



## **5.0 TRENDING AND STATISTICAL USAGE OF DATA**

### **5.1 Purpose**

This section provides members of the DOE community with consistent guidelines for trending and analyzing internal occurrences, conditions, and events, and to provide support for the analysis of trends. An effective trend analysis program should have the ability to identify conditions or trends adverse to quality and to provide management with recommendations for meeting customer requirements through improved performance (i.e., more safely, quickly, and less expensively). It is recognized that effective trending programs may already exist that contain many of the same elements described in this section; therefore, it is intended to serve as a guideline rather than provide prescriptive direction.

### **5.2 Scope**

This section describes recommended programs for reviewing ORs, collecting pertinent data from each report for trending, and identifying key issues for senior management attention. Recommendations are directed toward onsite reports, with considerations for current ongoing DOE programs mandated by DOE Order 210.1, as well as for related activities such as root cause analysis. Recommended subjects for trending are included along with an explanation of considerations employed in the selection of methods for data analysis.

### **5.3 General**

The DOE community represents a wide variety of sites and functional responsibilities. Some sites include thousands of personnel spread out over hundreds of square miles, while others include only several hundred people on a few hundred acres. Certain facilities are specific to defense-related activities, whereas others are considered to be environmental projects, which may operate within a somewhat different set of rules.

However, because there are many regulatory requirements that form a common core for all of the DOE community, it is essential that a somewhat standardized set of guidelines exist for classifying and evaluating occurrence-related activities. DOE Manual 232.1-1 contains provisions for standardized coding of causes.

Selecting specific aspects of occurrences to group and evaluate, however, may be less obvious. The selection of how to group, categorize, and statistically analyze data will, in many cases, be as unique as the site or facility within the site. Certain statistical tools lend themselves to situations with high volumes of data with little expected variation. Conversely, smaller organizations performing less routine or production-oriented activities would not obtain meaningful output by applying the same statistical approach. Therefore, it is essential that careful consideration be given to the selection of methods employed for the gathering, processing, and analyzing of data.



## **5.4 Guidance**

### **5.4.1 Determination of Trend Program Objectives, Sources, and Information**

Your organization must first determine the type of data to be gathered and trended, as well as how it will be obtained. A facility and/or site trending program is most effective when it contains similar information from as many available sources as possible. Trending should not be limited to ORs, but also should include input from documents such as audits, assessments, surveillances, nonconformances, requests for corrective action, and critiques. Data compiled for trending may include direct causes, root causes, quality assurance program elements, level of severity, activity type, responsible organizations, or anything else that may be tracked for identifying common or recurring problems and which can ultimately lead to the generation of effective solutions or preventive measures. At a minimum, the cause codes specified in DOE Manual 232.1-1 should be included in a trending program to help in the identification of early indications of deteriorating conditions.

Trending can provide several different types of information. Examples follow.

- Trending of performance relative to DOE Manual 232.1-1 requirements, including the number of late reports, the number rejected and the number with overdue corrective actions. This data is available from ORPS by using predefined or ad hoc reports.
- Trending of specific topics or areas, such as the number of halon discharges, the number of occurrences involving inadequate freeze protection, or the number of occurrences involving continuous air monitors. This is more difficult to get from ORPS because a narrative search must be used. Each report must then be reviewed to ensure that it is valid. The data could be presented as a function of time or by facility.
- Trending of recurring events or conditions tracks a known problem to ensure that corrective actions taken were effective in preventing recurrence. It can also identify deteriorating conditions that may lead to an occurrence. The use of control charts is a good method for this type of trending. Control charts allow comparisons of current performance to historical baselines. They are most effective when a large amount of data is available and it is reported frequently (i.e., weekly or monthly).
- Trending by cause requires careful data screening. Some screening can be done using a computer (i.e., ORPS or a facility/site database that has ORPS data loaded). For example, to track the cause of equipment failures for facility "X," the following could be the criteria:
  - Only look at this facility
  - Only look at the ORs that dealt with equipment failure
  - Only look at failures above a certain threshold



- Only look at ORs that had a formal, detailed root cause analysis  
Without more detailed information, it will be very difficult to assign causes and corrective actions for an apparent trend without at least some review of specific ORs. The computer screen or sort is only a portion of the process. Because there is not a detailed convention for assigning titles or event types to occurrences, it is difficult to sort simply by key words.

