

Organizational Epidemiology and Energy Facilities

Statistical Techniques for Empirical Analysis and
Prediction of Human Performance Trends

Technical Report

Organizational Epidemiology and Energy Facilities

Statistical Techniques for Empirical Analysis and
Prediction of Human Performance Trends

1004669

Final Report, July 2002

EPRI Project Manager
M. Gross

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

THIS DOCUMENT WAS PREPARED BY THE ORGANIZATION(S) NAMED BELOW AS AN ACCOUNT OF WORK SPONSORED OR COSPONSORED BY THE ELECTRIC POWER RESEARCH INSTITUTE, INC. (EPRI). NEITHER EPRI, ANY MEMBER OF EPRI, ANY COSPONSOR, THE ORGANIZATION(S) BELOW, NOR ANY PERSON ACTING ON BEHALF OF ANY OF THEM:

(A) MAKES ANY WARRANTY OR REPRESENTATION WHATSOEVER, EXPRESS OR IMPLIED, (I) WITH RESPECT TO THE USE OF ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT, INCLUDING MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, OR (II) THAT SUCH USE DOES NOT INFRINGE ON OR INTERFERE WITH PRIVATELY OWNED RIGHTS, INCLUDING ANY PARTY'S INTELLECTUAL PROPERTY, OR (III) THAT THIS DOCUMENT IS SUITABLE TO ANY PARTICULAR USER'S CIRCUMSTANCE; OR

(B) ASSUMES RESPONSIBILITY FOR ANY DAMAGES OR OTHER LIABILITY WHATSOEVER (INCLUDING ANY CONSEQUENTIAL DAMAGES, EVEN IF EPRI OR ANY EPRI REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES) RESULTING FROM YOUR SELECTION OR USE OF THIS DOCUMENT OR ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT.

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

Thomas J. Ayres

ORDERING INFORMATION

Requests for copies of this report should be directed to EPRI Orders and Conferences, 1355 Willow Way, Suite 278, Concord, CA 94520, (800) 313-3774, press 2 or internally x5379, (925) 609-9169, (925) 609-1310 (fax).

Electric Power Research Institute and EPRI are registered service marks of the Electric Power Research Institute, Inc. EPRI. ELECTRIFY THE WORLD is a service mark of the Electric Power Research Institute, Inc.

Copyright © 2002 Electric Power Research Institute, Inc. All rights reserved.

CITATIONS

This report was prepared by

Thomas J. Ayres, Ph.D.
Human Factors Consultant
PO Box 60591
Palo Alto, CA 94306

Principal Investigator
T. Ayres

Part of the work described in this report was performed while he was employed by Exponent, Inc., Menlo Park, California.

This report describes research sponsored by EPRI.

The report is a corporate document that should be cited in the literature in the following manner:

Organizational Epidemiology and Energy Facilities: Statistical Techniques for Empirical Analysis and Prediction of Human Performance Trends, EPRI, Palo Alto, CA: 2002. 1004669.

REPORT SUMMARY

This report describes an empirical study of the prediction of human performance and energy facility performance based on analysis of historical data relating to worker-, workplace-, management-, and organization-centered factors. Findings verify the relationship between such antecedent conditions and subsequent human performance and facility productivity, reliability, and safety.

Background

The most frequently used approach to study human performance and reduce the impacts of human error is retrospective, i.e., the review of an accident, intended to identify its cause(s), leads to identification of corrective action(s) intended to prevent repetition of the accident. Although the retrospective approach has yielded significant benefits, the ability to predict human performance could support more efficient allocation of resources toward error prevention and human performance optimization. A literature and experience review (EPRI report 1004668) led to the concept of *organizational epidemiology*, in which data relating to human error, human performance, and facility performance are combined with contextual information that could reveal antecedent conditions for performance changes.

Objective

To explore the feasibility and practicality of applying organizational epidemiology in energy industry contexts by collecting historical data from energy facilities and assessing the use of various analytical methods to predict human and facility performance.

Approach

Previous EPRI strategic work (1004668) revealed that exploration of organizational epidemiology should involve the collection of historical data relating to measures of potential antecedent conditions (predictors) and of human and facility performance (outcomes). In this study, a preliminary list of measures of interest potentially available from energy facilities was developed in conjunction with management and personnel from two nuclear plants. Through collaborative work with two fossil plants, investigators obtained suitable data for various measures of interest. They next applied statistical methods based primarily on regression techniques to find reliable relationships between possible predictor variables and four outcome measures.

Results

Statistically significant predictive models emerged for each outcome measure, despite limitations imposed by the restricted data sets ultimately obtained. These data sets were restricted in the number of years of data archived, in the usability of some data, and, even more, by the lack of other measures that would more directly tap training, safety efforts, and other factors that had

been deemed of interest. Based on analyses of these data sets, predictor measures related to injuries, sickness, absenteeism, corrective and preventive maintenance activity, and, possibly, seasonality hold particular promise for reflecting the antecedent conditions associated with changes in human performance and consequent facility performance. Several of the models accounted for more than 50% of the variance in the outcome measures.

Predictive models could have real value for energy facilities that take advantage of their existing data resources. Of perhaps equal importance is the conclusion that it will not be possible to specify the models (particularly model coefficients and optimum lags) in advance. Although some connections identified in this study between predictors and outcomes were not surprising, the details were discovered only through statistical analysis. This finding supports the long-held suspicion that performance may be shaped by many factors operating at multiple organizational levels and interacting over varied time scales. It also shows that analytical tools are needed to quantify such relationships and optimize predictive models.

Organizational epidemiology appears promising for energy industry facilities and additional industrial and organizational settings, but several lessons learned and important caveats should be considered. Practical recommendations are provided to help energy company analysts and researchers apply lessons regarding the acquisition of data on potential measures of interest. Major caveats are as follows: 1) the models developed in this study are predictive only after the fact, describing relationships found in historical data; and 2) statistically predictive models do not necessarily help analysts understand the bases for performance changes. Follow-on studies could help determine whether truly predictive models can be developed as well as examine causal relations between predictors and important performance outcomes.

EPRI Perspective

This report was prepared as part of the “Human Performance Management: Database and Analysis” project under the Strategic Human Performance Program. The project sought to improve understanding of how varied factors affect human and facility performance. Such knowledge would enable development of methods and tools for predicting performance and informing design of prospective or proactive interventions. The study described in this report—the first major attempt to explore the use of organizational epidemiology in energy industry settings—provides proof of concept for this type of approach. Compatible findings were produced in a complementary strategic study (1004670), which investigated a predictive approach based on work by EPRI’s Nuclear Sector to develop leading indicators of organizational health (reports TR-107315, 1000647, 1003033). Further work is needed to develop predictive methodologies, test their applicability in real-world environments, and employ their findings for guiding efforts to optimize human performance and improve the reliability, safety, security, productivity, and environmental performance of the energy infrastructure.

Keywords

Human Error

Accidents

Human Performance

Organizational Epidemiology

Intervention

Prediction

ACKNOWLEDGMENTS

This report benefited from the advice, contributions, and comments of many people. These included representatives from energy companies and industry organizations, members of the research community, personnel from Exponent, Inc., and EPRI technical and management staff.

Management and staff of the collaborating energy company and of the two participating fossil generation facilities played a central role in this project, and they gave generously of their time in order to supply data and information. Dorcas Moloi and David McGhee, both Eskom employees seconded to EPRI for a significant part of the project period, served as liaisons to the participating company and did much of the work of obtaining and pre-processing the data. Considerable effort and support were also provided by personnel of two additional energy companies, including facilitation of visits to two nuclear power plants by the research team in order to learn about data resources.

Members of the Data Analysis group at Exponent, Inc., carried out the initial organization and processing of the fossil plant data. Jack Tomskey, Ph.D., a statistical consultant to the research team, performed the final quantitative analyses described in this report.

Madeleine Gross, manager of the EPRI Strategic Human Performance Program, provided essential management and technical direction, both for this project and for its synergy with related strategic projects. Gail Kendall, former director of EPRI Strategic Science & Technology, gave early impetus to this work. Over the span of this project, several additional EPRI technical and management staff afforded significant assistance. In particular, John Stringer, executive technical fellow, provided invaluable management support for the continuing research, and he was of great assistance in defining the agreement between EPRI and the collaborating energy company. Revis James contributed ongoing programmatic support, and Jack Haugh collaborated in integrating this strategic project with related work conducted by the Nuclear Sector's Human Performance Technology Program. Barry van Brunt and Art Kenny facilitated the timely completion of this work.

Christopher R. Powicki of Water Energy & Ecology Information Services provided valuable assistance in editing and producing this report.

CONTENTS

1 INTRODUCTION	1-1
Research Context.....	1-1
Report Organization	1-2
2 HUMAN ERROR—CAUSES AND ANTECEDENTS	2-1
Human Error.....	2-1
Antecedent Conditions	2-2
Deductive and Inductive Approaches	2-2
Organizational Epidemiology in the Energy Industry.....	2-4
3 METHODOLOGY—DATA IDENTIFICATION	3-1
Data Sought	3-1
4 DATA COLLECTION.....	4-1
Site Visits	4-1
Data Acquisition	4-2
Confidentiality of Plant Data	4-4
5 METHODOLOGY—ANALYTICAL METHODS CONSIDERED AND APPLIED	5-1
Methods Applied.....	5-1
Additional Methods.....	5-2
6 ANALYSES AND FINDINGS.....	6-1
Pairwise Correlations.....	6-1
Stepwise Regressions	6-5
Analytical Procedures	6-5
Results	6-7
Seasonality.....	6-10
Multivariate Visualization	6-11

7 PRACTICAL LESSONS LEARNED FOR FUTURE APPLICATION AND RESEARCH	7-1
Lessons for Energy Companies.....	7-3
Lessons for Researchers.....	7-5
8 CONCLUSIONS AND RECOMMENDATIONS	8-1
Conclusions.....	8-1
Recommendations	8-2
9 REFERENCES	9-1
A FACIAL DISPLAYS	A-1

LIST OF FIGURES

Figure 6-1 Plant A: Correlation of Trips and Sick Days.....	6-2
Figure 6-2 Plant A: Monthly Trips and Sick Days	6-3
Figure 6-3 Plant B: Correlation of Trips and Engineering Sick Days.....	6-5
Figure 6-4 Plant A: Trips: Stepwise Regression	6-7
Figure 6-5 Plant A: Chemical Index: Stepwise Regression.....	6-8
Figure 6-6 Plant A: Monthly Chemical Index and Ratio of Corrective to Preventive Maintenance Orders Outstanding	6-9
Figure 6-7 Plant B: Thermal Index: Stepwise Regression	6-10
Figure 6-8 Plant B: Seasonal Variation in Trips	6-11
Figure 6-9 Plant A: Selected Facial Plots	6-12

LIST OF TABLES

Table 3-1 Indices of Interest for Organizational Epidemiology That Are Likely to be Available at Energy Facilities.....	3-2
Table 4-1 Indices Obtained with Adequate Detail and Number of Observations for Analyses.....	4-2
Table 6-1 Significant Pairwise Correlations at Plant A (for at least one lag)	6-4
Table 6-2 Significant Pairwise Correlations at Plant B (for at least one lag)	6-4

1

INTRODUCTION

Human performance is fundamental to the planning, design, procurement, installation, management, analysis, operation, and repair of energy infrastructure, as well as to the introduction and acceptance of new technology. Managing and improving human performance are critical to optimize the performance of the energy infrastructure and achieve economic, environmental, and social goals.

