

CHAPTER 1

Health Care Data

Introduction	1
People and Health Care	2
Recipients/Users/Patients	2
Providers.....	3
Health Care Language	4
Diagnoses	4
Treatment Protocols.....	5
Combinations of Diagnoses and Treatments	5
Types of Administrative Files.....	5
HCFA-1500	6
Pharmacy	6
UB-92	6
Claims and Encounters	7
Conclusion.....	8

Introduction

From birth certificate until death certificate, a person accumulates health care data. Each time a person has contact with a health care professional some type of record is produced, be it paper, electronic, or both. In one year, a normal healthy American may visit the dentist twice, see a doctor for a wellness visit, call a physician for a bad cold, and refill a dozen allergy prescriptions. Each one of these encounters becomes a part of the many records stored somewhere in the archives of health care data. Much of the data is merely entered into a computer system so that the provider can receive payment, and then the data are added to a massive “warehouse.” Each one of these records can provide insight not only into an individual's well-being but may, in fact, affect health care across similar groups based on age, gender, or location, and provide an analytic base to determine insurance rates, study disease trends, and modify treatment protocols.

This administrative type of information — claims and encounter, eligibility and enrollment — is still vastly underused. This is due in part to the massive amounts of administrative data, in part to their complexity, and in part to their misuse in the past. Studies of this type of data are only now beginning to receive proper time and attention. The process of converting health care data into useful information is the focus of this book.

Where does one begin to learn about and work with these data?

2 *Health Care Data and the SAS System*

Analysis of health care data without an awareness of the accompanying peculiarities, eccentricities, and limitations is risky business. A possibility always exists for variability and discontinuities in medical coding practices and in personal, demographic, time, and place variables. This is further complicated by the passage of time, by regional variances, by similarities and differences within and among demographic groups, and by vast numbers of treatment protocols and practices. Still, there are some basic clues that help solve the puzzle.

In a perfect world, accurate, standardized health care information would flow freely among patients, providers, and payers. Health care administrators, planners, researchers, and reviewers would have quick and easy access to the data necessary for their activities. Instead, health care data are collected and delivered in diverse formats, most without common identifiers. It would seem an easy task to identify each person or group with a unique code. Unfortunately, there is no one identification system either within or across recipients, providers, or facilities. Each payer organization has developed unique forms of coding. Therefore, it is difficult to link across databases, across time, and across treatment sources. This problem is faced regularly in the health care industry.

Health care data are necessary for many activities, such as studying disease patterns, analyzing treatment trends, and setting payment rates. Health care data files are large, cumbersome, and complex, but none of these problems is insurmountable. This book illustrates ways that you can use SAS to perform some of these activities.

People and Health Care

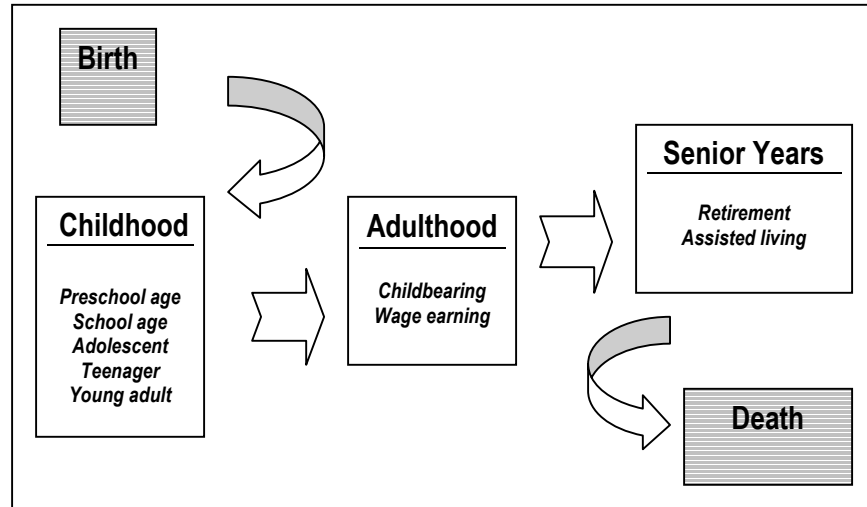
People can receive or provide health care. It is important to have a basic understanding of the similarities and differences between and among the receivers and providers of health care. This book discusses the different types of folks involved in or affected by health care. The people who receive care are called recipients, users, or patients. Providers of care can be individual practitioners, groups of caregivers (e.g., health maintenance organizations, or HMOs), or institutional facilities.

Recipients/Users/Patients

Throughout a person's life, data are being collected about health care and related costs. Over the lifetime of a typical individual, health care administrative records are created from enrollment into a selected coverage system to conditions serviced by various providers either within or outside that system. These conditions include both preventive care and treatment of illness or injury.