5.4.2 Statistical Methods - Selection/Requirements

As previously stated, careful consideration should be given to determining what type of statistical tools will be most effective for the facility and/or site application. This should include considerations such as volume or number of data points, lot size, distribution of data, range of data and time frame represented for each analysis. For example, control charting may be an effective indicator in areas of at least 18 data points, however, it may be meaningless with only five. Part of selecting appropriate statistical tools is also determining at what point those tools will send up a "flag" and just what will be done at that point.

Analysis techniques may include the use of Pareto charts to identify which areas are the largest contributors to adverse trends. The trend analysis data system has sufficient information to develop charts based upon categories such as defect code, organization, NQA-1 program element or other identifiers as provided by the event coding system. For a starting point, charting based upon defect code is recommended. Pareto analysis will facilitate optimal resource allocation by directing attention to areas, which, if properly addressed, will contribute most towards diminishing trends adverse to quality. An example chart (Figure 5-1) is provided at the end of this section.

Another statistical tool effective in the early detection of negative (or positive) trends is control charting. This is basically the comparison of current data and data patterns to a historical baseline consisting of a mean and upper and lower control limits. Because the total number of events to be trended varies from period to period, the use of a "p" (percentage) type control chart is recommended. Control charts, based upon the eight primary cause categories as identified in DOE Manual 232.1-1, will allow early detection of significant increases or decreases within, as well as between, the seven cause areas. Once a historical baseline consisting of a mean and upper control limit has been established, comparisons to new data will help identify gradual, as well as sudden, changes. Data and data patterns warranting the need for further sorting and review to determine if a common relatable problem may exist include the following: any point outside the upper control limit, five points in a row steadily increasing or decreasing, or five consecutive points above the mean. Based upon these "flags," a data sort should be initiated to determine if a correlation exists between two or more of the following: organization, cause code, report type, NQA-1 element, or date. Control chart baselines should be based upon at least 18-25 data points. New data should be plotted as frequently



as possible (at least monthly). An example of a control chart (Figure 5-2) is provided at the end of this section.

