# Forecasting Weekly Outpatient Demands at Clinics within a Large Medical Center


**Stephen DeLurgio** *(Corresponding author)*

*University of Missouri–Kansas City, Bloch School of Business, 5100 Rockhill Road, Kansas City, MO 84110, delurgios@umkc.edu*

**Brian Denton**

*North Carolina State University, Edward P. Fitts Department of Industrial & Systems Engineering, 376 Daniels Hall, Campus Box 7906, Raleigh, NC 27695-7906, bdenton@ncsu.edu*

**Rosa L. Cabanela**

*Mayo Clinic, Health Services, 200 First Street, SW, Rochester, MN 55905, cabanela.rosa@mayo.edu*

**Sandra Bruggeman**

*Mayo Clinic, Health Services, 200 First Street, SW, Rochester, MN 55905, bruggeman.sandra@mayo.edu*

**Arthur R. Williams**

*Children's Mercy Hospitals and Clinics, Center for Health Outcomes and Health Services, 2401 Gillham Road, Kansas City, MO 64108, awilliams2@cmh.edu*

**Sarah Ward**

*Mayo Clinic, Finance, 200 First Street, SW, Rochester, MN 55905, ward.sarah@mayo.edu*

**Ned Groves**

*Mayo Clinic, Systems and Procedures, 200 First Street, SW, Rochester, MN 55905, groves.net@mayo.edu*

**John Osborn**

*Mayo Clinic, Dept. of Surgery, 200 First Street, SW, Rochester, MN 55905, osborn.john@mayo.edu*



## ABSTRACT

Large regional health care systems face challenges in planning physician schedules and appointments to achieve patient-access goals and efficient use of resources. We describe the development of a planning and forecasting system that predicts outpatient visits (OPVs) for 23 primary and specialty care clinics at a large medical center in Rochester, Minn. We develop and compare univariate, multivariate, and combined methods for forecasting 12-week horizons of OPVs. The multivariate method provides forecasts and information that can be shared among the 23 clinics, a distinct advantage of this approach. The combined method—the average of the univariate and multivariate forecasts—is the most effective method for forecasting OPVs. Univariate, multivariate, and combined methods have the lowest forecast root mean squared errors (RMSEs) for 26.1%, 30.4%, and 43.5% of the clinics, respectively. In addition, when the symmetric median absolute percent error (SMdAPE) is used to evaluate models, the combined, univariate, and multivariate methods yield forecasts with medians of 6.8%, 7.1%, and 7.9%, respectively—again confirming the effectiveness of the combined method.



***Keywords:*** Sales and operations planning, health services, forecasting, time series, combining forecasts


## INTRODUCTION

Accurately forecasting patient demand often means the difference between success (effective patient care and financial viability) and failure. Medical clinics face the difficult task of forecasting patient demands and the resources needed to provide high-quality care while minimizing inefficiencies. To allocate clinical resources efficiently and effectively, clinics use formal and informal systems to forecast patient demands and acuity levels and thereby provide better personnel schedules. The setting of this study is the Mayo Clinic in Rochester, Minn., one of the largest integrated group medical practices in the world. It constitutes scores of divisions, including primary care, internal medicine, medical subspecialties, ancillary services, laboratory medicine, and surgical subspecialties. The Mayo Clinic provides a full range of medical care, from primary care services to quaternary-level care, to more than 500,000 patients each year, accounting for about 2.7 million annual outpatient visits (OPVs) across all services. This study of forecasting methods for predicting OPVs was motivated by the practical need to coordinate capacity across multiple health services to better serve patient demand.

With the emergence of large medical systems, providers are benefiting from improved logistics due to the pooling of common resources. While these benefits are significant, another potential benefit of integration arises from information sharing. For example, in many health systems, a primary care physician is the first point of contact for patients. Depending on the outcome of a medical examination, there may be referrals to one or more medical specialties or subspecialties. Upon referral to a specialist, the patient may be further referred to other specialties based on the outcome of the initial referral. For example, a surge in primary care visits can signal an increase in demand for downstream resources. In this article, we present results that test several methods for forecasting future OPVs at a large medical center. We show that by combining univariate forecasting methods with multivariate methods that incorporate information sharing, providers can more accurately predict resource requirements.

Three different approaches to forecasting OPVs are presented here: (1) univariate methods, (2) multivariate methods based on stepwise regression, and (3) the simple average of the univariate and multivariate forecasts (hereafter referred to as *combined*).

Univariate methods are commonly used in many industries when a single time series is available. However, health systems exhibit considerable dependencies among different specialty areas. A large forecasted demand for primary care OPVs may signal an increase in downstream specialty clinic visits. Similarly, certain specialty care clinics commonly refer patients to one another. In general, at least four types of influences can cause variations in OPVs:

- trends in diseases (e.g., the increasing prevalence of diabetes) and treatments such as the development of new health care services or the impact of new drugs or therapies
- seasonal influences, including climatic (cold and flu seasons), holidays (New Year's, religious holidays), and other personal conventions (summer vacations, new product/service introductions)
- cyclical/irregular influences, including flu epidemics, natural disasters, and terrorist acts
- causal influences such as changes in demands at upstream clinics, number of physicians, or the clinic's marketing mix

The forecasting system described herein is designed to model three of these four influences (all except cyclical/irregular ones). By combining univariate methods, which effectively model the first two influences listed above—trends in diseases and treatments, and seasonal influences—with methods of a multivariate stepwise regression approach, that effectively models causal influences, we achieve a more effective overall forecasting system. (Cyclical-irregular influences are generally difficult to forecast using automated processes; thus, these influences can be included using managerial overrides based on analyses this forecasting system doesn't include.)

