FORECASTING THE DEMAND FOR NURSES IN NEW JERSEY

Geri L. Dickson, Ph.D., RN

March 2002

Introduction. In order to develop a dependable, quantitative means to anticipate and forecast requirements for the New Jersey nursing workforce, we convened a Work Group to determine the data needed and the data available. A labor economist provided consultation to the Work Group. Together, we reviewed the most recent version of the Nurse Demand-Based Requirements Forecasting Model developed for the Division of Nursing by Vector, Inc. Initially, we attempted to replicate this model at the state level, but found this impossible. We then decided to develop our own Nursing Demand Forecasting Model for New Jersey.

The New Jersey Model is an econometric, longitudinal, multiple regression model that forecasts the demand for RNs and LPNs at the state and the county level. Demand is an economic term that is defined as the number of nurses that employers would hire, given their availability. Although we have forecasts at both the state and county level for RNs and LPNs, only the state forecasts for RNs will be presented here.

Two different nurse employment databases were used to derive forecasts of the demand for nurses. In both cases, the datasets are longitudinal, in that they follow each of New Jersey's 21 counties through the years 1986, 1990, 1994, and 1996. In addition, we obtained data about factors that can influence the demand for nurses, such as health statistics, HMO penetration, economic indicators, salaries, and population demographics. The Work Group identified those factors that in their expertise they believed would affect the demand for nurses. Available databases that best matched those identified were sought and obtained. These databases were also available for the 21 New Jersey counties for the years 1986, 1990, 1994, and 1996.

The first nursing database consists of data for the total employment of RNs and LPNs in all sectors combined (e.g., hospitals, community/public health, long term care, etc.) in each of New Jersey's 21 counties in 1986, 1990, 1994, and 1996. These data were obtained from the New Jersey Department of Labor. The data on total employment of nurses are somewhat "soft," in the sense that they are based partly on surveys and projections rather than on "nose-counts" alone. On the other hand, these statewide data do provide complete coverage of all nurse employment in New Jersey.

The second nursing database is derived from the American Hospital Association's (AHA) survey of New Jersey hospitals. The AHA data provide information on employment of the number of full time equivalent (FTE) positions for RNs and LPNs in each of New Jersey's 21 counties, but only for the acute care sector for the same four years: 1986, 1990, 1994, and 1996. These data are collected annually and are quite reliable, but they do not provide information about nurse employment outside of the acute care sector.

Next, for both the statewide and AHA data, we constructed a panel or longitudinal database following nursing employment (in the total or acute care sector[s], respectively) in New Jersey's 21 counties over time. As described in more detail below, we first analyzed each database and then used the results of our analyses to construct forecasts for nurse demand.

Perhaps the most important feature of our analyses and forecasts is that they are based on true longitudinal (sometimes called panel) data for each of the 21 counties for each of the four years. Using these data, one can hold constant many differences across counties that cannot adequately be taken into account in analysis of a single cross-section. In particular, these data make it possible to see how each individual county's employment of RNs and LPNs changes over time, which is what is needed in order to forecast how employment will change in the future. Most important, having the four sets of data (1986, 1990, 1994, and 1996) for each of the New Jersey 21 counties give us 84 data points, the sample size.

The method is illustrated in the next two sections. It begins with a discussion about the basic approach in terms of our statewide database (total RN demand) and continues in the second section in terms of our approach with the acute care RN demand database. Tables with detailed forecasts of our analyses are included. As these data become available for additional years, the new data for nurse employment and the independent variables for each county will by added to the databases. Likewise, to the extent that it is possible to add new variables to our forecasting equations, we will do so.

Analysis and Forecasts Based on the Statewide Data. The analyses based on the statewide data refer to the total employment of RNs in all sectors in each of New Jersey's counties for 1986, 1990, 1994, and 1996. In addition to information on nurse employment, we used measures of each county's health factors, HMO penetration, economic indicators, nurse salaries, and population demographics for the independent variables. In several steps, we used these data to construct forecasts for the nursing demand reported here for 2006.

First, we used regression analysis techniques to determine the relation between RN or LPN employment, the dependent variable or outcome of interest, and the independent variables above. The data in the regression analyses refer to all 21 counties in each of the four years. Thus, this constitutes a moving cross-section or longitudinal database.

