Case-Mix Measures for Monitoring System Performance:

Use of Risk-adjusted Change in Health Status to Assess the Performance of Medicare Managed Care Plans

Lewis E. Kazis, ScD, Alfredo Selim, MD, MPH,
Sheikh Usman Iqbal, MD, MPH, William Rogers Ph.D.,
Shirley Qian MA, Xinhua S. Ren, Ph.D.

Center for the Assessment of Pharmaceutical Practices (CAPP),
Department of Health Policy and Management,
Boston University School of Public Health
and
The Center for Health Quality Outcomes and Economic Research,
A VA Health Services Research and Development Field Program,
VAMC, Bedford MA.

Case-Mix Measures for Monitoring System Performance

Background

- As the Centers of Medicare and Medicaid Services (CMS) move towards objective appraisal of outcome and other clinical performance measures of Medicare managed care organizations (MCOs), risk adjustment has become increasingly important in making informed clinical, administrative and economic decisions

- Risk adjustment is a method of accounting for differences in patient characteristics that may affect health care treatment outcomes. These adjustments level the playing fields by controlling for patient characteristics such as demographics and other measures of co-morbidities that characterize the patient and their illnesses when they enter a health care encounter.

Case-Mix Measures for Monitoring System Performance

Background

- Only after risk adjustment can we accurately measure and assess the effects of clinical processes and their associated interventions on patients.

- A variety of measures are currently applied to adjust for risk across ambulatory populations. Although these applications represent significant advances in the measurement of case-mix, they focus primarily on cost and health care utilization. Studies have shown that both the socioeconomic background and individual clinical status of patients influence performance measurements

Case-Mix Measures for Monitoring System Performance

Background

- The Medicare Advantage Plan (MAP), formerly Medicare + Choice, provides comprehensive health services to 4.6 million enrollees through Medicare managed care plans across the US.

- In order to assess the quality of care provided by MCOs, CMS developed the Medicare Health Outcomes Survey (formerly the Health of Seniors Survey) as a longitudinal evaluation of the health outcomes for MAP enrollees.

- Since 1998, the SF-36 and most recently the Veterans RAND 12 item health Survey (VR-12) is the primary longitudinal outcome measure in Medicare managed care plans and has served as one of the indicators for assessing reimbursement and over all quality of the health plans.

Case-Mix Measures for Monitoring System Performance

Objectives

(1) to determine whether clinically credible and statistically reliable risk-adjusted models can be developed.

(2) to examine whether case- mix differences exist across MCOs.

(3) to assess whether risk adjustment alters judgments of MCO performance.

Case-Mix Measures for Monitoring System Performance

Methods

- The Medicare Advantage Program population was from the Medicare Health Outcomes Survey (HOS) Cohort 7 (2004-2007).
  - HOS randomly sampled a cohort of 1,000 beneficiaries continuously enrolled for at least 6 months in each of the Medicare managed care plans.
  - With exception of a few contract types, almost all Medicare managed care plans (MCO's) participated, a total of 150 MCO's. The population was limited to those beneficiaries 65 years and older.

Case-Mix Measures for Monitoring System Performance

Methods – (1) Outcome Measures

- The following analysis is based upon the MOS SF-12 at baseline and the Veterans RAND 12 Item Health Survey (VR-12) at follow-up two years later. Both 12 item health surveys measure physical to mental health functioning.
- The two health surveys were summarized into physical (PCS) and mental summary (MCS) scores by a linear t-score transformation to have a mean of 50 based upon a U.S population
- Validated conversion formulas allow for direct comparisons of the VR-12 scores with the MOS SF-12.
- The Social Security Administration - Death Master File was used to ascertain vital status.

Case-Mix Measures for Monitoring System Performance

Methods – (2) Outcome Measures

- The study outcome measures included:

(1) the probability of being alive with the same or better PCS score (than would be expected by chance) at 2 years.