**5.4.3 Database and Software Selection/Development**

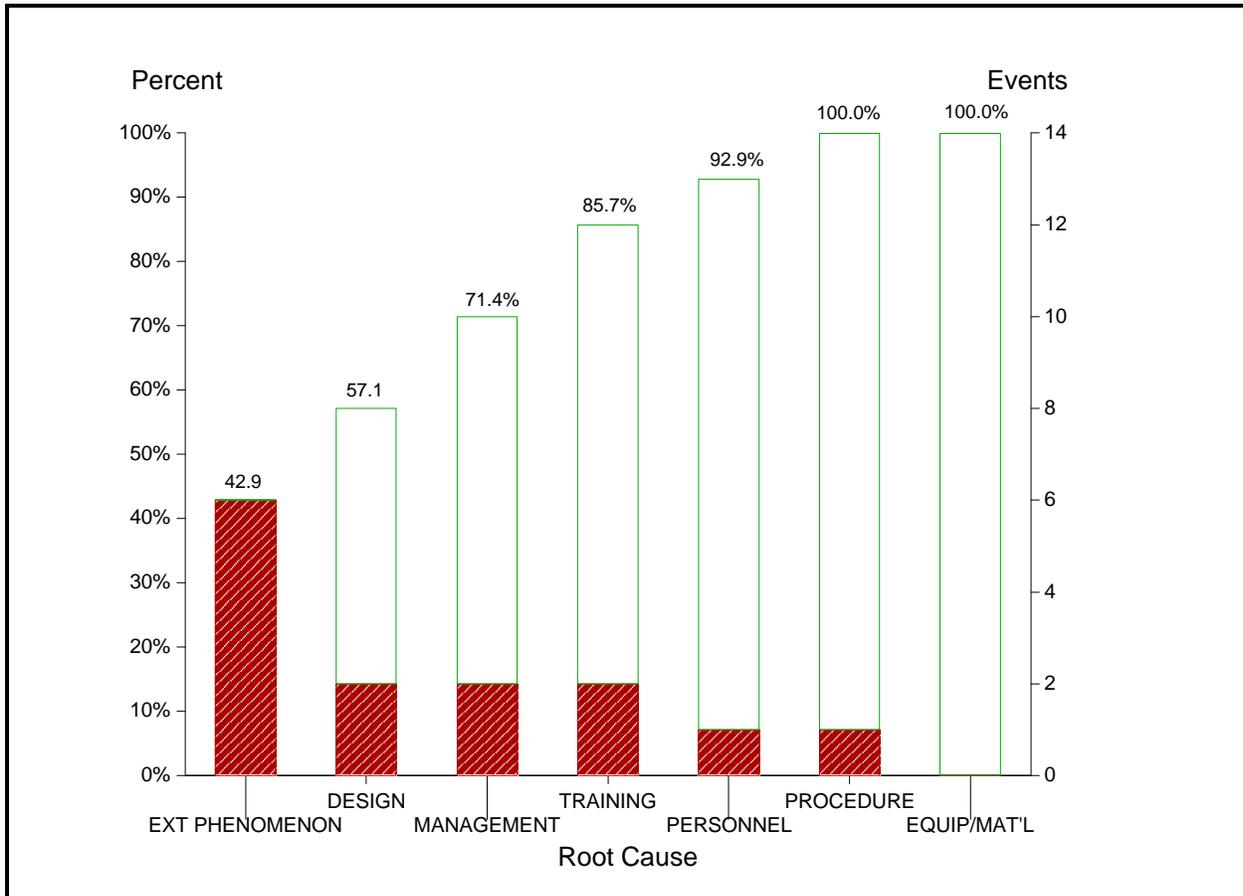
Once the most useful data has been identified, a database must be developed. When selecting software to process the information, there are several key abilities that must be evaluated. It must be interactive or compatible to the maximum extent possible. This will minimize the reentry of data and reduce the risk of errors. Ideally, selected data could be directly downloaded from ORPS. The database should lend itself to sorting or query for a given set of Fields or conditions. Also, a statistical graphics package should be selected based on the statistical methods determined to be most effective for your facility. It is important to have a package that has the ability to perform basic statistical activities such as deriving and plotting control charts, histograms, and pie charts. Valuable information may be wasted if it is not effectively presented. There are many such packages available that are relatively simple to use with a variety of different databases. Again, it should be compatible with your database to minimize reentry.

**5.4.4 Results/Recommendations - Follow-up**

A means of assigning recommendations, in terms of authority and responsibility, should be determined. It is essential that information derived from these analyses be reported and acted on in a timely manner to be most effective. Some type of reporting format should be established along with a distribution system to ensure that information is transmitted not only to organizations where specific corrective actions are necessary, but also to groups that may benefit in a lessons learned capacity. When corrective actions are identified, provisions should be included for tracking and status as necessary along with verification of completion and effectiveness of the action as appropriate.



