DRAFT

Applying a Group-based Trajectories Methodology to Measures of Patient Safety

April 27, 2009

J. Scott Ashwood
H. John Heinz III College
Carnegie Mellon University

Committee Members:
Martin Gaynor, H. John Heinz III College
Amelia Haviland, RAND Corporation and H. John Heinz III College
Daniel Nagin, H. John Heinz III College

1. Introduction

In 2000, the Institute of Medicine released a report on patient safety that included an estimate that adverse events occur in 2.9 to 3.7 percent of hospitalizations (Institute of Medicine, 2000). Every year, 50,000 to 100,000 people die from preventable adverse events that occur during hospital stays in the United States. The estimated cost of these events is $17 billion (Institute of Medicine, 2000). Since the IOM report, measuring and reporting patient safety has become an important component of healthcare policy and research.

Despite the attention and funding that studies of patient safety have received, there are still gaps in our understanding of the scope of the problem and the causes (Leape and Berwick 2005; Zhan et al 2005). These gaps are due in part to the difficulty in identifying medical errors using existing data as well as the reliance on analytic and reporting methods that may be misleading. Government agencies report patient safety levels and trends in the U.S. (CDC 2004; AHRQ 2007c), and public and private organizations provide annual grades and rankings of hospitals in the nation based on their patient safety practices and records (The Leapfrog Group 2008; Hospital Compare 2009). These reports and rankings are either cross-sectional or present an average trend, both methods that can miss important information about the distribution of patient safety rates in the population of hospitals.

Cross-sectional approaches mask trends in patient safety rates by focusing on a single year. Time series approaches can identify a single average trend but may miss multiple, distinct trends in patient safety that exist within a population of hospitals. A key assumption of average trend analysis is that there is a single patient safety level or trend that all hospitals follow. Individual hospitals may be higher or lower than the average, but they are all, in the end, following the same trajectory. Group-based trajectory models reveal important aspects of the distribution of patient safety rates that cross-sectional and simple time series models miss. I use group-based trajectory models on the AHRQ Patient Safety Indicators (PSI's) and find evidence that there are, in fact, multiple trends in patient safety followed by groups of hospitals.

When I apply the group-based trajectories method to individual PSI's, I find evidence that there are large groups of hospitals that follow trajectories of low, constant rates over time and smaller groups that follow trajectories of constant, high rates. There are also groups of hospitals that appear to be improving or getting worse on individual indicators. This suggests that not only is there an important time component to patient safety rates, but that there is more than one trend for a given patient safety measure.

In addition to looking at individual PSI's, I looked across PSI's for consistent patterns in safety. This is a way to form a comprehensive picture of patient safety that combines the dimensions measured by the individual indicators. If there are patterns that emerge when all of the PSI's are considered together, then policy makers can identify hospitals that are performing well overall or identify possible problems in the population as a whole. A few hospitals appear to be doing well on all indicators, and a few appear to be doing poorly on all indicators, but most hospitals are doing well on some and poorly on others. When I group PSI's together by type of care, I find that about one third of the hospitals I study are experiencing an increase in the surgical PSI's.

2. Background

An important outcome of hospital care is patient safety. In this case, patient safety is defined as the absence or low probability of medical errors. Until the 2000 report by the Institute of Medicine, patient safety was largely overlooked. In response to this report, the U.S. Congress directed the Agency for Healthcare Research and Quality (AHRQ) to fund research into improving patient safety. Measuring and reporting patient safety has become an important component of healthcare research and policy (Leape and Berwick 2005).

Researchers characterized the impact of preventable adverse events as well as looked for characteristics of healthcare organizations that contribute to higher error rates. For example, Chunliu Zhan and Marlene Miller found that discharges with a patient safety event are longer, cost more, and have higher mortality (Zhan and Miller, 2003). Eric Thomas and colleagues found that discharges from for-profit hospitals had slightly higher odds of having a preventable adverse event and that government-owned and minor teaching hospitals had the highest odds (Thomas, et al., 2000). A large body of research has looked for systems explanations for errors, borrowing from research into errors in other complex industries (Bogner 1994).

In addition to use by researchers, government and private organizations have defined measures and used them to assess the level of and trends in patient safety in the U.S. They have also used them to classify and rank hospitals. The Agency for Healthcare Research and Quality (AHRQ) releases annual reports on quality in the U.S. healthcare system that include a chapter on patient safety. In 2007 they report trends in six patient safety measures, some of which appear to be improving (e.g. deaths following complications of care), some that are getting worse (e.g. central venous catheter placements with bloodstream infections), and some that are relatively constant (e.g. postoperative complications) (AHRQ 2007c).

There are several databases tracking patient safety in the U.S. that are maintained by public and private organizations (Zhan et al. 2005; Faber et al 2009; Greenberg et al. 2009). Public reporting of healthcare provider quality has grown at the state and national level. Individual hospitals are classified and ranked by both public and private groups. For example, the Leapfrog Group provides annual rankings of hospitals based on patient safety measures. Their 2008 list includes the top 35 hospitals that meet their criteria. Leapfrog believes that by highlighting the top performers, they are “helping drive significant improvements in hospital care” (The Leapfrog Group, 2008, p.3).

These rankings were initially designed to provide consumers with information about the quality of hospitals. The idea behind public reporting is that better informed consumers will choose higher quality hospitals. That in turn will stimulate quality improvement as hospitals compete for patients. However, the evidence on the extent to which consumers use the publicly reported information is mixed (Faber et al 2009). Rather than rely on public reporting and market competition alone to drive quality improvement, the Centers for Medicare and Medicaid Services (CMS) has stated that it will stop paying for hospital stays that include medical errors. They also are developing Pay-for-Performance (P4P) policies that reward the highest quality hospitals and punish the lowest, where highest and lowest are identified on an annual basis (Khan et al 2006).

Identifying the best performers or worst performers in a single year is misleading. Hospitals can move into and out of these groups from year to year due to random fluctuation in their rates. It is more useful to identify hospital performance over time. The average trends reported by AHRQ account for changes over time in patient safety rates, but they assume a single trend. If there are distinct trends within the population, then the average trend will mask them. Group-based trajectories modeling can help to identify and model multiple trends.

3. Methodology

The group-based trajectories modeling technique (Nagin, 2005) takes advantage of longitudinal encounter data for individual hospitals, to determine whether groups of hospitals appear to follow distinct trends in safety outcomes. Trajectory modeling supports both an estimate for the number of distinct groups (or clusters) of hospitals implied by the data, as well as an estimate for the actual trajectory in outcomes for each group. As a result, groups of hospitals with stable rates of outcomes can be distinguished from those with either increasing or decreasing rates.

3.1 Maximum Likelihood

The group-based trajectory methodology is based on maximum likelihood estimation (MLE): A likelihood function is specified and the values of the parameters are chosen to maximize the likelihood function. The likelihood function includes parameters for the time trend, an estimate of the probability of belonging to each group, and parameter estimates for predictors of group membership. It has the following form:

$$L = \Pi\ P(Y_i), \qquad (1)$$

where

$$P(Y_i) = \Sigma\ \pi_j P^j(Y_i) \qquad (2)$$

is the unconditional probability of observing hospital i's sequence of PSI rates over time. In equation 2, π_j is the estimated probability of membership in group j, and $P^j(Y_i)$ is the probability of observing an individual hospital's sequence of outcomes conditioned on membership in group j. The probability of group membership is estimated by

$$\pi_j = e^{\theta j} / \Sigma e^{\theta k} \qquad (3)$$

This specification allows for the addition of individual level characteristics in the estimate of group membership. For example, the ownership type of a hospital may be related to the group membership. This form for estimating the probability of group membership allows for testing whether for-profit hospitals may be more likely to belong to a given group than non-profit, for example.

For a given group, the model assumes independence of observations over time, so

$$P^j(Y_i) = \Pi p^j(y_{it}), \qquad (4)$$

where $p^j(y_{it})$ is the probability distribution function of y_{it} given membership in group j. This is the probability of observing hospital i's sequence of PSI rates over time conditional on membership in group j. This specification assumes that the distribution of PSI rates for hospital i in period t is independent from the rate in period t-1. This may seem like an implausible assumption if this were the unconditional probability; however, this is the probability conditional on membership in group j. Group j's estimated trajectory has a structure over time, so the dependence among observations is captured at the group level. Individual variation around that group trend is assumed to be independent from year to year for a given hospital.

The specific form of $p^j(y_{it})$ depends on the type of outcome that is modeled. Three forms are most common: normal/censored normal, Poisson, and binary logit. Since I am modeling the rate per 1,000 of events, a continuous measure, I am assuming that the PSI rates for each hospital are best approximated with a normal model.

The conditional probability for each hospital's sequence of PSI rates can be specified as:

$$p^j(y_{it}) = 1/\sigma * \phi\ ((y_{it} - \beta^j t)/\sigma), \qquad (5)$$

where ϕ is the normal density function for a random variable with mean $\beta^j t$ and standard deviation σ, and y_{it} is hospital i's observed PSI rate at time t. $\beta^j t$ is the relationship between time and the PSI rate for group j. The SAS procedure I used to estimate the models allows for up to a cubic relationship between y_{it} and time.

3.2 Group Choice

Note that in equation 2, the number of groups is given. That requires an appropriate choice of the number of groups. For each PSI that I studied, I first chose the number of groups through a combination of statistical scoring (i.e., Bayesian Information Criterion score), direct observation of trajectory plots, and application of subject-matter knowledge (i.e., regarding safety outcomes in hospitals).

The Bayesian Information Criterion (BIC) score, derived from the maximized likelihood score, is recommended as one device for helping to select the appropriate number of groups to use in trajectory models (Nagin, 2005).

$$\text{BIC} = \log(L) - 0.5k \log(N), \qquad (6)$$

where L is the model's maximized likelihood, k is the total number of parameters in the model, and N is the sample size. The number of parameters, k, includes the polynomial order used to estimate each trajectory as well as the number of groups. The BIC score increases as the maximized log likelihood increases, but imposes a penalty for increasing the number of parameters. Recall that a separate set of parameters is included for each group that is added to equation 2 above. That means the BIC score will eventually decrease as the number of groups increases.

Selecting the number of groups that corresponds to the highest BIC score is one way to choose the number of groups. However, adding more groups to improve the BIC score may not improve the understanding of the patterns of patient safety in the data. It is important that the choice is meaningful in the context of the problem. For example, a trajectory model with two groups of hospitals might identify a large group that has a low, constant rate of adverse safety outcomes, and a second group that has a high and increasing rate. A trajectory model with three groups might have a higher statistical BIC score than does the two-group model. Nevertheless, if the third group is simply identified by splitting the low, constant group into two pieces, then little insight has been gained, especially if the new third group is estimated to have very few members. The important conclusion is captured by the two group model: namely, the fact that some hospitals have a low and stable rate of adverse outcomes, while others have an increasing rate of outcomes on the same measure. In sum, relying solely on the BIC score to estimate the appropriate number of groups in trajectory modeling may lead to choosing too many groups to be useful.

