DISEASE MANAGEMENT
Volume 8, Number 5, 2005
© Mary Ann Liebert, Inc.

Using Visual Displays as a Tool to Demonstrate Disease Management Program Effectiveness

ARIEL LINDEN, Dr.P.H., M.S.,[1,2] and NANCY ROBERTS, M.P.H.[3]

ABSTRACT

With the rapid introduction of new medical and information technologies, there are much more data available today than ever before. Translating data into useful information for a wide variety of audiences is a challenge for health care in general and disease management programs in particular. This paper addresses these issues by introducing several visual displays that illustrate important data elements in an unencumbered fashion. Examples are provided using the various stages of a hypothetical congestive heart failure (CHF) disease management program (ie, patient identification, program enrollment, intervention process, and outcomes evaluation). (Disease Management 2005;8:301–310)

INTRODUCTION

"ONE PICTURE IS WORTH a thousand words" applies to health care today more than ever before, with the field of disease management (DM) being no exception. With the rapid introduction of new medical and information technologies such as the electronic medical record (EMR), remote physiological monitors, and computer-controlled medical systems and equipment, there are more types of data available to DM programs than in the past. In addition, DM programs must speak to a wide range of audiences, from health care providers, to insurers, to patients, and to peer-professional groups. To effectively convey strategies and outcomes to extensive audiences using varied data sources, DM providers must carefully select their communication tools. A simple, but often underused, method for translating data into meaningful information is by way of visual displays or illustrations. Edward Tufte, considered by many to be the foremost authority on visual display of statistical data, suggests that graphical displays should[1]:

- Show the data
- Induce the reader to think about the substance of what is displayed
- Avoid distorting what the data have to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several levels of detail, from a broad overview to the fine structure
- Serve a reasonably clear purpose: description, exploration, tabulation or decoration

[1]Linden Consulting Group, Portland, Oregon.
[2]Oregon Health and Science University, School of Medicine, Department of Public Health/Preventive Medicine, Portland, Oregon.
[3]Integrated Performance/Six Sigma Champion, Providence Health System, Portland, Oregon.

- Be closely integrated with the statistical and verbal descriptions of a data set

Keeping within this conceptual framework, this paper will introduce several types of graphical displays and illustrate, using a hypothetical congestive heart failure (CHF) program, how they can be applied in DM. The intent is to provide DM programs with practical tools by which their programs' effectiveness can be demonstrated in the most evident way.

TYPES OF GRAPHICAL DISPLAYS

Often, the best way to understand data is to literally see the answers to the questions you pose. The concept of effective visual display of data will be illustrated using ten simple-to-use tools. While certainly not all inclusive, this set represents commonly used graphs and plots and has wide applicability for DM programs.

To select an appropriate data display, one first must understand the question being asked. The starting point is most often an analysis of the population (eg, demographics, disease prevalence, utilization patterns). Five tools—histogram, Pareto analysis, Venn diagram, pyramid graph, and pie chart—are all used in various ways to describe the distribution of data within a population.

Histogram

The *histogram* is a type of bar chart used for showing the range and depth of variation in a group of data. The histogram displays continuous or interval data with increasing quantity on the horizontal (x) axis and the frequency of occurrences/observations on the vertical (y) axis.

Pareto analysis

Vilfredo Pareto (1848–1923) was an economist who recognized the existence of the 80/20 rule in most naturally occurring situations (later referred to as the Pareto principle).[2] The 80/20 rule states that 80% of the occurrences, incidents, and costs are caused by 20% of the population, categories, and bins. By charting the frequency of each category against the cumulative percentage, a *Pareto analysis* identifies the most prevalent categories, as well as their percent contribution to the total amount.

Venn diagram

The *Venn diagram*, named for English mathematician John Venn (1934–1923), is comprised of two or more overlapping circles and is meant to illustrate relationships between subsets in a collection or population.

Pyramid graph

The stacked *pyramid graph* allows a visual depiction of narrowing subsets within a population. Starting at the base, each layer in the pyramid represents a smaller subset of the prior layer.

Pie chart

A common way to display the contribution of each value to a total is with a *pie chart*. This type of chart shows the proportional size of items that make up a data series and is useful when you want to emphasize a significant element.

Run chart

A different set of display tools are required to address changes in the observed variable over time. Initially developed by Walther Shewhart (1891–1967)[3,4] to complement statistical process control (SPC) indicators, *run charts* are the most basic method for displaying time series observations. Run charts plot the process measure of interest against a time interval on the x axis. Trends, seasonal factors and one-time events are easily detected upon visual inspection.

Control chart

An enhanced form of the run chart is called a *control chart*. In general, control charts include additional metrics to enhance the analysis. Typically, the mean or median is provided as the basic statistic, and upper/lower control limits are established by incorporating standard deviations, standard errors, or confidence intervals. A process is considered "out of control" when a predetermined number of observations fall outside of the control limits. There are two basic types of control charts, those using continuous or scalar data, and those for use with attribute or counts data.