It is easy to follow the path of life from birth through death shown in Figure 1.1:

Figure 1.1: Life Pathway



Throughout the continuum of life, the type and level of care vary dramatically. For example, people at all ages need some type of prevention and wellness care. Preventive immunizations are found throughout childhood but may be targeted to only specific age and functional groups in adulthood. Health care workers and international travelers receive vaccinations for illnesses not often encountered by others. Acute care hospital services can be provided to people across all ages and sexes, dependent on need. Chronic care, long-term care, and hospice care services increase as a person ages but cannot be excluded from any age group.

Birth data are specific to a newborn; this is probably the only category that has no exceptions. Criteria can be set to study specific groups (e.g., pregnancy for women ages 15–44), but the possibility of either younger or older outliers cannot be excluded. Understanding the common occurrences for specific age/gender groupings can help move forward the validation of administrative data sets.

Providers

Thus far, the discussion has centered on the recipients of health care. Consideration must also be given to the variety of those providing health care services. Providers may be individuals, groups of practitioners, or institutions and facilities. Each provider group may have its own method of collecting, coding, storing, and sharing data.

4 Health Care Data and the SAS System

Physicians can service genders, all age groups, and all conditions in various places of service. In order to clearly define the physician, the patient, and the treatment, complex coding schemes have been developed. Dentists have various coding methods to identify exactly which tooth and what specific tooth surface was treated. Pharmacists have methods of collecting information on each individual prescription, including the type of drug, the quantity, costs, and co-payments. Other health care professionals, nurses, home health workers, social workers, suppliers of medical equipment, and ambulance companies also have their own terminology. Therefore, coding schemes are complicated and require unique identifiers.

Facilities may include:

- ◆ inpatient hospitals providing acute care
- ◆ chronic, rehabilitative, and psychiatric centers providing long-term care services to sicker patients
- ◆ nursing homes providing full-time care to patients unable to care for themselves

As the facilities are varied, so facility data are voluminous, complex, and often difficult to analyze. Each individual facility may have a specific set of standards and practices to define collection and storage of data. Information on surgery performed in an acute care hospital is quite different from the information collected on the care of an Alzheimer's patient in a skilled nursing facility (SNF). Most of these facilities likely have different storage and reporting formats. Add to this the various claims and utilization reporting requirements among federal and state agencies and third party payers, and the problems and complexities grow.

Health Care Language

Every community has its own language, its own dialect. The health care industry has a very special style and jargon. There are hundreds of acronyms; just read the proceedings of a meeting of health care administrators or policy analysts and the mind boggles. Moreover, there are codes, codes, and more codes. Luckily, many of these coding systems are centrally organized and well documented. The trick is to know where to find the meaning of the codes and when clinical expertise is needed for correct interpretation of study results.

Diagnoses

In order to track a patient's care, it is necessary for a provider to classify his or her condition. The simplest note in a chart indicates whether a person is healthy or unhealthy. This would have been minimal but sufficient information in the past when a single doctor birthed a baby and cared for this person throughout life. In today's society, one person may see a number of practitioners in a short amount of time and the medical chart needs to be verbose.

Lists of diagnoses are not static; new illnesses, accident types, and casualties are discovered or defined each year and, thus, classifications of these data grow. Understanding how to take this diagnostic information and manipulate it to create useful analyses is therefore quite complicated.

Treatment Protocols

Depending on the diagnosis or lack of diagnosis of a condition, a provider must define a treatment regimen for each patient. In order to receive payment for each service rendered, this treatment must be specific and readily identifiable. In general, the sicker the person is, the more likely these services are complicated and numerous. Additionally, since treatment rarely rests in the hands of only one provider, it is important that this information is readily available to all those caring for the patient.

Combinations of Diagnoses and Treatments

Studying a population or investigating a particular health care area often requires both diagnostic information and treatment or procedure information. Combining these two related data items may require the assistance of a clinician or epidemiologist who has been trained in this area. Presenting undocumented results may be dangerous.

Types of Administrative Files

This book covers the three major administrative file types:

- ◆ HCFA-1500
- ◆ Pharmacy
- ◆ UB-92

All three files include basic and necessary information, but not all file types may contain all information elements. These data include:

- ◆ patient information including name, address, some type of identifier, date of birth, gender, race
- ◆ provider information including name, address, some type of identifier, and data concerning place and type of service
- ◆ billing data including both charges and payments

HCFA-1500

Health care services provided by single practitioners or practitioner groups are entered on paper forms or in electronic format called HCFA-1500, the Health Insurance Claim Form. This is the common claim form for noninstitutional providers that was developed and is updated by the Health Care Financing Administration (HCFA) and approved by the American Medical Association (AMA) Council on Medical Services.