To choose the number of groups for each PSI, I started by generating ten models for each with the number of groups increasing from one to ten. I then compared the BIC scores from these ten models and focused on the models with higher BIC scores. Next, I examined the plots of the trajectories for the corresponding models and chose the number of groups that identified meaningful patterns in the data without including superfluous trajectories. I used a SAS procedure, PROC TRAJ, to apply the group-based trajectories models (Jones et al 2001).

3.3 Trajectory Shape

In addition to choosing the number of groups, the form of each trajectory must be estimated. That is, the pattern of PSI rates over time for each group must be approximated ($\beta^j t$ from equation 5). Nagin recommends choosing the number of groups first and then estimating the trajectory form. The form is important because it describes whether and how PSI rates change over time. After choosing the number of groups for each PSI, I used a combination of viewing the observed group averages over time and looking at the significance of the coefficients for each order of time. I chose the trajectory form that was supported by statistically significant coefficient estimates and that followed the average observed rates.

4. Outcomes

4.1 Patient Safety

In order to apply the group-based trajectories methodology, I need a measure that can be computed for the same hospital over multiple years. There are three broad categories of patient safety measures: structural measures, process measures, and outcome measures (Zhan et.al. 2005). Structural measures include the design of the hospital environment and the staffing levels. Process measures include errors in the process of care (e.g. diagnosis mistakes, prescription errors). Outcome measures are measures of actual injuries or adverse events. I focus on outcome measures in this paper in part because they are commonly analyzed in the patient safety research literature and because it is easier to gather and compute outcome measures than it is to gather the other types of measures (Romano et al. 2003; Zhan et al. 2005).

There are several outcome measures to choose from (Zhan et al. 2005; Greenberg et al. 2009). The Sentinel Events statistics reported by the Joint Commission and Nosocomial Infections Surveillance Measures collected and maintained by the Centers for Disease Control are reported for all hospitals in their samples combined either as annual summaries or as trends. The adverse drug events statistics based on the Medical Expenditure Panel Survey (MEPS) data and the Patient Safety Monitoring Program (MPSMS) recently implemented by the Centers for Medicare and Medicaid Services are based on nationally representative samples of patients, not hospitals. The Utah-Missouri measures and the AHRQ Patient Safety Indicators (PSI's), which are based on hospital discharge data can be computed for individual hospitals over time. I chose the AHRQ PSI's because they are focused more directly on patients at risk for each type of event and because they are risk adjusted. The Utah-Missouri measures consider all surgical or all medical discharges as eligible for events. The AHRQ PSI's define a unique set of eligible discharges for each type of event (Greenberg et al. 2009, p 743).

4.2 AHRQ Patient Safety Indicators

The AHRQ Patient Safety Indicators (PSI's) are designed to "screen for problems that patients experience as a result of exposure to the healthcare system and that are likely amenable to prevention by changes at the system or provider level." (AHRQ, 2007a, p.3) They rely primarily on diagnosis codes (ICD-9 CM classifications) to identify possible adverse events. They were developed to be used with hospital discharge abstract data, the kind of data that all states routinely collect from hospitals. Thus, the indicators can be applied to data from multiple states to define regional and national benchmark levels (AHRQ, 2007a, p.6).

AHRQ originally developed a set of patient safety indicator measures in the 1990's, called the Healthcare Cost and Utilization Project (HCUP) Quality Indicators. After advances in both quality measurement and risk adjustment, AHRQ asked the UCSF-Stanford Evidence-Based Practice Center to develop a new set of measures. The Center followed an extensive process that included a literature review, expert panel review, and empirical analysis. In the end, the UCSF-Stanford team came up with the current set of indicators (see Table 3 below).

Like the original HCUP indicators, these were designed to be used with routine discharge data. This has advantages and disadvantages. The main advantage is that the administrative data are easy to collect and analyze, and they cover large populations for continuous periods of time (Romano et al. 2003; Zhan and Miller 2003). Hospitals generate discharge records every time a patient is discharged in order to bill the insurance company or government agency that pays for the patient's coverage. In addition, many states require hospitals to submit discharge records for all discharges regardless of who pays (or does not pay in the case of some uninsured).

The disadvantages are that some information about each hospital stay may not be captured in discharge records or may be coded incorrectly, and there is limited information to use for patient level risk adjustment (Romano et al. 2003; Zhan and Miller 2003; Grobman et al. 2006). Discharge records are designed for reimbursement for services, not identifying adverse events. A major disadvantage of most discharge data is the inability to identify when a condition is present on admission versus occurring during the hospitalization. Some states, including California, have present on admission (POA) flags that greatly improve the ability of the PSI's to flag events that occur during hospitalization. Using the primary diagnosis alone tends to lead to overestimates of patient safety events (Bahl et al. 2008).

As a result of the limitations of administrative data, the quality indicators analyzed here do not provide a definitive measure of patient safety. They can only be used "to provide *indicators* of health care quality that can serve as the starting point for further investigation." (AHRQ, 2007a, p.12). However, they are well suited to be used as screening tools for individual hospitals or to define trends or benchmarks in populations of hospitals (Romano et al. 2003; Zhan and Miller 2003).

<u>Table 1. AHRQ Inpatient Quality Indicators</u>

1 - Complications of Anesthesia
2 - Death in Low-Mortality DRGs*
3 - Decubitus Ulcer
4 - Failure to Rescue
5 - Foreign Body Left During Procedure*
6 - Iatrogenic Pneumothorax
7 - Selected Infections Due to Medical Care
8 - Postoperative Hip Fracture
9 - Postoperative Hemorrhage or Hematoma
10 - Postoperative Physiologic and Metabolic Derangements*
11 - Postoperative Respiratory Failure*
12 - Postoperative Pulmonary Embolism or Deep Vein Thrombosis
13 - Postoperative Sepsis*
14 - Postoperative Wound Dehiscence
15 - Accidental Puncture or Laceration
16 - Transfusion Reaction*
17 - Birth Trauma – Injury to Neonate
18 - Obstetric Trauma – Vaginal with Instrument
19 - Obstetric Trauma – Vaginal without Instrument
20 - Obstetric Trauma – Cesarean Delivery*

* Unable to compute rates using CA discharge data

4.3 Computing Raw Rates

There are 20 hospital level measures (Table 1) that cover a variety of adverse events from the relatively minor (decubitus ulcer) to fatal (death in low-mortality DRG's). AHRQ has created a set of programs for the statistical package SAS that can generate hospital level rates for each of the 20 PSI's. The software for computing the measures has gone through several revisions as AHRQ has made changes to the methods for identifying adverse events as well as to the risk adjustment process. One of the most important revisions to the process involves adding consideration of whether a condition is present on admission (added for the March, 2007 release of Version 3.1). In general, the primary diagnosis code is supposed to represent the reason for admission. The PSI process ignores the primary diagnosis code under the assumption that it is present on admission and not likely to have occurred during the stay. Early versions of the software look through the secondary diagnosis fields in order to identify possible diagnoses that occur during the stay. Some states (California among them) include "present on admission" flags for each diagnosis in their discharge data. It is possible that the primary diagnosis code is not present on admission and should not be excluded. It is also possible that diagnoses in the secondary diagnosis fields are present on admission but not the primary reason for admission. Early versions of the PSI software would mistakenly identify these as adverse events. I use version 3.1 of the AHRQ software to generate hospital level PSI rates and am therefore using the POA flags in the California discharge data.

To calculate the PSI's, each discharge record is examined for eligibility for each of the indicators (added to the denominator) as well as whether it appears one of the events actually occurred (added to the numerator). Each PSI has its own set of eligibility and event requirements. For example, PSI 3 (decubitus ulcer) identifies all cases that have a diagnosis code indicating decubitus ulcer. Cases that have a length of stay less than 5 days or are admitted from a long term care facility are excluded since decubitus ulcers generally take longer than 5 days to develop and it is possible that an ulcer developed at the long term care facility rather than the hospital.

A raw rate for each indicator is computed by summing the eligible discharges and events for each hospital in each year and then dividing the number of events by the number of eligible discharges.

4.4 Risk Adjustment Process

Once raw PSI rates are computed for each hospital, they can be risk-adjusted to account for the risk profile of each hospital's patients. Some hospitals see sicker patients on average than others. Those hospitals might have higher rates of accidents than hospitals with healthier patients but may still be doing well compared to hospitals with similar, sick populations. In order to compare rates among hospitals, AHRQ recommends risk adjustment (AHRQ 2007a, p65).

Risk adjustment is based on a combination of demographic and comorbidity factors. The PSI software includes coefficient estimates for each PSI that correspond to age, gender, an interaction between age and gender, diagnosis related group (DRG), and several comorbidities (AHRQ 2007a, p65). These coefficients are derived from a logistic regression of patient demographics and disease profiles on the occurrence of an event for a national sample of hospital discharges (the AHRQ State Inpatient Data). Separate logistic models were run on each PSI.[1] Once the coefficient estimates are applied to the individual discharges, a predicted value is generated for each discharge that reflects the likelihood that an event would have occurred if that discharge had been from an average U.S. hospital. All of the predicted values are summed for the hospital. This is the hospital's predicted number of total events.

In addition to the coefficients, the PSI software includes average national rates for each PSI.[2] The hospital's predicted rate is compared to this national average. If the hospital's predicted rate is higher than the national average, then the hospital has a more severe case mix than average, and their observed rate is adjusted downward. If their predicted rate is

[1] Furthermore, there are two sets of logistic models, one for events that are defined without reference to present on admission (POA) flags and one for events whose definition does consider the POA flags. The POA flags have a large enough effect on the number of events that are identified that they merit a separate risk adjustment process. Since California and New York were the first states to include POA flags, the set of POA logistic models were run on a sample of hospitals from these two states alone. The non-POA logistic models were run on the full national sample.

[2] Once again, there are separate averages for POA and non-POA definitions.

lower, then they have a less severe case mix than average, and their observed rate is adjusted upward.[3] This new rate is the risk-adjusted rate and represents "the rate the provider would have if it had the same case-mix as the reference population given the provider's actual performance." (AHRQ, 2007a, p.65)

Finally, the risk-adjusted rates are smoothed by applying shrinkage factors in a process called multivariate signal extraction (AHRQ, 2007b, p.19). This is done in order to remove excess variation for smaller hospitals and for PSI's that are less reliable. These final, smoothed rates are used for the analyses reported in this paper.

5. Data

5.1 California Discharge Data

The primary data for this paper are the annual hospital discharge data for the state of California, 1997-2006. This public dataset is maintained by the California Office of Statewide Health Planning and Development (OSHPD). It includes a record for each hospital discharge during a single calendar year from a California licensed hospital. Table 2 includes the key fields present in the discharge data. The Inpatient Quality Indicators are designed to be used on discharge data and make use of many of the fields listed here.