In addition to its application in time series analysis, the control chart can also be used to compare the performance of a set of individuals/categories over a number of observations. While in the earlier example the horizontal axis is increasing units of time, in this application the individuals/categories for comparison are arrayed along the *x* axis.

Line graph, box-plot, and scatter-plot

Three additional tools can be employed to demonstrate differences between groups (line graph, box-plot, and scatter-plot). While a statistical analysis should be used to verify conclusions drawn by visual inspection of these graphs, they can provide an understanding of the relationship between the groups not easily gained by examination of the numbers alone.

The *line graph* is the archetypical method used to depict comparative outcomes data in a summarized way. While it is straightforward and unencumbered from a visual perspective, it requires the addition of several supplemental features to make this one picture worth a thousand words. For example, vertical error-bars can be added to indicate standard errors, or more illustratively, 95% confidence intervals (CI). Similarly, the analyst may choose to insert text indicating the level of significance noted between the two variables.

A *box-plot* (also referred to as box-and-whisker or five-number summary plot) is also an excellent means of displaying such summary data in a concise manner. It is suited for comparisons between two groups or between one group pre- and post-treatment. Each box-plot provides 5 summary statistics: the high and low extreme values (the whiskers), and the

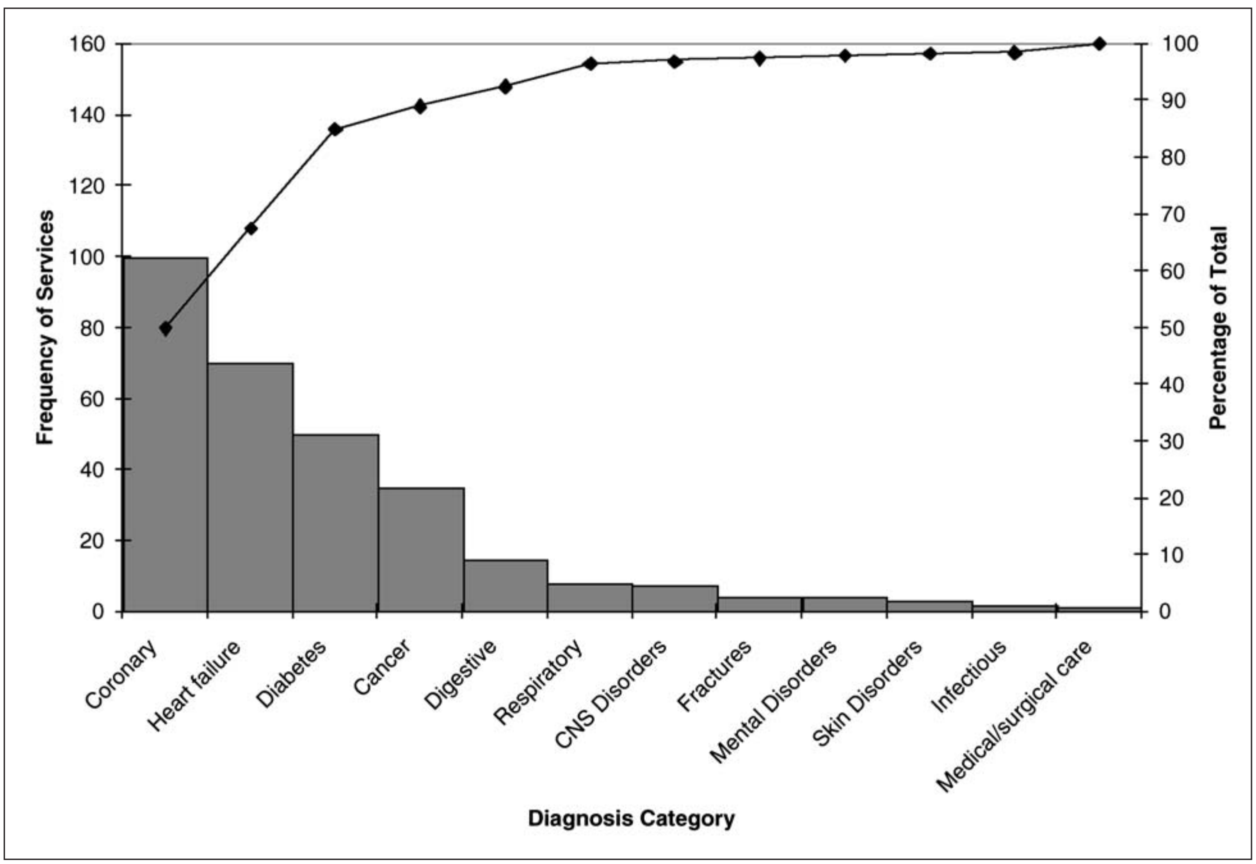

FIG. 1. A Pareto chart illustrating the frequency of health services utilization (hospital admissions and emergency department visits) by diagnosis category in a hypothetical Medicare population, and the percent contribution to the total.

25th, 50th (median), and 75th quartiles. Taken together, one can visually inspect the distribution of values for each variable. Similarly, informal inferences can be made as to whether a significant difference exists between the two values.

