity: stage 1 (32% of patients), stage 2 (40%), stage 3 (21%), and stage 4 (7%).

‡In constructing the overall process scores, we standardized domain-level process scores to mean 0 (SD1). The mean overall process score was generated as a mean of each patients' domain-level process score with equal weighting across domains. Because not all patients had process measures for each domain, the sample size associated with domain-level scores varied, resulting in a nonzero mean overall process score.

SF-12-PCS, SF-12 physical component summary scores.

We used the augmented test (Davidson and MacKinnon 1993) as a way to evaluate whether instrumental variables would be useful in addressing a potential bias.

Instrumental Variables

To gain unbiased estimates of the influence of process on change in SF-12-PCS, we used the *structure of care* associated with study patients as an instrument for process. Structure of care meets the two essential criteria for an instrument (McClellan et al. 1994; McClellan and Newhouse 2000): *structure* theoretically is a major determinant of process (Donabedian, 1980; Ann Arbor, Michigan) and *structure* influences outcomes *only* as mediated by process. Using structure as the instrument allows us to exploit treatment variation across *structure*-specific patient cohorts. This allows us to evaluate the effect of process on health-related quality of life outcomes for patients whose process might change at the margin if they were engaged in a different structural arrangement (Angrist et al. 1996).

We use indicator variables for each of the physician organizations as the instruments for structure when predicting outcomes using two separate models. With instrumental variables, observed process in Stage 1 is modeled using patient-level burden of illness (defined as comorbidity score, the severity of heart, lung and diabetes diseases, and BMI); demographics; the frequency of clinical visits; and an indicator for each of the physician organizations. In Stage 2, we model change in SF-12-PCS score as a function of predicted process from Stage 1, patient-level burden of illness, demographics, and the frequency of clinical visits. To account for the unique ways in which adherence to individual process measures vary by comorbidity, disease-specific stage, and BMI, we supplement the three patient-level burden of illness variables in Stage 1 with the aggregate predicted process criteria. This detailed modeling allows us to improve the prediction in the first stage model to improve the estimation of the process-outcome link.

The first stage equation for instrumental variable uses OLS to predict observed aggregate process as follows:

$$\begin{aligned}&\text{observed aggregate process}\\&= f\,(\text{aggregate predicted process, patient demographics, burden of illness,}\\&\quad\text{frequency of clinical visits, a dummy for each medical organization})\end{aligned}$$

where *aggregate predicted process criteria* is specified as the aggregate of predicted adherence for each explicit process criteria. It is introduced into the model for

the same reason that standard burden of illness measures such as comorbidity, severity, and BMI are included in the prediction of process. However, because we have both clinical and empiric evidence that providers consider the importance of burden of illness uniquely for each individual process measure, we have included in the model a measure of predicted process criteria. This is generated for each individual process measure by a regression:

$$\text{Adherence to process measure}_{1-120} = f\,(\text{comorbidity, severity, body mass index})$$

The aggregate predicted process criteria is the aggregate version of these 120 individual predicted variables. In summary, this variable is a supplemental measure of burden of illness included in Stage 1 of the instrumental variables model to represent the ways providers uniquely consider burden of illness as they approach the implementation of individual process measures. Note that at the end of this process, the aggregated predicted process criteria is only a function of comorbidity, severity, and BMI.

The second stage equation for instrumental variables uses OLS to predict change in SF-12-PCS using aggregate predicted process criteria to account for the unique ways in which adherence to individual process measures varies by comorbidity, disease-specific stage, and BMI as follows:

$$\text{Delta SF-12-PCS} = f\,(\text{Y-hat for predicted process from Stage 1, patient demographics, burden of illness, frequency of clinical visit})$$

where *predicted process* is defined as Y-hat from instrumental variables Stage 1. This is the model resulting from the regression of aggregate observed process on the full set of predictors in Stage 1 of instrumental variables.

Statistical Analysis and Weighting

We used *SAS* 8 and *Stata* 8 for analyses. We present regression results adjusted for clustering of patients within physician organizations using Huber–White correction (White 1980) displaying regression coefficients, 95 percent confidence intervals, and *p*-values from OLS and both stages of the instrumental variables models. Augmented tests for endogeneity are presented (Davidson and MacKinnon 1993; Baum, Schaffer, and Stillman 2003). Analyses are weighted as the product of sampling weight in 1996, survey nonresponse weight in 1996, disease sampling, survey nonresponse weight in 1998, and medical record abstraction nonresponse weight.

