HealthGrades Hospital Quality in America Study

Sixth Annual
September 2003

HEALTHGRADES®
THE HEALTHCARE QUALITY EXPERTS®

The Sixth Annual HealthGrades Hospital Quality in America Study

Executive Summary

Since 1998, HealthGrades has been studying the quality of care at the nation's nearly 5,000 hospitals and publishing the results of its annual research on the Web to assist consumers in choosing a hospital in their community.

As in the past, the sixth annual *HealthGrades Hospital Quality in America Study* analyzed the most recent three years of risk-adjusted mortality and complication rates available and provides each hospital with a one-, three- or five-star quality rating for each of more than 25 procedures and diagnoses, from heart attack to knee replacement to pneumonia. Beyond these specific quality ratings, HealthGrades analyzed the quality of hospital care on a state level. This analysis found that the quality of care ranged from high to low across many procedures and diagnoses.

Individual hospital results of the sixth annual *HealthGrades Hospital Quality in America Study* are available at www.healthgrades.com. There continues to be considerable variation in quality amongst the nation's hospitals. Other key findings include:

1. Despite previous findings that processes of care are improving, we found substantial differences in outcomes between hospitals, and at an aggregate level, between states. Consistent with Jencks, et al (*JAMA 2003;289:305-312)*, better performance was concentrated in northern states and less populated states while worse performance was concentrated in southern states. See Figure 1.

2. The greatest difference in outcomes at the state level was found with Percutaneous Coronary Intervention (PCI), which includes angioplasty, stent, and atherectomy. In the continental US, New York was the best performing state and Mississippi was the worst. Adjusting for differences in patient populations, a patient who had a PCI in Mississippi was more than one and a half times more likely to die than if they had the procedure performed in New York. See Table 1.

3. Although still significant, the smallest difference in outcomes at the state level was found with Acute Myocardial Infarction (AMI), or Heart Attack. Colorado was the best performing state and Mississippi was the worst. Adjusting for differences in patient populations, a patient who was treated for a heart attack in Mississippi was approximately 50% more likely to die than if they had been treated for the heart attack in Colorado. See Table 1.

4. Many of the worst states ranked among the worst consistently across all procedures and diagnoses (AR, AL, OK, TN, MS). See **Table 1**. However, despite this poor performance at the state level, these states have some 5-star rated hospitals that contributed quality at the hospital level. A few include, but are not limited to: **University of Alabama-**

Birmingham, AL; **University of Tennessee**-Knoxville, TN; **St. Thomas Hospital**-Nashville, TN.

5. Many of the states that ranked the best overall were ranked among the best consistently across all procedures and diagnoses studied (FL, CO, OH, PA, MI). See **Table 1**. Many 5-star rated hospitals are responsible for the aggregate performance of these states. A few of these hospitals include, but are not limited to: **The Pennsylvania Hospital**-Philadelphia, PA; **Nazareth Hospital**-Philadelphia, PA; **Cleveland Clinic Foundation**-Cleveland, OH; **Lanwood Regional Medical Center**-Ft. Pierce, FL; **Parma Community General Hospital**-Parma, OH; **St. Vincent Health System**-Erie, PA; **William Beaumont Hospital**-Royal Oak, MI. Although not located in the above-mentioned states, several well-known academic centers also contributed positively to their states' performance that ranked in the top quartile. A few of these hospitals include, but are not limited to: **Mayo Clinic**-Rochester, MN; **Yale New Haven**- New Haven, CT; **Brigham and Womens Hospital**-Boston, MA; **Massachusetts General Hospital**-Boston, MA; **UCSF Medical Center**-San Francisco, CA.

6. States with well-known and publicized quality-improvement efforts around, or hospital and physician specific public profiling of, Coronary Artery Bypass Graft Surgery (CABG) outcomes are ranked among the best states for performance of CABG surgery (NJ, NY, PA, MI, MA, VA) See **Table 1**. Many 5-star rated hospitals are responsible for the aggregate performance of these states. A few include, but are not limited to: **St. Vincent Health System**-Erie, PA; **St. Peter's Hospital**-Albany, NY; **CJW Medical Center**-Richmond, VA.

Introduction

During the past 5 years, HealthGrades has consistently seen significant variation in the quality of care provided by the nation's hospitals. HealthGrades provides to consumers, for each hospital and each procedure and diagnosis, a star rating indicating whether that hospital has performed "best," "as expected" or "poor". Hospital ratings are assessed based on patient outcomes, specifically, mortality or complications. Because no two hospitals or their patients' risk profiles are alike, HealthGrades has developed extensive risk-adjustment algorithms to ensure that it is making fair, apples-to-apples comparisons.

Consumers are becoming increasingly knowledgeable about quality differences and are using quality data to make better informed health care choices. In a study conducted by VHA in 2000, 87% of respondents said that a poor or below average clinical quality report would persuade them to choose a different hospital. 52% of respondents to a survey done by Endresen Research in 2002 said that having information that a hospital was top-rated would have a major impact on selecting a hospital. The primary goal of the 6th Annual *Hospital Quality in America Study* was to meet this consumer need for additional hospital quality information. HealthGrades' Web site has approximately 4,000,000 users per year and provides quality ratings to 15 million people via its subscription-based sites.

Several studies have consistently demonstrated persistent quality gaps, including both unexpected mortality and complications from both poor processes of care and medical errors. As a result of the identification of this well known and well publicized “quality chasm”, national efforts have focused on the substantial opportunity for improvement in the effectiveness of care processes. Recent studies by Jencks et al. and Burwen et al. (JAMA 2003;289:305-312 and Archives of Internal Medicine 2003;163:1430-1439, respectively) indicate that while medicine still has a long way to go, adherence with standards of practice is improving. No national study has been done to evaluate the outcomes performance subsequent to this improvement finding. A second important goal of this study was to determine if outcomes of care are improving or if there continues to be substantial variation in the quality of care in America.

While conducting this study, we also looked at state-level performance (outcomes) for 5 high-volume inpatient procedures and diagnoses that represent the most common and well studied of quality improvement projects: Coronary Artery Bypass Graft surgery (CABG), Percutaneous Coronary Interventions (PCI), Acute Myocardial Infarction (AMI), Congestive Heart Failure (CHF) and Community Acquired Pneumonia (CAP). This is a national, state-based study that aims to identify whether quality of care gaps still exist among Medicare beneficiaries and attempts to identify characteristics and trends that might contribute to these differences. As part of this study, HealthGrades also examined the relative or comparative increased risk of dying or surviving between the best and worst performing states to further highlight the persistent variation in quality of care throughout the United States and to create a sense of urgency. It should be noted that there are excellent hospitals in both the best and the worst performing states, however, Medicare patients, on average, received better care in the “best” states.