Based on this analysis, we find the combined method is the best forecasting approach overall because it is, on average, more accurate than either the univariate or multivariate method alone. The combined method has the advantage of modeling univariate influences such as trends, seasonality, and nonstationarity while simultaneously including other causal influences of the multivariate method such as estimated demands of upstream clinics or aggregate past actual and forecasted demands for all clinics.

The remainder of this article is organized as follows: The second section briefly reviews related research and applications. The third section describes the forecasting methods and the methodology used for univariate, multivariate, and combined forecasts. The fourth section discusses the results of these forecasts. Finally, the fifth section provides conclusions and limitations of this study.

## RELATED RESEARCH AND APPLICATIONS

Forecasting demand for hospital services has been a popular topic of research for decades. Early studies addressed forecasts of monthly (Helmer, Opperman and Suver 1980) or quarterly demand (Kwon, Eickenhorst, and Adams 1980). Several studies have forecasted the daily patient census in hospitals. For instance, Kao and Pokladnik (1978) used multiple regression and Fourier analysis to capture seasonality and the influence of exogenous variables on demand. A number of studies have been devoted to forecasting admissions and estimates of average lengths of stay (ALOS) using ARIMA methods (Farmer and Emami 1990, Jones and Joy 2002). Lin (1989) provides analysis of monthly hospital patient forecasts for two years using ARIMA, Holt–Winters and multiple time-series methods. Mackay and Lee (2005) present analysis and forecasting of hospital bed occupancy, which includes models of patient flow throughout a hospital. Hussain et al. (2005) study a time-series approach focused on forecasting a particular type of hospital admission; they study the influence of seasonal weather patterns and primary care provider visits on influenza patient admissions.

Forecasting emergency department arrivals has also received attention in the literature. A recent study of monthly forecasts for an emergency department using ARIMA and exponential smoothing models was completed by Champion et al. (2007). Several studies have explored time-series forecasting models for hospital admissions, including predictions of emergency visits (Cerrito 2005, Tandberg and Qualls 1994). Littig and Isken (2006) present a comprehensive system of managing short-term hospital occupancy by eight-hour shifts using forecasts of four inflows of patients: emergency arrivals, scheduled arrivals, direct arrivals (non-emergent, non-scheduled), and transfers into the unit from other areas. The inflows and outflows of these sources are predicted using multinomial logistic regression models.

Proudlove, Black, and Fletcher (2007) provide a recent update on the literature devoted to improving inpatient flows to hospitals, including a considerable discussion of emergency departments. Also included is a detailed review of the inpatient forecasting and modeling literature. They note that improving inpatient flow models is “a particularly urgent issue in the National Health Service (NHS).” They discuss and recommend improvements in the management of inpatient flows through hospitals. Problems addressed by Proudlove et al. relate to the application of this article, albeit in the hospital setting.

The methods and techniques of clinical demand forecasting are not unique when compared to other service applications (Armstrong 2001). As illustrated herein and as reported by Makridakis and Hibon (2000) and Armstrong (2001, page 417), combining forecasts can be an effective approach to reduce large forecasting errors. Minimizing errors in forecasting is particularly important in health care, as large errors are not only costly to the health system but also can lead to patient harm (Murray and Berwick 2005).

The contribution of this article is as follows. First, the above literature focuses on forecasting of hospital services and emergency departments. In contrast, we present a weekly forecasting model for a large integrated group practice of outpatient services, including primary care and specialty care services such as neurology, cardiology, endocrinology, gastroenterology, and so on. The rapid discovery of preventive medical treatments in recent years has led to significant decreases in hospital services and corresponding increases in demand for outpatient services. This has represented a significant and growing portion of the health system since the 1980s (CDC 1995).

With the exception of the aforementioned articles, the majority of literature on forecasting health services considers univariate approaches that capture trends and seasonality. On the other hand, we investigate combined forecasts of weekly OPVs that leverage the strength of both univariate and multivariate methods. In particular, our multivariate regression-based approach captures referral dependencies among different health services to improve forecasts by capturing natural referral patterns between related outpatient clinics. Such patterns have recently been recognized by health care providers who model flow patterns among multiple health services needed to treat complex, chronic diseases such as diabetes, cancer, and cardiovascular disease (Berenson, Bodenheimer, and Pham 2006).