RESULTS

The final study cohort includes 963 patients from 39 physician organizations with baseline 1996 and 30-month follow-up survey data, as well as abstracted medical record data spanning 30 months after the baseline patient self-report survey. The study cohort has a mean age of 60 years (SD 9), with 41 percent at least 65 years and 52 percent female. Forty-four percent of patients had no more than a high school education; 30 percent reported annual income less than $30,000. This cohort of patients with at least one chronic illness had frequent clinical encounters with a mean of one clinical visit per month.

Burden of Illness

Patients had a mean of 7.90 (SD 3.49) comorbidities of 39 studied (Table 1). The severity of patients' heart, lung, and diabetes conditions ranged from 0 to 1.0 with a mean of 0.49 (SD 0.29). The mean BMI was 29 (SD 6) with 38 percent overweight, and 34 percent obese. As a test of the construct validity of the comorbidity and severity scores, we evaluated the relationship between them and the number of medications used by patients. We found a positive relationship between the patient's overall medication count and the number of comorbid conditions ($p < .0001$) and BMI ($p = .0003$); and the disease-specific medication count and the disease severity scores for heart disease ($p < .001$), lung disease ($p = .049$), and diabetes ($p = .097$). The mean 1996 SF-12-PCS was 42 (SD 12), 8 points lower than the national average for adults (Ware et al. 1996), reflecting our cohort having been defined as patients with at least one chronic condition. In 1998, the mean SF-12-PCS was 41 (SD 12) with a mean change of -1.20 (SD 9.83).

Burden of Illness and Process

Overall, we note the relationships between overall process and the dimensions of severity, comorbidity, and BMI are all positive with statistical significance ($p < .0001$) for comorbidity and severity. The pattern is reproduced with domain-level process scores and severity and comorbidity. To better understand the nature of the relationship between adherence to process indicators and burden of illness, we compared patients who passed and failed individual process criteria with respect to severity, comorbidity, and BMI. Table 2 presents examples of one explicit process criterion (columns B–D) from each domain of process with documentation regarding the number of patients for whom the criterion was applicable (column C), and the number of applicable patients for whom the criterion was met (column D). Table 2 (columns G–K)

Table 2: Examples of Process Criteria with Adherence Rates Stratified by Burden of Illness Measures and by Domain

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
Domain	*IF*	*% cohort to whom IF criteria is applicable*	*THEN*	*% applicable patients who pass the THEN*	*Burden of illness measure*	*Mean (SD) burden of illness score for patients who pass versus fail the process criterion*		*Mean (SD) difference in burden of illness score for patients who:*	*Sign of the difference and p-value associated with comparison of patients who pass versus fail process criterion**	
						Pass the process criterion	Fail the process criterion	Pass versus fail criterion	Bivariate comparison	Multivariable comparison
Cognitive	Patient with known diabetes	387 (100%)	Provider should query the patient about foot pain burning, numbness, tingling, sores, or ulcers at least once during a 15-month observation window (Hunt and Gerstein 1999)	250/387 (65%)	Severity Comorbidity Body mass index	0.6 (0.21) 7.9 (3.4) 31.3 (6.7)	0.5 (5.20) 7.1 (3.4) 29.7 (6.9)	0.1 (0.21) 0.9 (3.4) 1.6 (6.8)	Plus ($p<.001$) Plus ($p=.02$) Plus ($p=.03$)	Plus ($p<.001$) Minus ($p=.87$) Plus ($p=.19$)
Physical exam	Patient with known diabetes	387 (100%)	Provider should examine the patient's feet with shoes and socks off at least once during a 15-month observation window (American Diabetes Association 1998)	279/387 (72%)	Severity Comorbidity Body mass index	0.5 (0.21) 7.9 (3.5) 31.2 (6.9)	0.5 (0.19) 7.0 (3.3) 30.0 (6.5)	0.1 (0.21) 0.92 (3.4) 1.5 (6.8)	Plus ($p<.001$) Plus ($p=.02$) Plus ($p=.06$)	Plus ($p=.01$) Plus ($p=.68$) Plus ($p=.26$)
Lab	Patient is using a potassium depleting or potassium sparing diuretic	393/963 (41%)	Potassium value should be checked at least once during a 15-month observation window (Knight and Avorn 2001)	315/393 (80%)	Severity Comorbidity Body mass index	0.6 (0.26) 8.9 (3.6) 30.5 (7.0)	0.5 (0.29) 8.2 (3.3) 30.9 (8.7)	0.1 (0.26) 0.7 (3.6) −0.4 (7.4)	Plus ($p=.001$) Plus ($p=.14$) Minus ($p=.72$)	Plus ($p=.005$) Plus ($p=.47$) Minus ($p=.76$)