When an understanding of the relationship between two variables is desired the *scatter-plot* is helpful visual tool. Paired observations or variables for a given individual are plotted as an *x-y* axis coordinate on the graph. After several points are plotted, a pattern may emerge that indicates the strength of the relationship between those variables.

IDENTIFICATION PROCESS

The most logical first step an organization must take before choosing to develop or purchase DM services is to assess the prevalence of various diseases in its population.[5] A simple tool available for this type of analysis is the Pareto chart. As an example, Figure 1 represents the use of the Pareto diagram in a hypothetical Medicare population. As shown, approximately 20% of the diagnoses (coronary disease, heart failure, and diabetes) account for over 80% of the total number of hospital admissions and ED visits. This exemplifies the 80/20 rule and provides direction regarding which diseases should be considered for DM strategies (assuming that high utilization is a guiding principle in choosing a DM program). The Pareto chart is useful in various other strategic and operational areas of an organization as well, allowing administrators to identify and prioritize areas of concern according to the frequency with which they appear.

Given that only limited resources are made available for DM activities, the next type of visual display—the Venn diagram[6]—may assist in identifying where those resources should be allocated to achieve maximum benefit. Figure 2 depicts a population of patients with coronary artery disease (CAD), diabetes, and CHF (encompassing the three largest disease categories identified in the Pareto chart shown in Fig. 1). As illustrated, there are several overlapping areas between the disease groupings, indicating that some patients have more than one condition. The most prevalent overlap between any two conditions is CHF and diabetes (7%), whereas 1% of this population has all three conditions.

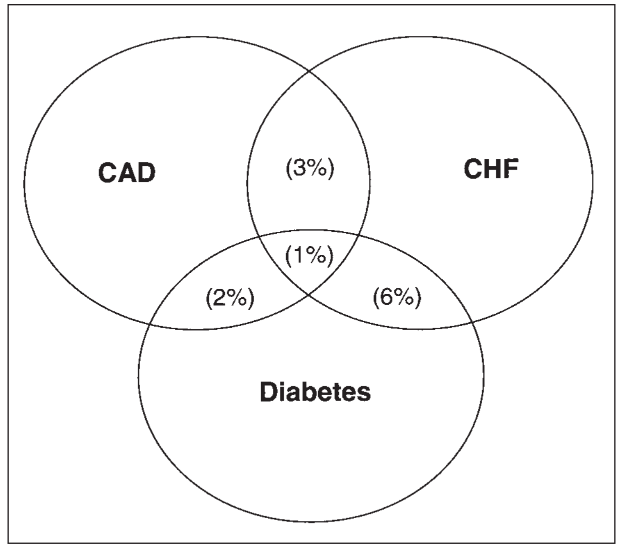

FIG. 2. A Venn diagram illustrating the overlap between patients with coronary artery disease (CAD), congestive heart failure (CHF), and diabetes.

In our hypothetical CHF program, patients identified in the overlapping areas between CHF and the other diseases may be the preferred candidates for targeted program enrollment strategies. The logic is that this set of individuals may represent the highest risk group, utilizing the most health care services, incurring the most costs, and possibly presenting the best opportunity to make an immediate impact (eg, by getting one of these patients in control of their blood pressure, cholesterol, and body composition, risk is reduced in all three disease processes concomitantly).

PROGRAM ENROLLMENT

The enrollment process may be considered the fulcrum on which program success ultimately hinges. Program implementation must be well thought out and executed according to the characteristics of the population, their health condition, and their healthcare providers.[5] Delays in enrollment or an inability to convince eligible individuals to participate will seriously hamper the program's ability to achieve targeted outcomes. Figure 3 uses a pyramid graph to illustrate the proposed implementation and enrollment plan for our hypothetical CHF program. The bottom tier or block represents the entire CHF population, estimated to be 3000 people (5% prevalence; 60,000 total Medicare members). Forty-five percent of those patients are thought to be the sickest, or the New York Heart Association (NYHA)[7] level IV. (For the purposes of this discussion it will be assumed that the most benefit will be derived by targeting the sickest patients. In reality, however, there are many factors outside of disease severity that determine who should be targeted for enrollment in the intervention.) The DM program assumes that it will be able to contact, enroll and initially screen 90% of those 1350 patients, resulting in 1215 program participants. Of those enrollees, 80% will receive the intensive nursing intervention (n = 972) and 20% of the participants (n = 194) will require home health visits as well. This graph is visually appealing and informative, and can be used in many different contexts. For example, two side-by-side pyramids would be an excellent depiction of the "before and after" effect, in which the proposed program implementation plan can be compared to that actually achieved.

A pie chart is an alternative method for displaying the data elements presented in Figure 3. However, including too many variables within a pie chart renders the categories too small and thus indecipherable. Figure 4 illustrates a pie chart with only four variable levels making it easy to read and interpret. Using the hypothetical CHF program, this graph illustrates that of those patients preliminarily predicted to be classified as NYHA level IV, only 45% were truly determined to be at that level after the initial telephonic screening call from a nurse manager. Similarly, it is shown that as many as 14% of those initially classified as level IV were later stratified at level I. This simple display provides the program administrators with an indication that their preliminary classification method is not valid for this purpose. The algorithm should be changed to improve both sensitivity and specificity,[8] so that limited program resources can be applied to more useful endeavors.

