Chapter 3

Data Collection and Analysis

Although Grace Hospital collects a large amount of information on its day to day operations there is little central coordination of this data. Each department collects any utilization data it needs, but often the form of this data varies from department to department making comparisons difficult. For example, Postpartum counts the number of days the treatment rooms are used and Antepartum calculates total patient days. The Management Information Systems (MIS) department at Grace has implemented a system that stores some of the most important information on computer, but at present it is a labourious task to obtain anything other than simple summary statistics. To create the Grace Hospital simulation model it was necessary to combine two data files— one containing patients’ personal statistics, medical history and diagnosis, and the other file containing a log of all transfers through the hospital. Transfers entered into the computer are assumed to take place at the present time. However, in practice transfers are often not entered until a later time when the unit clerk is free. Also, the computer is down every day from 2 a.m. to 4 a.m. As a result, the patient transfer data was of poor quality and needed to be corrected. Handwritten records were superior because of the greater flexibility of entry. The computer is operated only by the unit clerk, whereas anyone can fill out the handwritten log. At present, the partial computerization causes some duplication in the data collection. Once a reliable and flexible computer system is in place many of the current problems in data collection will be reduced. When compared with the handwritten records, about 20% of the patients in the computer log were found to
require some modification. As a result of this time consuming correction, only 2 months of data were used to create the model, namely July and August of 1988. This amount of data reflected over 1200 births and was enough to create the model. However, there is an implicit assumption in using this data to create a model. The length of stay (LOS) and transfers are assumed to reflect the ideal operation of the hospital, i.e. they are assumed to not be subject to capacity constraints. Fortunately July and August 1988 were relatively slow months. Therefore, the limited capacity effect is minimized in the data.

The first step in analysing the data was to determine which characteristics of the patients affect their transfers and length of stay in the hospital. Brooks et al. [4] using multiple regression and 117 factors to group patients predicted bed needs. Many studies, especially those concerned with generating cost estimates, group patients according to diagnosis related groups (DRG). DRGs were developed by Fetter to produce groups of patients homogeneous with respect to some measure of resource usage. More recently, Schachtman et al. [26] suggested a case mix adjustment index for increased flexibility. After consulting with Dr. Effer, head of Maternal and Fetal Medicine at Grace, the following factors were considered likely to be important. For adults, antepartum (AP) complications, delivery type, postpartum (PP) complications, location of residence, and whether or not a sterilization was performed were considered potentially significant. For babies, the characteristics examined were birth weight, gestational age, APGAR score (a medical measure of newborn health), and whether the baby is a newborn or a pediatric patient. It was also of interest to determine which of the adult characteristics are correlated, and which characteristics are the best predictors of their baby's health. To illustrate the manner in which the categories were chosen, consider AP complications. Dr. Effer initially suggested 9 major categories of AP complications. After examining the data it became clear that 2 of the categories, renal, and multiple births, were too small
and needed to be combined with the "Other AP complication" category. In addition, a reworking of the "no significant complication", "other complication" distinction was needed due to the large number of minor complication initially classified as "other". In the following analysis the Cramér's V is used as a measure of the degree of association between characteristics. Cramér's V is based on the chi-square value, but minimizes the influence of sample size and degrees of freedom. As a result, it can be used to compare the strength of association of various characteristics. Cramér's V always lies between 0 and 1, with values close to 1 showing a strong association.

3.1 Adults

Since the preregistration system at Grace (see section 2.4 for more information) differentiates between patients who deliver and those admitted only for AP or PP complications it was best to generate arrivals in the model by delivery type. Examining the arrival pattern (time of day Table A.16, day of week Table A.17) it is evident that day of surgery (DOS) patients have a special arrival pattern, whereas all the other patients arrive uniformly throughout the day and week. The DOS patients arrive for scheduled caesarean sections, and they usually arrive Monday to Thursday between 6 a.m. and 2 p.m. A chi-square test strongly suggests that the interarrival time of all the patients, except the DOS mothers, is exponential. For each delivery type except DOS, the hypothesis of similarity could not be rejected at the 0.20 level. This duplicates past results [31],[2],[32],[7]. In other words, this means the arrival rate is a Poisson process, and arrivals are random events. This is a reasonable conclusion since the delivery process is very unpredictable, and labour does not wait until a convenient time to start.