(2) the probability of being alive with the same or better MCS score (than would be expected by chance) at 2 years.

Cut points for better were based on 2 SE's of the measurement which were change of more than 6 points for PCS and 7 points for MCS

Outcome formula defined as:

[(Probability (1-death) + probability (PCS (or MCS) the same or better)

* (1-probability of death)]

Case-Mix Measures for Monitoring System Performance

Methods – (3) Risk Adjustment

- Risk adjustment used three domains of risk: sociodemographics, co morbidities and baseline health status.

- Sociodemographics: are included since risk of health outcomes differs by demographic groups (age, gender, race/ethnicity, marital status, education and income).

- Diagnoses: selected for those commonly encountered in clinic visits and known to be indicators of outcomes (acute MI, coronary artery disease, congestive heart failure, stroke, hypertension, diabetes, COPD, asthma, cancer (other than skin cancer), GI disorders, arthritis (hip and hand), and sciatica).

Case-Mix Measures for Monitoring System Performance

Methods – (4) Risk Adjustment

- Baseline Health Status
  - Used the physical (PCS) and mental (MCS) summaries from the MOS SF-12 which summarizes the 8 domains of health.
  - Included the baseline physical (PCS) and mental (MCS) summary scores.

Case-Mix Measures for Monitoring System Performance

Methods – (5) Statistical Analysis

Objective one (develop risk adjusted model)

- Model developed with a derivation sample - 2/3 of the study population (n=98,637).
- Retained variables significant at $p \leq 0.05$ in the final risk adjustment models for the derivation sample.
- Applied regression coefficients from those models onto the remaining 1/3 of the sample (n=49,318).

Case-Mix Measures for Monitoring System Performance

Methods – (6) Statistical Analysis

Objective one (continued- development of risk adjusted model)

• Performance of the models assessed using a c-statistic (0.5 indicates chance) to evaluate the predictive power of the model to discriminate among patients by ordering them according to the rates of the outcome event.

• The Hosmer Lemeshow statistic to evaluate the calibration of the model. Patients divided into deciles based upon the expected risk of improvement or no change in PCS or MCS. Within each decile, the expected rate of improvement or no change was compared with the observed rate. A p value greater than 0.05 indicates a good fit among the deciles.

Case-Mix Measures for Monitoring System Performance

Methods – (7) Statistical Analysis

Objective two (Do case mix differences exist across MCO's?)

- Using the calibrated models developed, we applied expected risk adjusted rates of the outcome (for PCS or MCS) using the multivariate regression models to calculate the expected outcomes for each patient in every MCO.
- Conduct analysis of variance to test for differences in case mix or expected rates among the 150 MCO's.

Case-Mix Measures for Monitoring System Performance

Methods – (8) Statistical Analysis

Objective three (Do risk adjustments alter judgement of MCO performance?).

- Calculate adjusted rates for PCS and MCS.
- Defined as: Observed rate/expected rate * mean of the rates observed for all MCO's.
- Observed the MCO's that changed rank after adjustment with special attention to the identification of outliers.