One example of how the histogram can be used is in plotting participants' length of time in program (LOTIP). This is an important indicator to observe since program tenure is expected to correlate with outcomes.[9] In other words, to achieve the desired psychosocial behavioral change in a participant, a significant amount of attention and support must be provided over the necessary period of time.[10] Figure 5 presents an LOTIP analysis for the hy-

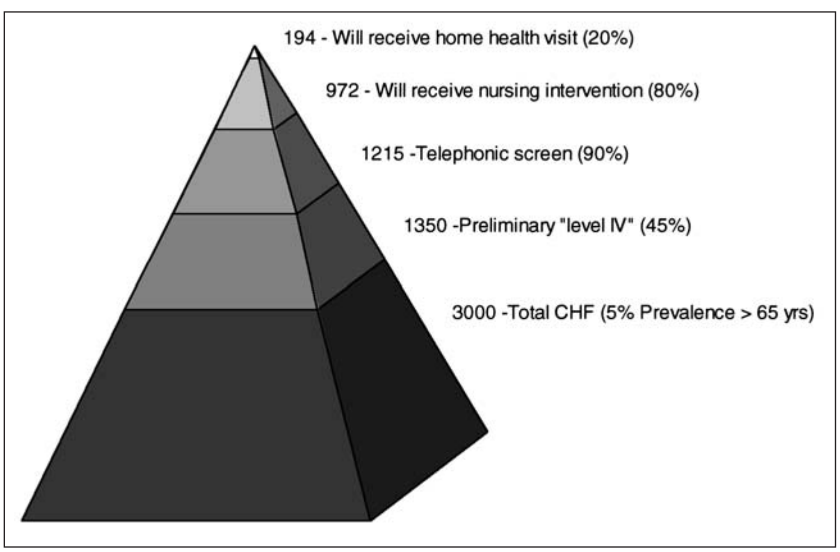

FIG. 3. A pyramid graph displaying the enrollment and intervention process of a hypothetical congestive heart failure (CHF) disease management program.

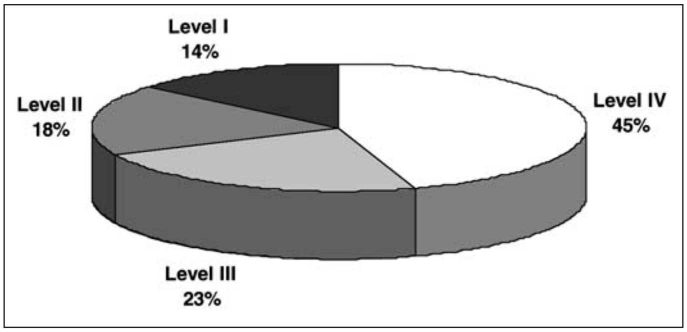

FIG. 4. A pie chart illustrating the percentage of congestive heart failure (CHF) patients by New York Heart Association (NYHA) level following nurse telephonic screening (in a hypothetical CHF disease management program).

pothetical CHF program over the course of 25 months. As illustrated, there is an incremental decline in the number of participants as time goes on. Nevertheless, it is readily apparent that there are a sufficient number of participants at almost any time period to conduct a statistically meaningful program analysis. Similar displays can be created to depict utilization or cost experience as a function of LOTIP.

INTERVENTION PROCESS

As discussed above, it takes time (in some cases, a significant amount) for a participant to acquire the skills necessary to better self-manage his or her disease state. The intervention process typically is multifaceted, including several instructional and treatment components given in tandem. The coordination of these efforts is paramount if the desired outcome is to be achieved. Observing these procedural measures over time allows the investigator to identify anomalies in the process, some of which may be normal point-over-point variability, while others may be a true function of the intervention.[11]

Using a run chart, Figure 6 plots the percentage of program eligible individuals who were actually enrolled in the hypothetical program on a monthly basis. This variable is useful to determine if enrollment targets are being