Methods

6th Annual Hospital Quality Ratings Methods

HealthGrades rated nearly 5,000 hospitals in the following categories (ratings available at www.healthgrades.com):

1. Aspiration Pneumonia
2. Atrial Fibrillation
3. Back and Neck Surgery (except Spinal Fusion)
4. Back and Neck Surgery (Spinal Fusion)
5. Bowel Obstruction
6. Carotid Endarterectomy
7. Cholecystectomy (gallbladder surgery)
8. Chronic Obstructive Pulmonary Disease
9. Community Acquired Pneumonia
10. Coronary Bypass Surgery
11. Gastrointestinal (GI) Bleed
12. Heart Attack
13. Heart Failure
14. Hip Fracture Repair (ORIF)
15. Partial Hip Replacement
16. Percutaneous Coronary Intervention (PTCA/Angioplasty, Stent, Atherectomy)
17. Peripheral Vascular Bypass
18. Peripheral Vascular Interventional Procedures (Angioplasty and/or Stent)
19. Prostatectomy
20. Resection / Replacement of Abdominal Aorta
21. Respiratory Infection except Aspiration Pneumonia and Tuberculosis
22. Sepsis
23. Stroke
24. Total Hip Replacement
25. Total Knee Replacement
26. Valve Replacement Surgery

Data Acquisition

HealthGrades used MedPAR data for 2000 – 2002 to perform this study. The MedPAR data was selected for several reasons. First, it included virtually every hospital in the country, with the exception of military and Veterans Administration hospitals. Second, hospitals were required by law to submit complete and accurate information with substantial penalties for those that report inaccurate or incomplete data. Third, the Medicare population represented a majority of the patients for all of the clinical categories studied, with approximately 55% to 60% of all cardiac patients and 75% to 80% of all joint replacement surgeries, for example.

To preserve the integrity of the HealthGrades ratings, HealthGrades conducted a series of data quality checks. Based on the results of these checks, we excluded a limited number of cases because they were inappropriate for inclusion in the database or miscoded. Examples of excluded patient records were:

- Patients under the age of 65
- Patients who left the hospital against medical advice or who were transferred to another acute care hospital
- Patients discharged alive with a length of stay equal to or less than one day for Coronary Artery Bypass Graft Surgery, Valve Replacement Surgery, Aspiration Pneumonia, Stroke, Resection/Replacement of Abdominal Aorta, Hip Fracture Repair (ORIF), Partial Hip Replacement, Total Knee Replacement, Total Hip Replacement, and Sepsis.
- Patients who were still in the hospital when the Medicare claim was filed
- Patients with an invalid gender

Data Analysis

The Inhospital data on the HealthGrades Web site represent three years of patient discharges (2000-2002). The Inhospital +1 Month and Inhospital +6 Months Mortality data on the HealthGrades Web site represent two years of patient discharges (2001-2002) due to inaccuracies in the discharge to date of death field in the 2000 MedPAR file.
In the initial analysis of the data, a separate data set was created for each group of patients having a specific procedure or diagnosis based on ICD-9-CM coding (e.g., coronary bypass surgery, total hip replacement), which can be found in **Exhibit A**. Each group of patients was defined by using the information on diagnoses and procedures coded in the patient records. See **Exhibit A** for a list of the diagnosis and procedure codes that define each patient cohort. The quality measure for some cohorts was mortality, whereas, for other cohorts, the quality measure was major complications.

For each patient cohort, we developed a list of specific procedures (e.g., quadruple bypass surgery), a list of risk factors, and a list of post-surgical complications. These latter two lists were developed in two steps:

(1) We identified all diagnoses occurring in more than 1% of the patients for the current analysis and the previous analysis.
(2) We used a team of clinical and coding experts to identify the complications in the list created in Step One.

Some diagnosis codes were merged together (e.g., primary and secondary pulmonary hypertension) to minimize the impact of coding variations.

Outcomes were binary, with documented major/minor complications either present or not, and patients recorded as either alive or expired. Contact HealthGrades for a list of complications included in the quality measure "Major Complications". In cohorts where the quality measure is major complications, mortality is considered a complication.

Risk-Adjustment Methodology

The purpose of risk-adjustment is to obtain fair statistical comparisons between disparate populations or groups. Significant differences in demographic and clinical risk factors are found among patients treated in different hospitals. Risk-adjustment of the data is needed to make accurate and valid comparisons of clinical outcomes at different hospitals.

Fair and valid comparisons between hospital providers can be made only to the extent that the risk-adjustment methodology considers important differences in patient demographic and clinical characteristics. The risk-adjustment methodology used by HealthGrades defines risk factors as those clinical and demographic variables that influence patient outcomes in significant and systematic ways. Risk factors may include age, sex, specific procedure performed, and comorbid conditions such as hypertension, chronic renal failure, congestive heart failure, and diabetes. The methodology is disease-specific and outcome-specific. This means that individual risk models are constructed and tailored for each clinical condition or procedure, and also for each outcome.

Developing the HealthGrades ratings involved four steps for each cohort (e.g., coronary bypass surgery) and quality measure (e.g., inhospital mortality). First, the predicted value (e.g., predicted mortality) was obtained using logistic regression models discussed in the next section. Second, the predicted value was compared with the actual, or observed, value (e.g., actual mortality). Third, a test was conducted to determine whether the difference between the predicted and actual/observed values was statistically significant. This test was performed to make sure that differences were very unlikely to be caused by chance alone. Fourth, a star rating was assigned based upon the outcome of the statistical test.

Statistical Models

Unique statistical models were developed for each patient cohort and each outcome using multivariate logistic regression.

Comorbid diagnoses (e.g., hypertension, chronic renal failure, anemia, diabetes), demographic characteristics (e.g., age and sex), and specific procedures (where relevant) were classified as potential risk factors. We used multivariate logistic regression to determine which of these were actually risk factors and to what extent they were correlated with the quality measure (e.g., mortality). A risk factor stayed in the model if it had a positive coefficient and was also statistically significant ($p<0.0001$) in explaining variation. Complications were ***not*** counted as risk factors as they were considered a result of care received during the admission.