In order to determine the predicted value of the independent variables, we made some assumptions about predicted value of each variable based on past patterns and current information. For example, based on the trends in HIV/AIDS rates between 1994 and 1996 and in relation to the forecasted population for 2006, we assumed that the total number of HIV/AIDs would increase by 7% between 1996 and 2006. Similarly, to forecast for the rate of HMO penetration for each county, we estimated a logistic non-linear regression model using data from 1986 to 1996. We set an upper limit of 75% for the HMO penetration based on our assumption that the HMO penetration rate would never be 100%. Using these results, we estimated the HMO penetration rates for 2006.

For the personal per capita income, the variable we used was the specification of the natural logarithm of the deflated (dollars reduced by the CPI) personal per capita income for each county. To forecast this variable we assumed that between 1996 and 2006 the "real" personal per capita income would grow at the same rate as between 1986 and 1996. For other variables, e.g., employment rate and population growth, the forecasts were used from the New Jersey Department of Labor.

Each set of regression results provides a formula or equation for use in forecasting: simply plug in the predicted value of each independent variable, and compute the implied or predicted value

of the dependent variable – i.e., nursing employment – by multiplying the assumed value of each independent variable times its regression coefficient, and then cumulating the results. Table 1. presents the basic regression results of the significant variables. The county data available for the following independent variables were not statistically significant and, therefore, not included in the following table: the per capita income, the mortality rate, and nurses' salaries.

Table 1. Dependent Variable – Total Employment of Registered Nurses in All Sectors

Variable	Unstandardized Coefficient	Std. Dev.	T-statistic	P-value
Constant**	2.743	0.892	3.073	0.002
HIV/AIDS rate**	0.002	0.001	3.908	0.000
Employment/Population Ratio**	3.891	1.315	2.958	0.003
HMO Penetration Rate*	0.021	0.012	1.807	0.071
Population Over 65 Rate**	0.075	0.036	2.099	0.036
More or Less Medical Facilities **	2.280	0.313	7.294	0.000

R square	0.699
Adjusted R square	0.680

**significant at the 95% confidence level
* significant at the 90 % confidence level

These results indicate that the RN total New Jersey workforce will grow by 17.2% from 1996 to 2006. In numbers, this represents a demand of almost 11,000 RN positions in all sectors of employment. This compares to a 25% increase in the demand for total RNs from 1986 to 1996.

Further, one can conduct "what if" exercises for each of the individual factors considered in the regressions. For example, in the logarithmic regression, the coefficient of the variable of the ratio of employment to the population is 3.891 (see table 1.). That variable is positively and significantly related to the demand for nurses. If the employment to population ratio does not increase as forecasted, it can significantly affect the demand for nurses. Thus, a "what if" exercise can pose the question: What if the ratio of employment to population decreases, then what effect would the decrease have on the demand for RNs? If the ratio of employment to population decreases by 1%, instead of the projected increase of 2.76%, this would translate into a decreased growth of RN employment of about 1.76 x 3.891 = 6.86 percentage points, i.e., the percentage decrease of the change multiplied by the coefficient of the variable. Thus, under our revised assumption about growth in the employment/population ratio, the revised forecast is calculated by taking the forecasted percentage of growth, 17.2%, minus the result of the above equation, 6.86%: 17.2 – 6.86 = 10.34. This represents a change in the RN employment of 10.34% rather than the previous 17.2%. This forecasts an increase of 6576 additional RN positions, instead of the previous forecast of almost 11,000 additional RN positions.

It should be noted that the "what if" exercises, relating to the effect of a change in some individual variables, are necessarily subject to a relatively large "margin of error," because the individual regression coefficients have relatively large standard errors. This problem, however, should be alleviated, as the additional years of data we are now adding are included.

Analyses and Forecasts Using the AHA Data. The method underlying our analyses and forecasts based on the AHA data on the employment of nurses in acute care is similar to the method used in performing the analyses and forecasts based on the statewide data. The basic AHA database

gives employment of FTE RNs and FTE LPNs in each of New Jersey's hospitals in each of the 21 counties. The first step was to aggregate these data up to the level of the county. This made it possible to construct a database for acute care employment of FTE RNs and FTE LPNs for the years 1986, 1990, 1994, and 1996 in each of New Jersey's 21 counties. This aggregation up from the level of the individual hospital to the level of the county not only facilitates comparison with the analyses of the statewide data; it also is a natural way to deal with hospital closings, opening, and mergers.

