INTRODUCTION

Medpar Data is used for billing Medicare; it typically is “packed” into different sheets of data. The first step required to use the data is to “unpack” it into its component sheets and then to merge the necessary data. Once that is finished, we can examine patients in sequence to investigate the total care of wounds, especially patients who have diabetic foot ulcers. This particular dataset was received from a clinic that specialized in the treatment of wounds. It contains detailed diagnosis information for inpatient and outpatient care. One of the problems is that there are two different coding sets that are used in the Medpar data. Inpatient data are coded using the ICD9 codes; outpatient data are coded using HCPCS codes.

The advantage of such a dataset is that the wound care is highly concentrated within it, so that we can focus and drill down into wound care. The disadvantage is that the different coding systems have to be matched for analysis purposes. In this section, we want to look at the type of outpatient care prior to inpatient treatment, and to see if there is a progression of treatment that leads to better patient outcomes.
BACKGROUND

A detailed study of diabetes foot ulcers and osteomyelitis concluded that physicians do not use antibiotics in the proper dose and the proper duration, often resulting in the recurrence of the infection. This earlier study examined 2 years of data from the Thomson MarketScan database. (Zahedi, 2007, 2009) This database contains information from 100 different insurers and approximately 40 million individuals. However, antibiotics remain a recommended treatment for the infection, although not used as often in practice as recommended in guidelines. (Concia, et al., 2006; Mader, Norden, Nelson, & Calandra, 1992; Wimalawansa & Wimalawansa, 2008) Nevertheless, surgical management has become more common and has been suggested as the optimal treatment. (Henke, et al., 2005)

The CPT-4 (Current Procedural Terminology, 4th Edition) coding system is most prevalent. CPT-4 codes are five-digit numeric codes. ICD-9-CM procedure codes are found on hospital claims. These codes are three to four digits in length and are all numeric. There is an implied decimal between the second and third digits. HCPCS (HCFA Common Procedural Coding System) procedure codes are found less often than CPT and ICD procedure codes in the MarketScan data. These codes are five digits in length. The first character is alpha; all other characters are numeric. CPT-4 codes are available online at www.ama-assn.org. An alternate website is www.docofficerx.com. Both of these sites allow for a trial use of the code database. The remaining HCPCS codes are available at http://icd9cm.chrisendres.com/icd9cm/. However, since we are working with data that are 2005 or older, we want codes that are related to that time period. To do this, we found an inexpensive, used 2005 CPT coding book. These codes are not as readily available because the CPT-r coding system is proprietary to the American Medical Association. Therefore, they are not posted on the internet. However, older copies of the CPT coding book are available in the used book market through amazon.com.

METHOD

Medpar gives the billing information for Medicare. It is publicly available, although there can be a small cost affiliated with data acquisition. In fact, many of the public companies that define patient severity indices rely upon Medpar because it is so readily available. In addition, The Centers for Medicare and Medicaid maintain a data warehouse consisting of patients with a defined chronic disease, and this information, too, is available for a fee. There are multiple forms in Medicare data and we provide the SAS code on how to “unpack” the different forms for use in analysis. We are using the 2005 version of the data. The preprocessing to combine the data was given in Chapter 3. We need to unpack the different sheets of the Medpar data in order to use the information. Once that is done, we can first focus on the inpatient datasets.

It is our intention to examine the data using a variety of techniques. These include market basket analysis and survival data mining.

Market Basket Analysis

We will examine the data using Market Basket Analysis, or association rules. (anonymous-market, 2009; Berry & Linoff, 2004) We want to determine which treatments are associated with each other during both inpatient and outpatient treatment. Market Basket Analysis is a modeling technique based upon the theory that if a patient has treatment A, then the patient is also likely to have treatments B, C, and/or D. For example, in wound care, amputation may be linked to debridement with debridement used in an attempt to preserve the limb.

The set of related items is referred to as an itemset, and the purpose of market basket analysis is to find relationships between purchases. Typi-