Figure 5-1 - Example of a Pareto Chart

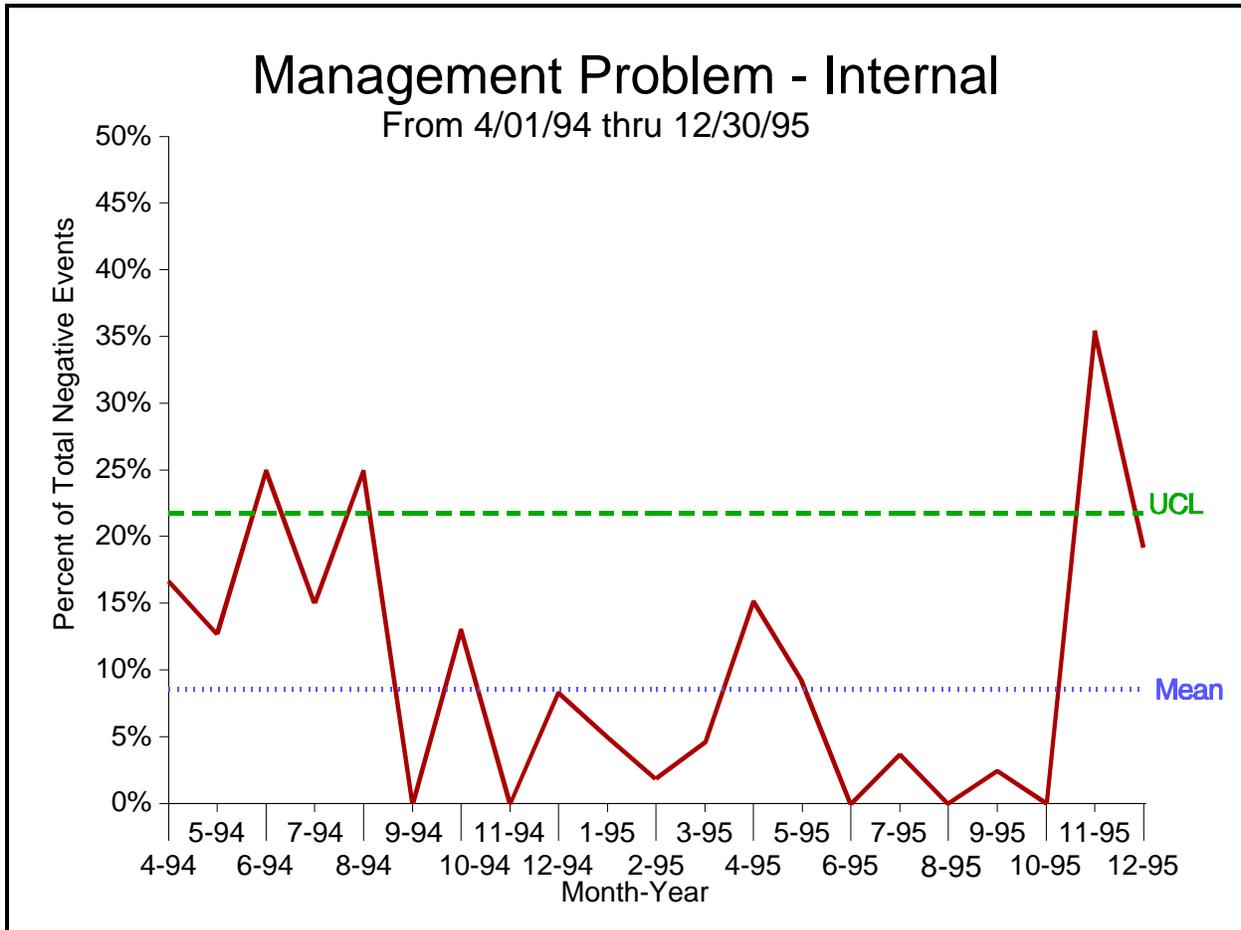


**Performance Indicator 2.5 - Unplanned Shutdowns**  
**Distribution by Root Cause Code from Final Occurrence Reports**





Figure 5-2 - Example of a Control Chart



**Legend:**

**UCL** - Upper Control Limit

**Mean** - Average of the data points

**Y axis** - Events or documents coded as negative, such as ORs, Critiques, Nonconformance Reports, etc., having a root cause of Management Problem.