The statistical models were checked for validity and finalized. All of the models were highly significant, with C-statistics ranging from ~ 0.6 to ~ 0.9. These cohort and outcome specific models were then used to estimate the probability of the outcome for each patient in the cohort. Patients were then aggregated for each hospital to obtain the predicted outcome for each hospital. Statistical significance tests were performed to identify, by hospital, whether the actual/observed and predicted rates were significantly different. We used a binomial distribution to establish an approximate 90% confidence interval. To test the fit of a binomial distribution to the data, we performed tests on each model for 20% of the hospitals whereby we included statistical significance, for each hospital individually, as an independent variable in the logistic regression model. We subsequently used a two-tailed z-test to again determine statistical significance. The match between the binomial distribution results and the test sample within the logistic regression models themselves was nearly 100%.

Assignment of Star Ratings

The following rating system was applied to the data for all procedures and diagnoses:

★★★★★ Actual performance was better than predicted and the difference was statistically significant.

★★★ Actual performance was not significantly different from what was predicted, or "as expected".

★ Actual performance was worse than predicted and the difference was statistically significant.

In general, 70% to 80% of hospitals in each procedure/diagnosis are classified as three stars, with actual results statistically the same as predicted results. Approximately 10% to 15% were one-star hospitals and 10% to 15% were five-star hospitals. The data fell out in a fairly well structured bell shaped curve.

State Level Performance Study Methods

The purpose of this part of the study was to evaluate state-level performance by measuring the outcomes of five key procedures and diagnoses. Risk-adjusted outcomes (inhospital mortality) performance was calculated at the state level for each cohort: Coronary Artery Bypass Graft surgery (CABG), Percutaneous Coronary Interventions (PCI), Acute Myocardial Infarction (AMI), Congestive Heart Failure (CHF) and Community Acquired Pneumonia (CAP). These 5 procedures and diagnoses were chosen because they represent some of the most studied procedures and diagnoses for quality improvement and because they rank high among the most common diseases for hospital admission among Medicare beneficiaries.

Using the MEDPAR 2000-2002 data and the risk-adjustment methodology discussed in the last section, HealthGrades calculated the actual (observed) and predicted (expected) number of deaths by state by cohort. A ratio of observed (O) to expected (E) deaths was then calculated for each state for each cohort. An O/E ratio of less than 1 means that the state had fewer deaths than expected given their patient population. An O/E of greater than 1 means that the state had more deaths than expected given their patient population. These state performance ratios were then rank ordered for each cohort in ascending order of their respective O/E ratio (lowest to highest). Finally, the average of the 5 cohort ranks for each state was calculated and then the states were arrayed in ascending order according to their average final rank. When ties occurred among the averaged ranks, state volume was used to break the tie.

States were also designated Best or Worst in performance. "Best" was defined as an O/E ratio of less than 1 and the difference between the observed and expected was statistically significant ($p<0.1$). "Worst" was defined as an O/E ratio greater than 1 and the difference between observed and expected was statistically significant.

The relative mortality risk increase (decrease) was calculated for each state for each cohort compared to the performance of the best performing state and the US average. This was calculated by subtracting the O/E ratio of the benchmark (best state or US average) from each state's O/E ratios. A positive value is associated with a relative risk increase where a negative value is associated with a relative risk decrease given the medical complexities of each population.

The absolute number of extra deaths (lives saved) was calculated by subtracting the observed from the expected for each state in each cohort. Similar to the relative mortality risk, a positive value denotes extra or unexpected deaths and a negative value (shown in parentheses) denotes extra or unexpected lives saved given the medical complexities of each population.
Average yearly hospital volume by state was calculated by first identifying the total number of patients for each year for each state by cohort and then dividing it by the total number of hospitals for each year in that respective state by cohort. Finally, all three respective year's average hospital volumes were averaged. The average yearly hospital volume for each state that ranked among the best and worst was averaged to determine

the average yearly hospital volume for the Best and Worst performing state group in each cohort.

Results

HealthGrades' ratings of nearly 5,000 hospitals, based on the sixth annual HealthGrades Hospital Quality in America Study, can be found at www.healthgrades.com. For all of the specific procedures and diagnoses rated, 10 – 15% of hospitals stand out as "best" performers (5 star rated), while another 10 - 15% stand out as "poor" performers (1 star rated). The remaining hospitals are "as expected" (3 star rated). Past studies done by HealthGrades showed that a substantial number of lives could be saved if Americans simply did not go to hospitals rated as "1 star."

Results of the State Level Performance Study:

Table 1 shows for Medicare patients the 2000-2002 risk-adjusted inhospital O/E ratio and rank performance for each cohort and the average rank for the overall aggregated performance of the total of 5 cohorts evaluated in each state. **Figure 1** shows geographically the national pattern of overall performance in 2000-2002. Both Table 1 and Figure 1 clearly highlight the substantial variation in national performance. Consistent with the findings from the previously cited study, **"**Change in the Quality of Care Delivered to Medicare Beneficiaries 1998-1999 to 2000-2001." Jencks, et. al. *JAMA* 2003;289:305-312, which looked at process measures of care and was co-authored by the United States Assistant Surgeon General, we found that better performance is concentrated in northern states and less populated states while worse performance is concentrated in southern states.

Risk-adjusted outcomes vary significantly from state to state and region to region. **Table 2** shows that the greatest difference in outcomes was observed in the state level performance of Percutaneous Coronary Intervention (PCI). After accounting for patient severity of illness and other medical conditions in the risk-adjustment process, compared to an average patient in New York, the best performing state, an average patient in Mississippi, the worst performing state in the continental US had a relative inhospital mortality risk increase from PCI of **87.05**%. This means that patients undergoing PCI in Mississippi were more than one and a half times more likely to die than patients undergoing the same procedure in New York.

Compared to the US average, this same patient in Mississippi had a **47.07**% relative inhospital mortality risk *increase* while a patient in New York had a **39.97**% relative inhospital mortality risk *decrease*. The relative risk increase over the national average was associated with **75** unexpected deaths from PCI in Mississippi from 2000-2002. This number does not seem too alarming until compared to the best performing state, New York, whose practices and resultant quality outcomes saved **287** additional lives over the same time period. New York also ranked among the best states for performance of CABG surgery while Mississippi also ranked among the worst for CABG surgery.