Human performance is also a critical determinant of safety in the energy and other industries. It is often estimated that a majority of significant events (accidents or mishaps) in commercial nuclear power plants and other industrial facilities—perhaps as high as 70-80%—involves human error or inappropriate action as a critical element (e.g., Ayres *et al.*, 1993; Muschara, 1997). In addition to contributing to significant events, human error also plays a frequent role in exacerbating the severity of consequences (Fujimoto, 1994; Heyes, 1995). Successful reduction of human error problems yields clear benefits with respect to both safety and cost (e.g., Lanoie & Trottier, 1998; Smith & Larson, 1991).

In regulated, high-risk industries such as electricity production and delivery, the potential economic, environmental, and social costs of events—including the remote but real possibility of a catastrophic outcome—motivate prospective efforts to identify antecedent conditions associated with human errors and performance deficiencies. Ideally, selection of corrective actions and planning and implementation of preventive measures could be implemented based on the appearance of antecedent conditions of problems; similarly, detection of antecedent conditions that foreshadow good performance could guide the allocation of resources for human performance optimization. Methods and tools that enable prospective or proactive human performance intervention thus could prove extremely valuable in the energy industry and other sectors.

Numerous tools based on the management of organizational factors have been proposed and applied in industry settings for human performance improvement (see review in EPRI, 2001a); selection of site- or situation-specific tools would benefit from identification of antecedent conditions. This report describes a methodology for developing predictive models of human and facility performance based on statistical analysis of historical (archived) data from energy facilities. Application of this method could be expected to improve the effectiveness of human performance management efforts.

Research Context

The “Human Performance Management: Database and Analysis” (HPM) project, a major element of EPRI’s Strategic Human Performance Program (1999-2001), sought to improve

understanding of how aspects of worker-, workplace-, management-, and organization-centered factors affect human and facility performance. The overall strategic program is providing tools, capabilities, and services to optimize human productivity and reliability in specific workplace environments, as well as to anticipate and address factors with adverse impacts on human performance and on the productivity, reliability, and safety of energy and other facilities.

An extensive literature and background review, conducted as a lead-in to the HPM project, was published in 2002, with a substantial annotated bibliography incorporated in the appendices (EPRI, 2002a). It surveyed relevant technical human performance literature and state-of-the-art human performance data collection, handling, and analysis practices developed for and applied in industry and government contexts. The literature and experience review was intended in large measure as support for empirical research.

Two empirical studies were initiated as the most substantial part of the effort carried out under the strategic HPM project. The central study, described in this report, sought to apply what may be called *organizational epidemiology* (Rosenthal, in Hale *et al.*, 1997). The second, the “Predictive Validity of Leading Indicators of Human Performance” (PV) study, took an initial look at a related approach (EPRI, 2001b).

The organizational epidemiology concept involves linkage of data regarding human and facility performance with information on workplace conditions in order to explore possible relationships between corporate performance measures, such as events or accidents, and organizational attributes. The study that is the central topic of this report took a broad perspective on the types of antecedent conditions that could influence both human and facility performance.

Preliminary findings from the study were described in an interim report, *Organizational Epidemiology: Analytical Approaches for Predicting Human and Energy Facility Performance* (EPRI, 2002b). Interim findings from the literature/background review and the two empirical studies were briefly summarized in several conference presentations (Murray *et al.*, 1999; Gross *et al.*, 2000; Ayres *et al.*, 2001; Gross *et al.*, 2001).

Report Organization

Section 2 of this report reviews the background for this research study, based primarily on the review conducted earlier (EPRI, 2002a). It discusses characteristics of human error and describes how discovery of the factors contributing to human error, i.e., the antecedent conditions, may enable development of proactive human performance improvement measures.

Section 3 details the methods used to identify data for the study. This includes discussion of the empirical approach chosen for this work, as well as description of the types of data considered most likely to be useful. Section 4 describes the actual data collection process and the final data set. Section 5 identifies analytical tools considered for trial use in the project, as well as methods that were actually applied. Section 6 details the analyses and findings. Section 7 covers the practical lessons learned for conducting this type of research, both from this study and from the previous PV study. Finally, Section 8 provides conclusions and recommendations, and references are listed in Section 9. Appendix A details results from trial use of a multivariate visualization technique.

2

HUMAN ERROR—CAUSES AND ANTECEDENTS

Much has been written about human error. Numerous reviews are available (e.g., Moray, 1992; Park, 1997; Reason, 1990, 1997), and a variety of cognitive models have been developed for error generation. The literature and background review conducted for the strategic HPM project guided the research directions pursued in this empirical study of human performance prediction in energy facilities. Conclusions and recommendations from that review are summarized in this section; more details, including an annotated bibliography, can be found in the full literature and background review report (EPRI, 2002a).

Human Error

The term *human error* is itself controversial. It implies fault and invites blame even though, in many instances, the person(s) who acted improperly did not do so intentionally. The general term *human error* is used in this report for the sake of convenience (since it is the most common term in the field), but public pronouncements and workplace safety programs might do well to adopt more neutral language for referring to situations in which human action or inaction is judged to be less than adequate.

Human error is sometimes distinguished from deliberate inappropriate action or inaction. For this project, *human error* is understood to include both unintentional and intentional errors of commission or omission, while it is recognized that categorizing errors by type or circumstance can be useful for understanding, prediction, and prevention. For example, one common error typology distinguishes between skill-based, rule-based, and knowledge-based errors (Rasmussen, 1986), and several researchers add a fourth category for judgment- or attitude-based errors (Lehto, 1991; Ayres *et al.*, 1993). Intentional errors, including deliberate sabotage, could be regarded as extreme cases of judgment/attitude errors.

The preponderance of judgment/attitude errors in many contexts—and their resistance to training programs (Ayres *et al.*, 1993)—creates a quandary for human performance improvement efforts. Problems with traditional enforcement solutions for such errors in one area of human behavior in an industrial context are discussed in a report produced under the EPRI Strategic Human Performance Program (EPRI, 2001c), as are suggested directions for human error reduction through implementation of behavioral safety programs. Error types that are more difficult to eliminate, however, are not necessarily more difficult to predict. Indeed, the very difficulty of improving certain aspects of human performance by traditional solutions, such as training and enforcement, increases the need for tools that can help predict such problems in advance.

Antecedent Conditions

In order to find ways to reduce error frequency (or the frequency of nonoptimal performance) in the workplace and to mitigate the associated consequences, it is important to identify antecedent conditions for human errors. Strictly speaking, it is not necessary to identify *causes* of human error, nor even the facilitating factors that allow errors to occur (although knowledge of both causes and facilitators would be very helpful, and this approach receives considerable research attention). Rather, considerable safety gains should be possible if *antecedent conditions* can be discovered, allowing managers and supervisors to predict general error trends and to spot trouble in advance—even if the causal links are not understood. In this project, the search for antecedent conditions concentrates on *antecedent conditions for human error* (ACHEs). It is just as reasonable, however, to try to find antecedents for human performance improvements or, more generally, antecedents for human performance changes.

For antecedent conditions to be useful for predicting future performance, they need to be causally related in some direct or indirect way to the performance; otherwise there would be no reason to believe that the antecedent conditions would have predictive value. It is important to note that the causal relation can take several forms. An antecedent condition can *cause* a human performance change, either directly (e.g., hot weather could cause people to have mishaps) or indirectly (e.g., hot weather could cause high absenteeism, in turn leading to mishaps). Another possibility is that an antecedent condition can *reflect* or *be affected* by some factor that also causes a human performance change, either directly (e.g., a rise in minor injuries could reflect hot weather, which might lead to mishaps) or indirectly (e.g., a rise in minor injuries could reflect hot weather, which might lead to maintenance errors that would show up later as facility performance problems). Thus, the search for ACHEs includes but is not limited to the search for causes of human performance problems. For practical purposes, it would be useful to find measures that are *correlated* with future performance, even if the causes of the correlations are not understood.

Deductive and Inductive Approaches

Two general approaches may be used to find antecedent conditions of human performance changes or human error: deductive and inductive. Deduction involves reasoning from principles to specific conclusions; induction involves generalization from data to general rules. The contrast between these approaches may be used to distinguish much of the previous work on error prevention from the study described in this report.

Deduction makes use of human intuition and insight to develop models of human error and organizational behavior. As a prime example, the search for *leading indicators* of human performance has involved review of both research literature and accident data. Work for the U.S. Nuclear Regulatory Commission (NRC) has produced proposed leading indicators for nuclear power plant safety based on experience in other industries; the NRC's proposed indicators include significant incidents, reportable incidents, precursor incidents, equipment-forced downtime, safety system unavailability, and unrelated contained releases (Connelly *et al.*, 1990; Van Hemel *et al.*, 1991; see also American Society for Quality, 1999).

Recently, in work funded by the EPRI Nuclear Sector's Human Performance Technology (HPT) Program, review of available models of human performance led to the identification of seven "recurrent themes" that could form the basis for proposing leading indicators of human performance (EPRI, 1999a, 2000, 2001d; Wreathall & Jones, 2000). The seven themes represent a high-level or very general synthesis of cultural or organizational factors. Through a process of deduction and discussion, the themes could be tied to industry-specific issues and eventually to potential indicators that would reflect conditions at a site or facility relevant to the corresponding themes. Since the themes were intended to represent factors that are widely believed to affect human performance in work settings, it was hoped that related indicators could serve as leading indicators, giving advance notice of human performance changes.