In addition to the information on FTE nurse employment in hospitals, the AHA data also provide information on the number of hospital staffed beds, which were also aggregated up from the level of the individual hospital to the county level. We then merged into the AHA data the information on the other variables considered in our analyses of the statewide data (e.g., the HIV/AIDs rate, the HMO penetration, etc.). The number of staffed beds was not significantly related to the outcome: the demand for RNs. Therefore, three other independent variables were identified as indicative of hospital activity—birth rate, surgeries, and inpatient days – and, were added to reflect the level of hospital activity. They were all statistically significant.

In this hospital forecasting model, the above variables describing the level of hospital activity were forecast in the following way: the average rate of change in the number of births per 1,000 population between 1990-1996 was used to forecast the 2006 level of births; the average of change in the number of surgeries per 1,000 population between 1986 and 1996 was used to forecast the number of surgeries in 2006; and the average rate of the number of inpatient days per 1,000 population between 1986 to 1996 were used to forecast the number of inpatient days in 2006.

Then, each of the independent variables used in the forecasting model for the total RN employment, plus the above hospital activity variables, were entered using traditional multiple regression techniques. In this forecast for the demand of FTE positions for RNs in the acute care sector, all but the employment to population ratio and the dummy variable of more or less medical facilities in counties were significant; see table 2 for the regression statistics. In this case the R^2 is 0.881 indicating a high degree of correlation between the independent variables and the outcome – the demand for FTE positions of RNs in the acute sector.

Table 2. Dependent Variable – Hospital Employment of Registered Nurses (FTEs)

Variable	**Unstandardized Coefficient**	**Std. Dev.**	**T-statistic**	**P-value**
Constant**	-0.947	0.415	-2.278	0.023*
HIV/AIDS rate**	0.001	0.000	4.180	0.000*
Employment/Population Ratio**	-0.543	0.732	-0.742	0.458**
HMO Penetration Rate*	0.024	0.006	4.304	0.000*
Population Over 65 Rate**	0.064	0.017	3.743	0.000*
Birth Rate*	0.093	0.018	5.103	0.000*
Surgery Rate*	0.010	0.003	3.753	0.000*
Inpatient Days Rate*	0.853	0.162	5.278	0.000*

R square	0.861
Adjusted R square	0.848

**significant at the 95% confidence level
* significant at the90 % confidence level

These results indicate a growth in the demand for RNs in New Jersey hospitals of nearly 15% or about an additional 4,400 RNs. This brings the FTE RN workforce in hospitals to 34, 070. The percentage of growth in demand for the previous 10 years, 1986 to 1996, was 24%, representing almost 5,800 additional RN full time positions.

Multiple regression analyses allow researchers to address important questions about the relationships among a dependent variable and multiple independent variables. In this case, the dependent variable was the demand for RNs in New Jersey, both on a statewide and at the hospital levels. The above tables give the statistics generated in the forecasts. However, the major value of a forecasting demand model lies in the ability to relate nurse demand to the corresponding supply data. To incorporate the two, supply and demand, we have collected primary supply data, as well as projections from the Division of Nursing Supply Forecasting Model. When exploring the relationship of New Jersey nurse supply and demand, we determined that unless strategic actions are taken, these data indicate a shortage of nearly 14,000 RNs or a vacancy rate of 18% by 2006.

Conclusions. This econometric, longitudinal, multiple regression model does provide forecasts that are useful in anticipating the number of RNs and LPNs demanded in the future. In this scenario, we used data for the years 1986 to 1996 to forecast the demand for 2006 in the total RN population in the state, as well as in the FTE RN demand in acute sector only.

However, any forecasting regression model relies on existing databases to predict the future. Therefore, it is as reliable as the databases that it relies on. Currently, there are no databases on a county and state level for New Jersey that identify the educational levels of nurses or for other individual sectors other than the acute care.

Forecasting represents one model of predicting the future. The business guru Peter Drucker suggests another: “The best way to predict the future is to create it.” We, in Colleagues in Caring, are doing both: predicting and creating!