The smallest difference in outcomes was observed in the state level performance of AMI. Although this represented the smallest performance gap in this study of performance of 5 cohorts, this difference is still substantial, as AMI is responsible for so many deaths each year in America. Compared to an average patient in Colorado, the best performing state, an average patient in Mississippi, the worst performing state, had a relative mortality risk increase of **49.35**%. Compared to the national average, this relative risk increase was associated with **448** unexpected deaths from AMI in Mississippi from 2000-2002. Colorado's relative risk reduction was associated with **232** additional lives saved during the same time period as compared the national average performance. Colorado also ranked among the best states for performance of CHF while Mississippi, Arkansas, Oklahoma, Tennessee, and Alabama consistently ranked among the worst states for all cardiac cohorts.

Table 1. State Performance Rankings 2000-2002

		Quality Targets									
		CABG		PCI		AMI		CHF		CAP	
State	Avg Rank	Rank	O/E Ratio	Rank	O/E Ratio	Rank	O/E Ratio	Rank	O/E Ratio	Rank	O/E Ratio
AK	39	30	1.025	51*	1.796	36	1.072	51*	1.552	8†	0.852
AL	50	46*	1.293	45*	1.204	46*	1.100	36*	1.106	42*	1.091
AR	49	44*	1.258	47*	1.253	50*	1.141	48*	1.328	40*	1.066
AZ	13	31	1.036	34	1.081	6†	0.869	1†	0.760	2†	0.774
CA	25	29	1.021	17	0.958	22†	0.967	14†	0.952	29	0.991
CO	6	20	0.881	15	0.934	1†	0.800	4†	0.805	3†	0.779
CT	8	10†	0.840	18	0.960	13†	0.906	21	0.987	26	0.966
DC	22	34	1.053	31	1.062	9†	0.893	5†	0.821	10†	0.865
DE	35	40	1.099	37	1.152	35	1.065	8†	0.850	25	0.952
FL	2	27	1.001	6†	0.881	10†	0.896	2†	0.766	4†	0.797
GA	37	48*	1.370	9†	0.915	31	1.022	22	0.996	36*	1.041
HI	44	51*	1.483	26	1.032	44*	1.086	41*	1.166	44*	1.143
IA	40	41*	1.155	42*	1.175	25	0.976	40*	1.164	14†	0.893
ID	24	33	1.049	3†	0.778	5†	0.869	39*	1.160	31	1.000
IL	17	14†	0.915	23	0.988	23†	0.967	12†	0.893	11†	0.882
IN	20	26	1.001	21	0.969	21†	0.956	26	1.019	27	0.977
KS	47	39	1.078	49*	1.437	41*	1.065	45*	1.212	21†	0.959
KY	34	37	1.071	40*	1.133	38*	1.046	20	0.987	23†	0.970
LA	28	38	1.077	33	1.075	26	0.991	15†	0.963	19†	0.950
MA	26	4†	0.778	25	1.008	20†	0.956	18	0.975	38*	1.054
MD	5	25	0.983	27	1.036	17†	0.942	3†	0.791	7†	0.851
ME	16	19	0.875	2†	0.732	2†	0.833	32*	1.090	17†	0.934
MI	4	11†	0.851	10†	0.920	18†	0.944	10†	0.875	15	0.900
MN	12	7†	0.806	19	0.962	4†	0.849	11†	0.881	6	0.839
MO	27	36	1.071	38*	1.096	42*	1.075	25	1.014	20	0.953
MS	51	45*	1.289	50*	1.471	51*	1.294	50*	1.474	51*	1.334
MT	14	3†	0.709	11	0.846	7†	0.887	24	1.014	16†	0.910
NC	29	24	0.982	28	1.041	43*	1.083	33*	1.090	48*	1.195
ND	1	8†	0.817	12	0.920	14†	0.913	6†	0.833	1†	0.682
NE	32	32	1.043	35	1.112	33	1.050	38*	1.154	35	1.033
NH	23	17	0.846	13	0.929	8†	0.888	42*	1.166	32	1.009
NJ	15	12†	0.864	4†	0.781	40*	1.058	31*	1.059	33	1.017
NM	31	1†	0.664	20	0.966	30	1.010	47*	1.303	46*	1.165
NV	36	50*	1.389	43*	1.181	32	1.039	16	0.937	12†	0.892
NY	33	9†	0.833	1†	0.601	49*	1.134	46*	1.227	50*	1.272
OH	3	13	0.882	8†	0.913	12†	0.901	7†	0.847	9†	0.853
OK	45	47*	1.332	44*	1.202	47*	1.110	44*	1.190	34	1.028
OR	21	21	0.888	24	0.997	16†	0.938	35*	1.102	30	0.997
PA	7	5†	0.793	7†	0.907	28	0.998	13†	0.942	22†	0.964
RI	18	16	0.785	14	0.930	15†	0.919	9†	0.868	24	0.943
SC	42	22	0.932	30	1.060	48*	1.129	43*	1.177	49*	1.249
SD	10	18	0.854	16	0.953	11†	0.898	27	1.019	13†	0.892
TN	48	43*	1.172	46*	1.248	45*	1.087	37*	1.123	43*	1.104
TX	30	42*	1.160	41*	1.150	37*	1.036	19	0.984	28	0.990
UT	9	6†	0.801	29	1.051	3†	0.843	29	1.071	5†	0.822
VA	11	2†	0.700	22	0.978	29	1.004	23	1.005	37*	1.043
VT	46	49*	1.377	32	1.064	34	1.058	49*	1.342	45*	1.151
WA	19	28	1.019	39*	1.115	19†	0.945	17	0.964	18†	0.935
WI	38	35	1.057	5†	0.850	27	0.992	34*	1.095	39*	1.055
WV	41	23	0.952	36	1.120	39*	1.052	28	1.021	41*	1.069
WY	43	15	0.768	48*	1.425	24	0.944	30	1.094	47*	1.168