## HEALTH SYSTEM PLANNING BACKGROUND

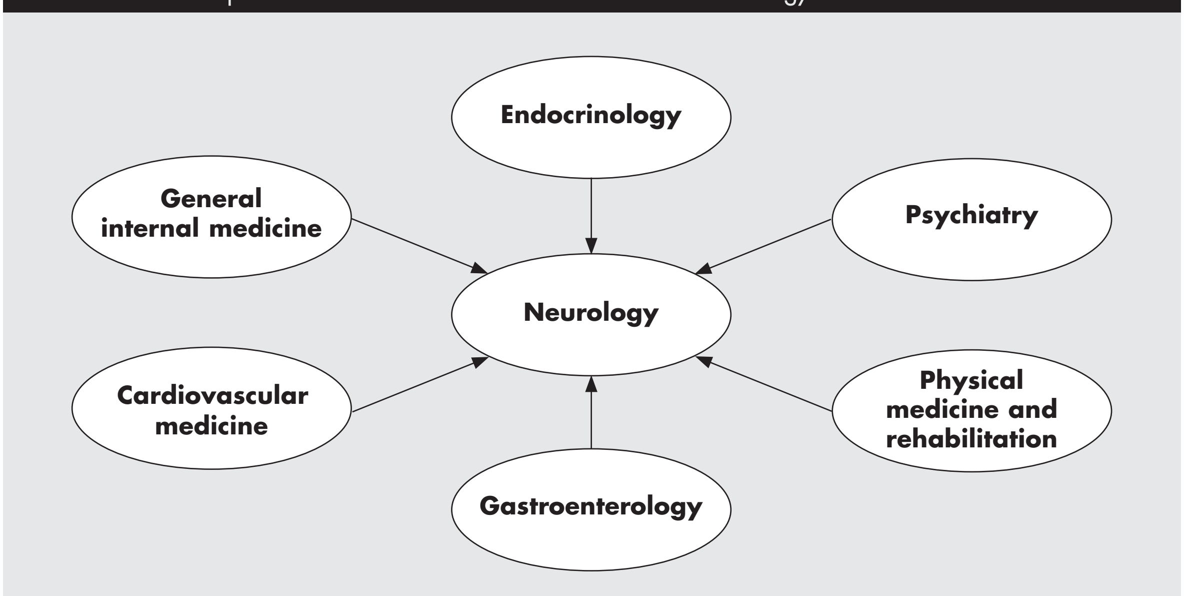
The Mayo Clinic serves regional, national, and international patients. Forecasting demand in advance for health services is particularly important for the latter two groups since they must travel long distances and therefore are particularly sensitive to long waits for services. The clinic's goal is to provide the majority of patients with access to the health services they require within one week.

Our study focuses on outpatient services, which, due to recent advances in medical treatment, are a large and growing proportion of health services. The most expensive resources in outpatient clinics are typically physicians. To enable advance bookings of patients and to give patients time to make travel plans, the design of physicians' calendars begins 12 weeks in advance. The goal of outpatient clinic chairs and administrative managers is to balance two important criteria: (1) timely access of patients to care and (2) efficient use of resources. Clinics have very limited recourses to influence appointment-slot supplies or patient demand for OPVs on short notice. Thus, they must rely on forecasts (often rule-of-thumb) to make decisions about the number of appointment slots in a particular week.

Individual physician's calendars define times when they are available to see patients for scheduled OPVs. At regular intervals, physicians release their calendars so appointment schedulers can book future patient appointments. The estimates of future OPVs are critical to the planning process. Before setting expectations for individual physician calendars, clinic chairs and administrative managers estimate the aggregate number of physician hours required in a particular week. The units of measure used for this estimate are referred to as staff in office practice (SOP). One SOP is equivalent to an eight-hour block of a single physician's time.

Each week, all clinics are required to post estimates of their anticipated SOP for the next 12 weeks (see Table 1 for a list of the clinics considered in this study). The purpose is to provide feedback to upstream and downstream clinics about potential demand fluctuations resulting from changes in SOP supply. For instance, a large general internal medicine (GIM) professional conference or meeting may cause a significant decrease in GIM SOPs. This may in turn affect downstream clinics

**FIGURE 1:** Example of Clinics that Generate Referrals to the Neurology Clinic



that receive referrals from GIM. Figure 1 provides an example of referral patterns for a neurology clinic. Because GIM is an upstream producer of referrals for many clinics (e.g., cardiovascular medicine, endocrinology, and neurology), such a drop in GIM SOP may cause a reduction in total demand for OPVs for the downstream clinics. Clinic chairs and administrative managers can react by encouraging physicians to take research time on funded grants, administrative time in service to committees, and/or vacation time.

## FORECASTING METHODS

This section presents the methods used for forecasting OPVs of each outpatient clinic. The three methods are (1) univariate methods using the historical time-series of OPVs at each clinic, (2) multivariate autoregressive methods utilizing a number of predictor variables, and (3) a combined method, which is the average of the univariate and multivariate forecasts. To make the seasonal univariate models more effective, all data of the 23 clinics have been adjusted so that holidays are 52 weeks apart. In a normal calendar, the year-to-year chronology of Thanksgiving, Easter, and other holidays can be more or less than 52 weeks apart; however, in this database, these holidays are adjusted to be 52 weeks apart. This is a simple process of moving a holiday one week before or one week later so that 52 weeks follow the previous holiday.

### *Univariate Models*

We test the 42 univariate models identified in Table 2 using a time series of historical data for OPVs for each of the 23 clinics in this study. Results for a full spectrum of smoothing, random walk, trending, and seasonal models are provided; these include naïve, seasonally naïve, several preselected ARIMA, exponential smoothing (simple, double, seasonal, Holt's linear, damped trend, additive and Winters' multiplicative), and logarithmic models. (See DeLurgio 1998 or Makridakis, Wheelwright, Hyndman 1998 for further descriptions).