Procedures	Patient with coronary artery disease AND new or worsening: angina, congestive heart failure, dyspnea, myocardial infarction, or a new coronary revascularization procedure	112/268 (42%)	At least one of the following tests should be performed at least once during a 15-month observation window: stress ECG, resting ECG, coronary angiography, or cardiac catheterization (ACC/AHA 1997)	73/112 (65%)	Severity Comorbidity Body mass index	0.8 (0.12) 11.2 (4.0) 27.9 (4.9)	0.8 (0.16) 11.3 (3.3) 29.2 (6.0)	0.1 (0.13) – 0.1 (3.8) – 1.3 (5.3)	Plus ($p = .06$) Minus ($p = .95$) Minus ($p = .22$)	Plus ($p = .04$) Plus ($p = .73$) Minus ($p = 0.12$)
Medication	Patient with myocardial infarction and none of the following: hypotension, bradycardia, LV dysfunction, decreased LVEF, asthma or reactive airways disease	26/268 (10%)	Patient should use a β-blocker (Ryan et al. 1999)	12/26 (46%)	Severity Comorbidity Body mass index	0.6 (0.24) 11.9 (3.4) 29.8 (4.8)	0.5 (0.23) 10.9 (3.4) 27.6 (5.2)	0.1 (0.24) 1.1 (3.4) 2.2 (5.1)	Plus ($p = .54$) Plus ($p = .44$) Plus ($p = .29$)	Plus ($p = .61$) Plus ($p = .88$) Plus ($p = .39$)
Counseling	Patients with any of the following risk factors for coronary artery disease: hypertension; overweight or obese by body mass index; elevated LDL; known	916/963 (95%)	Patient should receive at least one of the following diet or nutrition interventions: counseling or recommendation for diet or nutrition program, OR recommendation	635/916 (67%)	Severity Comorbidity Body mass index	0.5 (0.28) 8.2 (3.5) 30.2 (6.9)	0.5 (0.30) 7.6 (3.3) 26.7 (4.6)	0.0 (0.29) 0.6 (3.5) 3.5 (6.3)	Plus ($p = .26$) Plus ($p = .02$) Plus ($p < .001$)	Plus (0.74) Plus ($p = .012$) Plus ($p < .001$)

Table 2. *Continued*

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
Domain	*IF*	*% cohort to whom IF criteria is applicable*	*THEN*	*% applicable patients who pass the THEN*	*Burden of illness measure*	*Mean (SD) burden of illness score for patients who pass versus fail the process criterion*		*Mean (SD) difference in burden of illness score for patients who:*	*Sign of the difference and p-value associated with comparison of patients who pass versus fail process criterion**	
	coronary artery disease; or known diabetes		for patient to visit a specialist or health care organization program directed toward diet or nutrition (U.S. Preventive Services Task Force 1996)							

shows the bivariate (column J) and multivariable (column K) relationships between burden of illness scores (severity, comorbidity, and BMI) and process measure adherence varies from measure to measure (shown as rows). This provides support for the clinical observation that providers adhere to process measures after consideration of the patients' burden of illness, with their consideration of burden being individualized for each process measure. This also provided support for our analytic strategy of adjusting adherence to each individual process measure for comorbidity, severity, and burden of illness before grouping adherence to the measures together into a single aggregate predicted process score. Note in the Stage 1 regression, the three individual components of burden of illness are included as separate independent variables. The small negative relationships associated with these variables in the prediction of aggregate observed process criteria (Stage 1), should be interpreted in the context of these variables already being in the model in the alternative form of aggregate predicted process criteria.

Process Scores

After standardization, the mean observed aggregate process score was (– 0.02), SD 0.62 [minimum (– 2.34), maximum (1.40), 25th percentile (– 0.38), 75th percentile (0.45), interquartile range (0.83)]. The mean predicted overall process score was – 0.03, SD 0.62, [minimum (– 1.84), maximum (1.36), 25th percentile (– 0.45), 75th percentile (0.43), interquartile range (0.89)].