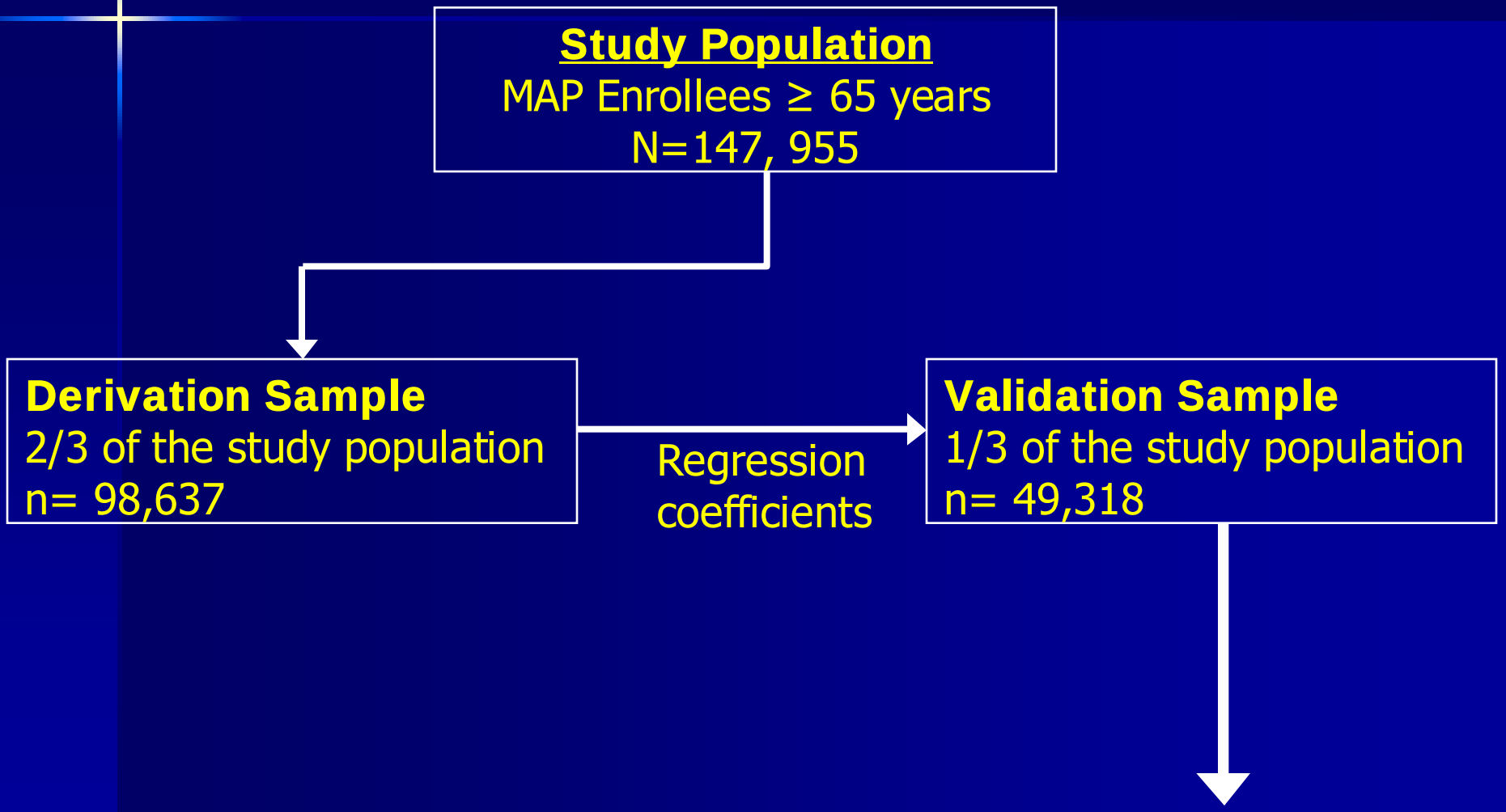
Case-Mix Measures for Monitoring System Performance

Methods – (9) Statistical Analysis

Objective three (Do risk adjustments alter judgement of MCO performance?)

- Statistical significance of an outlier defined as t-statistics valuation (significance of plan differences from the average results).
- Calculated for every plan as (actual minus expected rate/ standard error of the deviation).
- Plans with t statistic >= 2 were designated as better than expected;
- plans <= -2 were significantly worse than expected.