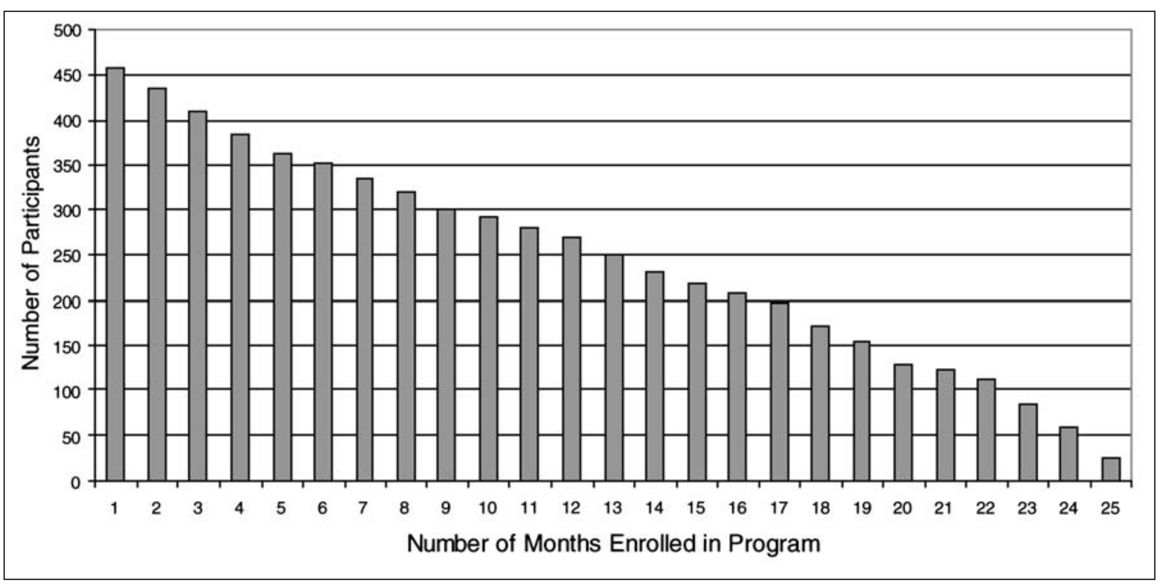

FIG. 5. A histogram illustrating the number of participants in a disease management program by the number of months enrolled.

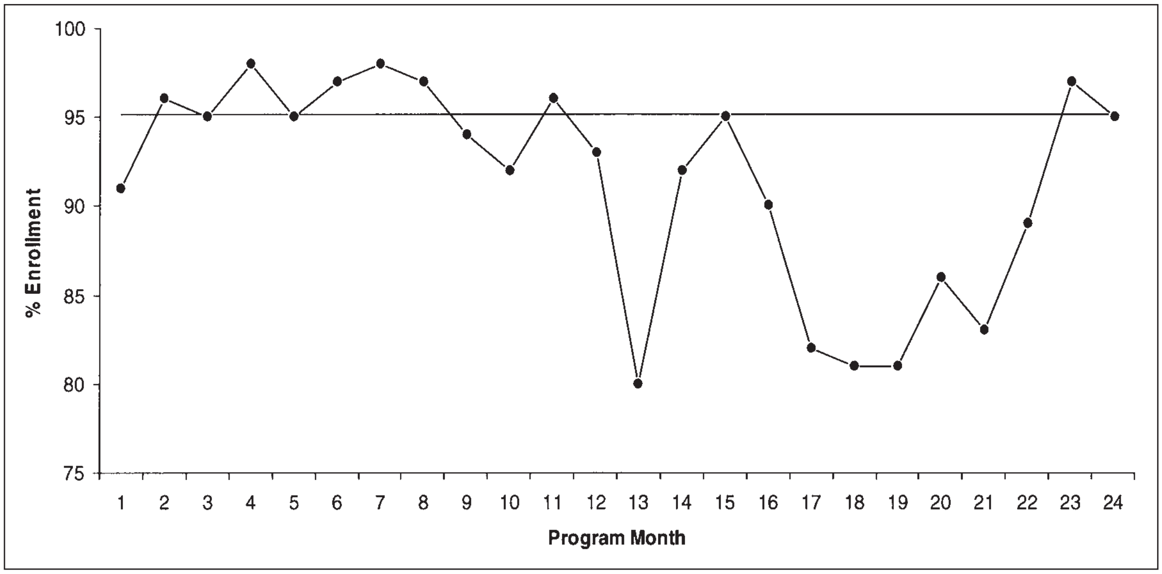

FIG. 6. A run chart indicating the percentage of eligible program candidates who were enrolled every month up to 24 months of the program. The mean, or average (solid line), was based on the first 12 months of the program.

met consistently. A solid line representing the process mean for the initial 12 months was added as a marker to assist the visual examination. Upon closer inspection one can see that in the second program year (months 13–24) there was an apparent decline in the enrollment rate. In fact, all but one observation were positioned below the process mean during that period. If monitored contemporaneously, this decline in performance would have been noted early and the root cause should have been investigated.

Figure 7 presents a control chart that plots average weekly patient contacts (including outbound and inbound calls to and from participants, respectively), by disease manager. As