The univariate results reported here are based on the automatic forecasting model selections of the SAS TSFS system (SAS 2004). Specifically, the

**TABLE 1:** 23 Outpatient Clinics Studied

| | |
|---|---|
| Allergy (ALG) | Cancer oncology (ONCOL) |
| Breast (BRS) | Primary care internal medicine (PCIM) |
| Cardiovascular medicine (CV) | Pediatrics (PED) |
| Dermatology (DERM) | Physical medicine and rehabilitation (PMR) |
| Endocrinology (ENDO) | Preventive medicine (PREV) |
| Family medicine (FAMMED) | Psychiatry (PSYCH) |
| Gastroenterology (GI) | Pulmonary medicine (PULM) |
| General internal medicine (GIM) | Rheumatism (RHEUM) |
| Hematology (HEM) | Sleep disorder (SLEEP) |
| Immunology (ID) | Obstetrics and gynecology (OBGYN) |
| Nephrology (NEPH) | |

**TABLE 2:** 42 Univariate Models

| | |
|---|---|
| Mean | Linear trend |
| Linear trend with autoregressive errors | Linear trend with seasonal terms |
| Seasonal dummy | Simple exponential smoothing |
| Double (brown) exponential smoothing | Linear (Holt) exponential smoothing |
| Damped trend exponential smoothing | Seasonal exponential smoothing |
| Winters method – additive | Winters method – multiplicative |
| Random walk with drift | ARIMA(0,1,1)(1,0,0)s No constant |
| Airline model ARIMA (0,1,1)(0,1,1)s No constant | ARIMA(0,1,2)(0,1,1)s No constant |
| ARIMA(0,1,1)s No constant | ARIMA(0,2,2)(0,1,1)s No constant |
| ARIMA(2,0,0)(1,0,0)s | |
| ARIMA(2,1,0)(0,1,1)s No constant | |
| ARIMA(2,1,2)(0,1,1)s No constant | |

Plus 21 additional models using the above models with logarithmic dependent variables.

automatic forecast model selection procedures of TSFS were used to select univariate models with minimum out-of-sample root mean square error (RMSE) for 12 weeks of out-of-sample forecasts.

It is worth noting that the TSFS system has a number of diagnostic tools to automatically pre-select models, including the augmented Dickey-Fuller test for trends/random walks, the use of seasonal dummies to identify seasonal models, and methods to detect and model variance nonstationarity. In our analysis, these tests were not used because the out-of-sample RMSE for 12 weeks was used to select the best-performing model of the 42 possible models, and these models were comprehensive in terms of trends and seasonality. Thus, a hold-out sample of 12 weeks at the end of the data was used to select the structure of the best forecasting model.

### Multivariate Autoregressive Models

A number of approaches were used to determine a best multivariate method for forecasting OPVs. Several experiments were performed to determine if analyst-intensive, interactive multivariate analysis was superior to automated stepwise procedures. While interactive multivariate analysis was found to be slightly superior to automated stepwise procedures for a few clinics, it was not superior for all; therefore, the additional accuracy from interactive multivariate modeling was not found to be cost-effective. Consequently, the multivariate procedures and reported results are those of an automated system.

The automatic multivariate procedure is as follows: A forward stepwise procedure was used to estimate models where the statistical significances of the entering and exiting variables are set at .03 and .04, respectively, to assure model parsimony. If the Durbin-Watson statistics of the stepwise procedure denotes significant serial correlation, models were re-estimated using an autoregressive procedure (specifically we use SAS's AUTOREG procedure). If a significant coefficient exists for the first-order error term (i.e, rho is significantly different than zero), then the results of AUTOREG are used to forecast the series. Otherwise, the original OLS model is used. In addition, if an independent variable becomes insignificant in the process of autoregressive estimation, it is deleted, and the relationship is refitted using the reduced number of independent variables.

The OPVs of clinics are affected by many of the variables shown in Table 3. For example, one important variable in determining OPVs is the initial schedule of health care providers (physicians and clinicians). This variable, previously introduced and termed staff in office practice (SOP), is denoted by $S_{i,t}$ and $S_{i,t}(\mathrm{m})$ in Table 3. The variable $S_{i,t}$ are actual past values of SOP, while the $S_{i,t}(\mathrm{m})$ is a m-period-ahead forecasted value. The SOP variable is a useful predictor of OPVs because it shows a clinic's estimated demand for staff. In addition to the SOP variable, dichotomous variables were used to model specific holidays, including New Year's, Easter, Memorial Day, July Fourth, Thanksgiving, and Christmas. These individual holiday variables were not found to be effective predictors of OPVs; however, we had greater success with aggregated dichotomous variables representing the week before, during, and after all holidays—$H_t^B$, $\mathrm{H}_t$, and $\mathrm{H}_t^A$, which are shown in Table 3. All of these independent variables (including their lags or forecasts) are known in period *t*, the period in which OPV forecasts are generated.

The two general models used in the multivariate method are as follows:

- *OLS Multiple-Regression Model*

$O_{i,t}(\mathrm{m}) = \mathrm{b}_0 + \mathrm{b}_1 X1_{\mathrm{i,t}} + \mathrm{b}_2 X2_{\mathrm{i,t}} + \ldots + \mathrm{b}_n Xn_{\mathrm{i,t}}$ *[Equation 1]*

- *First Order Autoregressive Model*

$O_{i,t}(\mathrm{m}) = \mathrm{b}_0 + \mathrm{b}_1 X1_{\mathrm{i,t}} + \mathrm{b}_2 X2_{\mathrm{i,t}} + \ldots + \mathrm{b}_n Xn_{\mathrm{i,t}} + \rho \mathrm{e}_{\mathrm{i,t-1}}$ *[Equation 2]*

In the equations above, $O_{i,t}(\mathrm{m})$ is the dependent variable, *X*i's are the previously defined variables of Table 3, and ρ is estimated using the Prais-Winsten autoregressive method of SAS (2004).