Strategy - I



Strategy - II

Expected rates (over all sample)
- Mortality
- PCS/MCS same or better

Expected rates (plan level)
- Mortality
- PCS/MCS same or better
- Analysis of variance

- Calculation of adjusted rates at the plan level

(Observed rate / expected rate) x national average plan rate

- Comparison of observed vs. adjusted rates of "PCS/MCS same or better" at the plan level
- Outliers identified

Sociodemographic Characteristics of Patients in the Medicare Advantage Program

(N=147,955)

Age, years (Mean ± SD)	76.4 (± 6.9)
65 – 74	42.60%
75 – 84	37.00%
85+	12.50%
Gender – male	40.8%
– female	59.2%

Sociodemographic Characteristics of Patients in the Medicare Advantage Program

(N=147,955)

Race – White	85.1%
– African American	9.3%
– Hispanic	2.5%
– Others	3.1%
Married	54.9%
Education <12	29.7%
Income <$20,000	47.3%
Medicaid	7.4%

Clinical Features of Patients in the Medicare Advantage Program

(N=147,955)

Comorbidity (Mean ± SD)	2.87(± 2)		
Hypertension	62.00%	Cancer	14.90%
Arthritis Hip	44.90%	COPD	13.70%
Arthritis Hand	38.60%	Myocardial Infarction	11.20%
Sciatica	23.10%	Stroke	9.30%
Other Heart Condition	22.40%	Congestive Heart Failure	9.00%
Diabetes	20.10%	Gastrointestinal Disorder	5.10%
Angina/CAD	16.20%		

Functional Status at Baseline for Patients in the Medicare Advantage Program

(N=147,955)

SF-12 Physical And Mental Component Summary Scores

Baseline PCS (Mean ± SD)	39.6 (± 12)
Baseline MCS (Mean ± SD)	51.9 (± 10)

PCS and MCS scores standardized to a US population with a mean of 50

C-statistic and Hosmer-Lemeshow Statistic results for sequential models predicting PCS same or better and MCS same or better in the derivation sample (Validation sample results were comparable and are not shown on the other 1/3 of sample).

Models[1]	Mortality		PCS Same or Better		MCS Same or Better	
	C-statistic	Hosmer-Lemeshow	C-statistic	Hosmer-Lemeshow	C-statistic	Hosmer-Lemeshow
Sociodemographics	0.71	P = 0.033	0.53	P = 0.018	0.57	P = 0.087
Sociodemographics + Medical Conditions	0.77	P = 0.018	0.54	P = 0.066	0.56	P = 0.013
Sociodemographics + Medical Conditions + baseline PCS, MCS	0.80	P = 0.903	0.67	P = 0.214	0.69	P = 0.076

[1]Covariates used in the models: Sociodemographics [age, gender, race/ethnicity, marital status, level of education (<12 years), and income (<$20 000)], Hypertension, Congestive Heart Failure, Stroke, COPD, Gastrointestinal Disorder, Arthritis Hip, Arthritis Hand, Sciatica, AMI, Diabetes, Cancer, baseline PCS and MCS scores.