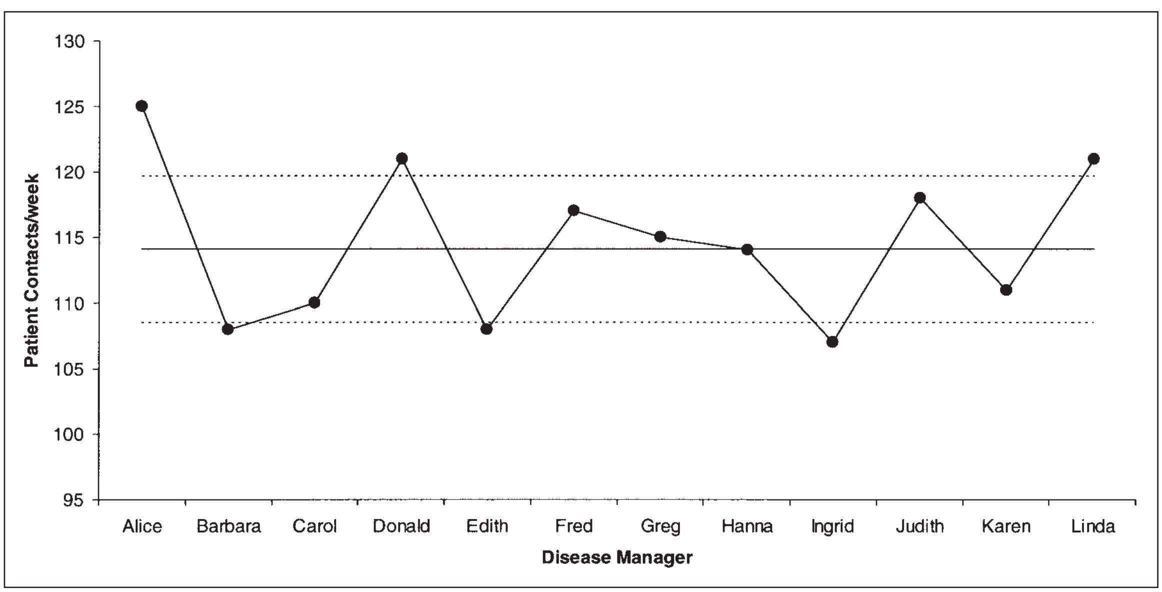

FIG. 7. A control chart showing hypothetical patient contacts per week (averaged over the course of a 12-week period) for each disease manager. The mean is indicated by the solid line, and upper and lower 95% confidence intervals are indicated by the dotted lines.

shown, three disease managers had average weekly contact rates that exceeded the upper control limit, and three disease managers had rates that fell below the lower control limit. Such information is useful in assisting program administrators to identify where a more focused review may be necessary. In this case, disease managers that were "out of control" would be compared to their "in control" peers to determine the cause of the disparity between them. There is an entire field of study devoted to statistical process control and other quality management methods. Since a comprehensive discussion of all these techniques is beyond the scope of this paper, readers interested in learning more should refer to Pyzdek[12] for a good introduction to quality management.

PROGRAM OUTCOMES

Figure 8 uses a box plot to display baseline and first program year compliance rates for angiotensin converting enzyme (ACE) inhibitor use for the intervention group in the hypothetical CHF program. Viewing the location of the whiskers, there does not appear to be an overlap in the level of the highest extreme value on the pre-program box-plot and the lowest extreme value on the post-program box-plot, thus, it can be preliminarily concluded that there was indeed a significant program effect. A statistical analysis will need to be performed to verify these findings.

Figure 9 displays a line graph in which differences in the average cost per patient are plotted for CHF program participants and controls across the pre- and post-treatment periods. Vertical bars were added to indicate the 95% confidence intervals. As shown, there appears to be no difference between the two cohorts at pre-treatment and an enormous difference at post-treatment. The fact that the CI do not overlap at the post-treatment measurement indicates that this difference is statistically significant.

Sometimes a more complete display of data is required by the evaluator than that provided by either the box-plot or line graph. One technique that can prove valuable under these circumstances is the scatter-plot diagram. For example, Figure 10 displays a scatter-plot of the individual's level data (as opposed to summarized data) that was used in Figure 9. Each individual's pre and post-treatment costs are plotted, with the pre-program cost variable on the *x*-axis and the post-program costs listed on the *y*-axis. As shown, the majority of matched values fall under $60,000 on either the *x* or *y* axes. However, upon visual inspection, it becomes apparent that many controls had substantial increases in cost in the post-program

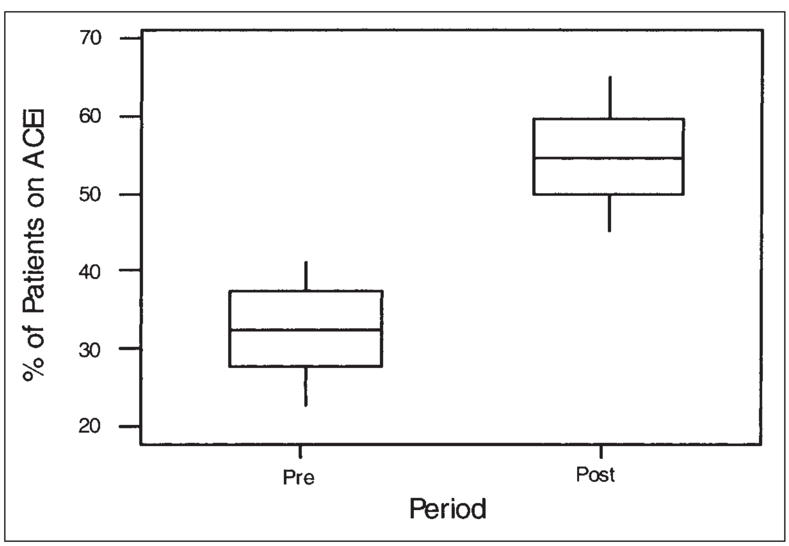

FIG. 8. A box-plot display in which pre- and post-program rates of patients on ACE (angiotensin converting enzyme) inhibitors are compared.