As an example of a typical model, consider the following:

$O_{i,t}(\mathrm{m}) = \mathrm{b}_0 + \mathrm{b}_1 O_{i,t-52+m} + \mathrm{b}_2 S_{i,t-52+m} + \mathrm{b}_3 AO_{i,t}(\mathrm{m}) + \rho \mathrm{e}_{\mathrm{t-1}}$ *[Equation 3]*

This is a model for predicting the OPVs of clinic *i* for period *t+m* made at the end of period *t*. The predictor variables are as follows:

$O_{i,t-52+m}$ = the actual OPVs of period *t-52+m* (52-*m* weeks ago),

**TABLE 3:** Notation and Variables Used in Multivariate Models

| | | |
|---|---|---|
| t | = | index for time periods (weeks), t =1, ..., T |
| $T_f$ | = | number of time periods from earliest historical observation to last forecasted period |
| $T_p$ | = | number of time periods from earliest historical observation to present period. |
| $i$ | = | index for clinics including primary care and medical subspecialties, i =1, ..., $N$ |
| $N$ | = | total number of primary care and medical subspecialties clinics |
| $m$ | = | number of periods being forecasted into the future as of period t |
| $S_{i,t}$ | = | actual observed SOP for period t for clinic i |
| $O_{i,t}$ | = | actual OPV demand for period t in clinic i |
| $S_{i,t}$(m) | = | the forecasted value of SOP for period t+m of clinic i made in period t |
| $O_{i,t}$(m) | = | the forecasted OPV demand for period t+m of clinic i made in period t |
| $AS_t$ | = | the actual aggregated value of the SOPs of all 23 clinics for period t, hereafter Aggregated SOP |
| $AO_t$ | = | the actual aggregated value of the OPVs of all 23 clinics for period t, hereafter Aggregated OPV |
| $AS_t$(m) | = | the forecasted value of Aggregated SOP for period t+m made in period t |
| $AO_t$(m) | = | the forecasted value of Aggregated OPV for period t+m made in period t |
| $H_t$ | = | dichotomous variable that is 1 if t is a 4-day holiday week and 0 otherwise |
| $H_t^B$ | = | dichotomous variable that is 1 if t is prior to a 4-day week and 0 otherwise |
| $H_t^A$ | = | dichotomous variable that is 1 if t is a after a 4-day week and 0 otherwise |
| $W_t$ | = | weather variable denoting ice and snow fall at the Rochester, MN, airport |

$S_{i,t-52+m}$ = the actual past value of SOP for period $t$-$52$+$m$.

$AO_t$(m) = the predicted value of aggregate OPVs for period $t$+$m$, which was made at the end of period $t$.

$e_{t-1}$ = the error of period $t$-$1$

Equation 3, as well as all other multivariate relationships, was estimated first using a forward stepwise procedure where the following variables were in the selection list:

$O_{i,t}(m)=f\ [S_{i,t-52+m},\ O_{i,t-52+m},\ S_{i,t}(m),\ AS_{t-52+m},\ AO_{t-52+m},\ AS_t(m),\ AO_t(m),\ H_t,\ H_t^B,\ H_t^A,\ e_{t-1}]$. All included variables are known as of the end of period $t$.

#### Combined Models

While several different approaches to combining forecasts were tested, it was found that the simple average of the univariate and multivariate forecasts was the most accurate. Simple averages of the forecasts of competing models have been found to be the most accurate approaches in many studies (Makridakis and Hibon 2000, and Armstrong 2001).

■ *Forecast Root Mean Squared Error (RMSE) and Symmetric Absolute Percent Error (SAPE)*

Two statistics are used to compare actual patient demands to out-of-sample forecasts, RMSE and SAPE. The RMSE for clinic $i$ using method j for t = 1 to 24 periods is

$RMSE_{i,j} = [\Sigma e^2_{i,j,t}/24]^{.5}$ *[Equation 4]*

The SAPE for clinic $i$ using method j for t = 1 to 24 periods is

$SAPE_{i,j} = \Sigma[|e_{i,j,t}|/((O_{i,t+m} + O_{i,t}(m))/2)]/24$ *[Equation 5]*

In the equation above, $O_{i,t}$(m) is defined in Table 3 as the forecast of period $t$+$m$ made at the end of period $t$, and $O_{i,t+m}$ is the actual of value for period $t$+$m$.

## FORECASTING EXPERIMENT

The three methods of interest here are used in the following process: Assume we are using 164 weekly observations for each of the 23 clinics. Denote the total observations for this clinic as $T_p$ and divide these 164 observations into three groups:

Group 1 data: $T_p - 24$ = first 140 observations.

Group 2 data: $T_p$-23 to $T_p$-12 = observations 141 to 152.

Group 3 data: $T_p$-11 to $T_p$ = observations 153 to 164.

### *Univariate Procedure*

Step I: Fit models to Group 1, in-sample data of 140 observations.

Step II: Forecast Group 2 withheld data and then calculate the out-of-sample RMSEs as shown in equation 4 and SAPEs as shown in equation 5 for each of the 42 different univariate models of Table 2.

Step III: Using the model with the lowest forecast RMSE from Step II for each clinic, forecast the values of the Group 3, out-of-sample data.

Step IV: Replicate Steps I to III for two different overlapping time periods, one with $T_p$ = 164 and the other with $T_p$ = 187, and report the error measures in the tables and figures below.