Table 2. List of fields included in OSHPD Annual Discharge Data

Hospital Identification Number
Type of Care
Age in Years
Age (20 Age Categories)
Age (5 Age Categories)
Sex
Ethnicity
Race
Patient ZIP Code

[3] The formula used for the adjustment is (observed rate / predicted rate) * national average rate

Patient's County of Residence
Length of Stay
Admission Quarter
Admission Year
Source of Admission
Type of Admission
Disposition of Patient
Pre-hospital Care and Resuscitation (Do Not Resuscitate)
Expected Source of Payment - Payer Category
Expected Source of Payment - Payer Type of Coverage
Expected Source of Payment - Payer Plan Code Number
Total Charges
External Cause of Injury - Principal E-Code
External Cause of Injury - Other E-Codes
Major Diagnostic Category (MDC)
Diagnosis Related Group (DRG)
Principal Diagnosis
Condition Present at Admission (Principal Diagnosis)
Principal Procedure
Days from Admission to Principal Procedure
Other Diagnoses
Condition Present at Admission (Other Diagnoses)
Other Procedures
Days from Admission to Other Procedures

5.2 Hospital Selection

The discharge files include discharges for four types of facilities: general acute care, acute psychiatric, chemical dependency recovery, and psychiatric health. This paper focuses on general acute care hospitals. This is the most common type of hospital in the state and the one most people are familiar with. The PSI’s are not designed to capture errors associated with psychiatric or chemical dependency care. There are 568 facilities of all types in the discharge data over the 1997-2006 time period, of these 434 are general acute care hospitals. The general acute care hospitals account for about 91% of the discharges in the data. I selected hospitals that are in the discharge data every year from 1997-2006. I wanted to focus on hospitals with complete data over the period I am analyzing. This excludes hospitals that open or close during the time period I analyze as well as handful of hospitals that do both. This leaves 318 hospitals and about 85% of the total discharges in my final analytic file.

5.3 PSI Rate Calculation and Adjustment

Not all of the PSI's can be computed for the California discharge data, and it requires some manipulation before running the AHRQ PSI software. Three of the PSI's require identifying whether a hospitalization involved an elective surgery. The California discharge data does not have the information necessary to identify whether surgery is elective. The type of admission field present in the California discharge data identifies whether an admission was scheduled (at least 24 hours in advance) or not, but does not distinguish between elective and non-elective admissions. As a result, I am not able to compute rates for Postoperative physiologic and metabolic derangement (PSI 10), Postoperative respiratory failure (PSI 11), and Postoperative sepsis (PSI 13). In addition to these three PSI's, I am unable to compute trajectories for four PSI's that are extremely rare. They are Death in low mortality DRG's (PSI 2), Foreign body left during procedure (PSI 5), Transfusion reaction (PSI 16), and OB Trauma – C-section (PSI 20). There is not enough evidence for distinct groups of hospitals following different trajectories for these four PSI's. The seven indicators that I do not analyze are identified with an asterisk in table 1 above.

On average, 1.3 percent of eligible discharges appear to have had an adverse event in California 1997-2006. This is a lower rate than reported in the Institute of Medicine report (2.9-3.7 percent); however, the rates reported in the IOM report are based on earlier methods for identifying events that do not include consideration of conditions that are present on admission. Furthermore, the rates in the IOM report and those here are based on different time frames and different states.

The age categories in the 1997 and 1998 discharge data do not match the categories used for risk adjustment by the AHRQ software. The risk adjustment coefficients described above assume a continuous age field that gets coded into categories by the AHRQ software. I created a temporary, "continuous" age field by assigning age in years based on lowest age in the age category range. For 1999-2006, these age assignments matched the categories assigned by the software. For example, if the age category was 35-44, I set the age in years equal to 35.   the AHRQ software assigns any age in the 35-44 range to a single category. For 1997 and 1998, the lowest age categories in the discharge data crossed the categories used by the AHRQ software. In those cases, I randomly picked an age to match the AHRQ ranges.[4]

Some discharges have more than one PSI event. In those cases, I picked one for risk adjustment based on the population prevalence. I chose the single PSI that occurred the most in the entire population of discharges.

I used the risk adjustment coefficient estimates, national average rates, and smoothing parameters from the 2000 population. That means that all of the risk adjustment and smoothing I use is based on the population of hospitals in 2000. Since the California data has present on admission flags, I used the version of the risk adjustment and smoothing process that incorporates these flags. That means that the hospitals I analyzed are normed to the population of hospitals in California and New York in 2000.

Table 3 below describes the change in rates from raw to smoothed for the 13 PSI's.

Table 3. Comparison of raw PSI rates to smoothed rates

[4] For example, if the age category in the discharge data was 25-34, there are two possible categories in the AHRQ software that cover this range. I randomly assigned half of the records in this age range to each of the AHRQ categories.

PSI	Raw Rates (per 1000 eligible discharges) Mean	Median	IQR	Smoothed Rates (per 1000 eligible discharges) Mean	Median	IQR
Complications of anesthesia (PSI 1)	0.17	0.00	0.00	0.08	0.001	0.01
Decubitus ulcer (PSI 3)	2.93	0.63	3.02	2.56	0.98	2.95
Failure to rescue (PSI 4)	257.80	236.60	173.33	252.86	251.68	61.14
Iatrogenic pneumothorax (PSI 6)	0.31	0.00	0.49	0.35	0.16	0.53
Selected infections due to medical care (PSI 7)	1.30	0.38	1.29	1.06	0.57	1.46
Postoperative hip fracture (PSI 8)	0.09	0.00	0.00	0.04	0.00	0.00
Postoperative hemorrhage or hematoma (PSI 9)	1.93	1.44	2.87	1.96	1.77	1.98
Postoperative pulmonary embolism or deep vein thrombosis (PSI 12)	3.08	1.89	3.89	2.79	2.25	3.59
Postoperative wound dehiscence in abdominopelvic surgical patients (PSI 14)	1.60	0.00	1.85	1.96	0.85	2.13
Accidental puncture and laceration (PSI 15)	1.52	0.59	2.38	1.96	1.28	2.85
Birth trauma -- injury to neonate (PSI 17)	6.47	2.48	5.21	4.39	1.70	3.49
Obstetric trauma -- vaginal delivery with instrument (PSI 18)	167.35	154.64	122.57	144.52	134.80	100.47
Obstetric trauma -- vaginal delivery without instrument (PSI 19)	38.54	35.59	27.50	32.53	30.19	21.97

6. Results

I estimated trajectories for thirteen PSI's using the OSHPD data. In general, the results below provide evidence for the existence of distinct groups of hospitals following different trajectories of patient safety. This provides evidence that cross-sectional rankings of hospitals or reporting single average trends are misleading. I will focus on two PSI's to illustrate the strengths of the group-based trajectories approach over cross-sectional or average trend analysis. After estimating trajectories for the individual PSI's, I then looked for patterns across the PSI's.

6.1 Evidence Against Cross-sectional Rankings

The Premier Hospital Quality Incentive Demonstration and the Medicare Payment Advisory Commission have Pay-for-Performance (P4P) proposals that reward hospitals in the top two deciles of safety measures with financial bonuses (Khan et al 2008). In both proposals, the top deciles are identified each year based only on the scores for that year. The Leapfrog Group provides annual rankings for hospitals and has started a program designed to reward financially hospitals that are in the top decile of safety measures annually.

While none of these P4P programs rely specifically on the AHRQ PSI's, I looked for consistency in the definition of the top deciles from year to year for two PSI's. PSI 7 (Selected infections due to medical care) has been suggested as a canary measure due to it's correlation with other PSI rates within hospitals (Yu et al 2009). Of the 64 hospitals in the top two deciles for PSI 7 in 1997, only 4 remain in the top two deciles for all ten years I examined. 45 are in the top two deciles for at least five years. Two of the 64 "top" hospitals in 1997 are in the bottom decile at least once during the following nine years, and seven are in the bottom two deciles at least once. PSI 17 (Injury to neo-nate) has a more extreme pattern. Of the 54 hospitals in the top two deciles in 1997, none remain in these deciles for all ten years. Only 21 are in the top two deciles for at least 5 years. Eight wind up in the bottom decile at least once, and 10 wind up in the bottom two deciles at least once.

Figures 1 and 2 below show the plots of PSI 17 and PSI 3 (Decubitus Ulcer) rates over time for four hospitals. All four of these hospitals are in the top decile for the PSI at least once during the 1997-2006 time period. Clearly there is a lot of variation from year to year in ranking for individual hospitals. Cross-sectional ranking is sensitive to this year to year variation.