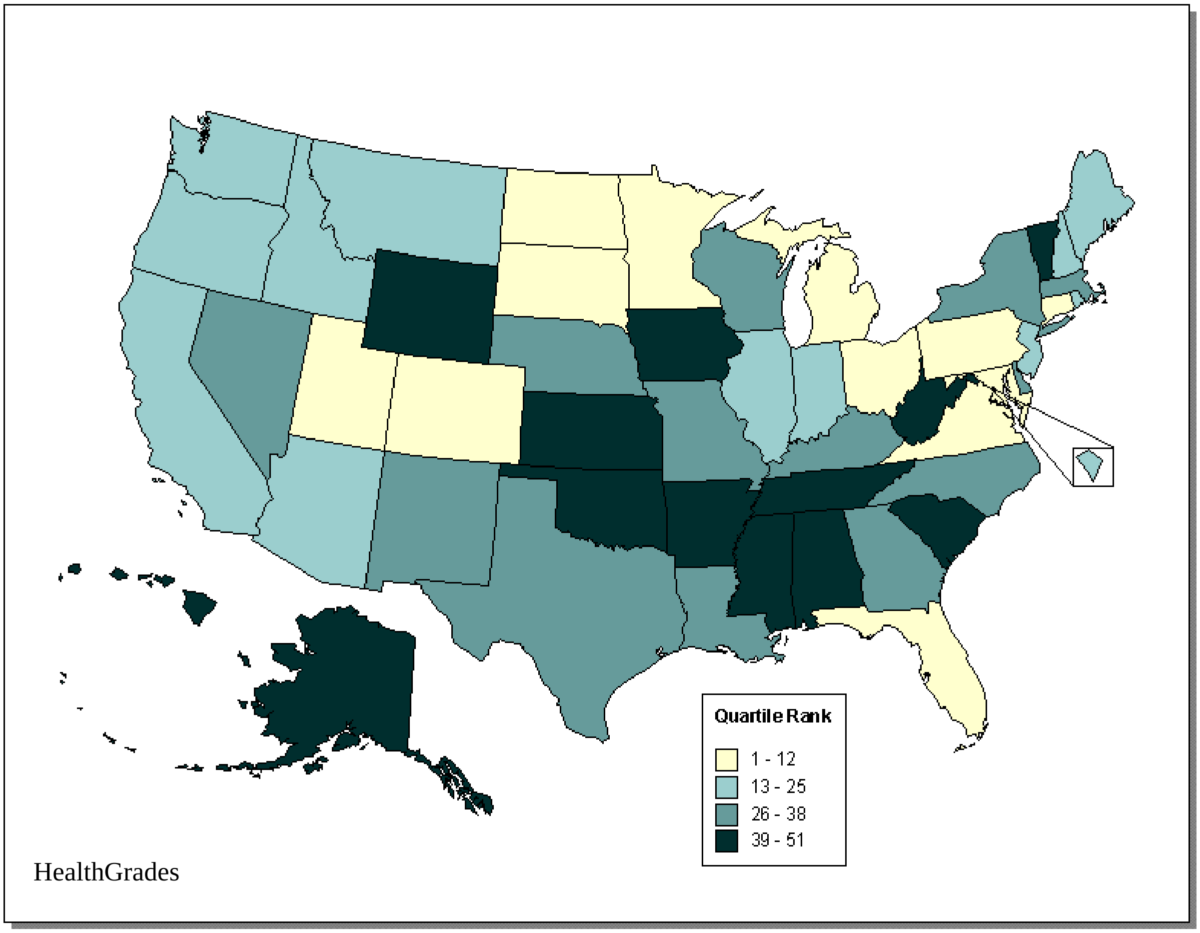
† Ranked among the Best states-performance was statistically significantly better than expected
* Ranked among the Worst states-performance was statistically significantly worse than expected
Puerto Rico, Guam, Virgin Islands and American Samoa were analyzed, but their results are not included in this table.

Table 2. State Relative Risk Increase (Decrease) and Number of Extra Deaths (Lives Saved) 2000-2002

