Using Time Series ARIMA Modeling for Forecasting Bed-Days in a Medicare HMO

Ariel Linden, Dr.P.H., M.S., Stuart O. Schweitzer, Ph.D.

Presented by: Ariel Linden, Dr.P.H., M.S., Department of Health Services, School of Public Health, University of California, Los Angeles, CA 90095-1772, Tel: 310-825-2595; Fax: 310-825-3317; E-mail: ariellinden@yahoo.com

Research Objective: Managed care was established under the premise that cost-saving strategies (e.g., capitation, gate-keeping, pre-authorization, peer-review, case-management, pharmacy formularies, etc.) could contain costs and utilization of health services in a controlled environment. Recently though, it has become apparent that both costs and utilization have increased despite these strategies. Since hospital spending continues to consume the largest portion of the health care dollar (38 percent of spending on personal health services in 1997), bed-days have remained one area of medical care to undergo continued intense scrutiny. Yet the inability to control for such variables as chronic disease, normal aging, advances in medical technology, and market pricing for goods and services, make forecasting difficult. Unfortunately, most HMOs typically establish budget bed-day goals prospectively, using either external benchmarks or subjective valuations, rather than forecasting hospitalization rates using mathematical models based on actual past utilization. This paper reports the results of an Auto-Regressive Integrated Moving Average (ARIMA) time series model that used retrospective Medicare HMO data for forecasting future bed-days.

Study Design: 72 months of Medicare hospital bed-day data adjusted per thousand members, was provided by a medium-sized HMO in Southern California. Time series analysis is predicated on serial dependency (autocorrelation) of data points. An Auto-Regressive (AR) model represents a process in which the observation at time t is a weighted average of the most recent observations in a series $t - 1$ and is designated as (p). A Moving Average (MA) model represents a process in which an observation is a weighted average of the most recent random disturbances (variables that affect the previous series of observations), and is designated as (q). Both AR and MA models require that the mean and variance of the time series data be constant over time (stationary). The order of differencing, known as Integration (I), refers to the number of times each previous observation is subtracted from each successive observation until the data become stationary, and is designated as (d). Thus, an ARIMA model is typically presented as ARIMA (p, d, q), where the p is the number of autoregression parameters, d is the order of differencing, and q is the number of moving average parameters. Similarly, if the data indicates a seasonal trend, the model is typically presented as ARIMA (p, d, q) (p, d, q), where the second grouping indicates the seasonal component of the model. Our model was developed using the first 60 months of observations as the historical period, and the last 12 observations (January through December 1999) as the validation period.

Population Studied: Aggregate data used in this study represents the hospitalization experience of an HMO's Medicare population between January 1994 and December 1999.

Principal Findings: Following the iterative process of identification, estimation and diagnosis of the time-series data, an ARIMA (1,0,1)(0,1,1) model was developed. A strong seasonal component was identified, indicating spikes in hospital bed-days occurring yearly in the month of December. The difference between the predicted and actual bed-days for the historical period (January 1994 to December 1998) was only 0.5 bed-days per 1000 (98.5 versus 98.0 bed-days/1000 respectively), indicating that retrospectively, this model fit the data very well. For the validation period, the difference between the predicted and actual bed-days was 2.4 bed-days/1000 members (106.3 versus 103.9, respectively). As illustrated graphically, all values were well within the 95% confidence intervals, indicating that bed-days were within acceptable control limits. When comparing the yearly HMO budget bed-day goals with the actual bed-days reported for that year, the differences ranged from 1 to 11 percent. Similarly, when comparing the yearly budget goals to the ARIMA predicted values, the differences ranged from 2 to 14 percent.

Conclusions: The ARIMA model developed in this study for forecasting Medicare HMO hospital bed-days proved to be very accurate, much more so than the subjectively derived budget bed-day goals established yearly by the health plan. Additionally, the model uncovered a strong seasonal trend in hospitalizations occurring every December in this senior population.

Implications for Policy, Delivery, or Practice: ARIMA time-series modeling is ideal for forecasting Medicare HMO bed-days. This type of analysis takes into account the serial dependency of observations in an uncontrolled setting. Thus, health plans can accurately forecast hospitalizations without first having to identify individual variables that may account for the variation in monthly bed-days. Similarly, outlier observations are easily identified as those that fall beyond the 95% confidence levels, allowing the plan to focus attention on specific occurrences as they happen.

Primary Funding Source: None.