Figure 1: Rates of injury to neo-nate (PSI 17) for two hospitals 1997-2006

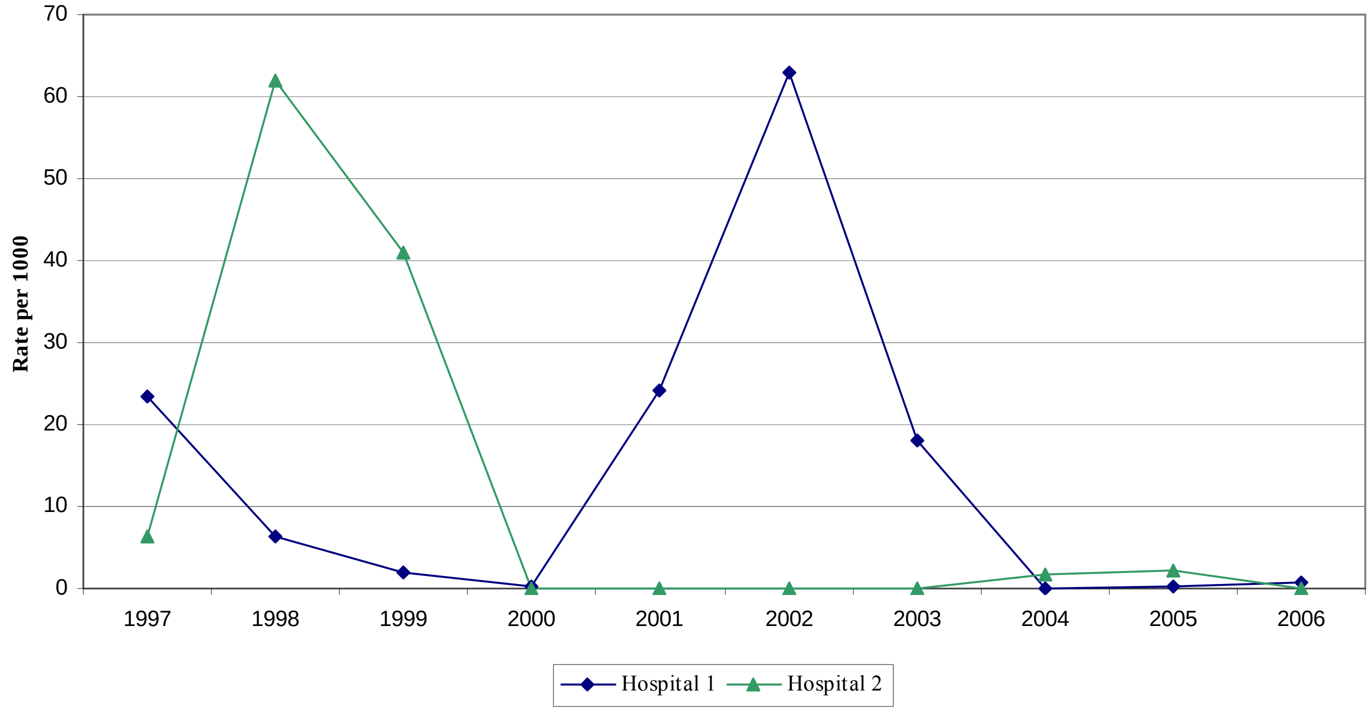
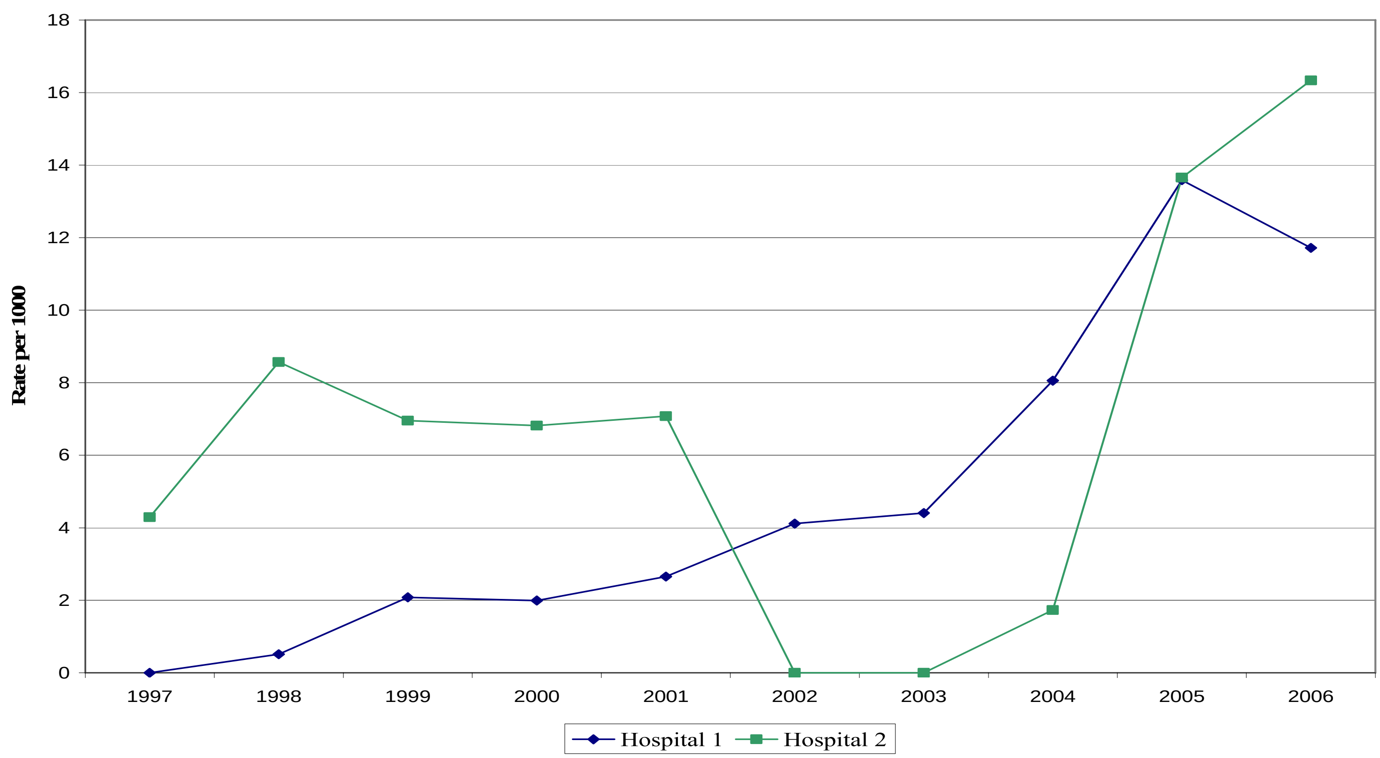


Figure 2: Rates of decubitus ulcer (PSI 3) for two hospitals 1997-2006



6.2 Evidence Against Single Trends

The annual National Healthcare Quality Reports (AHRQ 2007c) provide trends in selected patient safety measures and draw conclusions about whether patient safety is improving or getting worse over time. The assumption behind these reports is that there is a single trend in the measure that is described and that all hospitals are following this single trend. After applying the group-based trajectories methodology to the AHRQ PSI's , I find evidence against this assumption.

Figure 3 below contains the average trend in PSI 17 for California hospitals 1997-2006. The trend is improving dramatically on average from a rate of around 9 per 1,000 eligible discharges to a rate of just over 1 per 1,000. While this is good news, it gives no indication of the multiple trends that hospitals may actually be following.

Figure 3: Plot of average trend for "Injury to Neonate" (PSI 17)

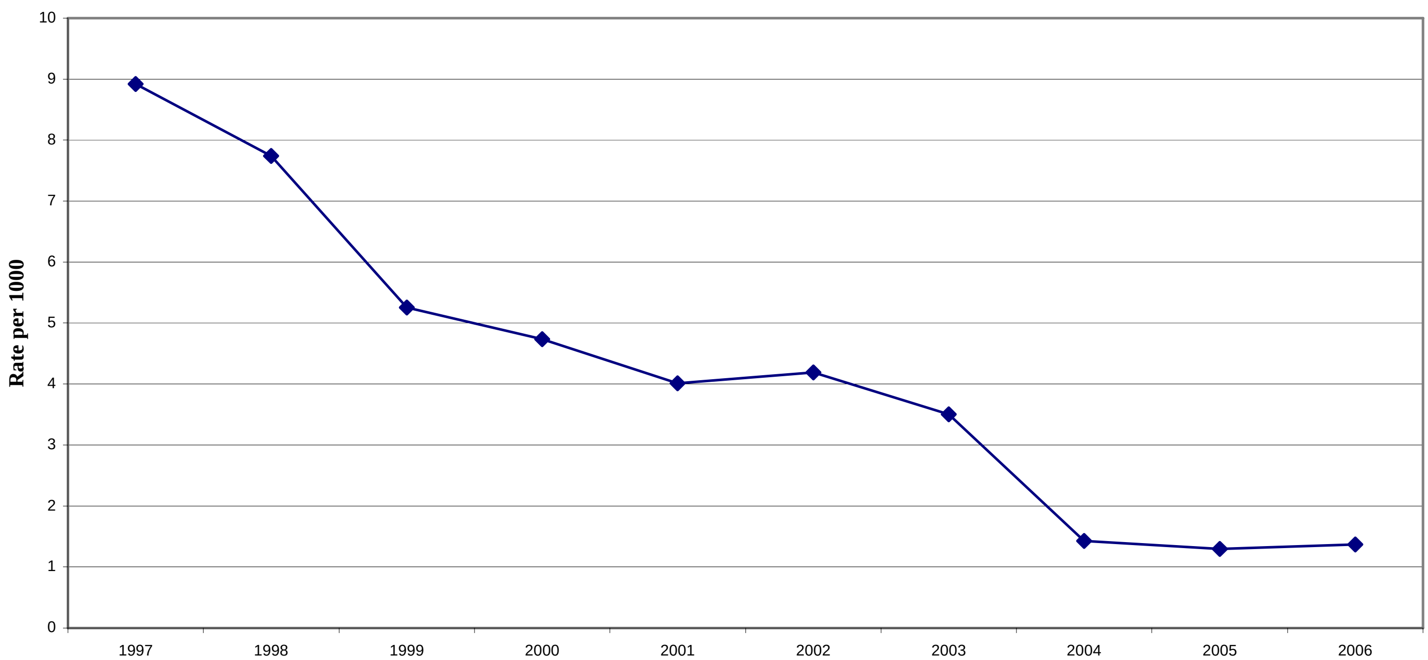
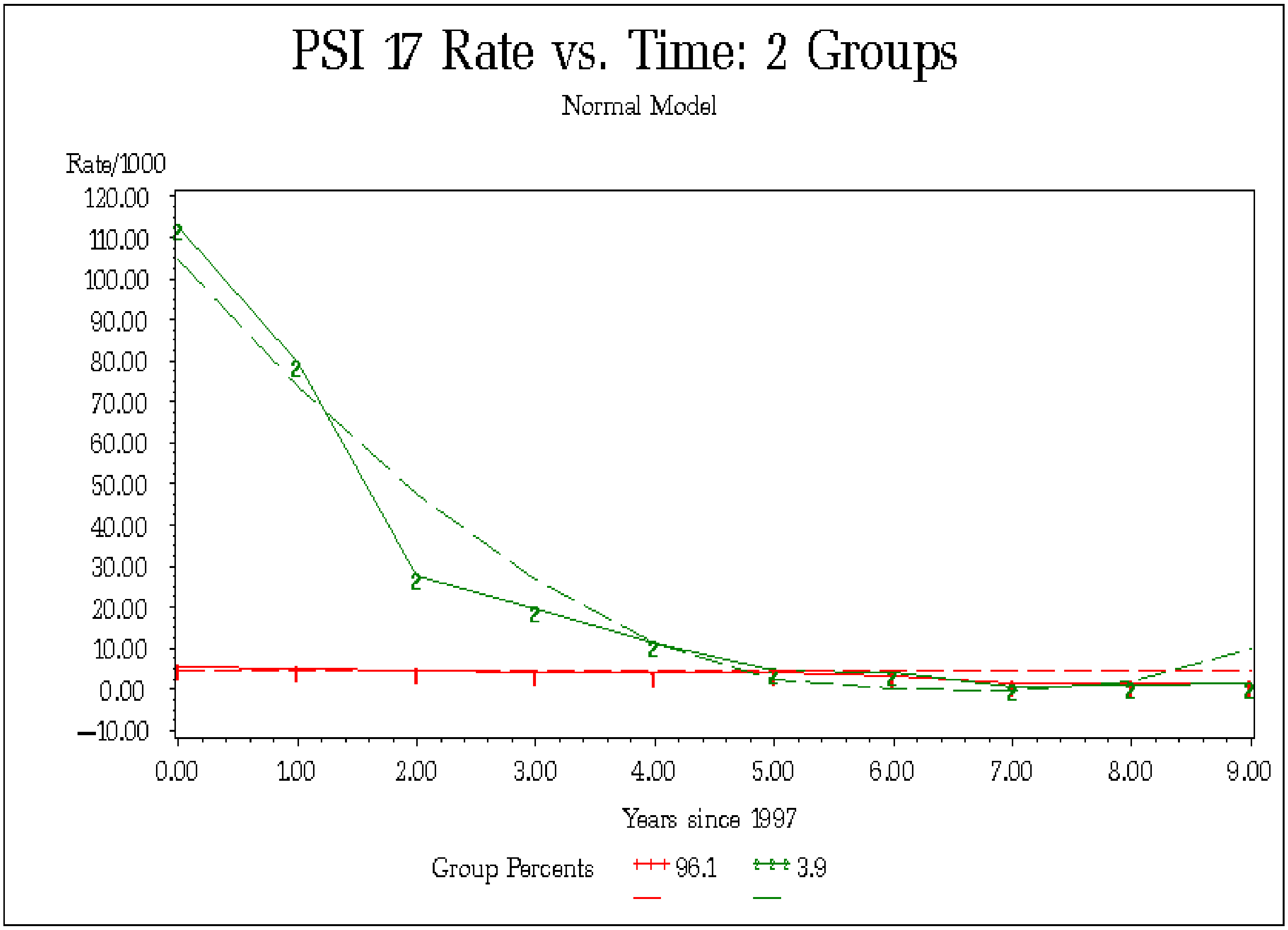


Figure 4, below, provides evidence that there appear to be two trajectories, a low, constant trajectory, and a smaller trajectory of improvement. This plot is derived from fitting a group-based trajectories model to PSI 17 rates.