	CABG			PCI			AMI			CHF			CAP		
State	%higher risk of dying compared to Best Ranked State (NM)	%higher risk of dying (surviving) compared to US average	number of extra deaths (lives saved)	%higher risk of dying compared to Best Ranked State (NY)	%higher risk of dying (surviving) compared to US average	number of extra deaths (lives saved)	%higher risk of dying compared to Best Ranked State (CO)	%higher risk of dying (surviving) compared to US average	number of extra deaths (lives saved)	%higher risk of dying compared to Best Ranked State (AZ)	%higher risk of dying (surviving) compared to US average	number of extra deaths (lives saved)	%higher risk of dying compared to Best Ranked State (ND)	%higher risk of dying (surviving) compared to US average	number of extra deaths (lives saved)
AK	36.10	2.34	0	119.56	79.59	9	27.16	7.20	8	79.19	55.18	39	16.91	(14.84)	(17)
AL	62.87	29.12	108	60.30	20.33	84	29.93	9.96	277	34.62	10.62	247	40.83	9.08	243
AR	59.37	25.61	48	65.24	25.26	54	34.06	14.09	265	56.87	32.87	484	38.33	6.58	122
AZ	37.26	3.51	8	48.03	8.05	24	6.89	(13.08)	(207)	NA	(24.00)	(241)	9.15	(22.60)	(331)
CA	35.72	1.96	18	35.77	(4.20)	(48)	16.69	(3.28)	(273)	19.19	(4.82)	(320)	30.86	(0.89)	(92)
CO	21.75	(12.01)	(13)	33.33	(6.65)	(13)	NA	(19.97)	(232)	4.49	(19.51)	(140)	9.63	(22.12)	(239)
CT	17.65	(16.11)	(28)	35.96	(4.02)	(9)	10.56	(9.41)	(171)	22.76	(1.24)	(17)	28.33	(3.42)	(67)
DC	38.89	5.13	5	46.11	6.14	7	9.30	(10.67)	(56)	6.17	(17.84)	(65)	18.21	(13.54)	(50)
DE	43.56	9.80	4	55.17	15.19	10	26.50	6.53	29	9.06	(14.95)	(57)	26.98	(4.78)	(20)
FL	33.72	(0.04)	1	28.02	(11.95)	(173)	9.53	(10.44)	(1161)	0.58	(23.43)	(2020)	11.43	(20.32)	(1712)
GA	70.63	36.88	114	31.41	(8.56)	(32)	22.15	2.18	73	23.64	(0.36)	(11)	35.81	4.05	168
HI	81.94	48.18	19	43.11	3.13	1	28.56	8.59	31	40.61	16.61	33	46.06	14.31	41
IA	49.09	15.33	30	57.43	17.45	54	17.60	(2.36)	(47)	40.41	16.40	204	21.10	(10.66)	(200)
ID	38.52	4.77	2	17.70	(22.28)	(14)	6.86	(13.11)	(59)	40.01	16.00	42	31.76	0.01	0
IL	25.14	(8.61)	(60)	38.71	(1.26)	(10)	16.71	(3.26)	(207)	13.29	(10.72)	(570)	19.97	(11.78)	(809)
IN	33.71	(0.05)	0	36.84	(3.14)	(19)	15.60	(4.37)	(179)	25.89	1.88	58	29.45	(2.30)	(81)
KS	41.47	7.71	12	83.66	43.68	81	26.50	6.53	85	45.22	21.21	219	27.65	(4.11)	(66)
KY	40.77	7.01	24	53.26	13.28	41	24.53	4.56	128	22.70	(1.31)	(30)	28.73	(3.02)	(98)
LA	41.31	7.55	16	47.41	7.43	26	19.11	(0.86)	(18)	20.32	(3.69)	(86)	26.77	(4.98)	(136)
MA	11.39	(22.36)	(53)	40.75	0.77	3	15.59	(4.38)	(141)	21.57	(2.44)	(56)	37.20	5.44	170
MD	31.89	(1.87)	(4)	43.53	3.55	11	14.14	(5.82)	(162)	3.12	(20.88)	(489)	16.81	(14.94)	(368)
ME	21.10	(12.65)	(10)	13.12	(26.86)	(27)	3.28	(16.69)	(189)	32.98	8.97	49	25.11	(6.64)	(44)
MI	18.71	(15.05)	(87)	31.93	(8.04)	(70)	14.33	(5.64)	(351)	11.55	(12.45)	(612)	21.77	(9.98)	(510)
MN	14.23	(19.52)	(41)	36.10	(3.88)	(14)	4.84	(15.13)	(392)	12.17	(11.83)	(182)	15.62	(16.13)	(347)
MO	40.68	6.92	29	49.50	9.53	60	27.46	7.49	275	25.44	1.43	37	27.09	(4.66)	(156)
MS	62.56	28.80	41	87.05	47.07	75	49.35	29.38	448	71.44	47.43	710	65.14	33.39	606
MT	4.51	(29.25)	(14)	24.56	(15.42)	(15)	8.67	(11.30)	(52)	25.40	1.39	4	22.75	(9.00)	(42)
NC	31.87	(1.89)	(7)	44.09	4.11	24	28.25	8.28	367	33.05	9.05	304	51.30	19.55	815
ND	15.33	(18.43)	(18)	31.95	(8.02)	(7)	11.26	(8.71)	(54)	7.37	(16.63)	(50)	NA	(31.75)	(161)
NE	37.94	4.18	6	51.12	11.14	16	24.94	4.97	42	39.47	15.47	88	35.08	3.33	32
NH	18.26	(15.49)	(12)	32.85	(7.13)	(6)	8.72	(11.25)	(79)	40.63	16.62	63	32.63	0.88	4
NJ	20.00	(13.75)	(51)	18.07	(21.91)	(102)	25.81	5.84	285	29.90	5.89	280	33.47	1.72	82
NM	NA	(33.76)	(15)	36.57	(3.41)	(3)	21.01	1.04	5	54.36	30.35	91	48.24	16.48	119
NV	72.53	38.77	36	58.08	18.11	20	23.91	3.94	23	17.74	(6.27)	(29)	20.92	(10.83)	(66)
NY	16.88	(16.88)	(114)	NA	(39.97)	(287)	33.38	13.42	1139	46.75	22.74	1601	58.93	27.18	2694
OH	21.79	(11.97)	(79)	31.27	(8.71)	(77)	10.09	(9.88)	(681)	8.73	(15.27)	(919)	17.10	(14.65)	(1030)
OK	66.79	33.03	65	60.12	20.15	55	30.96	10.99	237	43.02	19.01	269	34.57	2.82	61
OR	22.42	(11.34)	(13)	39.69	(0.29)	0	13.76	(6.21)	(78)	34.23	10.23	68	31.44	(0.31)	(3)
PA	12.92	(20.83)	(171)	30.59	(9.38)	(92)	19.81	(0.16)	(13)	18.22	(5.79)	(403)	28.16	(3.59)	(230)
RI	12.10	(21.66)	(7)	32.96	(7.01)	(3)	11.89	(8.08)	(47)	10.82	(13.18)	(60)	26.08	(5.67)	(30)
SC	26.81	(6.94)	(14)	45.96	5.98	16	32.88	12.91	272	41.68	17.67	298	56.67	24.92	489
SD	19.06	(14.69)	(9)	35.24	(4.74)	(3)	9.74	(10.23)	(62)	25.91	1.91	5	20.93	(10.83)	(59)
TN	50.81	17.06	78	64.78	24.81	120	28.67	8.70	337	36.29	12.29	378	42.17	10.42	452
TX	49.63	15.87	172	54.91	14.93	193	23.55	3.58	315	22.45	(1.56)	(115)	30.73	(1.02)	(98)
UT	13.73	(20.02)	18	45.06	5.08	7	4.25	(15.72)	103	31.08	7.07	25	13.95	(17.81)	(133)
VA	3.59	(30.16)	(88)	37.78	(2.20)	(8)	20.38	0.41	14	24.53	0.53	15	36.10	4.35	145
VT	71.36	37.60	10	46.38	6.40	2	25.72	5.75	20	58.24	34.24	57	46.82	15.07	41
WA	35.50	1.74	4	51.44	11.47	32	14.48	(5.49)	(116)	20.45	(3.55)	45	25.24	(6.52)	(116)
WI	39.28	5.53	21	24.99	(14.99)	(60)	19.20	(0.77)	(22)	33.51	9.50	178	37.23	5.48	127
WV	28.86	(4.90)	(8)	51.95	11.97	16	25.13	5.16	90	26.12	2.11	29	38.66	6.91	116
WY	10.38	(23.38)	(3)	82.48	42.51	12	14.41	(5.56)	(10)	33.44	9.44	11	48.60	16.84	37

Puerto Rico, Guam, Virgin Islands and American Samoa were analyzed but their results were not included in this table.

Figure 1. State Ranking on Overall Performance of 5 Inpatient Medicare Quality Targets 2000-2002



States are ranked according to their average performance across 5 cohorts.
Puerto Rico, Guam, Virgin Islands and American Samoa were analyzed, but their performances were not included in the map.

States with well-known and publicized Centers of Medicare and Medicaid Services' Quality Improvement Organization (QIO) efforts or with hospital and physician-specific public profiling of CABG outcomes, are ranked among the best states for performance of CABG surgery. With the exclusion of Maryland which did not make the best states' list, New York, Pennsylvania and New Jersey all publicly profile hospital and physician specific performance for CABG surgery to consumers. Massachusetts, Virginia, Michigan and Illinois have participated in major quality initiatives and/or have very active QIOs focused on initiatives to improve the outcomes of patients who receive this surgery.