In a study performed under the strategic HPM project, based on the prior work performed by the Nuclear Sector's HPT Program, the predictive validity of the leading indicators methodology was explored (EPRI, 2001b). Results provided support for the premise that such indicators ultimately might help predict facility performance outcomes and guide human performance interventions. In the present report, the practical lessons regarding research methodology issues (Section 7) and the conclusions regarding performance prediction (Section 8) draw in part on the findings of the predictive validity study.

At the opposite end of the spectrum, a purely inductive approach would start with data on human performance and the workplace context, and atheoretical analyses would be used to look for patterns or relationships. In principle, given enough data about the background or context in a workplace—along with information about observed human errors—it should be possible to discover predictive relationships (if any exist) between background antecedents and the errors.

In practice, a purely inductive approach does not make sense; some initial decisions need to be made about the data to be collected, based in part on intuition, convenience, and similarity of possible predictors to factors that appear relevant based on prior research. The models and findings of the deductive approach can be used here to suggest potentially interesting measures. Alternating the complementary processes of induction and deduction—observations lead to generalizations that lead to hypotheses or research questions to be tested or examined with further observations—is a normal feature of the experimental sciences.

Most attempts to identify potential ACHEs involve (or begin with) deduction. For example, if accidents seem to have occurred when complicated tasks were performed across a shift change, it makes sense to suggest that the scheduling of such tasks across shift changes might be a predictor or antecedent condition for human performance problems (logical deduction); a hypothetical causal mechanism would be inadequate information transfer between personnel at a shift change. Thus, a combined approach—reasoning from observations to hypothesized principles (induction) and then back to potential specific indicators (deduction)—is useful for suggesting potential antecedent conditions.

The distinguishing feature of a strongly inductive approach to human error precursors (or an inductive phase of investigation) is the effort to collect a wide range of measures about context or background (and thus about possible antecedent conditions) without second-guessing the nature of any relationships that may show up. Rosenthal (in Hale *et al.*, 1997) proposed the pursuit of *organizational epidemiology*, linking databases to explore possible relationships between accidents and the attributes of organizations and regulatory systems in the chemical

industry. The hope of such an approach is to find emergent and perhaps unanticipated patterns of relationships. Instead of starting from a proposed causal relation and then seeking confirmation, the inductive approach would discover a consistent relationship between conditions and performance and, thus, invite speculation, analytical consideration, or research to understand the causal basis.

Organizational Epidemiology in the Energy Industry

Conducting organizational epidemiology in the energy industry involves three components. First, detailed data on errors (or other outcomes to be predicted) need to be collected and organized. Second, extensive related data reflecting background conditions must be identified and gathered. Finally, analytical tools for discovering patterns and predictive relationships need to be applied. Issues related to data are discussed below and in Section 3; analytical tools are discussed in Section 5.

The most useful error data are likely to involve reports of incidents. The term *incident* is used here to cover a wide spectrum, from major accidents and events at one end, through minor mishaps to near misses, to inconsequential errors at the other extreme. Catastrophic accidents need to be studied in great detail because of their severe consequences, but (fortunately) they are too rare to permit systematic, quantitative (statistical) study of the causes and likelihood of human error. Accident frequency, of course, is at least partly a function of *exposure*, or how often a given activity takes place; it is not surprising that there are more fatal accidents in the trucking industry than in the power industry, given that far more annual person-hours are spent working in the former than in the latter.

Review of incident reports can provide insights into commonly attributed causes and into the apparent success of various safety interventions, such as with regard to errors involving selection of the wrong unit or train in a nuclear power plant (EPRI, 1994). Data regarding workplace injuries (and other occupational health and safety factors) in the energy industry are being studied in a continuing project by the EPRI Environment Sector's Occupational Health & Safety Program (EPRI, 1999a, 2001e). Accident reports, however, have limitations for studying human error: Accidents are rare and uncontrolled events, reporting tends to be inconsistent and biased towards more serious events (Thompson *et al.*, 1998), and accidents tend to be complex and to raise concerns about liability and punishment within organizations (Tamuz, 1994). Near-miss reports seem to offer a desirable alternative, but such events are rarely reported with adequate consistency and detail for analysis; many energy companies have programs in place for near-miss reporting (e.g., EPRI, 2001e), but these programs typically lack mechanisms and policies to ensure frequent reporting and centralized archiving. Thus, in this study, incident reports were selected as a source of more direct and available data on human errors at energy facilities.

Selecting and obtaining data related to the context within which errors occur represent major hurdles for an inductive or empirical approach to human error analysis in complex organizations. The number of variables that might be studied in a large work setting such as an energy facility is too large to justify any attempt to exhaustively collect, organize, and analyze associated data. Even in the restricted environment of a control room, many types of measures of work conditions, operator conditions, and human and system performance have been considered, such as in the ongoing research program at the Halden Reactor Project (e.g., Haugset, 1997).

From a practical standpoint, it would be ideal if useful contextual measures could be found among data streams that are already collected on a routine basis. This would allow researchers to study the predictive value of the various measures using historical (archived) data; it would also allow organizations to make human performance predictions without additional and possibly cost-prohibitive data collection efforts. Discovery of predictive relationships requires considerable historical data. Thus, finding predictive value in the data that have been collected for years at energy facilities would help companies obtain practical results more quickly and more economically than if new data collection procedures had to be instituted.

On the other hand, although complex energy facilities already collect a staggering amount of data, there is no assurance that they collect the best data for examination of potential antecedent conditions for human performance and safety. Growing interest in human performance prediction and improvement is leading some nuclear plants to begin collecting new measures, such as survey responses following task completion (EPRI, 2001g) or periodic observations of the frequency of safe behavior by employees, as well as measures intended to bear on selected leading indicators of organizational health (EPRI, 2001b, 2001d). Eventually, when sufficient data accrues, studies of organizational epidemiology may be enhanced by the availability of such measures.

An empirical inductive approach to studying context-behavior relationships needs to cast a rather broad net for data, but some tentative guidelines also need to be adopted for what is most likely to be useful. Rather than relying too heavily on intuition or educated guesswork, it is appropriate to consider research on factors that influence human performance. Indeed, that was a major goal throughout the literature review conducted for this project (EPRI, 2002a): to learn from past efforts to identify worker-, workplace-, management-, and organization-centered factors that influence human and facility performance and to use these factors for predicting future performance.

There have been several attempts to review and consolidate previous work on factors that influence human performance in organizational and industrial settings. For example, Wilpert & Miller (1999) reduced 160 potential influencing factors at nuclear power plants that had been proposed in 13 organizational factor models to just over 60 factors, grouped in seven categories. Similarly, participants at a 1998 workshop on organizational factors related to nuclear power plant safety reached a consensus on 12 major factors important for safety (Committee on the Safety of Nuclear Installations, 2000):

- External influences (from outside the boundary of an organization)
- Goals and strategies
- Management functions and overview
- Resource allocation
- Human resources management
- Training
- Coordination of work
- Organizational knowledge

- Proceduralization
- Organizational culture
- Organizational learning
- Communication

The report from that workshop provides definitions and “aspects” for each factor, as well as suggested techniques for gathering data. Unfortunately, most of the data, with the exception of some types of performance data, would be unusual to find routinely in energy industry or other organizational settings. Many of the items listed as aspects of the major safety factors have no clear counterparts in routinely archived facility data. The possibility exists, however, that routine data may provide surrogate measures—ones that indirectly reflect the influence of these or other important characteristics. It is not essential to understand connections between measures and underlying factors, although theory (or models or hunches) could lead to exploration of measures that would otherwise not be obvious. For example, possible ACHEs such as stress and workload (which are likely to influence performance) may not be measured directly in typical energy industry settings, but they may be captured indirectly in available measures such as overtime hours or corrective maintenance activity.

There is general agreement that operational safety depends on a wide variety of organizational, environmental, task, and worker factors. Wreathall *et al.* (1991) summarize this state of affairs in the “onion model” of human performance influence factors, with the worker at the center of rings of influence from the team and work environment, the surrounding organizational and corporate factors, the facility and site conditions, and the outside public and regulatory environment. Such factors have been codified as *performance-shaping factors* for probabilistic risk assessment in the nuclear power industry (e.g., Cooper *et al.*, 1996).

Selection of appropriate and useful variables to include for organizational epidemiology requires an iterative process. For example, if preliminary analyses indicate that certain factors seem to have no predictive value, they may be given lower priority for future data collection. On the other hand, ruling out variables on the basis of intuition, past research, or failure to find an interesting pattern may compromise the chances of finding new and nonintuitive patterns when a larger data set becomes available.

Finally, it bears repeating that the goal of this research project is *not* to understand the causes of human error and human performance problems in the energy industry (although that would be a valuable aim). Indeed, such an understanding may in principle be impossible within the framework of current science. The recent work on computation and complexity by Wolfram (2002) suggests that complex behavior—including that exhibited by even simple systems, let alone that of multi-agent interconnected organizations such as energy facilities—may arise from fairly simple rules and yet be impervious to analysis. Even if computational models can ultimately be developed to mirror the complex behavior of facilities and organizations, both understanding and detailed prediction may be beyond analytical reach. Nevertheless, statistical models that accept probabilistic or random (i.e., not understood) components in complex systems frequently provide a useful degree of predictability, and the same may be true for energy facility performance.

3

METHODOLOGY—DATA IDENTIFICATION

As noted in the previous section, the literature and experience review (EPRI, 2002a) indicated that an exploration of organizational epidemiology for energy industry facilities should involve the collection of data relevant to human and facility performance (outcomes) as well as to potential ACHEs (antecedent conditions of human error or, more generally, predictors of human performance). This section describes the basic methodology employed to identify appropriate data for the HPM project.

Data Sought

To pursue the primarily inductive, or bottom-up, approach planned for this strategic project, an effort was made to review the types of information normally collected and archived at energy facilities (for more details, see EPRI, 2002a). At two nuclear power plants (owned by separate U.S. companies) and two fossil power plants (owned by one non-U.S. company), several days of interviews were conducted per site with a variety of technical and managerial personnel. Interviewers spoke with senior management and with other personnel responsible for various areas within the organization. Based on those interviews and discussions and a synthesis of results across the four sites involved, *indices of interest* were identified as shown in Table 3-1. It was not expected that all of these types of data could be obtained at a single site. It was also not known whether there would be a great deal in common between this list and what might be available at other types of energy industry facilities.