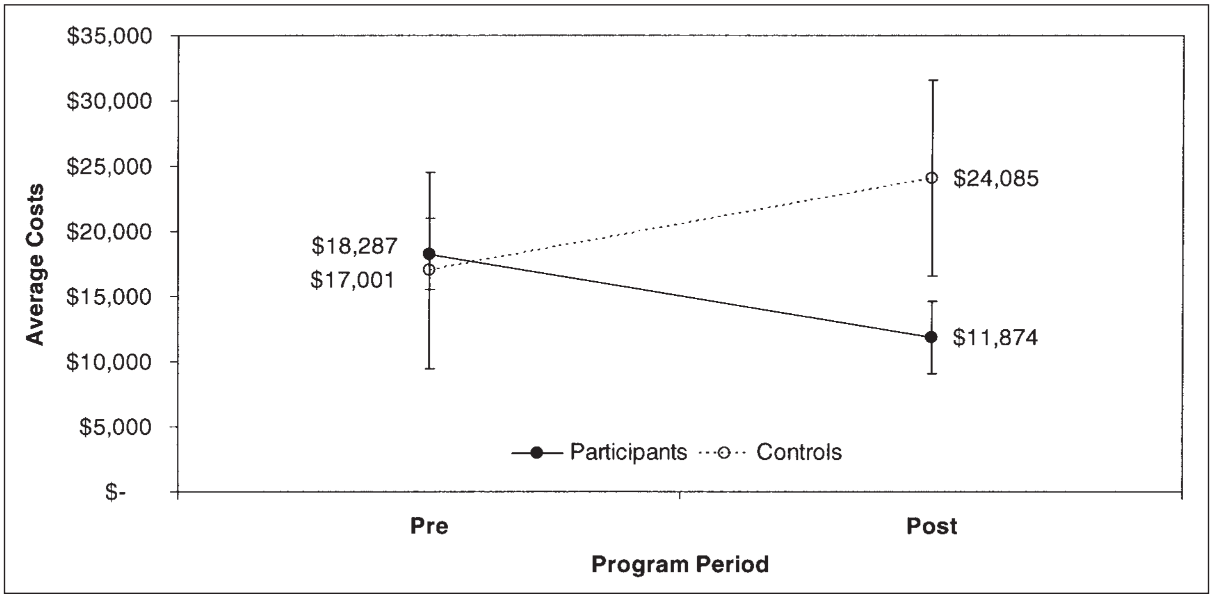

FIG. 9. A line graph illustrating average costs between program participants and controls at baseline and at post-treatment. Vertical bars represent 95% confidence intervals.

period while the program participant group remained relatively homogeneous. The addition of trend or regression lines for each group assists in determining whether these outliers impact the overall model. The regression lines are relatively flat for the two groups indicating that, on average, costs decreased over the period. However, the intercept (where the line crosses the *y* axis) for the control group is $19,477 while the intercept for the participant group was only $10,962. Informally, it is possible to infer from these statistics that the pro-