### *Multivariate Procedure*

Step I. Fit multivariate models to the first $T_p - 12$ observations using the stepwise method discussed in the previous section, "Multivariate Autoregressive Models."

Step II: Forecast Group 3 withheld data and then calculate the out-of-sample RMSEs as shown in equation 4 and SAPEs as shown in equation 5 for each of the models fitted in Step I.

Step III: Replicate Steps I to II for two different overlapping time periods of $T_p$ = 164 and 187 and report the error measures in the tables and figures below.

### *Multivariate Model Parsimony*

To ensure that multivariate models were not overfitted when using the forward stepwise procedure, the significance levels of entering and exiting variables were set at .03 and .04, respectively, which is considerably lower than the SAS default values of .15. We found the numerical reduction of the significance levels to be necessary to maintain forecasting accuracy (as opposed to in-sample fit accuracy). For the 23 clinics of this study, the number of selected predictor variables (excluding the autoregressive error term) varied from 2 to 6 as shown in Table 4. For example, 56.5% of the models had two or three predictor variables with the median number of three; only one relationship had six predictor variables.

**TABLE 4:** Variables in 23 Multivariate Models*

| Number of Variables | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Frequency | 4 | 9 | 7 | 2 | 1 |
| Percentage | 17.4% | 39.1% | 30.4% | 8.7% | 4.3% |

* not inclusive of first order autoregressive error terms

## RESULTS

Table 5 and Figure 2 report the RMSEs of Group 3 data where the minimum error statistics are shown in bold. As shown in Table 5, the combined model has the lowest median RMSE as well as the lowest statistics for all other percentiles. The superiority of the combined method over the other two methods is obvious.

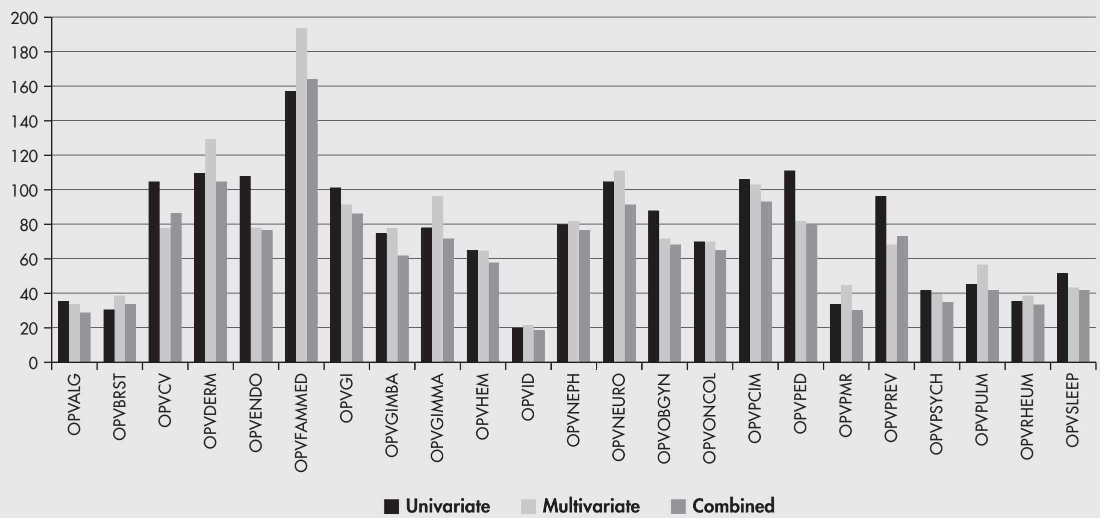
Figure 2 provides the forecast RMSEs for each of the 23 clinics of this study. As shown, the forecasts of the univariate and multivariate methods are more volatile than those of the combined methods. Despite their volatility, at times these forecasts are the most accurate for a specific clinic. By generating and reporting all three forecasts, individual clinic managers might be able to develop heuristics to better incorporate the information of a preferred method for his or her department.

Because of heteroscedasticity between clinics, few (if any) valid statistical significance tests can be performed on the results of Table 5 and Figure 2. The number of physicians and clinicians varies greatly by clinic, yielding great differences in their OPV means and variances. To better judge model accuracy, the symmetric median absolute percentage error (SMdAPE) has become a statistic

**TABLE 5:** Root Mean Squared Errors (RMSE) for Out-of-Sample Forecasts

| | Univariate | Multivariate | Combined |
|---|---|---|---|
| Mean | 76.8 | 75.2 | **66.8** |
| 5 Percentile | 22.9 | 25.6 | **21.9** |
| 25 Percentile | 40.7 | 45.3 | **38.9** |
| Median | 74.0 | 68.9 | **66.3** |
| 75 Percentile | 101.6 | 87.1 | **86.1** |
| 95 Percentile | 132.9 | 142.1 | **118.0** |
| Clinics | **23** | **23** | **23** |

**FIGURE 2:** Forecast RMSE for 24 Forecasts at Each of 23 Clinics




often used when judging the accuracy of different methods across different time series (Makridakis and Hibon 2000). Table 6 illustrates the accuracy of the three methods based on the SMdAPE.