Broadly speaking, two types of measures are needed for organizational epidemiology. On the one hand, outcome or performance measures are of practical interest within the energy industry, and they presumably reflect (to varying degrees) the influence of human performance and human error; thus, injuries, incidents, productivity, and unit trips are all important outcomes. On the other hand, the rest of the variables are hoped to characterize and influence the climate and conditions within which people work and thereby to harbor potential ACHEs.

The distinction between predicted (outcome) and predictor variables is not fixed in a correlational study such as the present one. Injuries to facility personnel are costly and unfortunate outcomes, to be reduced or avoided if possible. However, the rate of even minor injuries among staff may be useful as a barometer for less directly measurable factors (e.g., safety climate, morale, job stress), and it may allow prediction of subsequent problems having other outcomes.

For the purpose of analysis in the present study, the data on errors and on context need to cover a substantial time span, and they need to be collected on a regular basis. As with any study of events or changes across time, it is advantageous to cover as long a time span as possible. In the

course of interviews at the four plants, it appeared that 5 years was a reasonable span for which to seek data for many of the variables. A common time interval for observations is required for analyses, and it was found that observations on at least a monthly basis would be available for most types of data for which records were kept; variables recorded more often (e.g., daily or weekly) could be converted to monthly totals or averages in order to use a common monthly basis.

Table 3-1
Indices of Interest for Organizational Epidemiology That Are Likely to be Available at Energy Facilities

Measures	Indices of interest
Error-related	Events, incidents, injuries, errors Investigations, root cause codes, apparent cause codes Corrective actions Problem observations Positive behavior observations
Facility-related	Facility performance data, service records Equipment trip records Facility history
Worker-related	Total hours, overtime hours Shifts and work schedules Absenteeism, lost work days, voluntary departures Worker demographics, years of experience, promotions Training scores and records Hiring, retraining, job succession
Work/task-related	Project and budget overruns Operator workarounds Procedural changes (including temporary) Preventive maintenance actions Corrective maintenance Complaints, suggestions, human resource concerns Work orders, parts availability Clearance, tag-outs
Management/ Organization-related	Departmental self-evaluations Surveys Evaluations by external regulatory agencies Safety programs Human performance improvement efforts Audits, surveillance

Conceptually, the simplest use of incident data is for counting and trending. Most analysts, however, improve the diagnostic value of their trending through some form of categorization. Rare major incidents tend to be extensively studied; minor incidents receive varying degrees of attention and causal analysis. To the extent that some consistent causal coding has been done for a set of incident reports (or can be assigned later), this information could be used to look for relationships between ACHEs and specific types of incidents (or those with specific ascribed causes).

The list of potential outcome and predictor measures shown in Table 3-1 was developed to serve as a guide for data collection efforts in order to explore the usefulness of organizational epidemiology. Ideally, a data set containing measures of all of these variables would be organized and analyzed in order to look for interrelations among predictors and for predictive relations with outcomes. Furthermore, if such data were available from numerous facilities, along with characteristics that differentiate the facilities, it might be possible to do pooled or comparative analyses, increasing the analytical power to identify predictive antecedent conditions.

4

DATA COLLECTION

To perform organizational epidemiology research at energy industry facilities, close cooperation and active participation of technical and management personnel are essential. Initially, collaborative work was planned with two U.S. nuclear power plants owned by separate companies, and visits and interviews at those sites were invaluable for developing the list of indices of interest presented in Table 3-1. Despite strong expressed commitment from company management and extensive interaction with plant personnel, however, the EPRI project team was unable to acquire suitable data from these nuclear power plants within the specified time frame.

A non-U.S. energy company also expressed interest in the HPM project, and the project team was fortunate to have the opportunity to work closely with the company's management and staff. This and subsequent sections of this report focus on research performed using data collected from two fossil-fueled electric generation plants identified by company personnel. Both are large, multi-unit coal-fired plants with more than 15 years of operational history and no near-term plans for decommissioning. The company selected these two plants in part because one had experienced relatively more performance problems in recent years, especially with respect to unintended trips. By comparing results from analyses of these two plants, management hoped to gain useful insights for performance improvement planning.

Site Visits

Following initial discussions and planning with representatives from the non-U.S. company, with the list of desired data (Table 3-1) as a guide, EPRI-associated members of the project team visited both fossil plants, as well as the company's head offices. A total of 35 people (potential data providers) were questioned over a period of 3 days, with informal, semi-structured interviews used to address planned issues regarding data content and availability and to seek suggestions from the participants.

In some cases, the interviewees recommended data sets that the researchers had not anticipated; in other cases, interviewees were surprised to learn from the researchers about data collected elsewhere in the plants. Most interviewees appeared receptive to the interests of the researchers and supportive of the described goals (a brief description of the research project had been sent to all interviewees in advance, dealing with purpose, types of data sought, expected outcomes, possible benefits, and technical support needed).

The notes and materials gathered during the visit by the research team were used to draw up a list of specific data sets and specific data providers. It was determined that most of the variables of potential interest, if they were archived in a usable manner, were available on at least a monthly basis for 5 to 6 years (or more). As such, the research team decided to request 6 years of monthly

data for most variables and to try to convert requested data for other variables to monthly format. In addition to the monthly data, some historical or descriptive information was sought for each plant, such as its operational history, dates of any major changes or renovations, and whether/when there had been any management/labor disputes.

Data Acquisition

The next step was data acquisition. A member of the core research team served as the main liaison for this process. An additional visit to each plant by the liaison was followed by email and telephonic contacts (and another visit to one plant by the liaison) over a period of nearly a year to acquire data files and clarify details about the data. This follow-up phase was finally terminated in order to conduct final analyses of the data sets (to which no further changes would be made) and to initiate preparation of the project report; at that time, numerous types of data that had been identified as available had not yet been obtained (at least in a form usable for analyses).

Table 4-1 summarizes the types of data that proved suitable for analytical use once the data acquisition phase was terminated.

Table 4-1
Indices Obtained with Adequate Detail and Number of Observations for Analyses

Measures	Indices of Interest
Error-related	Major, significant, minor incidents (Plant A only) Injuries (Plant B only)
Facility-related	Facility performance data Unit trip records Facility history
Worker-related	Total hours, overtime hours Absenteeism, lost work days Worker demographics, education (Plant A only) Personal problem reports (Plant B only)
Work/task-related	Procedural and instructional changes (Plant A only) Preventive maintenance actions (Plant A only) Corrective maintenance Work orders (Plant B only)

Except as noted, indices in each category were available from both plants, although not always in the same form. Data on injuries, illness, and absenteeism (both counts and rates) were broken down by department (engineering, operations, mechanical/maintenance, and service areas).

The data sets included notable omissions from the list specified in the original study design (see Table 3-1), which necessarily compromised the research aims. For example, it was not possible to obtain detailed cause codes for a sufficiently large group of events. In general, detailed investigations are conducted at these facilities only for major events, which may occur once a year or less frequently. Unit trips, on the other hand, which may occur up to several times per month at each plant, are categorized as being plant-related (and thus possibly due to human action at the plant) or external; some effort is made to identify root causes, but they are not generally investigated in sufficient detail to be sorted readily and consistently by type of human error or less-than-adequate action. Training scores and records were available only in voluminous hard copy, and they could not readily be summarized or linked to employees in a way suitable for inclusion in the analyses. There was very little information directly related to management/organization-related issues, other than general observations (e.g., that efforts were being made in recent years to improve human performance; participation in this research project was one aspect of that effort).

Before the statistical analyses could be conducted, considerable time and effort were devoted to reviewing the data and organizing the data set from each energy facility. The first step was to learn about the individual variables, considering issues such as how each variable was measured, whether a consistent definition was used throughout the time of measurement in the data set, and whether all of the observations or data points should be treated as valid. This involved examining the values for each variable (in tables or in simple descriptive plots) to look for outliers (divergent values), missing values, or other anomalies, as well as communicating with people from the facility who were responsible for making the original measurements or who understood the data collection process and terms. In the absence of additional information, missing values were omitted from the analyses.

The final data set for Plant A consisted of 72 observations (6 years of monthly values) for most measures; several plant performance indices were only available for the most recent 5 years of the period (60 observations), and there were several missing values among the counts of sick days and injuries. For Plant B, numerous measures had only 4.5 to 5.5 years of data available; counts of reported personal problems were too incomplete to be included with confidence, and the associated data included multiple counts (e.g., if both a marital problem and a financial problem consultation were reported in a month, it was impossible to tell whether the reports were for two different employees or the same person).

Four outcome measures had sufficient data to be included in the analyses: the number of trips per month, the thermal index (a measure of how often and to what extent critical temperature readings exceeded desirable levels), the chemical index (a measure based on readings related to the material processed in the plant), and unplanned energy losses. Unfortunately, it was discovered during the data review period that the procedure used for measuring the thermal index might have changed during the period covered by the study. Thus, it became apparent that any results involving the thermal index would possibly be contaminated by the change in measurement procedure and, thus, difficult to interpret.

Lessons learned with respect to the process of data collection are discussed in Section 7.

Confidentiality of Plant Data

The management of the collaborating company and of the two participating fossil plants provided data for this strategic project with the understanding that attempts would be made to avoid the inclusion of identifying details (and proprietary or business-sensitive information) in any reports, either for EPRI funders or for the public. In this report, a concerted effort has been made to present important research findings, along with examples of quantitative results, without compromising the confidentiality of the data set, the participating facilities, or the collaborating company. Consequently, many quantitative details (which are not necessary for understanding the conceptual findings), such as the scale numbers for graphs and the specific means for calculating some of the measures, are omitted throughout this report and, especially, in the presentation of analytical results in Section 6.

5

METHODOLOGY—ANALYTICAL METHODS CONSIDERED AND APPLIED

As noted in Section 2, the third component involved in conducting organizational epidemiology in the energy industry is to examine the suitability of candidate quantitative analytical tools in light of collected data and to apply selected ones in an attempt to discover patterns and predictive relationships.

Methods Applied

It was decided that the primary focus of the analyses would be prediction of the monthly rate of unplanned trips. There was considerable interest on the part of the collaborating company (both expressed explicitly and demonstrated in its own internal efforts) in understanding the causes of (or precipitating factors for) trips and in reducing their frequency, since these involve significant economic losses. There was also a strong sense among company management, supported by their causal investigations, that human performance deficiencies contribute to a large portion of the trips. As noted, these occurred sufficiently often to be appropriate for monthly counts to be trended over a period of years.