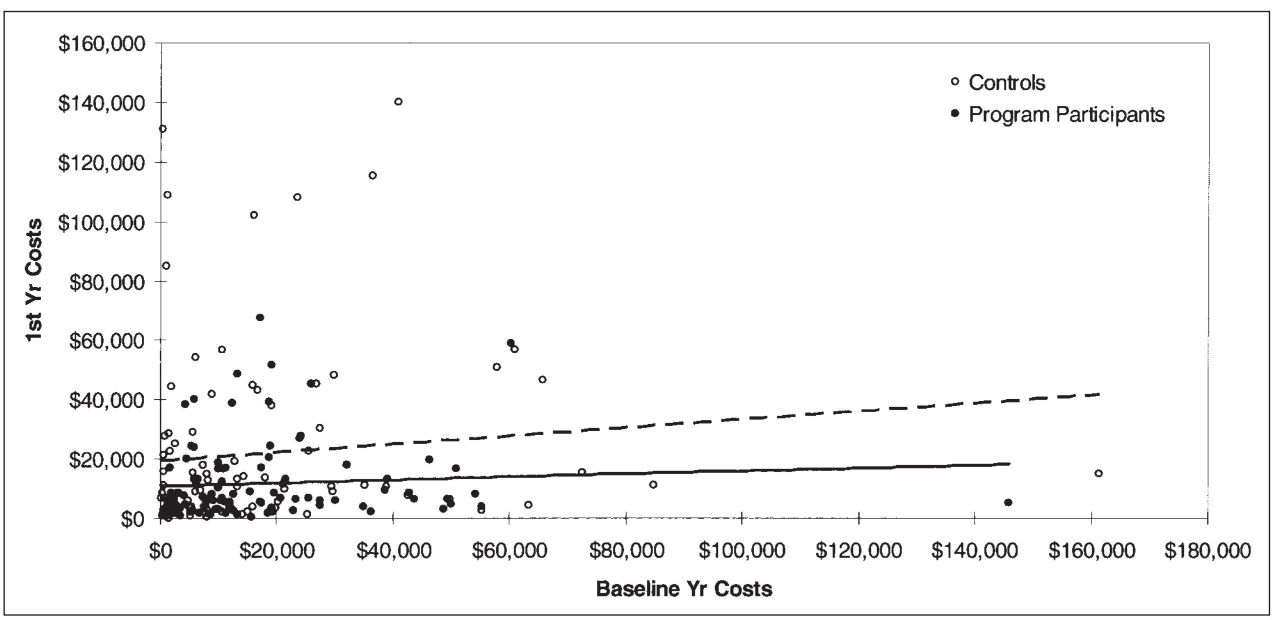

FIG. 10. A scatter-plot of baseline and first year costs of program participants and controls. Participants were enrolled in a congestive heart failure (CHF) program and the controls received no interventions during that period. The dotted line indicates the trend for the controls, and the solid line represents the trend for the participants.

gram group experienced significantly lower costs than the control group in the post treatment period. This difference appears to be attributed to the several extreme outliers in the control group. A review of the regression analysis would be required to draw a formal conclusion of whether the difference between the two groups is statistically significant.

CONCLUSION

This paper introduced several visual displays as a means of presenting data in a succinct and coherent fashion. Although they were organized according to DM program component (ie, identification, enrollment, intervention and evaluation), most of these tools can be used interchangeably in any other segment. Given the space limitations, this paper could not present all the various types of visual displays available for illustrating DM program or health services related data. Nonetheless, all the major graphs and plots were introduced. For readers interested in learning more about how to incorporate these displays in quality and process improvement activities, further reading should include: Pyzdek,[12] Carey and Lloyd,[13] Mears,[14] and Wheeler and Chambers.[15] Readers interested in learning the history of visual displays of data for many different applications can consult an excellent series of publications written by Edward Tufte,[1,16,17] as well as three recent publications by Wainer,[18] Miller,[19] and Best.[20]

REFERENCES

1. Tufte ER. The visual display of quantitative information, 2nd ed. Cheshire, CT: Graphics Press, 2001.
2. Schumpeter J. History of economic analysis. London: George Allen and Unwin, 1954.
3. Shewhart WA. Economic control of quality of manufactured products, 50th Anniversary Issue. Milwaukee, WI: American Society for Quality, 1980.
4. Shewhart WA. Statistical method from the viewpoint of quality control. Mineola, NY: Dover Publications, 1986.
5. Linden A, Roberts N, Keck K. The complete "how to" guide for selecting a disease management vendor. Dis Manag 2003;6:21–26.
6. Venn J. Symbolic logic, 2d ed. Bronx, NY: Chelsea Publishing Company, 1971.
7. The Criteria Committee of the New York Heart Association. Physical capacity with heart disease. In: Diseases of the heart and blood vessels, nomenclature and criteria for diagnosis, 6th ed. Boston: Little, Brown & Co., 1964:110–114.
8. Linden A. Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. J Eval Clin Pract (In press)
9. Linden A, Adams J, Roberts N. Evaluating disease management program effectiveness adjusting for enrollment (tenure) and seasonality. Res Healthc Fin Manag 2004;9:57–68.
10. Linden A, Roberts N. Disease management interventions: what's in the black box? Dis Manag 2004; 7:275–291.
11. Linden A, Adams J, Roberts N. Evaluating disease management program effectiveness: an introduction to time series analysis. Dis Manag 2003;6:243–255.
12. Pyzdek T. The handbook for quality management: a complete guide to quality management. Tucson, AZ: QA Publishing, 2000.
13. Carey RG, Lloyd RC. Measuring quality improvement in healthcare: a guide to statistical process control applications. New York: Quality Resources, 1995.
14. Mears P. Quality improvement tools and techniques. New York: McGraw Hill, 1995.
15. Wheeler DJ, Chambers DS. Understanding statistical process control, 2nd ed. Knoxville, TN: SPC Press, 1992.
16. Tufte ER. Envisioning information. Cheshire, CT: Graphics Press, 1990.
17. Tufte ER. Visual explanations. Cheshire, CT: Graphics Press, 1997.
18. Wainer H. Graphic discovery: a trout in the milk and other visual adventures. Princeton, NJ: Princeton University Press, 2005.
19. Miller JE. The Chicago guide to writing about numbers. Chicago: University of Chicago Press, 2004.
20. Best J. More damned lies and statistics: how numbers confuse public issues. Berkeley: University of California Press, 2004.

Address reprint requests to:
Ariel Linden, Dr.P.H., M.S.
Linden Consulting Group
6208 NE Chestnut St.
Hillsboro, OR 97124

E-mail: alinden@lindenconsulting.org