Of the five factors used to differentiate adult patients, delivery type was the best predictor of the total LOS in the hospital (see Tables A.20 to A.22). The Cramér's V for
delivery type was 0.4, whereas for AP complication, the next best predictor, it was 0.16. However, when making a physical distinction in Grace between the antepartum, delivery suite, and postpartum areas, the situation was more complicated. The AP complication had the strongest association measure with AP length of stay (Tables A.23 to A.27). AP complications, preterm labour, AP hemorrhage, and diabetes tended to have particular long lengths of stay. However, there is also a significant difference in AP LOS between undelivered and delivered patients (Table A.24). In addition, patients from B.C. also had longer AP lengths of stay, but this can be explained by their higher incidence of AP complications. In the delivery suite, the delivery type is clearly the best predictor of LOS (Cramér’s V of 0.32 versus 0.15 for AP complication, Tables A.28 to A.32). Patients who have elective C-section stay the least time (since they arrive for scheduled surgery), and patients who have emergency C-section stay the longest. In postpartum, delivery type has the strongest association with PP length of stay (Cramér’s V of 0.40, tables A.33 to A.37). However, PP complication also has an effect (Table A.36). Transfers among the 3 areas of the hospital are also affected by the classifications. For example, patients with some AP complication are far more likely than those with no complications to enter the AP. In the delivery suite, almost all vaginal births occur in low or high risk, whereas C-sections all deliver in the OR, and patients with a PP complication are more likely to visit the OR after delivering. Sterilizations are much more common for patients having an elective C-section or an instrumental birth (see Table A.43). But, since all C-section patients already have to recover from surgery (sterilizations are performed at the same time as the caesarean) sterilizations have little effect on LOS (Table A.22). For vaginal deliveries, the number of sterilizations is also reflected in the proportion of patients who transfer to an operating room after delivering.

The factors used to classify adult patients are quite correlated. A patient’s delivery type can be used as a predictor of all the other factors (Tables A.40 to A.43). A patient’s
geographical residence is correlated with her delivery type and AP category, but does not influence sterilizations or PP complications (Tables A.44 to A.46). Her AP category is correlated with her delivery type and geographical residence, but not with whether she will have a sterilization or a postpartum complication (Tables A.47 and A.48). Finally, whether a patient has a PP complication or not does not influence whether she will have a sterilization (Table A.49).

Discharge times of the adult patients was not random (Tables A.38 and A.39). The majority of discharges occur during the day, most from 11 a.m. to 3 p.m. This is no surprise since patients usually leave at their convenience.

3.2 Babies

The analysis of the data collected about babies was more straightforward. They were classified by their health. Because the Sick Children’s Nursery (the tertiary care nursery) is in another hospital, the most seriously ill babies are immediately discharged from Grace. To avoid this very short LOS from skewing the results, a special category of newborns was needed, namely “Short Stay” (less than 2 hours in Grace). In addition, a separate category for pediatric babies was needed. Pediatric patients have both a different arrival procedure and a much longer average LOS than newborns. (see Table A.50). Pediatric babies tend to stay longer because they must be ill to require re-admission. The remaining analysis used standard medical divisions of the health characteristics. Birthweight was divided into two classes: greater than 2500 grams and less than 2500 grams, gestational age was divided into two classes: greater than 37 weeks and less than 37 weeks, and the APGAR scores were divided into three categories: 0-3, 4-6, 7-10. All these health characteristics are highly correlated. For example, low birthweight babies tend to result from deliveries at an early gestational age, and also have a low
APGAR score. However, the analysis shows that birthweight (Cramér’s V of 0.21), and gestational age (0.27) are much better predictors of LOS than APGAR score (0.11). These are results are shown in Tables A.51 to A.53. This result was confirmed when presented at a meeting of the nurses’ executive. The nurses felt that the APGAR score was a very subjective measure and of little value. As a consequence, a combination of the birthweight and gestational age (combining small categories) were used to predict the LOS for the non “Short Stay” newborns. (Table A.54). All newborns are admitted at delivery, but pediatric babies arrive on their own. The pediatric babies’ inter-arrival time distribution fits an exponential distribution very well (like the adult patient’s arrival distribution).

After determining which baby health classifications to use it was important to determine which aspects of a mother could best predict their baby’s health. Trying both AP category (Cramér’s V of 0.35) and delivery type (Cramér’s V of 0.05) it is obvious that the AP category is by far the better predictor (see Tables A.55 and A.56).