Two plant health indices—the chemical index and the thermal index—and a measure reflecting unplanned energy losses were also singled out as important outcome measures. Although these were not regarded by the company as being as closely related to human performance as the unplanned trips, they had the advantage (for application of statistical analytical techniques) of being continuous measures rather than counts.

Possible predictive relationships between context measures and performance or outcome measures were explored through regression analysis. This began with examination of relationships between one predictor and one outcome measure at a time, using plots (e.g., plotting both measures across time on the same figure, so any strong, obvious relationships might be detected) and two-way (pairwise) correlations. Correlations were examined at various lags or time delays between the potential predictor and the outcome measure; for this study, lags of up to 12 months were examined. Any context variables found to be statistically associated (concurrently, or at one or more time lags) with an outcome measure were entered into a stepwise regression. This technique is designed to find the best-fitting mathematical model for predicting the outcome measure from one or more predictors. All calculations were performed using routines from MINITAB™ (available from Minitab, Inc., www.minitab.com).

Most of the analyses were conducted using the data from each plant separately. Although many of the variables were common across the two plants, there was no *a priori* basis for assuming that both plants would have the same ACHEs affecting performance in the same manner. Rather,

it would be very interesting, but not necessary, to find that similar influences were involved at both sites. If the same variables were found to be useful predictors (of the same outcome measures) at both plants, this would inspire confidence in the generality of those measures as ACHes.

In addition to the regression analyses, a tool for scientific visualization was explored. Large multidimensional arrays of data may not yield their secrets to routine analyses when no theory or prior findings are available to guide the search. Sometimes, however, representing the data in several spatial dimensions can allow a human observer to detect potential complex patterns, which then can be tested by directed analyses. A variety of such *icon plots* have been explored in previous research, including star glyphs and sunflower plots (e.g., Yu, 1995). These and related techniques involve depicting values of variables as physical dimensions of a figure; thus, 10 variables could be represented as the radius lengths for the points of a 10-pointed star or a 10-petaled flower. A person who views an icon plot of values over time may come to recognize certain important patterns quickly, such as whether one of the many radii (star vertexes or flower petals) is much shorter or longer than the others, which could be useful for diagnosing the situation represented by those values.

An elaborate form of icon plot that is intended to draw on extensive innate and learned human information processing involves transforming a series of data values into the dimensions of faces, typically Chernoff faces (Chernoff, 1973). Considerable previous research has been devoted to the ways in which various dimensions of schematic Chernoff faces are either easily separated or typically integrated when perceived, although it has been difficult to establish that the faces have an advantage over other icon plots (e.g., Morris *et al.*, 1999). In this study, several potential predictors were encoded as dimensions of Chernoff faces during the initial analysis of plant data.

Additional Methods

Various other candidate analytical tools exist for organizational epidemiology studies (see EPRI, 2002a). Described below are a few that were considered or explored for use on the data collected in this project:

- Neural nets could be useful for optimizing regression coefficients in a larger data set (with complete data over a substantially longer time period, and with more predictor variables), but were not appropriate for the final data sets.
- Multivariate stepwise regressions could be tried, treating a number of outcome variables as a vector (e.g., Kerlinger, 1979). However, this requires that data pass certain statistical tests for fitting a normal distribution, which in general were violated in the data sets obtained in this study.
- Another multivariate approach, canonical correlation, would search for linear combinations of the predictors as a group and the outcomes as a group in order to obtain maximum correlation, but this was not felt to be necessary in this proof-of-concept study.
- Exploratory regression analyses were performed based on a Poisson (rather than a normal) distribution for the trips (since the Poisson distribution is more appropriate for a count of rare events, such as the monthly count of unplanned trips). These analyses did not yield

substantially different results from those performed under the assumption of a normal distribution for the trips, in terms of the ability to create predictive models from the data. Thus, this approach was not pursued further for this study.

- In related EPRI strategic work to evaluate the predictive validity of leading indicators of human performance (EPRI, 2001b), the most promising findings emerged through comparison of data for two contrasting periods for which data were available: the period preceding a relatively successful planned outage versus the period preceding a more troubled planned outage at the same nuclear power plant. Several of the potential predictor variables were found to display significant differences between those two periods, lending hope that they might be useful indicators for the success of future planned outages. It was thought that similar comparisons could be conducted using the data from the two fossil plants in the present study. Unfortunately, there was nothing quite analogous to the major planned outages at the nuclear plant, nor could any discrete time periods be identified for meaningful, parallel “good versus bad” comparisons for these fossil plants.
- Another idea was to compare predictor as well as outcome measures between the two fossil plants. Since one plant was regarded by company management as having performed better than the other in recent years, it was thought that the bases of better performance might be uncovered. As it turned out, however, the two plants and their data sets (especially including the particular variables that were documented in this study) proved to be different in so many ways that it was felt that no meaningful comparisons could be made.

Some of these additional methods may merit further exploration for organizational epidemiology at energy industry facilities. With the possible collection of larger data sets in the future (i.e., more detailed taxonomies, more variables, more frequent observations, and longer time periods covered), it may be possible to improve upon the statistical predictive models that are produced by stepwise regression.

6

ANALYSES AND FINDINGS

As noted earlier, the confidential nature of the data set—and, consequently, of the quantitative findings specific to these data—prevents full disclosure of the detailed analytical results. The purpose of this strategic project, however, was to explore the possibility and practicality of identifying antecedent conditions of human performance changes and of developing predictive models for human and facility performance. Therefore, for the stated purpose of this proof-of-concept study, it is sufficient in this report to provide an overview of the results, along with examples of important findings, without including identifying details or the numerical outcomes of each analysis.

Key results from the regression analyses are described below. Sample results are provided from analyses performed using the multivariate visualization technique.

Pairwise Correlations

For each plant, pairwise parametric correlations were computed at lags (prediction intervals) of up to 12 months between predictor and outcome variables. The correlation coefficient is a measure of the extent and direction of the statistical relationship between two variables. Negative values indicate an inverse relationship, in which relatively lower values of one variable are associated with relatively higher values of the other (e.g., as temperature goes up, the solidity of ice cream goes down). A correlation coefficient close to 0 indicates that the relationship between the variables is very weak, whereas a value near -1 or +1 means the relationship is very strong.

Tables were prepared showing how well each outcome variable was correlated with (or statistically related to) the values for each predictor variable concurrently, 1 month ahead, 2 months ahead, and so on. In order to gauge the reliability of each of these correlations, one-sided p values were calculated under the normal distribution assumption. (Since the actual distributions may not be normal, these p values represent only approximations.) Correlations associated with $p < 0.20$ (or equivalently, correlations in the predicted direction with $p < 0.10$) were judged to be significant for the purpose of this analysis (i.e., a probability of less than 20% or 10%, respectively, that the relationship occurs solely by chance). Note that the criterion for significance (and the norm of specification of directionality) used here was less stringent than that typically demanded for published research in the behavioral sciences; this was done in order to explore possible predictive relationships that might not emerge strongly, on initial analysis, given that the number of observations was small.

The statistically significant correlations were not readily observable by visual inspection of the data in tables or in plots. That is, it was not obvious that predictors and outcomes were related, and certainly not that the correlation was best at specific time lags. However, the tables of

pairwise correlation results proved to be very useful for screening the data for relationships and patterns to be examined in predictive models.

At Plant A, for example, the number of sick days per month was found to be significantly correlated with the number of unplanned trips at all lags tested except 0 and 12 months, with the strongest correlations at 10, 5, and 4 months, as shown in Figure 6-1. Visual inspection of plots of unplanned trips and sick days across the 72 months of observations, as shown in Figure 6-2, would not have led easily to the conclusion that these two variables were significantly correlated—and certainly not to the specification of the optimum prediction lags. It should be noted that the correlations here, although statistically significant and even fairly strong, do not lend themselves to very reliable predictions: the R^2 values (reflecting how well the predictions fit the data) are all below 0.14, indicating that even at a 5-month lag the correlation accounts for less than 14% of the variance in the monthly values. (The variance is the mean or average of the squared variation or difference between the observed value and the average value; thus, it is a measure of how much variation there is among the observed values for a variable.)