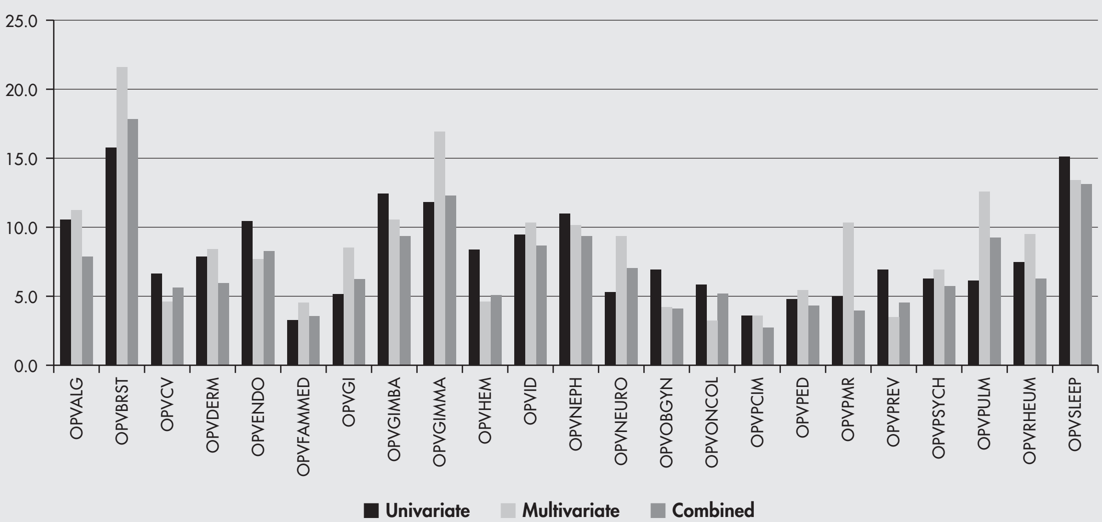
Figure 3 illustrates the SMdAPEs for each clinic for the out-of-sample Group 3 data. As shown by the means and percentiles of each of the 23 clinics in Table 6, the combined method is, on average, the most accurate. Also as shown in the median row of Table 6, the combined method has a SMdAPE of only 6.8%, while the univariate and multivariate methods have SMdAPE of 7.6% and 7.8%, respectively. Additionally, the combined method has the lowest SMdAPE for all percentiles except the 95%.

In addition to the above comparisons, it is useful to compare the RMSEs of the univariate to those of the multivariate. As shown in Table 7, the univariate RMSE is lowest for 41.3% of the forecasts, and the multivariate is lowest for 58.7%, when considering only these two methods. Also shown in Table 7, when the univariate and multivariate forecasts are combined into a single forecast by simply averaging the forecasts of each, the following occurs: The univariate method is best for 26.1% of all forecasts, the multivariate is best for 30.4% of forecasts, and the combined method is best for 43.5% of forecasts. Thus, on average, the combined method works better than either the univariate or multivariate method. We

**TABLE 6:** Symmetric Absolute Percent Error for 23 Clinics

| | Univariate | Multivariate | Combined |
|---|---|---|---|
| Mean | 8.0 | 8.8 | **7.3** |
| 5 Percentile | 3.4 | 3.2 | **2.7** |
| 25 Percentile | 5.1 | 5.0 | **4.7** |
| Median | 7.6 | 7.8 | **6.8** |
| 75 Percentile | 9.6 | 11.4 | **8.6** |
| 90 Percentile | 13.1 | 13.9 | **10.6** |
| 95 Percentile | 16.0 | 19.4 | **17.0** |
| Clinics | **23** | **23** | **23** |

**FIGURE 3:** SMdAPE for 24 Forecasts at Each of 23 Clinics




experimented with some other methods of combining forecasts (e.g., inverse weighted averages and weights based on neural network weights) without finding a more accurate way of combining forecasts (DeLurgio 1998, Makridakis and Hibon 2000).

## CONCLUSIONS

The lower RMSEs and SAPEs of the combined method reported in Tables 5 and 6 denote that the combined method can make schedules of clinical staff more accurate than either the univariate or multivariate methods. If the system were used to schedule clinicians to achieve a desired patient percentage fill rate using the RMSE, the mean RMSE of the combined forecasts of Table 5 was 66.8, while those of the univariate and multivariate methods were 76.8 and 75.2, respectively. These differences yield ratios of the combined mean to the univariate mean and the combined mean to the multivariate mean of .87 (66.8/76.8) and .89 (66.8/75.2), respectively. Thus, in the use of statistically determined clinician schedules, the level of “buffer clinicians” scheduled with the combined forecasts will be about 87% of those using univariate methods and about 89% of those

**TABLE 7:** Percentage of Times RMSE was Lowest

| | 2 Method Percentages | 3 Method Percentages |
|---|---|---|
| Univariate | 41.3% | 26.1% |
| Multivariate | 58.7% | 30.4% |
| Combined | N/A | 43.5% |

using multivariate methods. An 11% decrease (1-.89) in resources may seem modest; however, it is quite large considering that the vast majority of variable costs in health care clinics are clinician times. (The reductions in "buffer clinicians" from the combined method might be even higher depending on the underlying distribution of errors when using statistical inventory control methods. See Sridharan and LaForge 1989 or DeLurgio and Bhame 1991).