Case-Mix Measures for Monitoring System Performance

Results

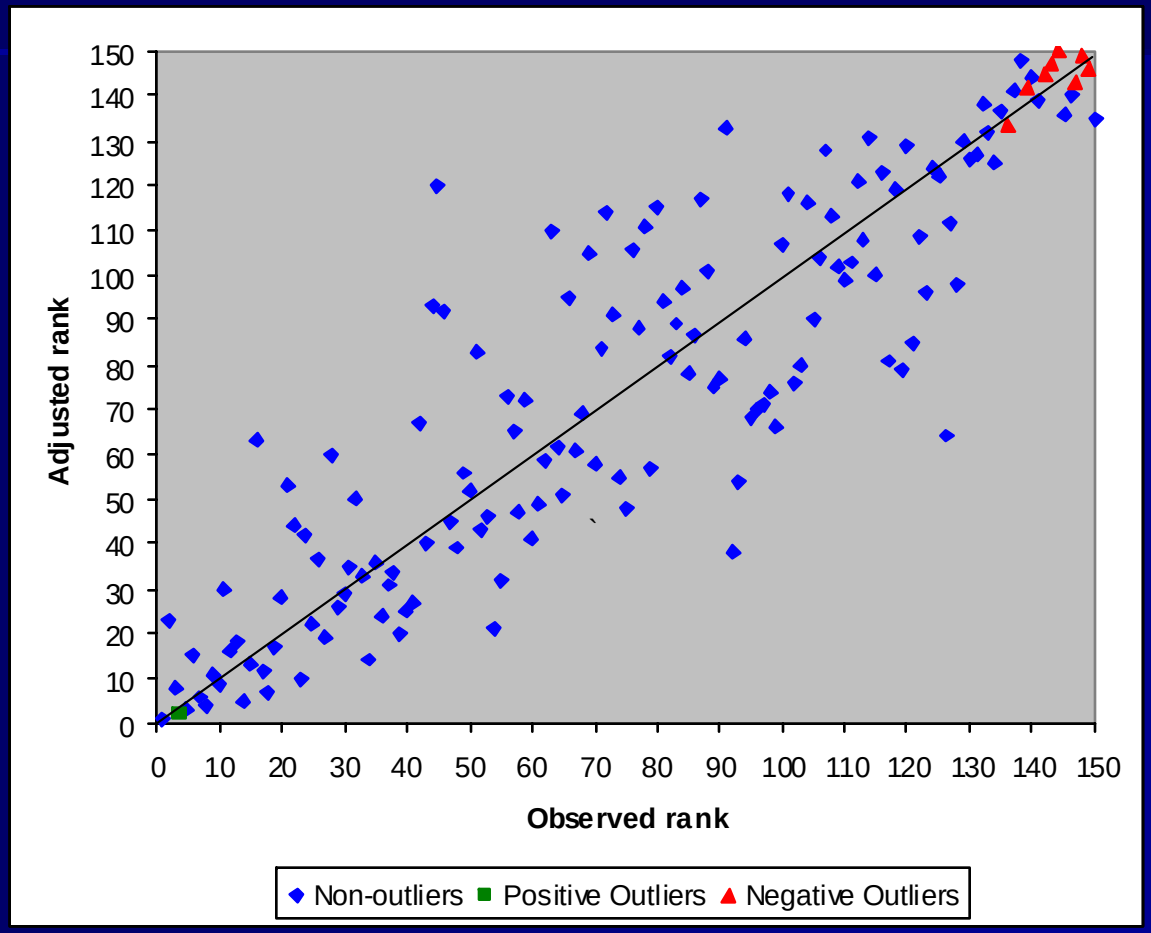
- Expected rates of improvement or no change for PCS ranged from 71.1% to 60.1% and for MCS ranged from 82.0% to 69.8% among the 150 MCO's.
- Analysis of variance to test for differences in case mix or expected rates among the MCO's using the validated model was highly significant for PCS and MCS ($p<0.0001$).

Identification of Outlier Plans

	Positive Outliers	Range (t statistics)	Negative Outliers	Range (t statistics)
PCS	1	(2.00, 2.52)	7	(-5.53, -2.00)
MCS	4	(2.79)	5	(-4.14, -2.02)

Pos. outlier (alive and same or better) >= 2 units (t-statistic)
Neg. outlier (alive and same or better) <= -2 units (t-statistic)