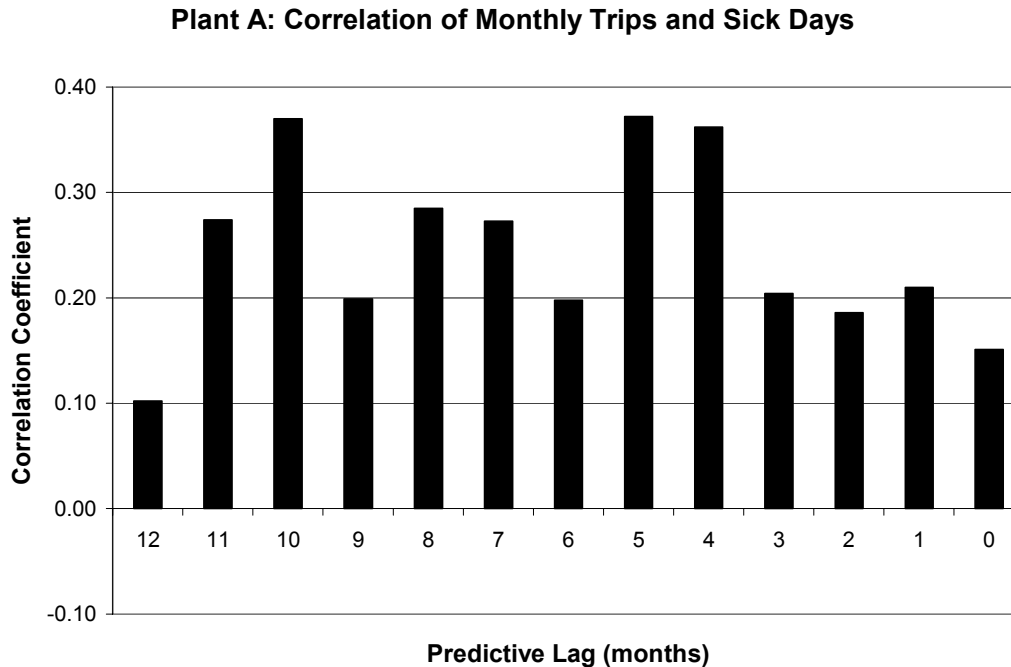


Figure 6-1
Plant A: Correlation of Trips and Sick Days

Plant A: Monthly Trips and Sick Days

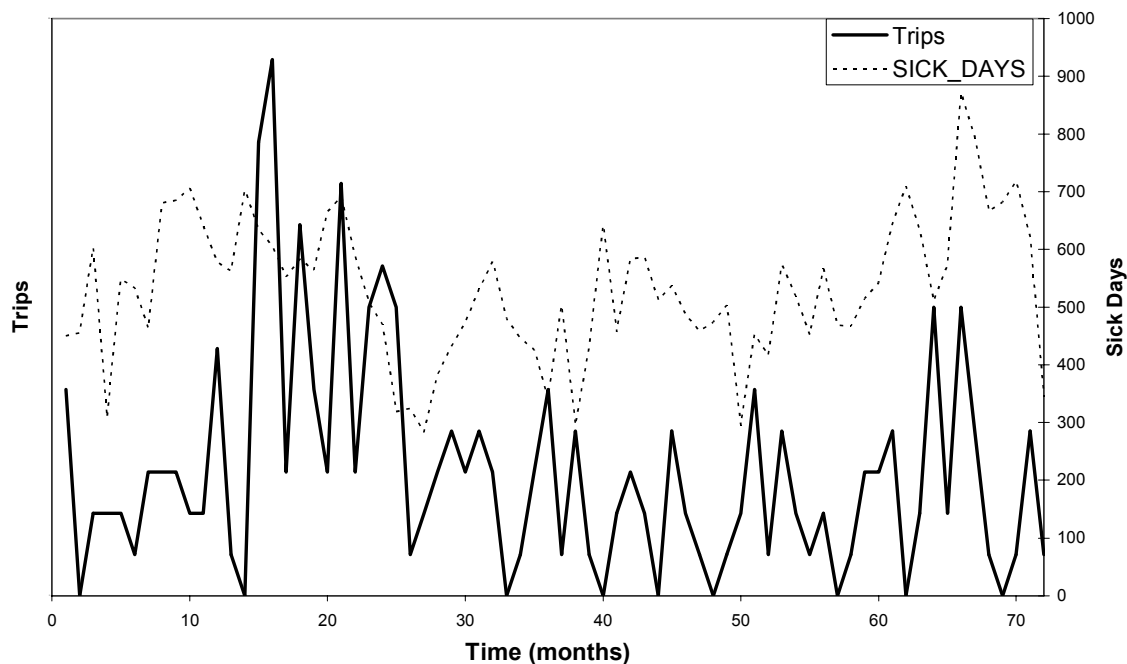


Figure 6-2
Plant A: Monthly Trips and Sick Days

Overall, for each plant, significant pairwise correlations were found (for at least one lag) between various predictors and each of the outcome measures (trips, chemical index, thermal index, and unplanned energy losses). Predictors included injuries, sickness, and/or absenteeism, separately and in combination, as well as corrective and preventive maintenance measures, separately and in combinations such as differences and ratios.

Table 6-1 indicates that, for Plant A, a significant correlation was found for at least one of the prediction lags (from 1 to 12 months) between variables related to sickness, absenteeism, and maintenance and each of the four outcome variables examined. In addition to correlations obtained for the workforce as a whole at Plant A, there were also numerous significant correlations when the predictor variables were limited to specific staff sectors. At Plant B, with data available only for the variables related to injuries, sickness, and absenteeism (as noted earlier, preventive maintenance data were not obtained from Plant B), correlations were found for most combinations, although sometimes only with predictors limited to specific staff sectors, as shown in Table 6-2.

Table 6-1
Significant Pairwise Correlations at Plant A (for at least one lag)

	Trips	Chemical Index	Thermal Index	Unplanned Energy Losses
Sickness	*	*	*	*
Absenteeism	*	*	*	*
Maintenance	*	*	*	*

Table 6-2
Significant Pairwise Correlations at Plant B (for at least one lag)

	Trips	Chemical Index	Thermal Index	Unplanned Energy Losses
Injuries	*	*	*	*
Sickness	Only Engineering, Operations	Only Engineering, Operations	*	*
Absenteeism	*		*	Only Engineering

The pattern of correlations, however, was far from consistent when results from the two plants were compared (to the extent that results could be compared, given differences in their data sets). For example, as described above (Figure 6-1), Plant A had significant correlations between trips per month and the overall days of sickness per month with various lags. At Plant B, on the other hand, trips were not significantly correlated with overall days of sickness for any lag. Instead, there were significant correlations at lags of 0 to 5 months between the number of trips per month and the number of days of sickness per month for the engineering staff, as shown in Figure 6-3. Significant correlations between monthly counts of trips and sickness days were also found for the operations staff (for lags of 0 and 1 months) but not for the maintenance staff.

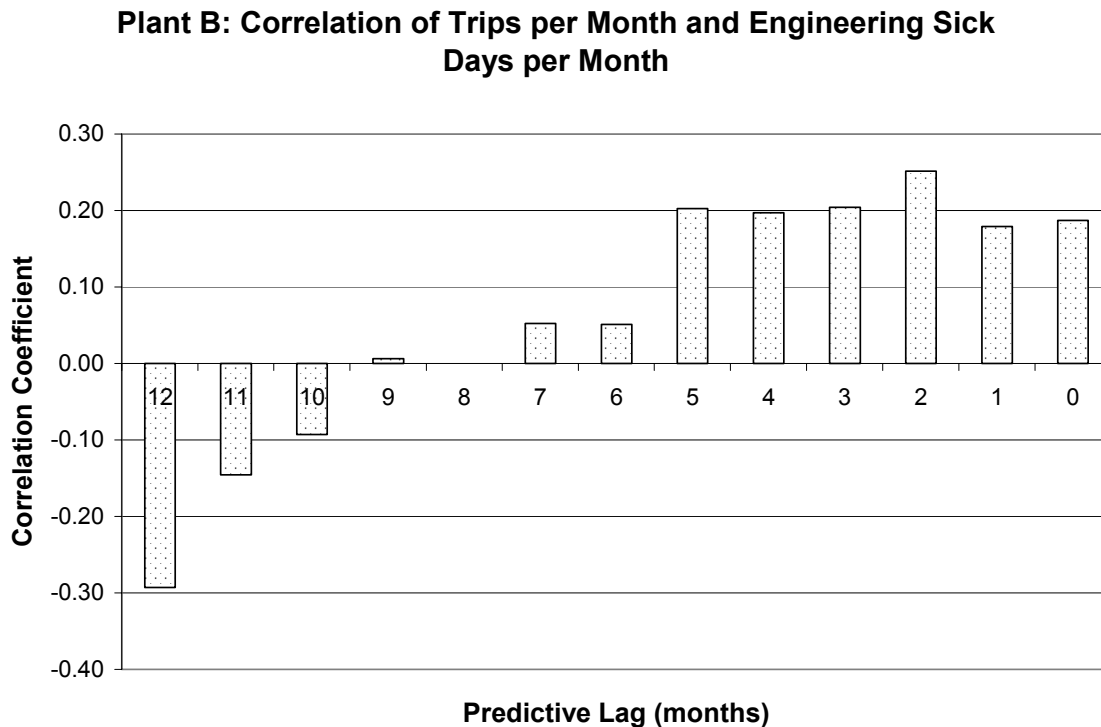


Figure 6-3
Plant B: Correlation of Trips and Engineering Sick Days

Stepwise Regressions

To explore predictive models using more than one predictor, stepwise regressions were used for each plant and each outcome variable. Detailed below are the analytical procedures and findings.

Analytical Procedures

For each plant and outcome case, the stepwise procedure begins with the predictor that had the highest pairwise correlation with the outcome (excluding 0-month lags, in the interest of being able to predict performance in advance, which was the underlying aim for this study). Then, one predictor is added at a time (using the remaining predictor that yields the largest increase, to that point, in the squared multiple correlation, R^2 , which is equivalent to adding the predictor that yields the largest remaining partial correlation)—provided that the increment (in R^2) is statistically significant. As in the case of pairwise correlations, statistical significance is here determined approximately by using the assumption of normality. The procedure is performed automatically by statistical software (MINITAB was used in this study, although other statistical packages, such as those available from Unistat, SAS, or BMDP, also perform these analyses); for these analyses, the criterion for including a predictor was that the improvement be statistically significant at approximately the $p < 0.05$ level.

It is important to note that the terms to be included in a stepwise regression model cannot be predicted just by looking at tables of pairwise correlations. Each time a predictor variable is added to the model, it may eliminate other possible predictors. In order to understand why this occurs, it is necessary to realize that various predictors may be correlated with each other. For example, suppose that both the number of sick days per month and the number of days absent per month initially are found to be correlated with the number of trips in the following month (i.e., at a 1-month lag). Sick day and absentee counts are probably also highly correlated with each other, since sick days account for many of the days that employees are absent. Once sick days are used as a predictor variable in a stepwise regression model, there may be little or no additional value in trying to use absenteeism as a second predictor in the model; this is because if sick days and absence counts are highly correlated, then the ability to predict or account for the variance in the trips per month (as measured by R^2) will quite likely not increase by much when absenteeism is added.

The stepwise regression calculations include only the time period for which all of the predictors have measurements; thus, if one of the predictors is entered in the calculations with a 6-month lag, then the first 6 months of the predicted variable are not included (because corresponding values of the predictor are not available). It would not make sense to try to predict the number of trips per month in January and February 1995, for example, if the regression model predicts the trip rate 6 months in advance and the predictor variables are only available starting in January 1995. For this reason, even though data for the trip rate may be available beginning with January 1995, the ability of the regression model to predict trip rate 6 months ahead can be evaluated only with data for trip rate starting 6 months after the time for which the predictor values are available.

The number of observations used in the analyses varied in order to ensure the greatest number of observations possible for each step. For example, when data were available for each predictor and outcome measure for a full 72 months, there would be 72 observations for pairwise correlations at a 0-month lag, 71 observations with 1-month lag, and so on down to 60 observations for a 12-month lag. Next, for the stepwise regression, 60 observations (out of 72) were used, so that predictors with lags of up to 12 months could be tried in the stepwise regression. Finally, when the resulting regression model for a given outcome measure was evaluated (by determining R^2 , the percent of variance accounted for), observations were added up to the maximum possible depending on the lags used in the model; for example, if there were two predictors in the model, with lags of 3 and 8 months respectively, then the regression model was evaluated using 64 (out of 72) months of data.