Table 3 shows some characteristics between the best and worst performing states. For all cardiac procedures and diagnoses, there was a positive association between performance and average yearly hospital volume consistent with previous studies that have identified a positive relationship between volume and outcome. Although most of the states that ranked among the best for CABG surgery performance had either the toughest or moderate certificate of need (CON) laws, most states that ranked among the worst also had the toughest or moderate CON laws. There was no identified trend observed between quality and average length of stay or hospital system ownership.

Table 3. Characteristics of Best and Worst states 2000-2002

	Average Yearly Hospital Volume		Percent higher comparative risk to die or (live) between Best & Worst states		Percent of hospitals in a system		Percent of states with tough or moderate CON laws*	
	Best	Worst	Best	Worst	Best	Worst	Best	Worst
CABG	152	144	(49.40)	49.40	56.72	61.01	64.28	90.90
PCI	338	200	(45.00)	45.00	56.62	55.12	NA	NA
AMI	77	72	(19.62)	19.62	53.07	59.9	NA	NA
CHF	185	111	(34.73)	34.73	58.06	54.87	NA	NA
CAP	103	118	(26.15)	26.15	57.29	59.87	NA	NA

* Source: *Modern Healthcare* April 21, 2003.

Interpretation of Results

Despite the improvement in process measures of care associated with many of these procedures and diagnoses observed by Jencks et al and Burwen et al., this study shows that there is significant variation in the quality of care delivered by different hospitals and at the aggregate level, in different states.

The greatest disparity in performance noted in PCI may be due to several concerning factors: 1) PCI was associated with the least improvement in processes of care, specifically time to reperfusion (source: Jencks et al., Burwen et al.); 2) ineffective transfer protocols resulting in reperfusion delays; 3) proliferation of cardiac intervention labs resulting in lower average volume per hospital and failure to adhere to established consensus guidelines that recommend >75 PCI procedures per physician per year and >200 PCI procedures per hospital per year.

In contrast, AMI performance across the nation was associated with the least disparity. This may represent a quality gap that is narrowing and may be due to several encouraging factors: 1) the treatment of AMI has seen the most significant attention to and improvements in the processes of care of any diagnosis or procedure affecting Medicare beneficiaries; 2) considerable consensus among physicians on the recommended management guidelines for AMI; 3) influential and important media coverage around various aspects of AMI has increased patient awareness; 4) advances in technology supporting the identification and treatment of AMI.

Despite previous studies (Vaughan-Sarrazin et al. JAMA. 2002;288: 1859-1866) that have demonstrated a correlation between the presence of restrictions on the development of new cardiac surgery centers (CON laws) and outcomes, this study did not find this to be a differentiating characteristic of quality. This may be due to the fact that more states have shifted away from continuous CON to intermittent or no CON regulation since 1999. At that time, the majority of states (27) had the toughest CON laws. Today, this number has decreased to six. The study by Vaughan-Sarrazin et al. found the best outcomes for CABG surgery were located in those states with the toughest CON laws. This finding warrants further research to evaluate the correlation with outcomes performance subsequent to the 1999 laws.

It is important to note that four of the seven states that have hospital-specific public profiling of CABG surgery performance rank among the best for this procedure (Maryland, has the newest public reporting system and ranked 25th; Texas and California also recently launched their sites in late 2001 and 2002 and ranked 42nd and 29th , respectively). This is an extremely important finding because this association may have national public policy and governmental implications. Several studies have documented both the effectiveness and lack of effectiveness of using public profiling to improve quality. Some researchers believe that publication of comparative performance data does not produce improved clinical performance. However, our finding suggests that it may and that it might be the very stimulus that will promote appropriate investment in and demand for quality. As consumerism of healthcare grows and public profiling becomes the norm, it is hypothesized that improvement in clinical quality will occur as an adaptation to demand in the marketplace. More research is needed to affirm and measure this potentially causal relationship.

Although important advances have been and continue to be made, these stark national and state-level trends highlight continued quality gaps and underscore the urgency to understand the systems that create the outcomes, remove road blocks to successful change, hold providers and payers accountable, and to identify methods to leap forward and close the quality chasm by reducing preventable morbidity and mortality. With the support and leadership of credible and experienced quality improvement organizations and the continued improvement in processes of care, it is hopeful that the adoption of diffusion of best practices will occur more quickly than it has historically, resulting in decreased variation in and improved outcomes for all patients. Physicians have been and will continue to be charged with leading this needed paradigm shift. As multitudes of organizations, from payers to employers, choose to pay for quality, physicians will have incentives to lead successful quality improvement in their hospitals.

Limitations of the Data Models

These models are limited by the following factors:

- Cases may have been coded incorrectly or incompletely by the hospital.
- The models can only account for risk factors that are coded into the billing data – if a particular risk factor was not coded into the billing data, such as a patient's socioeconomic status and health behavior, then it was not accounted for with these models.
- Although Health Grades, Inc. has taken steps to carefully compile these data using its proprietary methodology, no techniques are infallible, and therefore some information may be missing, outdated, or incorrect.

Second, although we discuss the general consistencies between the state performance ranks of the Jencks' study on processes of care and our study on outcomes, we did not perform a correlation test due to different methodologies and lack of key data from the Jencks' study. We evaluated the outcomes performance on two procedures (CABG, PCI) and did not evaluate two preventive measures that were included in the Jencks' study. Consequently, we cannot draw any conclusions with regards to the relationship between processes and outcomes of Medicare beneficiaries.

Third, ranking ties occurred because of taking the average of rank numbers, which were represented as whole numbers. This created a few ties, which were then broken by evaluating state volume; the higher-volume state was better ranked. Although there have been many studies correlating volume and outcome, we acknowledge the limitations of this method to accurately differentiate performance between two states with the same average rank.