Testing the Instrumental Variables Model Assumption

A test of potential bias using the augmented test (Davidson and MacKinnon 1993) shows OLS is not appropriate for predicting changes in SF-12. The *F*-test for the instruments in the first-stage regression is 349 with 38 degrees of freedom confirming a very strong joint effect. The model rejected the null hypothesis of no endogeneity ($p = .001$) suggesting the need for instrumental variables as we have done (Davidson and MacKinnon 1993).

Predicting Change in SF-12-PCS

Table 3 column one presents the OLS-model with no significant relationship between process and delta SF-12-PCS; the negative sign on process predicting delta SF-12-PCS is noted. In contrast, column 3 shows process is associated with a significant improvement in delta SF-12-PCS ($p = .014$) with instrumental variables, Stage 2. We repeated the analysis with the subset of the 120 explicit measures that are included in HEDIS scores (23 measures) and then

Table 3: Predicting Change in SF-12-PCS ($n = 963$)

		Using Instrumental Variable to Predict Changes in SF-12-PCS	
	Using Ordinary Least Squares to Predict Changes in SF-12-PCS	*Instrumental Variable-Stage 1 (Predicts Process)*	*Instrumental Variable-Stage 2 (Predicts Changes in SF-12-PCS)*
n = 963	*Coefficient (95% CI), (p-Value)*	*Coefficient (95% CI), (p-Value)*	*Coefficient (95% CI), (p-Value)*
Aggregate observed process criteria	− 1.41 (− 3.54,0.72), ($p = .188$)		
Aggregate predicted process criteria		0.54 (0.35,0.72), ($p = .000$)	
Predicted process (Y-hat from instrumental variable stage 1)			**7.48 (1.61,13.35), ($p = .014$)**
Comorbidity	**0.35 (0.02,0.67), ($p = .036$)**	**− 0.03 (− 0.05,0.01), ($p = .008$)**	**0.38 (0.03,0.73), ($p = .033$)**
Body mass index	− 0.13 (− 0.29,0.02), ($p = .093$)	**− 0.01 (− 0.02,0.00), ($p = 0.041$)**	− 0.09 (− 0.26,0.08), ($p = .275$)
Disease specific stage	− 3.06 (− 9.23,3.12), ($p = .322$)	− 0.21 (− 0.52,0.11), ($p = .201$)	− 7.46 (− 15.31,0.39), ($p = .062$)
Age in 1996	− 0.06 (− 0.16,0.04), ($p = .215$)	**0.01 (0.01,0.02), ($p = 0.000$)**	**− 0.19 (− 0.33, − 0.06), ($p = .006$)**
Female	− 1.07 (− 3.35,1.22), ($p = .350$)	− 0.04 (− 0.17,0.09), ($p = .523$)	− 0.06 (− 3.31,3.19), ($p = .972$)
African American	4.32 (− 0.63,9.28), ($p = .085$)	− 0.03 (− 0.38,0.31), ($p = .848$)	5.66 (− 1.16,12.48), ($p = .101$)
Asian	1.51 (− 3.78,6.80), ($p = .567$)	0.08 (− 0.14,0.30), ($p = .496$)	1.23 (− 4.28,6.74), ($p = .655$)
Hispanic	2.45 (− 1.50,6.40), ($p = .217$)	0.04 (− 0.11,0.19), ($p = .582$)	2.36 (− 2.14,6.86), ($p = .295$)
Other race	0.54 (− 4.45,5.53), ($p = .828$)	− 0.41 (− 0.96,0.14), ($p = .142$)	4.22 (− 3.91,12.35), ($p = .300$)
Missing race	3.82 (− 1.06.8.70), ($p = .121$)	0.06 (− 0.20,0.32), ($p = .656$)	3.63 (− 0.65,7.91), ($p = .094$)
Less than high school	− 0.30 (− 3.89,3.28), ($p = .865$)	0.04 (− 0.10,0.19), ($p = .539$)	− 0.88 (− 3.98,2.21), ($p = .566$)
College graduation plus	1.23 (− 2.78,5.24), ($p = .538$)	0.02 (− 0.16,0.20), ($p = .833$)	0.21 (− 4.28,4.69), ($p = .925$)
Missing education	3.31 (− 2.13,8.76), ($p = .226$)	**0.22 (0.00,0.43), ($p = .046$)**	0.95 (− 3.79,5.69), ($p = .687$)
Income 30 (<$30,000)	− 1.29 (− 4.35,1.77), ($p = .398$)	− 0.07 (− 0.21,0.08), ($p = .362$)	0.13 (− 2.75,3.01), ($p = .928$)
Income 61 (>$60,000)	− 0.18 (− 3.84,3.48), ($p = .922$)	− 0.04 (− 0.22,0.15), ($p = .684$)	− 0.54 (− 5.31,4.23), ($p = .821$)
Missing income	− 0.38 (− 5.29,4.54), ($p = .878$)	− 0.11 (− 0.38,0.15), ($p = .406$)	− 0.32 (− 6.91,6.28), ($p = .923$)
Number of clinical visits	0.08 (− 0.07,0.23), ($p = .299$)	**0.02 (0.01,0.03), ($p = .000$)**	− 0.24 (− 0.53,0.05), ($p = .098$)

Boldface indicates p-value < .05.