Each final stepwise regression model is constrained to non-negative values when the outcome measure cannot in fact be negative (e.g., the number of trips per month cannot be less than 0). This restriction does not affect the stepwise regression process or the R^2 values that are calculated to evaluate the models. Rather, it makes the predictions easier to understand and use. The outcome of a stepwise regression is a set of coefficients or multipliers for the predictor variables, along with a constant; these terms can be used to construct a linear equation for predicting the outcome variable.

Results

For Plant A, a two-predictor model emerged, in which the number of trips per month could be predicted by sick days in one staff sector at a 1-month lag, plus absentee frequency rate in another sector with an 8-month lag:

$$\text{Monthly trips} = 0.0188x + 18.9y - 1.903$$

Where

x = Sick days per month in Sector 1 at a 1-month lag

y = Absentee frequency rate in Sector 4 at an 8-month lag

This model accounted for 54% of the variance (just over half of the total mean squared variation of the monthly trip rate). The actual (data) and estimated (from the regression model) monthly numbers of trips are shown in Figure 6-4. Visual inspection suggests that the model captures some of the trends in trips but not the detailed month-to-month variations.

Plant A -- Trips: Stepwise Regression

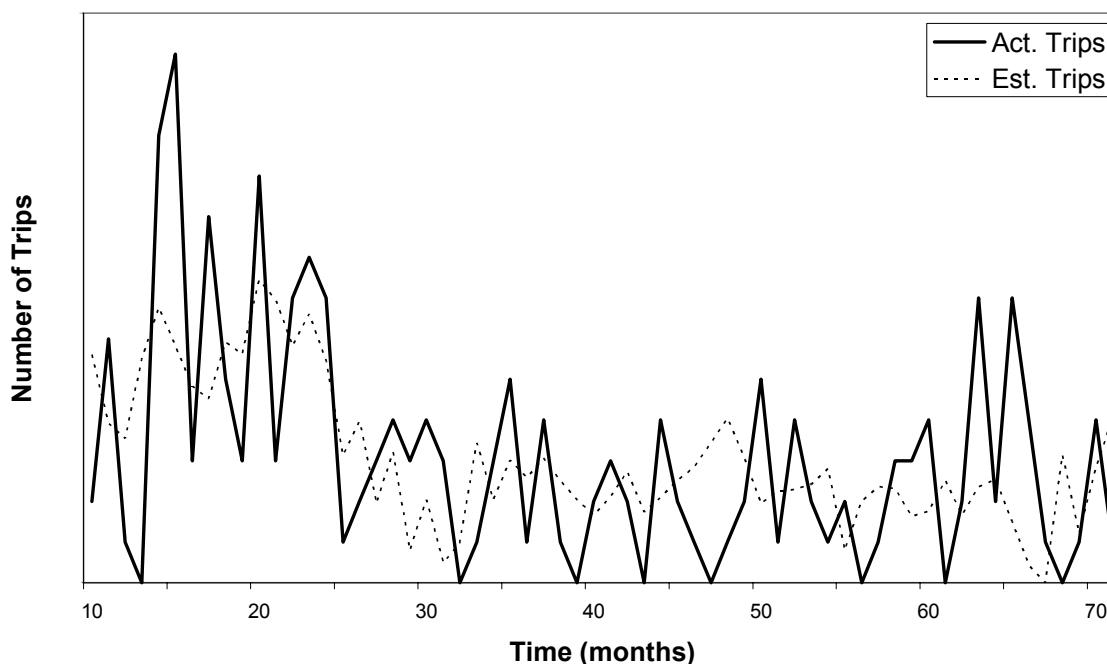


Figure 6-4
Plant A: Trips: Stepwise Regression

Somewhat better prediction was obtained for the stepwise regression of the chemical index at Plant A. In order of importance, the predictors were the ratio of corrective to preventive maintenance orders outstanding (CM:PM) at a 2-month lag, the gross sickness and absenteeism rate at an 8-month lag, and the number of sick days (a count of the days missed due to illness) and sick incidents (a count of the number of reports of illness, regardless of whether an illness

was associated with no time off or with one or more days off) for one staff sector at lags of 10 and 6 months, respectively. The resulting model accounted for nearly 60% of the variance. The actual (data) and estimated (from the regression model) monthly trips are shown in Figure 6-5. Visual inspection suggests that this model does a reasonably good job of capturing trends in the chemical index, including several peaks and valleys in the first year as well as the reduced level after the first year.

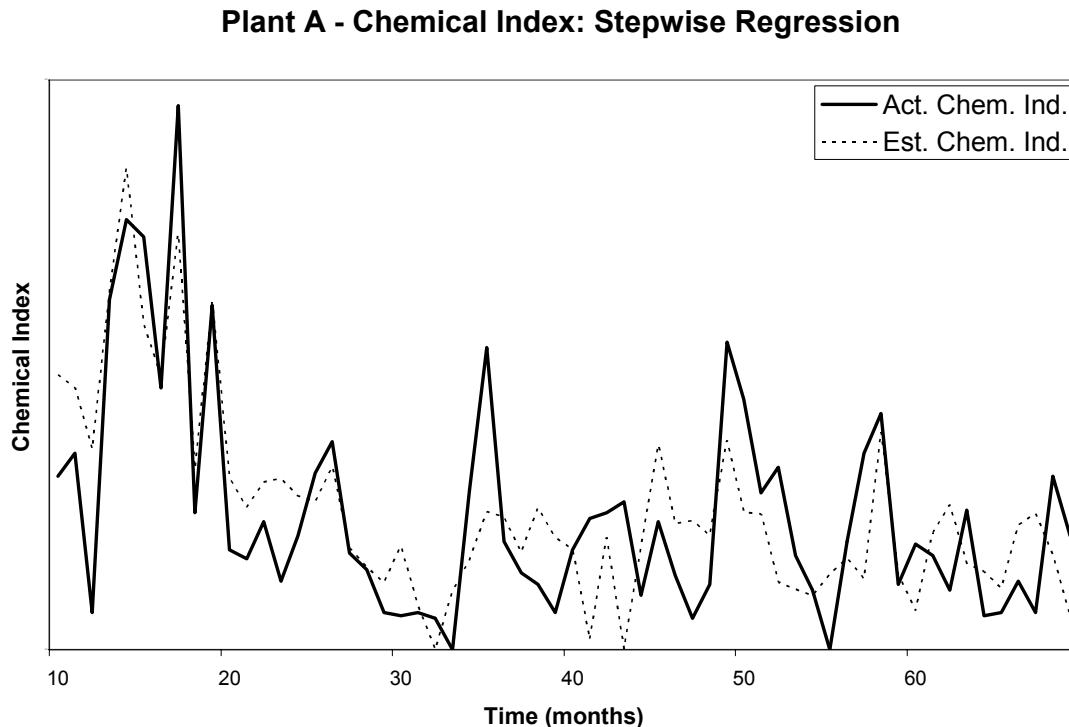


Figure 6-5
Plant A: Chemical Index: Stepwise Regression

It is again interesting to note that visual observation (in plots) of the values of the potential predictors and the chemical index would not readily lead to specifying the terms and lags likely to show up in a stepwise regression model. Figure 6-6 shows values for the first (strongest) predictor (the CM:PM ratio at a 2-month lag) and the chemical index across the 5-year period for which chemical index values were available. From inspection of this figure, it is not obvious that CM:PM ratio is fairly well correlated with values of the chemical index observed 2 months later.

Plant A: Monthly observations of chemical index and CM/PM orders outstanding

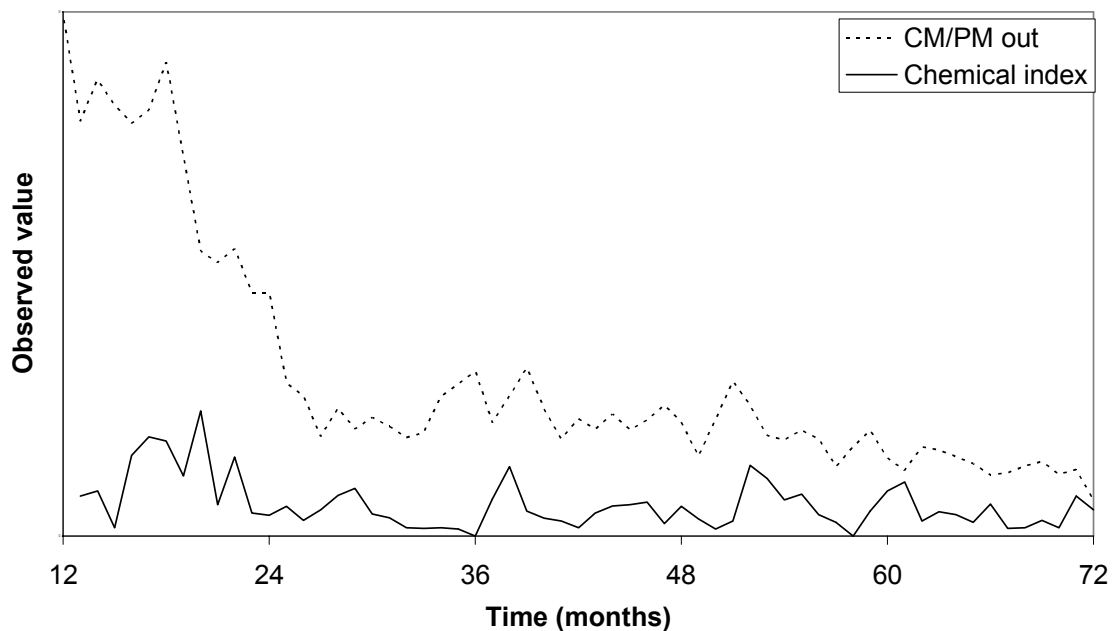


Figure 6-6
Plant A: Monthly Chemical Index and Ratio of Corrective to Preventive Maintenance Orders Outstanding

In general, the stepwise regressions were not as successful for Plant B, accounting for only a small portion of the variance in the outcome measures. The only exception was the model for the thermal index. A model accounting for more than 50% of the variance in this measure was developed, involving sick days (total sick days at a 2-month lag as well as sick days for the maintenance staff at a 1-month lag) and the injury rate (overall rate at an 8-month lag as well as the rate for power sector staff at a 10-month lag). The actual (data) and estimated (from the regression model) monthly trips are shown in Figure 6-7. Once again, visual inspection suggests a potentially useful degree of match between the estimated and actual thermal index.