EXHIBIT A

Patient Cohorts and Related ICD-9-CM Codes

Patient Cohort	ICD-9-CM Procedure/Diagnosis Codes and Criteria
Coronary Bypass Surgery	Principal Procedure Codes: 36.10 – 36.16 or 36.19, excluding patients in 35.2* or 35.1*
Valve Replacement Surgery	Procedure Codes: 35.20 – 35.28, excluding patients with ICD-9-CM diagnosis of 441.2
Percutaneous Cardiac Intervention	Procedure Codes: 36.01, 36.02, 36.05, 36.06, 36.07, 36.09
Acute Myocardial Infarction	Principal Diagnoses: 410.00 – 410.91 (where the fifth digit is "1") or DRGs: 121, 122, or 123
Heart Failure	Principal Diagnoses: 428.0 – 428.9, 398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93
Atrial Fibrillation	Principal Diagnosis: 427.31
Total Hip Replacement – Primary	Principal Procedure Code: 81.51, excluding patients with procedure 81.53
Total Knee Replacement – Primary	Principal Procedure Code: 81.54
Back and Neck Surgery (except Spinal Fusion)	Principal Procedure Codes: 03.09, 03.53, 80.50, 80.51, 80.59, excluding patients with procedures 81.00 – 81.09, 81.61, 81.3* or patients with diagnosis 722.83 and procedure 03.02, or patients with diagnosis V45.4
Back and Neck Surgery (Spinal Fusion)	Principal Procedure Codes: 81.00 - 81.08, 81.61, excluding patients with procedure 81.3* or diagnosis V45.4
Hip Fracture Repair (Open Reduction Internal Fixation)	Principal Procedure Codes: 79.25, 79.35
Partial Hip Replacement	Principal Procedure Code: 81.52
Stroke	Principal Diagnoses: 430, 431, 432.0, 432.1, 432.9, 433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434.01, 434.11, 434.91, 436
Aspiration Pneumonia	Principal Diagnoses: 507.0, 507.1, 507.8

Patient Cohort	ICD-9-CM Procedure/Diagnosis Codes and Criteria
Respiratory Infection except Aspiration Pneumonia and Tuberculosis	Principal Diagnoses: 003.22, 006.4, 020.3 – 020.5, 021.2, 022.1, 031.0, 039.1, 052.1, 055.1, 073.0, 095.1, 112.4, 114.0, 114.4, 114.5, 115.05, 115.15, 115.95, 121.2, 122.1, 130.4, 136.3, 482.0, 482.1, 482.4*, 482.81 – 482.84, 482.89, 484.1, 484.3, 484.5 – 484.8, 510.0, 510.9, 511.1, 513.0, 513.1, 519.2
Chronic Obstructive Pulmonary Disease	Principal Diagnoses: 491.1, 491.20, 491.21, 491.8 – 491.9, 492.8, 493.20, 493.21, 493.22, 494, 496
Community Acquired Pneumonia	Principal Diagnoses: 480.0 – 480.2, 480.8, 480.9, 481, 482.2, 482.30 – 482.32, 482.39, 482.40, 482.41, 482.9, 483.0, 483.1, 483.8, 485, 486, 487.0
Resection and Replacement of Abdominal Aorta	Procedure Code: 38.44 or 39.71 and in DRGs: 110, 111, 483. Excluding patients with procedures: 36.1*, 38.08, 38.16, 38.18, 38.36, 39.24, 39.25, 39.29, 39.50, 39.59, or with diagnosis 441.02
Carotid Endarterectomy	Principal Procedure Code: 38.12, 39.72, excluding patients with procedures: 36.1*, 38.08, 38.16, 38.18, 38.36, 39.24, 39.25, 39.29, 39.50, 39.59
Peripheral Vascular Bypass	Procedure Code: 39.29 and principal diagnoses of any of the following: 250.60, 250.61, 250.62, 250.63, 250.70, 250.71, 250.72, 250.73, 250.80, 250.81, 250.82, 250.83, 440.20, 440.21, 440.22, 440.23, 440.24, 440.29, 440.30, 440.32, 442.2, 442.3, 443.89, 443.9, 444.22, 444.81, 447.1, 681.10, 682.2, 682.4, 682.6, 682.7, 686.8, 707.10, 707.12, 707.13, 707.14, 707.15, 707.19, 707.8, 730.06, 730.07, 730.16, 730.17, 730.18, 730.26, 730.27, 785.4, 902.53, 904.0, 904.41, 904.7, 904.8 excluding patients with procedures: 39.25, 39.49, or with diagnosis 440.31
Peripheral Vascular Interventional Procedures (Angioplasty and/or Stent)	Procedure Code: 00.55, 39.50, or 39.90 and principal diagnoses of any of the following: 250.60, 250.61, 250.62, 250.63, 250.70, 250.71, 250.72, 250.73, 250.80, 250.81, 250.82, 250.83, 403.91, 405.01, 405.91, 440.1, 440.20, 440.21, 440.22, 440.23, 440.24, 440.29, 440.30, 440.31, 440.32, 443.89, 443.9, 444.22, 444.81, 447.1, 447.3, 593.81, 681.10, 682.6, 682.7, 707.10, 707.12, 707.13, 707.14, 707.15, 707.19, 730.06, 730.07, 730.17, 730.26, 730.27, 785.4, 996.1, 996.62, 996.73, 996.74, 997.60, 997.62, 997.69, excluding patients with procedure 39.49

Patient Cohort	ICD-9-CM Procedure/Diagnosis Codes and Criteria
Bowel Obstruction	Principal Diagnoses: 277.01, 532._1 (where the fourth digit can be anything, and the fifth digit is "1"), 534._1 (where the fourth digit can be anything, and the fifth digit is "1"), 537.2, 537.3, 550.10 – 550.13, 552.00 – 552.03, 552.1, 552.20 – 552.29, 552.8, 552.9, 557.0, 560.0 – 560.9, 751.1, 751.2, 777.1, 777.2, 777.4, 936, 937
Cholecystectomy	Principal Procedure Codes: 512.1 – 512.4
GI Bleed	Principal Diagnoses: 456.0, 456.20, 530.7, 530.82, 531.00, 531.01, 531.20, 531.21, 531.40, 531.41, 531.60, 531.61, 532.00, 532.01, 532.20, 532.21, 532.40, 532.41, 532.60, 532.61, 533.00, 533.01, 533.20, 533.21, 533.40, 533.41, 533.60, 533.61, 534.0, 534.00, 534.01, 534.20, 534.21, 534.40, 534.41, 534.60, 534.61, 535._1 (where the fourth digit can be anything, and the fifth digit is "1"), 537.83, 537.84, 562.02, 562.03, 562.12, 562.13, 569.3, 569.85, 569.86, 578, 578.1, 578.9, 751.0, 772.4
Prostatectomy	Principal Procedure Codes: 602.1 – 606.9
Sepsis	Principal Diagnoses: 003.1, 022.3, 027.0, 036.2, 036.3, 038.0, 038.10, 038.11, 038.19, 038.2, 038.3, 038.40 – 038.49, 038.8, 038.9, 054.5, 771.81, 785.59, 995.91 – 995.94, 998.59, 999.3