*p-value from F-test for omitting 38 medical organization dummies is .0000.

SF-12-PCS, SF-12 physical component summary scores.

again with the subset of measures (27 measures) meeting criteria for measures with Grade A evidence based upon randomized trials and found comparable results with virtually no difference in either the coefficient or the *p*-value for the instrumental variables model.

Estimating Effect Size

Using the instrumental variables model to estimate effect size, we note an improvement of 4.24 points in SF-12-PCS from 1996 to 1998 as process changes from the first (worst) quartile to the third quartile of process scores, of 2.21 as process changes from the worst quartile to the median value, and of 2.03 SF-12 points as process changes from the median to the fourth (best) quartile of process.

We observed a mean decrement of 1.20 (SD 9.83) in SF-12-PCS scores across the study's 30-month abstraction window (Table 1). One quarter of the patients dropped their score by 6 points, and 10 percent of patients dropped by 14 points. Almost 25 percent of the cohort dropped their SF-12-PCS by more than 5 (1/2 of the SD). The estimated effect size of 4.24 SF-12 points noted in association with a change in process from the first to the third quartile of process is associated with a substantial improvement in relation to the decrement observed for the cohort after aging 2.5 years. These data suggest the application of better process of care to patients currently receiving poor process would alter their SF-12-PCS change scores in a manner comparable to the eradication of aging three years.

DISCUSSION

Two important findings emerge from this work. Patients with more burden of illness had better process scores; and patients with better process of care sustain better health-related quality of life outcomes. Challenges in finding a link between better process and better patient outcomes have long been recognized (Sperl-Hillen et al. 2000; Kerr et al. 2001; Leatherman et al. 2003). Clinicians put patients in a higher venue of care (e.g., intensive care) to increase the clinician–patient ratio and facilitate the patient receiving a greater proportion of their many needed services. Yet, every clinician also knows that patients in the intensive care unit are more likely to die than patients cared for elsewhere. This *counter–clinical* relationship between the delivery of more needed process and higher death rates highlights the problem others and we faced in revealing the expected link between process and outcomes.

We postulated a three-stage clinically plausible sequence to account for the counter–clinical findings of worse health-related quality of life for patients with better-measured process. First, incentives exist to deliver needed process at higher rates for sicker as compared with less sick chronically ill patients. Second, patients with more measured burden of illness also had more unmeasured burden of illness (despite the rich set of burden of illness measures derived both from the patient and the medical record). Third, we postulate a correlation between more unmeasured burden of illness and process, as well as a correlation between more unmeasured (and measured) burden of illness and worse outcomes. We postulate this latter relationship is responsible for the observed *counter–clinical* OLS regressions which showed better process was associated with worse outcomes. We think the unmeasured (and measured) severity predict worse outcomes. However, as more burden of illness is associated with more process, we observe that more measured process (actually reflecting more burden of illness) is associated with worse outcomes. With this ongoing challenge, it is no wonder that researchers have struggled to find a relationship between process and outcomes. In contrast, clinicians note the dynamic nature of patient's burden of illness and how good patient care is defined by ongoing responsiveness to each patient's ever changing clinical need.

The reversal of the *counter–clinical* finding (that more process was associated with worse outcomes using OLS) with the instrumental variables approach, as well as the significance of the augmented test, supports our decision to use instrumental variables. Instrumental variables allow us to put more realistic standard errors around the relationship between burden and process.

We found evidence that providers consider different dimensions of burden of illness as they decide whether or not to intervene with specific recommended processes. Use of the instrumental variable model allowed us to account for the ways in which clinicians consider unique components of burden of illness as they approach each clinical process decision. The instrumental variables model closely represents clinical practice by adjusting providers' process scores uniquely for each criterion as a function of three dimensions of burden. This method revealed a statistically significant relationship between better process and health-related quality of life outcomes that patients' value so highly.