Figure 4: Plot of 2 trajectories for "Injury to Neonate" (PSI 17)



The dashed lines in Figure 4 represent the estimated trajectories for each group, while the solid lines represent the observed average rates. The "group percents" listed at the bottom of the Figure show the estimated fraction of hospitals falling within each group. Most hospitals appear to have very low and constant rates throughout the period (group 1). The rate for this group is estimated to be around 3.4 per 1,000 with a 95% confidence interval of 3.1 to 3.8 per 1,000. The estimated group size is 96.1% of general acute care hospitals, with a 95% confidence interval of 93.6% to 98.6%. There is no statistical evidence of anything other than a constant rate (see Appendix). A small group of hospitals (group 2) follow a trend of even more dramatic improvement than the average depicted in Figure 3. The estimated percentage of hospitals in this group is 3.9 (1.4, 6.4), and the rate improves from about 104.6 per 1,000 (100.5, 108.7) to a rate that is just above the rate for the rest of the hospitals: 9.6 per 1,000 (5.0, 14.2).

The decreasing overall trend in "Injury to Neonate" in California plotted in figure 3 was driven primarily by the small percentage of hospitals in group 2 that started with relatively high rates: the overall trend masks the fact that most California hospitals experienced no change in rates of "Injury to Neonate" at all from 1997-2006.

Figure 5 contains the average rate of "Decubitus Ulcer" (PSI 3) for all California hospitals 1997-2006. There is little evidence of a dynamic trend here. Hospitals appear to experience a relatively constant rate of PSI 3.

Figure 5: Plot of Smoothed Average Trend for "Decubitis Ulcer" (PSI 3)

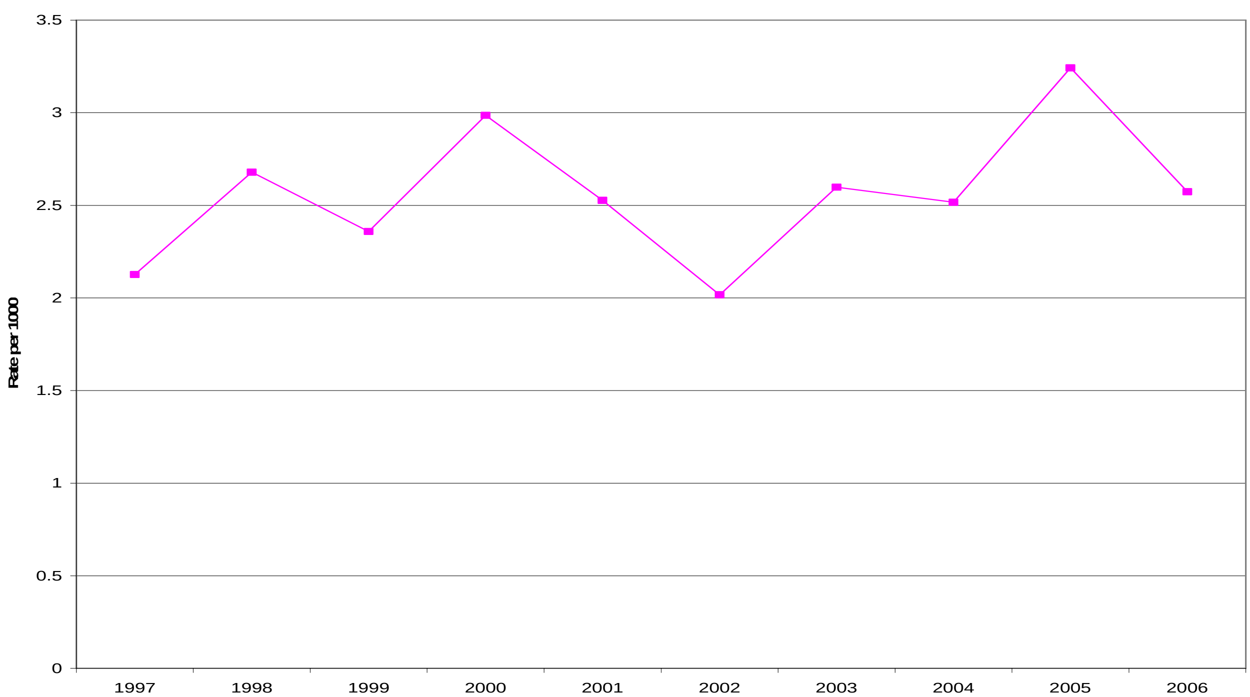
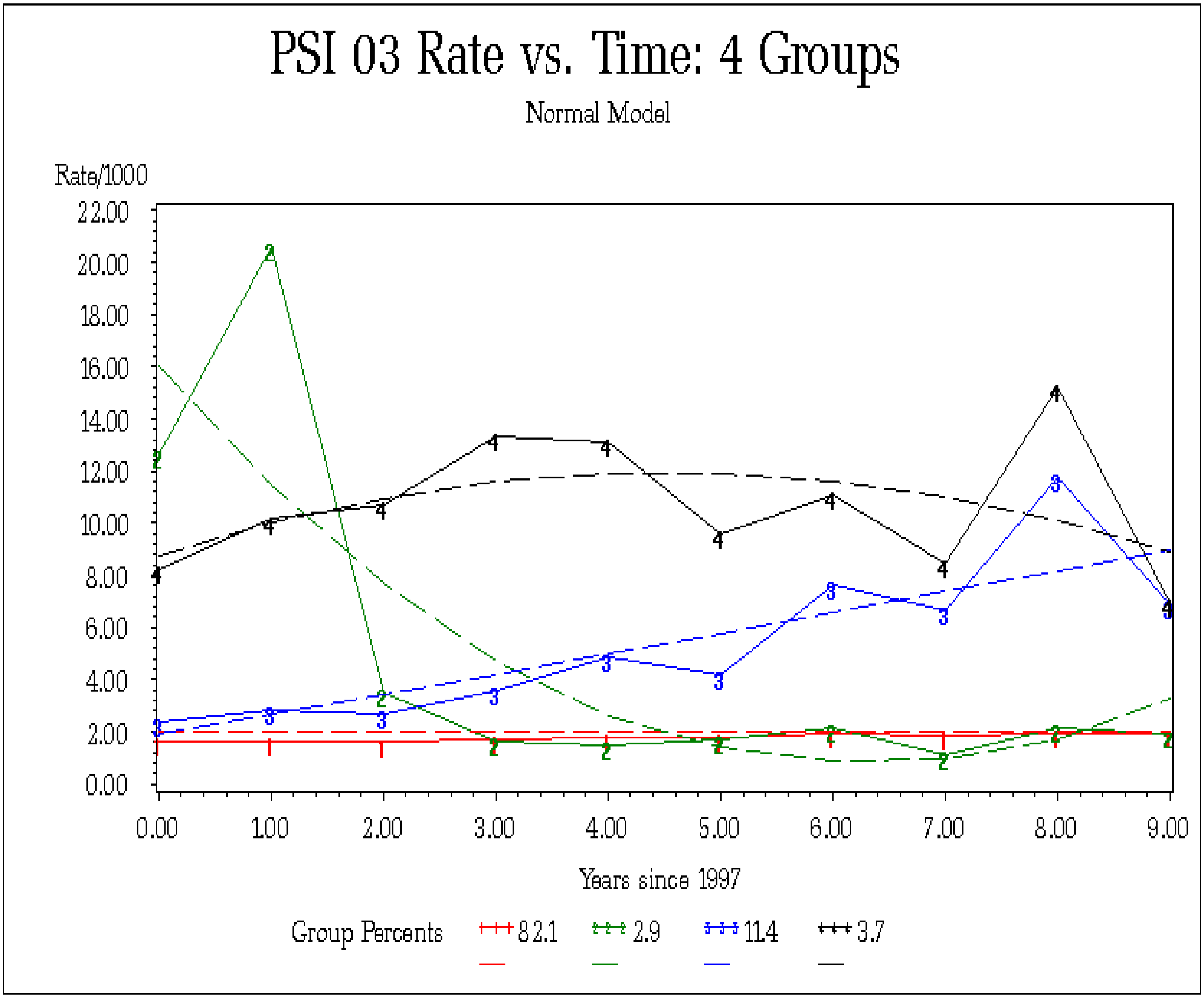


Figure 6 suggests that this relatively constant overall rate is masking a much more dynamic story for PSI 3. Many hospitals may, in fact experience constant rates over time, but there is evidence that three smaller groups of hospitals (estimated to include 18% of California hospitals) are getting better or worse.

Figure 6: Plot of 4 trajectories for "Decubitis Ulcer" (PSI 3)



The low, constant trajectory (group 1) is estimated to include 82.1% of the hospitals in California. This trajectory has an estimated rate of 1.8 per 1,000 (1.6, 1.9) throughout the period 1997-2006. Among the hospitals that did not have low, stable rates on "Decubitis Ulcer", one group (group 2, 2.9%) started with high rates that subsequently improved to about the same level as group 1. A second group (group 3, 11.4%) started at the same rate as group 1, but then got steadily worse. Finally, a third group (group 4, 3.7%) stayed relatively high the entire period.

In this case, California's state average "Decubitus Ulcer" rate is relatively low and constant, and represents the apparent rate for most of the hospitals in California. Nevertheless, the average state trend masks three small, but possibly important, alternative trajectories. There is one group of California hospitals that seemed to have a constant but higher rate throughout the period than reflected by the state average, and two other groups of California hospitals that experiencing significant changes in their rates of "Decubitis Ulcer" over time. A closer examination of the hospitals with increasing or decreasing rates of "Decubitis Ulcer" might reveal important hospital characteristics associated with different trajectories over the period.