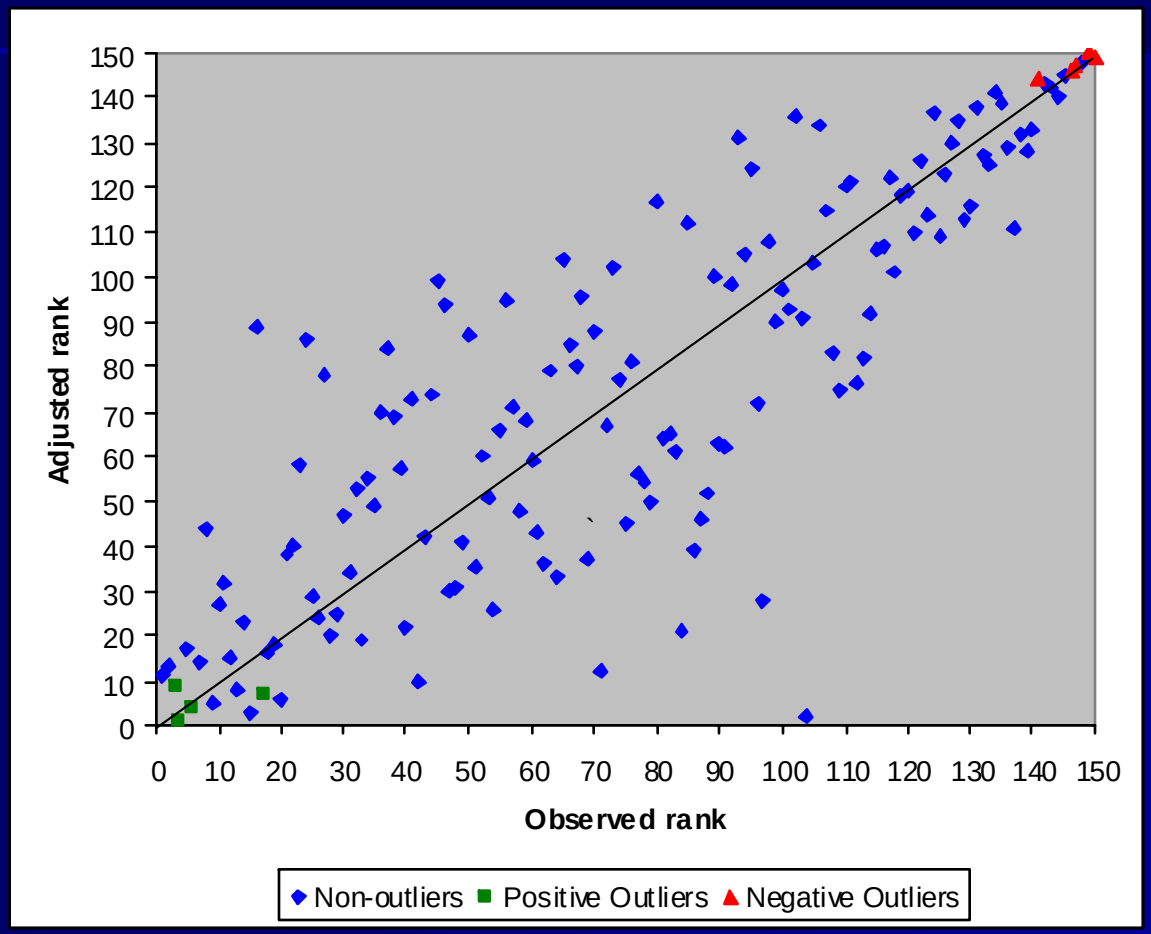
Comparison of Observed and Adjusted Ranks of Plans Physical Summary Outcomes (PCS)



Diagonal line represents no change between observed and adjusted ranks

R = 0.91, p<.0001

Comparison of Observed and Adjusted Ranks of Plans Mental Summary Outcomes (MCS)



Diagonal line represents no change between observed and adjusted ranks

R= 0.89, p<.001

Discussion

- A credible risk adjusted model was developed on the basis of the calibrated model.
- Using this model significant expected rates among the 150 plans strongly suggests that case mix differences exist across the plans.
- We found important differences between adjusted and unadjusted rates used to rank plans, where most plans change ranks to at least some degree after adjustment.

Conclusion

- This study shows that it is feasible to develop clinically credible risk adjustment models with good statistical properties for the health status outcomes using PCS and MCS among the plans in the Medicare Advantage program.
- The identification of the positive outlier plans can be examined for purposes of best practices, while the negative outlier plans can be identified with the purpose of a need for quality improvement.