Provision of Needed Process Relates to Patient Burden of Illness

These data suggest providers, systems, and patients are rising to the challenges associated with the care of sicker patients. We observe the proportion of

needed care that is delivered is greater for sicker patients than for patients with less comorbidity, less severity, or less obesity. We do not know whether providers, patients, or organizational structure is the main determinant of why patients with more (versus less) burden of illness receive a greater proportion of needed services than patients with less burden of illness. However, as process criteria are constructed to measure use of an intervention believed to improve outcomes for a selected set of patients (as defined by explicit conditional logic), good quality of care should deliver comparable rates of adherence to explicit process measures for patients regardless of burden of illness.

Advances in process for very sick patients are to be valued and emulated. In contrast, low process score for chronically ill patients who have not yet demonstrated major decline represents a missed opportunity pertinent to underuse of primary, secondary, and tertiary preventive strategies. This finding of a strong relationship between burden of illness and processes should stimulate efforts to improve the quality of care for patients regardless of whether they are acutely or chronically ill; both instances provide important opportunities to improve patient outcomes. Better understanding of the determinants of higher adherence rates for sicker patients may provide a valuable clue to the reorganization of the current health care system.

Process Predicts Outcomes

This demonstration that ambulatory process of care is a significant predictor of changes in health-related quality of life across 2 years should reassure patients, providers, and those involved with health care delivery that the net result of better process is realized by patients in terms that matter to them. The estimated effect size from changing process from a moderate (50th percentile) process to the next best quartile of process (75th percentile) was found to be associated with an improvement in physical health, roughly equivalent to the decrement observed with aging from 1996 to 1998. The application of the best quartile of process of care to patients currently receiving poor process is associated with a 4.24-point increment in SF-12-PCS change scores comparable to a change in function from New York Heart Association Class II to III or III to IV (Bennett et al. 2002). Our results suggest the delivery of evidence-based processes of care will greatly benefit patients by improving health-related quality of life.

Because the conduct of processes is cost-sensitive, it is important to develop methods for understanding the evaluation of the link between

processes and outcomes. This analysis is robust across a number of model assumptions suggesting it can provide the basis for important future analyses that will evaluate costs associated with observed improvements in outcomes, as a function of process.

Demonstrating a clinically important and statistically significant link between better process of care and better health status highlights the importance of clinicians delivering good process. For people interested in process measurement, this analysis provides a process-outcome link, traditional evidence that process matters. For those interested in outcomes measurement, identification of a substantial process-outcome link reinforces the need for taking specific actions to improve outcomes.

We need to explore how these lessons apply to patients across the burden of illness spectrum. Regardless of whether the incentive systems focus on outcomes or process measures, the actions of potential recipients of incentives will be to try to improve process of care. A better understanding of how patient characteristics such as burden of illness vary according to challenges providers face in delivering good process is likely to affect how providers and/or patients will be able to respond to incentives. This may advance our understanding of how organizational structure, process, and outcomes fit together even further.

This should reassure those with concerns about the value of measuring process of care for patients. Having established a link between process and health related quality of life outcomes, providers can use this analysis as both a motivation and a challenge to provide better process.

ACKNOWLEDGMENTS

The analysis of the relationships between burden of illness, process, and process and outcomes was supported by grants from the Agency for Healthcare Research and Quality, American Association of Health Plans, and Robert Wood Johnson Foundation. The Pacific Business Group on Health supported baseline data collection activities. Dr. Diana Tisnado was supported by a Ruth L. Kirschstein National Research Service Award (Training grant number T32-HS00046).

We are indebted to Dr. Arnold Milstein for thoughtful reviews of early drafts of this manuscript, to Sarah Gee for research assistance, and to Corinna Koehnenkamp for manuscript preparation.

Disclosures: There are no disclosures or conflicts of interest.

Disclaimers: None.

REFERENCES

ACC/AHA. 1997. "Guidelines for Clinical Application of Echocardiography: Executive Summary." *Journal of the American College of Cardiology* 29 (4): 862–79.

American Diabetes Association. 1998. "Position Statement: Foot Care in Patients with Diabetes Mellitus." *Diabetes Care* 21 (1): S54–5.

Angrist, J., G. Imbens, and D. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.

Antiplatelet Trialists' Collaboration. 1994. "Collaborative Overview of Randomised Trials of Antiplatelet Therapy—II: Maintenance of Vascular Graft or Arterial Patency by Antiplatelet Therapy. Antiplatelet Trialists' Collaboration." *British Medical Journal* 308: 159–68.