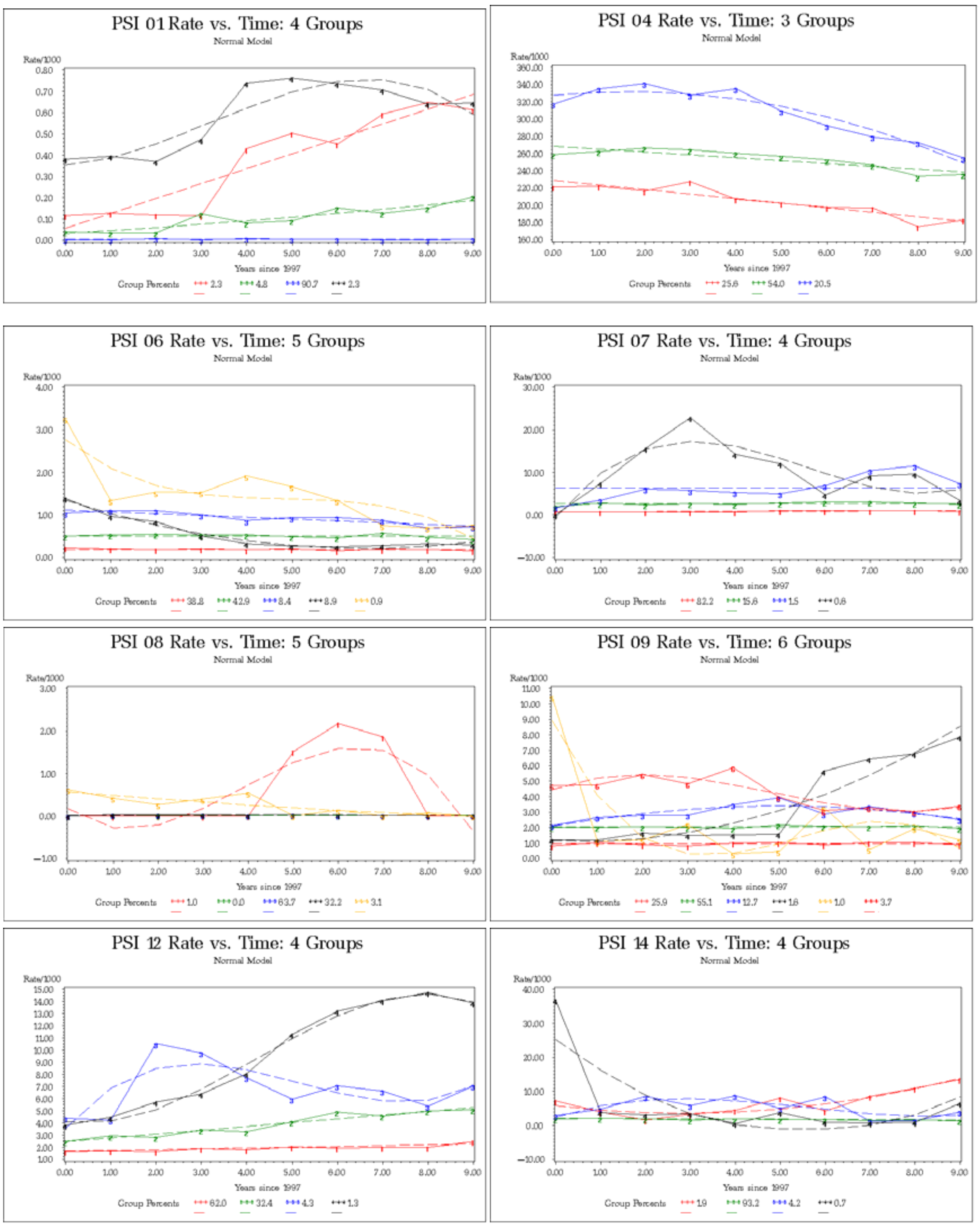
The two examples above demonstrate that relying on a single average trend to analyze patient safety in a population of hospitals may be misleading. There may in fact be multiple groups of hospitals following distinct trajectories.

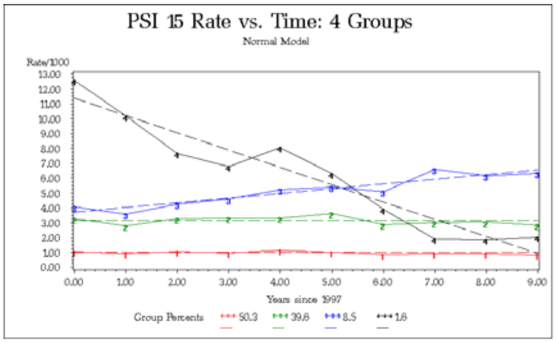
6.3 Evidence of Distinct Trajectories for 13 PSI's

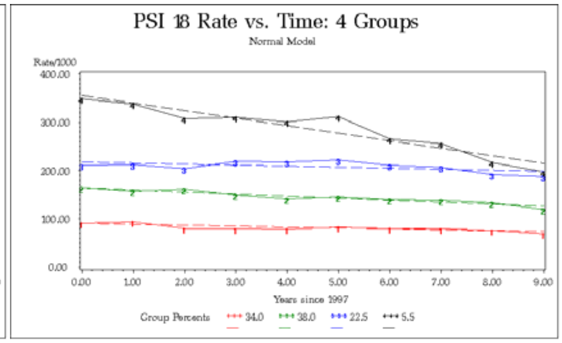
Figure 6 below contains the final trajectory plots for all 11 remaining PSI's that I was able to compute trajectories for. In general, applying group-based trajectory models to each of these individual PSI's revealed similar patterns to the sample above. Most of the PSI's have relatively constant average rates 1997-1006 (PSI's 3, 7, 8, 9, 14, 15, and 18). PSI's 1 and 12 have an increasing average rate, and PSI's 4, 6, and 19 have a decreasing average rate. The group-based trajectories models reveal richer trends, but the pattern is similar. Most hospitals follow low and constant rates while a small percentage follow a more dynamic trajectories. PSI 18 (Obstetric trauma – vaginal delivery with instrument) is an exception to this. Hospitals are more evenly distributed among the trajectories for

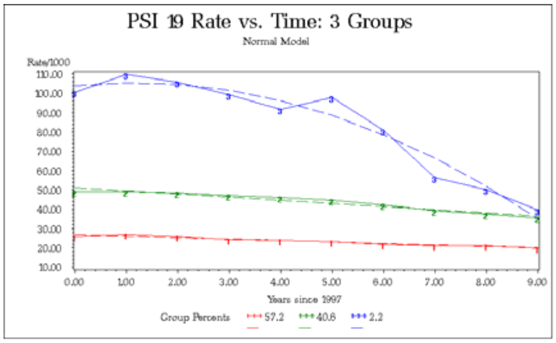
PSI 18. There is evidence for distinct trajectories for all but four of the 17 PSI's I can compute using the California data.

Figure 7: Trajectory plots of 11 PSI's









6.4 Combining Information from PSI's

The analyses above indicate that there is evidence of distinct patterns of patient safety across the PSI's. Hospitals can be classified as having high or low rates over time as well as improving or getting worse over time for each PSI. Is there a common message across the PSI's for a given hospital? I look at patterns across all 13 PSI's individually and by category.

I classified each individual PSI trajectory as high, low, increasing, or decreasing for each PSI. The high and low trajectories are identified based on their relative position. Not all trajectories are classified as high or low. For example, there are three trajectories for PSI 4 (see Figure 7 above). I classified the lowest trajectory as low and the highest as high, but I left the middle trajectory unassigned in the high/low classification. The increasing and decreasing trajectories were identified through a combination of both visual inspection of the trajectories and the coefficient estimates for the time trends (see Appendix). For PSI 4, all three trajectories are classified as decreasing.

After classifying all of the trajectories, I then assigned each hospital to the trajectory that PROC TRAJ gives the highest probability to. This is the default assignment that PROC TRAJ generates. Table 4 below contains the group classifications for each PSI as well as the percentage of hospitals that fall into each type of trajectory. I also categorized as PSI as either general, surgical, or OB.

Table 4. Assignment of groups to trajectory types

		High		Low		Decreasing		Increasing	
PSI	Category	Groups	%	Groups	%	Groups	%	Groups	%
1	Surgical	1, 4	4.6	3	90.7	na		1,4	4.6
3	General	2, 3, 4	17.9	1	82.1	2	2.9	3	11.4
4	General	3	20.5	1	25.6	1,2,3	100	na	
6	General	3, 4, 5	18.3	1, 2	81.7	4, 5	9.8	na	
7	General	3, 4	2.1	1, 2	97.8	na		3	1.5
8	Surgical	1,5	4.1	2,3,4	95.9	5	3.1	na	
9	Surgical	3, 4, 5, 6	19	1, 2	81	5, 6	4.7	4	1.6
12	Surgical	2, 3, 4	38	1	62	na		2,4	33.7
14	Surgical	1,4	2.6	2,3	97.4	4	0.7	1	1.9
15	General	2, 3, 4	49.7	1	50.3	4	1.6	3	8.5
17	OB	2	3.9	1	96.1	2	3.9	na	
18	OB	4	5.5	1	34	2,4	43.5	na	
19	OB	3	2.2	1	57.2	2,3	42.8	na	

Once individual hospitals were assigned to a type of trajectory for each PSI, I was able to look for patterns across PSI's. For example, 19 hospitals are on low trajectories for all 13 PSI's. Every hospital is on a low trajectory for at least one PSI. About 80% of hospitals are on low trajectories for at least seven of the 13 measures. Even though each PSI has a large percentage of hospitals on low trajectories, relatively few hospitals are always in this low trajectory group.

No hospitals are on a high trajectory for every PSI, and 61 hospitals are not on a high trajectory for any PSI. The maximum number of high trajectories for any hospital is six. 146 hospitals are on an increasing trajectory for at least one PSI, and 276 are on a decreasing trajectory for at least one PSI (there are 65 hospitals that have no discharges that are eligible for PSI 4). In fact, only 21 hospitals are not on any dynamic trajectory. 93.4% of the hospitals in California are predicted to be improving or getting worse on at least one of the 13 PSI's I can measure. 124 hospitals are getting better on some PSI while getting worse on another.

I can also combine the PSI's into categories and look for patterns among the categories. Table 5 contains summary information by category. Most hospitals are following a high trajectory on at least one general PSI and all are following a low trajectory for at least one general PSI. Most are also improving on at least one general PSI. A large percentage of hospitals are getting worse on at least one of the surgical PSI's. It appears that PSI 12, Postoperative PE or DVT, is the area where most of these hospitals are getting worse. 100 hospitals are estimated to be on an increasing trajectory for PSI 12.

Table 5: Trajectory types by category

	Any High		Any Low		Any Increasing		Any Decreasing	
Category	**Total**	**%**	**Total**	**%**	**Total**	**%**	**Total**	**%**
General	225	70.5%	319	100.0%	58	18.2%	262	82.1%
Surgical	159	49.8%	310	97.2%	121	37.9%	24	7.5%
OB	23	7.2%	229	71.8%			151	47.3%

Looking at the PSI's individually or by category provides the opportunity to identify possible trends in individual measures. By focusing in on the increase in surgical rates and even on a single PSI among the surgical PSI's, I can identify general and specific patient safety trends.