This format clearly identifies the patient, demographic information, payment source, the provider, and the services performed including dates, diagnoses, and type and place of service. Of course, all cost and charges are also included on this form; physicians must be paid! This paper form can then be converted to electronic format, usually storing one service with all accompanying patient and provider data on one record.

Pharmacy

Each time a prescription is filled, the data are entered into a database. This information is normally entered at the point of service (POS). It usually contains validated values. As a prescription is filled, it is often checked for drug interactions against other drugs that a person may be receiving. A drug record contains the basic information about the recipient, the provider, and the pharmacy, as well as all associated dates. In addition, there are specifics as to the drug, its formularies, the quantity, the packaging, the dosage, and the form.

Pharmacy data are relatively simple in format; one record is equal to one prescription. Because the information is entered at POS, the validity checking takes place immediately and additional checks are not normally necessary. There are certain fields that may not be valid and therefore not useful. For example, the prescription date is not always entered into the system, but rather defaults to the date the prescription is actually filled. Because the prescribing provider is often difficult to interpret at the pharmacy, a default provider identification number may be entered.

Prescriptions filled by facilities while a person is in an acute care hospital or a nursing home do not in fact create this type of drug record. This information may be included as a line item on the facility's billing form, the UB-92.

UB-92

Hospital inpatient and acute care outpatient services are submitted for payments in UB-92 (Uniform Billing) format. This uniform institutional provider bill is used for multiple third-party payers such as Medicaid, Medicare, and commercial insurance. UB-92 format is quite complex and can store vast amounts of data. UB-92 files are further complicated by payments, adjustments, denials, and resubmissions for one recipient's stay. A patient's medical record tracks each procedure, test, prescription, and other service provided during a stay. The UB-92 form and file does not include this specific type of information, but rather "rolled-up" items with charges and units entered.

Claims and Encounters

Access to health care has changed in the last decade of the twentieth century. In the past, payments to providers were simple to calculate. Practitioners were paid for each service that was provided, a practice called *fee for service* (FFS). Often, these payments were predefined by the payer organization, and fees were stored in a tablelike fee schedule. In order to receive payment, each provider would submit a bill, usually in HCFA-1500 format, for each recipient afforded a service. If the form was not prepared correctly or the data were not valid or complete, the bill was rejected and sent back to the physician's office. Insurance companies and government payers such as Medicaid or Medicare tracked these payments and thus could calculate expenses and count services.

Most people covered by FFS-type insurance select a primary care physician who handles preventive care for the patient. This provider then refers the patient for specialty care if necessary. Each service for which a payment is requested creates a claim. One HCFA-1500 service is equal to a claim line. Claims can be analyzed for many pieces of information such as costs of a specific disease or condition or rates of types of services for a particular demographic group. Because a payment is associated with each claim and, therefore, must meet certain standards, the validity and completeness of claims data are usually good.

In this new era of providing health care services, many people are assigned or choose care through HMOs or managed care organizations (MCOs). A large proportion of the U.S. population now receives health services from these organizations. Services provided by HMOs or MCOs are not paid in an FFS manner. Whether the person is self-insured or employer insured, the organization is paid a *capitated payment*. Generally defined, these are fixed payments regardless of services rendered. The payment may be per member per month (PMPM) and not dependent on services rendered. These fees are calculated by a variety of methods including demographic scales and risk-adjusted tables. This is called a *capitated payment arrangement*.

Services provided by HMOs and MCOs in one setting or timeframe are referred to as encounters. These encounters may be submitted to the organizations paying the capitation dollars. Encounter data can be used for further calculation of the capitated rates as well as quality assessment of the providers. Because the costs are not related specifically or directly to the services, it is more difficult to work with encounter data than FFS claims data.

The differences between claims and encounters throw an additional monkey wrench into the study of health care. These complexities, matched and merged with all the other variations of the data, create a massive web to unravel in order to create information.

Conclusion

It is easy for the health care analyst to become daunted by the volumes and complexities of data generated over the lifetime of just one individual. The language of health care is, at times, foreign. There are varieties of administrative data to be sifted through, filtering out the mixture of dimensions and facts. The array of data sources originating from and for a multitude of purposes is overwhelming at best.

This book provides some insight into health care data and helps with understanding the processes involved. Some of this material is simply background, while other information is included to assist in the actual reading and analysis of electronic sources of data. SAS code is used to demonstrate how to access the data in its basic format and process it through a system of checks and reports to a final endpoint of generated information.

For more information about the topics presented in this chapter, see Steinwachs, Weiner, and Shapiro 1995.