Asch, S. M., E. A. McGlynn, L. Hiatt, J. Adams, J. Hicks, A. DeCristofaro, R. Chen, P. LaPuerta, and EA. Kerr. 2005. "Quality of Care for Hypertension in the United States." *BMC Cardiovascular Disorders* 5 (1): 1–9.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2003. Report No. Working Paper No. 545, Boston College Department of Economics, Boston, MA.

Bennett, S. J., N. B. Oldridge, G. J. Eckert, J. L. Embree, S. Browning, N. Hou, M. Deer, and M. D. Murray. 2002. "Discriminant Properties of Commonly Used Quality of Life Measures in Heart Failure." *Quality of Life Research* 11: 349–59.

Berlowitz, D. R., A. S. Ash, E. C. Hickey, R. H. Friedman, M. Glickman, B. Kader, and M. A. Moskowitz. 1998. "Inadequate Management of Blood Pressure in a Hypertensive Population." *New England Journal of Medicine* 339 (27): 1957–63.

Brook, R. H., E. A. McGlynn, and P. D. Cleary. 1996. "Quality of Health Care. Part 2: Measuring Quality of Care." *New England Journal of Medicine* 335 (13): 966–70.

Brooks, J. M., E. A. Chrischilles, S. D. Scott, and S. S. Chen-Hardee. 2003. "Was Breast Conserving Surgery Underutilized for Early Stage Breast Cancer? Instrumental Variables Evidence for Stage II Patients from Iowa." *Health Services Research* 38 (6, part 1): 1385–402.

Brooks, J. M., M. McClellan, and H. S. Wong. 2000. "The Marginal Benefits of Invasive Treatments for Acute Myocardial Infarction: Does Insurance Coverage Vary?" *Inquiry* 37 (1): 75–90.

Damberg, C., and L. Bloomfield. 1997. *1996 Physician Value Check Survey Final Report.* Pacific Business Group on Health.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics.* New York: Oxford University Press.

Donabedian, A. 1980. "The Definition of Quality: A Conceptual Exploration." *The Definition of Quality and Approaches to Its Assessment*: 1–32.

———. 1982. *Explorations in Quality Assessment and Monitoring,* Vol. 1. The Definition of Quality and Approaches to Its Assessment. Ann Arbor, Mich.: Health Administration Press.

Field, A. E., E. H. Coakley, A. Must, J. L. Spadano, N. Laird, W. H. Dietz, E. Rimm, and G. A. Colditz. 2001. "Impact of Overweight on the Risk of Developing Common Chronic Diseases during a 10-year Period." *Archives of Internal Medicine* 161 (13): 1581–6.

Haynes, R. B., E. S. Gibson, D. W. Taylor, C. D. Bernholz, and D. L. Sackett. 1982. "Process versus Outcome in Hypertension: A Positive Result." *Circulation* 65 (1): 28–33.

Higashi, T., P. G. Shekelle, J. L. Adams, C. J. Kamberg, C. P. Roth, D. H. Solomon, D. B. Reuben, L. Chiang, C. H. MacLean, J. T. Chang, R. T. Young, D. M. Saliba, and N. S. Wenger. 2005. "Quality of Care Is Associated with Survival in Vulnerable Older Patients." *Annals of Internal Medicine* 143: 274–81.

Hunt, D., and H. Gerstein. 1999. *Foot Ulcers in Diabetes. Evidence. BMJ, A Compendium of the Best Available Evidence for Effective Health Care*, chapter 1, pp. 25–31. London: BMJ Publishing Group.

Iezzoni, L. I. 1994. *Risk Adjustment for Measuring Health Care Outcomes.* Ann Arbor, MI: Health Administration Press.

ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. 1988. "Randomised Trial of Intravenous Streptokinase, Oral Aspirin, Both, or Neither among 17,187 Cases of Suspected Acute Myocardial Infarction: ISIS-2." *Lancet* 2 (8607): 349–60.

Kahn, K. L., M. Dans, D. Tisnado, et al. 1999. *Medical Record Abstraction System for the Physician Value Check Validation Study.* Abstraction form. Los Angeles, CA: The Regents of the University of California.

Kahn, K. L., E. B. Keeler, M. J. Sherwood, W. H. Rogers, D. Draper, S. S. Bentow, E. J. Reinisch, L. V. Rubenstein, J. Kosecoff, and R. H. Brook. 1990. "Comparing Outcomes of Care before and after Implementation of the DRG-Based Prospective Payment System." *Journal of American Medical Association* 264 (15): 1984–8.