The picture of patient safety for a given hospital is mixed. All hospitals appear to be doing well on at least one measure, but very few are performing well on all measures. A large percentage of the hospitals in California are on low trajectories for at least half of the measures. Almost all hospitals are improving or getting worse on some measure, and many are doing both.

7. Conclusions and Possible Extensions

By focusing on cross-sectional rankings or average trends over time, researchers and policy-makers risk overlooking important dynamic trends followed by distinct groups of hospitals. The analyses presented here provide evidence that there are distinct groups of hospitals with different patterns of patient safety over time. This has implications for policy makers and government agencies who are trying to understand the levels and trends in patient safety. It has implications for hospital administrators who are trying to understand where their facilities stand in relation to others or what areas of safety within their hospitals may need attention. It has implications for consumers and consumer groups who are interested in identifying the safest hospitals (or the least safe). Finally, it has implications for researchers who are developing and using patient safety measures to study the causes of adverse events and how to prevent them.

There are three general topics for future research: aggregating information across individual measures, identifying outlier hospitals, and identifying characteristics of hospitals that are performing well and/or badly.

I have already presented and discussed one way to combine information across individual PSI's in section 6.4 above. Additional research could improve on this approach. The PSI measures developed by AHRQ are not perfect measures of patient safety. There are other types of measures out there. Future research might explore whether additional indicators of patient safety provide a similar picture of patient safety.

Identifying specific hospitals that merit further investigation is another interesting possibility. In addition to using some sort of aggregate measure, looking at individual hospital rates in relation to group average trajectories my be useful. The trajectories described above represent averages for the groups of hospitals. Individual hospital rates do not follow the trajectories perfectly. If a hospital moves too far from the group average, then it may warrant investigation.

Explaining the reasons behind the patterns across hospitals is another line of possible research. If groups of hospitals that perform well or badly in specific areas can be identified, then explaining those patterns would involve looking for characteristics or practices in those facilities that explain the performance. It may be possible to identify common characteristics or processes among the hospitals that are experiencing an increase in surgical safety events, for example. If so, then perhaps a way to reverse the trend can be identified. Similarly, it may be helpful to identify the characteristics of those few hospitals that appear to be doing well on all of the measures.

8. Acknowledgements

In addition to my advisors who have provided helpful comments on earlier versions of this draft, I would like to thank Donna Farley at the RAND Corporation for providing coverage to work on the analyses presented here as well as helpful advice. I would also like to thank Michael Greenberg and Regan Main at the RAND Corporation for helpful comments and assistance with processing the discharge data and generating the PSI measures and Robert Pritchard for helpful comments on earlier versions of this draft.

9. References

Agency for Healthcare Quality and Research (AHRQ). 2007a. *Guide to Patient Safety Indicators. Version 3.1.* Rockville, MD, U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. Downloaded from http://www.qualityindicators.ahrq.gov/downloads.htm Oct, 25, 2008.

Agency for Healthcare Quality and Research (AHRQ). 2007b. *Patient Safety Indicators: Software Documentation, Version 3.1 SAS.* Rockville, MD, U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. Downloaded from http://www.qualityindicators.ahrq.gov/downloads.htm April, 3, 2008.

Agency for Healthcare Quality and Research (AHRQ). 2007c. *2007 National Healthcare Quality Report.* Rockville, MD, U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality, 136pp.

Bahl, V., M.A. Thompson, T.Y. Kau, H.M. Hu, and D.A. Campbell, Jr.. 2008. "Do the AHRQ Patient Safety Indicators Flag Conditions That Are Present at the Time of Hospital Admission?" *Medical Care* 46 (5): 516-22.

Centers for Disease Control (CDC). 2004. "National Nosocomial Infections Surveillance (NNIS) System Report, data summary from January 1992 through June 2004, issued October 2004." Atlanta, Georgia, Division of Healthcare Quality Promotion, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Public Health Service, US Department of Health and Human Services. Downloaded from http://www.cdc.gov/ncidod/dhqp/nnis_pubs.html April 21, 2009.

M.S. Bogner (Ed.). 1994. *Human Error in Medicine.* Hillsdale, NJ: Lawrence Earlbaum Associates.

Greenberg M.D., A.M. Haviland, H. Yu, and D.O. Farley. 2009. "Safety Outcomes in the United States: Trends and Challenges in Measurement." *Health Services Research* 44 (2): 739-55.

Faber M., M. Bosch, H. Wollersheim, S. Leatherman, and R. Grol. 2009. "Public Reporting in Healthcare: How Do Consumers Use Quality-of-Care Information?" *Medical Care* 47 (1): 1-8.

Hospital Compare. 2009. U.S Department of Health and Human Services (HHS). Hospital Compare website. http://www.hospitalcompare.hhs.gov

Institute of Medicine. 2000. *To Err is Human: Building a Safer Health System.* Edited by L.T. Kohn, J.M. Corrigan, and M.S. Donaldson. Washington, DC: National Academy Press.

Jha, A.K., E.J. Orav, A.B. Ridgeway, J. Zheng, and A.M. Epstein. 2008. "Does the Leapfrog Program Help Identify High-Quality Hospitals?" *J. Quality and Patient Safety* 34 (6): 318-25.

Joint Commission. 2009. "Sentinel Event Statistics as of December 31, 2008." Downloaded from http://www.jointcommission.org/SentinelEvents/Statistics/ April 19, 2009.

Joint Commission. 2009. "Sentinel Events Trends Reported by Year." Downloaded from http://www.jointcommission.org/SentinelEvents/Statistics/ April 19, 2009.

Jones, B.L., D. Nagin, and K. Roeder. 2001. "A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories." *Sociological Research and Methods* 29: 374-393.

Khan III, C.N., T. Ault, H. Isenstein, L. Potetz, and S. Van Gelder. 2006. "Snapshot of Hospital Quality Reporting and Pay-for-Performance Under Medicare." *Health Affairs* 25 (1): 148-62.

Leape L.L. and D.M. Berwick. 2005. "Five Years After *To Err is Human*: What Have We Learned?" *JAMA* 293 (19):2384-90.

The Leapfrog Group. 2008. "2008 Leapfrog Top Hospitals." http://www.leapfroggroup.org/media/file/2008_Top_Hospital_Release.pdf. Downloaded April 19, 2009.

Nagin, Daniel. 2005. *Group-Based Modeling of Development*. Cambridge, MA: Harvard University Press.

Romano, P.S., J.J. Geppert, S. Davies, M.R. Miller, A. Elixhauser, and K.M. McDonald. 2003. "A National Profile of Patient safety in U.S. Hospitals." *Health Affairs* 22 (2): 154-66.

Thomas, E.J., E.J. Orav, and T.A. Brennan. 2000. "Hospital Ownership and Preventable Adverse Events." *J Gen Intern Med* 15:211-19.

Yu, H, M.D. Greenberg, A.M. Haviland, and D.O. Farley. 2009. "'Canary Measures' Among the AHRQ Patient Safety Indicators." *Am J of Med Qual*, forthcoming.

Zhan, C. and M.R. Miller. 2003. "Excess Length of Stay, Charges, and Mortality Attributable to Medical Injuries During Hospitalization." *JAMA* 290 (14): 1868-74.

Zhan, C. and M.R. Miller. 2003. "Administrative Data Based Patient Safety Research: A Critical Review." *Qual Saf Health Care* 12 (supp II): ii58-63.

Zhan, C., E Kelley, H.P. Yang, M. Keyes, J. Battles, R.J. Borotkanics, and D. Stryer. 2005. "Assessing Patient Safety in the United States: Challenges and Opportunities." *Medical Care* 43 (3, suppl): I42-7.

APPENDIX

PROC TRAJ Output

1

Normal model

PSI 01

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.05684	0.01323	4.297	0.0000
	Linear	0.07000	0.00248	28.249	0.0000
2	Intercept	0.03316	0.01231	2.695	0.0071
	Linear	0.01220	0.00642	1.901	0.0574
	Quadratic	0.00057	0.00070	0.802	0.4224
3	Intercept	0.00425	0.00113	3.749	0.0002
4	Intercept	0.35523	0.02043	17.389	0.0000
	Linear	0.01107	0.02079	0.533	0.5944
	Quadratic	0.02368	0.00555	4.264	0.0000
	Cubic	-0.00245	0.00040	-6.040	0.0000
	Sigma	0.05943	0.00076	78.363	0.0000
	Group membership				
1	(%)	2.25806	0.84555	2.671	0.0076
2	(%)	4.81433	1.23457	3.900	0.0001
3	(%)	90.66952	1.66739	54.378	0.0000
4	(%)	2.25808	0.84555	2.671	0.0076

BIC= 4171.85 (N=3100) BIC= 4187.97 (N=310) AIC= 4214.12 L=
4228.12

Normal model

2

PSI 03

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	1.78614	0.08070	22.133	0.0000
2	Intercept	16.04731	1.17034	13.712	0.0000
	Linear	-4.97186	0.60756	-8.183	0.0000
	Quadratic	0.39360	0.06154	6.396	0.0000
3	Intercept	1.69549	0.43920	3.860	0.0001
	Linear	0.81208	0.10012	8.111	0.0000
4	Intercept	8.72971	1.01176	8.628	0.0000
	Linear	1.42465	0.44337	3.213	0.0013
	Quadratic	-0.15654	0.04633	-3.379	0.0007
	Sigma	3.43521	0.04453	77.143	0.0000
	Group membership				
1	(%)	82.08998	2.65334	30.938	0.0000
2	(%)	2.85881	1.00485	2.845	0.0045
3	(%)	11.36129	2.30774	4.923	0.0000
4	(%)	3.68992	1.15196	3.203	0.0014

BIC= -8450.88 (N=3100) BIC= -8435.91 (N=310) AIC= -8411.62 L= -8398.62

Normal model
3

PSI 04

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	228.55626	4.70154	48.613	0.0000
	Linear	-5.25189	0.80174	-6.551	0.0000
2	Intercept	268.45547	3.78060	71.009	0.0000
	Linear	-3.31989	0.58416	-5.683	0.0000
3	Intercept	327.22857	7.03430	46.519	0.0000
	Linear	5.17507	3.21352	1.610	0.1074
	Quadratic	-1.54900	0.33824	-4.580	0.0000
	Sigma	52.10370	0.75519	68.994	0.0000
	Group membership				
1	(%)	25.57983	3.78662	6.755	0.0000
2	(%)	53.95654	4.16411	12.958	0.0000
3	(%)	20.46363	3.30417	6.193	0.0000