As in the M3 competition (Makridakis and Hibon 2000), combining methods, on average, outperforms specific methods such as the multivariate and univariate ones of this study. Despite the consistency of this principle, few applications of combined methods have been reported. In this application, the multivariate method performed better than the univariate method. This is true because its predictor variables and their lags provide causal explanations as well as trends and seasonality, a general characteristic of dynamic regression models. Specifically, the information provided by past SOP and OPV variables yield relationships that are useful in these predictions. Also, in the stepwise procedures, we found the use of more stringent entering and exiting significance levels to be essential in achieving parsimonious multivariate models. Although stepwise procedures have shortcomings, their intelligent use with correctly identified predictor variables in this application resulted in an effective forecasting method. In addition, while not developed here, stepwise procedures provided results that were comparable to interactive, analyst-intensive methods.

This study has several limitations. First, these results are based on only 24 forecasts for each clinic. We might ask whether these results will persist when another 24 or more weeks are analyzed. While several approaches were used to preselect clinic-specific univariate or multivariate models before forecasting, none outperformed the combined method. The performances of the univariate or multivariate methods may increase as the database of OPVs increases over time. As data increases or improvements are made in the univariate and multivariate methods, there is no guarantee that the combined method would remain superior.

After completion of this forecasting system, forecasts of OPVs for each division were made available via a central Web site at the Mayo Clinic. One hour of an operations research analyst's time and one hour of a financial analyst's time were needed each week to run the system. From initial committee meetings, users of the system confirmed that it offered value and improved forecasts. However, at this time, the system is temporarily on hold due to staff turnover in the areas supporting the system. Nevertheless, enthusiasm about the system from the Department of Neurology (Neuro), and from three divisions within the Department of Medicine—Cardiovascular (CV), Endocrinology (Endo), and Gastrointestinal (GI)—has led to efforts to further refine it. As technical personnel are replaced in central support areas, it is anticipated that the system will be further implemented and thereby continue to provide advantages and value.

## REFERENCES

Armstrong, J.S., ed. 2001. *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers.

Berenson, R.A., T. Bodenheimer, and H.H. Pham. 2006. Specialty-service lines: salvos in the new medical arms race. *Health Affairs* 25 (5): 337-343.

Centers for Disease Control. 1995. Trends in length of stay for hospital deliveries – united states, 1970-1992. *Morbidity and Mortality Weekly Report* 44 (17): 335-337.

Cerrito P.B. 2005. Visits to the emergency department as transactional data. *Journal of Healthcare Management* 50: 389–395.

Champion R., L.D. Kinsman, G.A. Lee, and K.A. Masman. 2007. Forecasting emergency department presentations. *Australian Health Review* 31 (1): 83.

DeLurgio, S.A. 1998. *Forecasting Principles and Applications.* Burr Ridge, IL: McGraw Hill Irwin.

DeLurgio, S.A. and C. Bhame. 1991. *Forecasting Systems for Operations Management.* Burr Ridge IL: Dow Jones Irwin.

Farmer R.D.T. and J. Emami. 1990. Models for forecasting hospital bed requirements in the acute sector. *Journal of Epidemiology and Community Health* 44: 307–312.

Helmer, F.T., E.B. Oppermann, and L.D. Suver. 1980. Forecasting nursing staffing requirements by intensity-of-care level interfaces. *Linthicum* 10 (3): 50.

Hussain, S., R. Harrison, J. Ayres, S. Walter, J. Hawker, R. Wilson, and G. Shukur. 2005. Estimation and forecasting hospital admissions due to influenza: planning for winter pressure, the case of west midlands, UK. *Journal of Applied Statistics*, 32 (30): 191-205.

Jones S.A. and M.P. Joy. 2002. Forecasting demand of emergency care. *Health Care Management Science* 5: 297–305.

Kao, E.P.C. and F. M. Pokladnik. 1978. Incorporating exogenous factors in adaptive forecasting of hospital census. *Management Science* 24: 1677–1686.

Kwon, I.W., K. Eickenhorst, and J. Adams. 1980. Short-term patient census forecasting models. *Hospital Financial Management* 34 (11): 38.

Lin, W.T. 1989. Modeling and forecasting hospital patient movements: univariate and multiple time series approaches. *International Journal of Forecasting* 5: 195–208.

Littig S.J. and M.W. Isken. 2007. Short term hospital occupancy prediction. *Health Care Management Science* 10: 47–66.

Mackay, M and M. Lee. 2005. Choice of models for the analysis and forecasting of hospital beds. *Health Care Management Science* 8: 221–230.

Makridakis, S. and M. Hibon. 2000. The M3-competition: results, conclusions and implications. *International Journal of Forecasting* 16 (4): 451-476.

Makridakis, S., S.C. Wheelwright, and R.J. Hyndman. 1998. *Forecasting: Methods and Applications.* Hoboken, NJ: John Wiley and Sons.

Murray, M. and D.M. Berwick. 2003. Advanced access: reducing waiting and delays in primary care. *Journal of the American Medical Association* 8: 1025-1040.

Proudlove N.C., S. Black, and A. Fletcher. 2007. OR and the challenge to improve the NHS: modeling for insight and improvement in in-patient flows. *The Journal of the Operational Research Society* 58 (2): 145-159.

SAS Institute Inc. 2004. SAS/ETS@ software: time series forecasting, version 9.1, Cary, NC: SAS Institute Inc.

Sridharan V. and R.L. LaForge. 1989. The impact of safety stock on schedule instability, cost and service. *Journal of Operations Management* 8 (4): 327-347.

Tandberg, D. and C. Qualls. 1994. Time series forecasts of emergency department patient volume, length of stay, and acuity. *Annals of Emergency Medicine* 23: 299–306.