Kahn, K. L., H. Liu, J. L. Adams, W. P. Chen, D. Tisnado, D. M. Carlisle, R. D. Hays, C. M. Mangione, and C. L. Damberg. 2003. "Methodological Challenges Associated with Patient Responses to Follow-up Longitudinal Surveys Regarding Quality of Care." *Health Services Research* 38 (6): 1579–98.

Kahn, K. L., W. H. Rogers, L. V. Rubenstein, M. J. Sherwood, E. J. Reinisch, E. B. Keeler, D. Draper, J. Kosecoff, and R. H. Brook. 1990. "Measuring Quality of Care with Explicit Process Criteria before and after Implementation of the DRG-Based Prospective Payment System." *Journal of American Medical Association* 264 (15): 1969–73.

Kerr, E. A., D. M. Smith, M. M. Hogan, T. P. Hofer, and R. A. Hayward. 2001. "Avoiding Pitfalls in Chronic Disease Quality Measurement: A Case for the Next Generation of Technical Quality Measurement." *American Journal of Managed Care* 7 (11): 1033–43.

Knight, E. L., and J. Avorn. 2001. "Quality Indicators for Appropriate Medication Use in Vulnerable Elders." *Annals of Internal Medicine* 135 (8 Part 2): 703–10.

Landis, J., and G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33 (1): 159–74.

Leatherman, S., D. Berwic, D. Iles, L. S. Lewins, F. Davidoff, T. Nolan, and M. Bisognano. 2003. "The Business Case for Quality: Case Studies and an Analysis." *Health Affairs (Millwood)* 22 (2): 17–30.

McClellan, M., B. J. McNeil, and J. P. Newhouse. 1994. "Does More Intensive Treatment of Acute Myocardial Infarction Reduce Mortality?: Analysis Using Instrumental Variables." *Journal of American Medical Association* 272: 859–66.

McClellan, M. B., and J. P. Newhouse. 2000. "Overview of the Special Supplement Issue." *Health Services Research* 35 (5, part 2): 1061–9.

Nobrega, F. T., G. W. Morrow, R. K. Smoldt, and K. P. Offord. 1977. "Quality Assessment in Hypertension Analysis of Process and Outcome Methods." *New England Journal of Medicine* 296 (3): 145–8.

Pacific Business Group on Health. "Pacific Business Group on Health Online" [accessed on June 22, 2004]. Available at URL http://www.pbgh.org/

Romm, F. J., and B. S. Hulka. 1980. "Peer Review in Diabetes and Hypertension: The Relationship between Care Process and Patient Outcome." *Southern Medical Journal* 73 (5): 564–8.

Ryan, T. J., E. M. Antman, N. H. Brooks, R. M. Califf, L. D. Hillis, L. F. Hiratzka, E. Rapaport, B. Riegel, R. O. Russel, EE III. Smith, W. D. Weaver, R. J. Gibbons, J. S. Alpert, K. A. Eagle, T. J. Gardner, A. Garson Jr., G. Gregoratos, T. J. Ryan, and S.C. Smith Jr. 1999. "1999 Update: ACC/AHA Guidelines for the Management of Patients with Acute Myocardiall Infarction. A Report of the American College of Cardiology/Amerian Heart Association Task Force on Practice Guidelines (Committee on Mangement of Acute Myocardial Infarction)." *Journal of the American College of Cardiology* 34 (3): 890–911.

Safran, D. G., D. A. Taira, W. H. Rogers, M. Kosinski, J. E. Ware, and A. R. Tarlov. 1998. "Linking Primary Care Performance to Outcome of Care." *Journal of Family Practice* 47 (3): 213–20.

Sperl-Hillen, J., P. J. O'Connor, R. R. Carlson, T. B. Lawson, C. Haltenson, T. Crowson, and J. Wuorenma. 2000. "Improving Diabetes Care in a Large Health Care System: An Enhanced Primary Care Approach." *Joint Commission Journal of Quality Improvement* 26 (11): 615–22.

US Preventive Services Task Force. 1996. *Guide To Clinical Preventive Services*, 2d Edition. Baltimore: Williams and Wilkins.

Ware, J. E., M. Kosinski, and S. D. Keller. 1996. "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity." *Medical Care* 34: 220–33.

White, H. 1980. "A Heteroskedasticity-Consistent Covariace Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48: 817–30.