BIC=-13881.76 (N=2540) BIC=-13870.24 (N=254) AIC=-13852.56 L=-13842.56

Normal model

4

PSI 06

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.19082	0.02219	8.598	0.0000
2	Intercept	0.50481	0.02696	18.723	0.0000
3	Intercept	1.11658	0.07754	14.399	0.0000
	Linear	-0.04237	0.01249	-3.392	0.0007
4	Intercept	1.35120	0.12006	11.254	0.0000
	Linear	-0.34533	0.05734	-6.022	0.0000
	Quadratic	0.02628	0.00555	4.734	0.0000
5	Intercept	2.76598	0.21803	12.686	0.0000
	Linear	-0.82603	0.22153	-3.729	0.0002
	Quadratic	0.16727	0.05917	2.827	0.0047
	Cubic	-0.01157	0.00431	-2.681	0.0074
	Sigma	0.41412	0.00543	76.280	0.0000
	Group membership				
1	(%)	38.75159	5.39557	7.182	0.0000
2	(%)	42.92202	5.42268	7.915	0.0000
3	(%)	8.44548	2.36543	3.570	0.0004
4	(%)	8.93981	2.65552	3.367	0.0008
5	(%)	0.94110	0.54207	1.736	0.0826

BIC= -2040.13 (N=3190) BIC= -2021.71 (N=319) AIC= -1991.59 L= -1975.59

Normal model

5

PSI 07

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.73862	0.05237	14.103	0.0000
	Linear	0.02976	0.00954	3.119	0.0018
2	Intercept	2.65279	0.08894	29.825	0.0000
3	Intercept	6.31335	0.21761	29.012	0.0000
4	Intercept	-0.65938	0.89149	-0.740	0.4596
	Linear	13.40239	0.90749	14.769	0.0000
	Quadratic	-3.01687	0.24240	-12.446	0.0000
	Cubic	0.17884	0.01767	10.119	0.0000
	Sigma	1.38668	0.01761	78.749	0.0000
	Group membership				
1	(%)	82.19398	2.45111	33.533	0.0000
2	(%)	15.63119	2.35095	6.649	0.0000
3	(%)	1.54787	0.69804	2.217	0.0267
4	(%)	0.62697	0.44270	1.416	0.1568

BIC= -5772.70 (N=3190) BIC= -5758.88 (N=319) AIC= -5736.29 L= -
5724.29

Normal model

PSI 08

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.16356	0.12025	1.360	0.1739
	Linear	-0.75223	0.12241	-6.145	0.0000
	Quadratic	0.34064	0.03275	10.402	0.0000
	Cubic	-0.02928	0.00239	-12.244	0.0000
2	Intercept	0.02311	0.32429	0.071	0.9432
3	Intercept	0.01929	0.01356	1.422	0.1550
	Linear	0.00515	0.00759	0.679	0.4975
	Quadratic	-0.00065	0.00083	-0.788	0.4307
4	Intercept	0.01929	0.01978	0.975	0.3296
	Linear	0.00515	0.01083	0.476	0.6342
	Quadratic	-0.00065	0.00117	-0.558	0.5772
5	Intercept	0.56658	0.21934	2.583	0.0098
	Linear	-0.08782	0.12475	-0.704	0.4815
	Quadratic	0.00265	0.01221	0.217	0.8280
	Sigma	0.22890	0.00304	75.358	0.0000
	Group membership				
1	(%)	0.97163	119.92402	0.008	0.9935
2	(%)	0.02527	3.75627	0.007	0.9946
3	(%)	63.66864	1686.86254	0.038	0.9699
4	(%)	32.23648	2192.32750	0.015	0.9883
5	(%)	3.09798	382.36780	0.008	0.9935

BIC= 40.57 (N=3090) BIC= 62.45 (N=309) AIC= 97.91 L=
116.91

Normal model

7

PSI 09

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.95058	0.10151	9.365	0.0000
2	Intercept	2.04868	0.09710	21.099	0.0000
3	Intercept	2.07620	0.34812	5.964	0.0000
	Linear	0.52898	0.18140	2.916	0.0036
	Quadratic	-0.05275	0.01729	-3.051	0.0023
4	Intercept	1.12508	0.54798	2.053	0.0401
	Linear	-0.14504	0.28270	-0.513	0.6080
	Quadratic	0.10763	0.03040	3.541	0.0004
5	Intercept	8.91115	1.25172	7.119	0.0000
	Linear	-6.05387	1.32956	-4.553	0.0000
	Quadratic	1.29547	0.33682	3.846	0.0001
	Cubic	-0.08054	0.02333	-3.452	0.0006
6	Intercept	4.43663	0.52145	8.508	0.0000
	Linear	1.09070	0.52467	2.079	0.0377
	Quadratic	-0.34552	0.14284	-2.419	0.0156
	Cubic	0.02342	0.01041	2.250	0.0245
	Sigma	1.52582	0.02045	74.606	0.0000
	Group membership				
1	(%)	25.92941	5.06738	5.117	0.0000
2	(%)	55.05108	5.97242	9.218	0.0000
3	(%)	12.69595	5.78073	2.196	0.0281
4	(%)	1.59597	0.72000	2.217	0.0267
5	(%)	1.02678	0.68532	1.498	0.1342
6	(%)	3.70082	1.46658	2.523	0.0117

BIC= -6032.16 (N=3100) BIC= -6006.84 (N=310) AIC= -5965.73 L= -5943.73

Normal model

8

PSI 12

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	1.60356	0.12564	12.763	0.0000
	Linear	0.06893	0.02270	3.037	0.0024
2	Intercept	2.43015	0.17608	13.801	0.0000
	Linear	0.31699	0.04046	7.834	0.0000
3	Intercept	3.56585	0.64908	5.494	0.0000
	Linear	4.22491	0.91227	4.631	0.0000
	Quadratic	-1.01838	0.24998	-4.074	0.0000
	Cubic	0.06576	0.01743	3.772	0.0002
4	Intercept	4.10285	1.12364	3.651	0.0003
	Linear	-0.58790	1.12376	-0.523	0.6009
	Quadratic	0.64719	0.30335	2.133	0.0330
	Cubic	-0.05110	0.02211	-2.311	0.0209
	Sigma	2.46813	0.03266	75.566	0.0000
	Group membership				
1	(%)	61.95781	6.17803	10.029	0.0000
2	(%)	32.37960	5.86426	5.522	0.0000
3	(%)	4.32123	1.37765	3.137	0.0017
4	(%)	1.34137	0.69145	1.940	0.0525

BIC= -7430.47 (N=3090) BIC= -7412.05 (N=309) AIC= -7382.19 L= -7366.19

Normal model

9

PSI 14

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	5.78180	0.94913	6.092	0.0000
	Linear	-1.52926	0.48357	-3.162	0.0016
	Quadratic	0.26868	0.05282	5.087	0.0000
2	Intercept	1.98221	0.10029	19.765	0.0000
	Linear	-0.06099	0.01856	-3.286	0.0010
3	Intercept	2.38584	0.77644	3.073	0.0021
	Linear	4.30235	0.79442	5.416	0.0000
	Quadratic	-1.01055	0.21327	-4.738	0.0000
	Cubic	0.05972	0.01556	3.838	0.0001
4	Intercept	25.17762	1.55071	16.236	0.0000
	Linear	-9.91161	0.80244	-12.352	0.0000
	Quadratic	0.89280	0.08583	10.402	0.0000
	Sigma	2.78224	0.03652	76.179	0.0000
	Group membership				
1	(%)	1.91421	0.82673	2.315	0.0207
2	(%)	93.20413	1.61097	57.856	0.0000
3	(%)	4.21274	1.33173	3.163	0.0016
4	(%)	0.66892	0.47259	1.415	0.1570

BIC= -7451.49 (N=2990) BIC= -7433.07 (N=299) AIC= -7403.47 L= -7387.47

Normal model
10

PSI 15

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.97304	0.04869	19.985	0.0000
2	Intercept	3.13178	0.07030	44.551	0.0000
3	Intercept	3.70191	0.19982	18.526	0.0000
	Linear	0.31571	0.04639	6.806	0.0000
4	Intercept	11.35670	0.42126	26.959	0.0000
	Linear	-1.16306	0.07880	-14.759	0.0000
	Sigma	1.59422	0.02029	78.583	0.0000
	Group membership				
1	(%)	50.32176	3.06486	16.419	0.0000
2	(%)	39.64198	3.04631	13.013	0.0000
3	(%)	8.46826	1.90464	4.446	0.0000
4	(%)	1.56801	0.69670	2.251	0.0245

BIC= -6337.08 (N=3190) BIC= -6325.56 (N=319) AIC= -6306.74 L= -6296.74

Normal model

11

PSI 17

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	3.43436	0.16956	20.254	0.0000
2	Intercept	104.58756	2.09386	49.950	0.0000
	Linear	-33.73393	1.08350	-31.134	0.0000
	Quadratic	2.57591	0.11590	22.226	0.0000
	Sigma	7.98067	0.11754	67.897	0.0000
	Group membership				
1	(%)	96.10384	1.27454	75.403	0.0000
2	(%)	3.89616	1.27454	3.057	0.0023

BIC= -8136.93 (N=2310) BIC= -8130.02 (N=231) AIC= -8119.70 L= -8113.70

Normal model

12

PSI 18

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	89.75643	4.06978	22.054	0.0000
	Linear	-1.89685	0.67160	-2.824	0.0048
2	Intercept	165.09552	4.40597	37.471	0.0000
	Linear	-4.36734	0.67508	-6.469	0.0000
3	Intercept	218.19727	5.73572	38.042	0.0000
	Linear	-2.15517	1.08988	-1.977	0.0481
4	Intercept	355.81252	9.03534	39.380	0.0000
	Linear	-15.66599	1.85960	-8.424	0.0000
	Sigma	50.07228	0.75993	65.891	0.0000
	Group membership				
1	(%)	33.98080	3.68448	9.223	0.0000
2	(%)	38.03208	3.72377	10.213	0.0000
3	(%)	22.47723	3.15651	7.121	0.0000
4	(%)	5.50989	1.62861	3.383	0.0007

BIC=-12290.65 (N=2250) BIC=-12276.84 (N=225) AIC=-12256.34 L=-
12244.34

Normal model
13

PSI 19

Maximum Likelihood Estimates
Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	26.87560	0.73058	36.786	0.0000
	Linear	-0.77464	0.12264	-6.317	0.0000
2	Intercept	51.11882	0.91728	55.729	0.0000
	Linear	-1.60591	0.14748	-10.889	0.0000
3	Intercept	103.93537	4.39316	23.658	0.0000
	Linear	2.60773	2.27292	1.147	0.2514
	Quadratic	-1.13930	0.24312	-4.686	0.0000
	Sigma	12.46755	0.18530	67.284	0.0000
	Group membership				
1	(%)	57.17888	3.64821	15.673	0.0000
2	(%)	40.64711	3.62664	11.208	0.0000
3	(%)	2.17401	0.96350	2.256	0.0241

BIC= -9267.31 (N=2300) BIC= -9255.80 (N=230) AIC= -9238.61 L= -
9228.61