Unfortunately, as discussed in Section 4, the possible change in the procedure used to collect thermal index data compromises the potential practical value of this finding.

Plant B - Thermal Index: Stepwise Regression

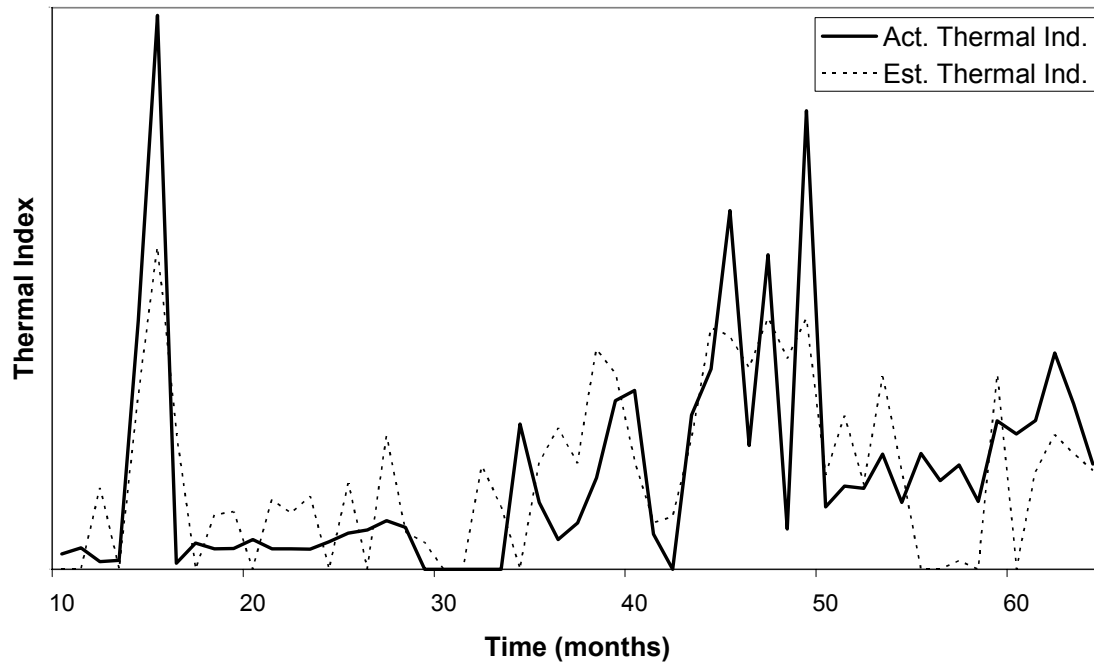


Figure 6-7
Plant B: Thermal Index: Stepwise Regression

Seasonality

An additional set of analyses was conducted to explore possible seasonality in the outcome measures. The residuals from the stepwise regression models (the difference between estimated and actual outcome for each month) were analyzed to determine whether there was a pattern based on either the 4 seasons or the 12 months of the year. In no case was there a statistically significant relationship. On the other hand, if the original outcome variables were analyzed directly, a statistically significant seasonal pattern showed up for trips at Plant B: Trips were most numerous in the summer, followed in order by spring, fall, and winter. Figure 6-8 shows the actual average number of trips per month in each season at Plant B, along with the estimated monthly average (from the regression model) in each season. It appears that each of the two predictors in the weak but significant stepwise regression model of trips for Plant B—the absentee frequency rate and the injury rate (at 6- and 11-month lags, respectively), both for the maintenance staff—had some seasonal variations, which were statistically associated, to some extent, with the seasonally varying trip rate.

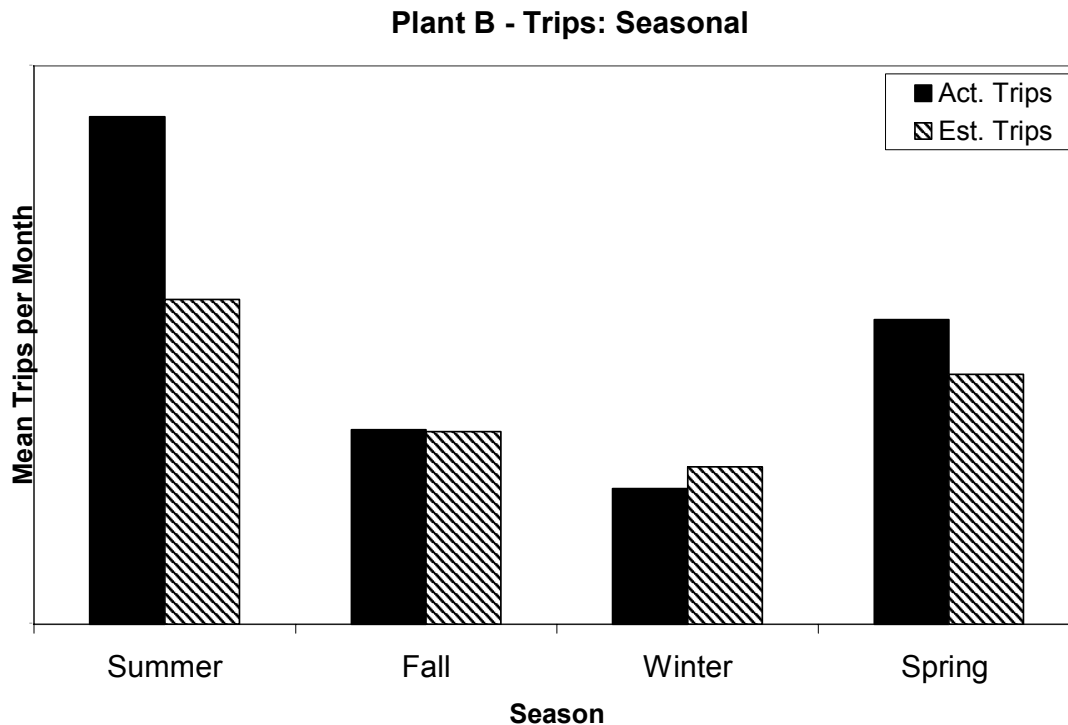


Figure 6-8
Plant B: Seasonal Variation in Trips

Multivariate Visualization

Relatively early during the initial analysis of the plant data, a multivariate visualization technique was tried with several measures from an early version of the data set from Plant A that were thought likely to be helpful as predictors. The intent was to try to discover patterns among the variables. It is possible that visualizing several variables together might allow prediction of an outcome measure in a way that would not be discovered by the pairwise correlations and stepwise regression process described earlier in this section.

In order to explore the use of multivariate visualization for discovering patterns, several potential predictors were encoded as dimensions of Chernoff faces (Chernoff, 1973). These predictors included measures involving absenteeism (the gross sickness and absentee rate, plus the absentee frequency rate), as well as a series of measures utilizing preventive and corrective maintenance data. In addition, facial area was used to encode the thermal index. For each of the 72 months of observations, a face was produced; these were then placed in order with respect to the number of trips per month. Examination of the facial plots suggested that, using this preliminary data set, the thermal index might be related to the trip rate at Plant A, with a higher thermal index associated with more trips; a staff member of the company indicated that such a relationship would not be surprising. There was also an indication that trips were higher with a higher number of preventive maintenance orders outstanding and with a higher ratio of corrective to preventive maintenance orders outstanding. (See Appendix A for more details and for the entire set of facial plots.)

The face plotting method did not appear to be useful for further exploration as a data analysis tool for this project. In the first place, these plots involve only concurrent (or 0-month lag) comparisons; in order to explore lags at 1 to 12 months systematically, allowing different lags for different variables, a prohibitively large number of arrays of facial plots would need to be generated and examined. In addition, the apparent patterns or relationships found in these analyses were identified only tentatively; there was no sensation of an emergent pattern jumping out at the viewer, which would be the kind of outcome that would make this technique most useful.

On the other hand, facial plots (or other multivariate displays) may prove useful for system monitors or analysts after important predictors have been identified. For example, as noted earlier, the number of trips per month for Plant A was significantly related to a model based on sick days in one staff sector at a 1-month lag, plus absentee frequency rate in another sector with an 8-month lag. If these two variables are encoded (or represented quantitatively) as two dimensions of a face, the difference between months with high and low trips can be quite noticeable, as illustrated in Figure 6-9. (These faces were generated using the final data set for Plant A and a simpler face configuration than those shown in Appendix A.)

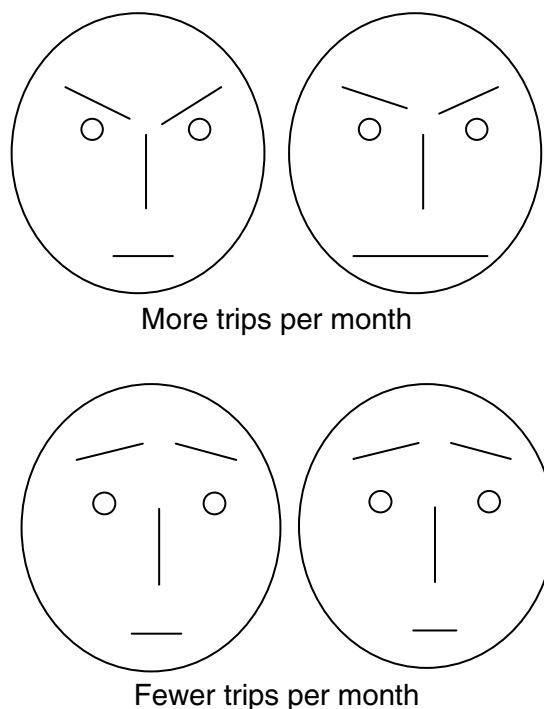


Figure 6-9
Plant A: Selected Facial Plots

In principle, a display such as this could be viewed a month ahead of the outcome, since it relies on input data with a 1-month as well as an 8-month lag. A recent EPRI report (EPRI, 2002c), which suggested directions for future strategic research related to automated tools for human performance management, identifies the integration of data analyses with iconic displays as a promising topic for exploration.

7

PRACTICAL LESSONS LEARNED FOR FUTURE APPLICATION AND RESEARCH

The aim of the strategic HPM research project has been to explore the feasibility of developing useful predictive models of human-performance-related outcomes in energy industry settings (e.g., generating plants or power coordination and control centers) based on antecedent conditions. Although an emphasis has been placed on exploring the use of statistical and other quantitative analytical techniques for this purpose, assessing the feasibility of such an approach involves other issues as well.

One of the primary lessons learned is the difficulty involved in acquiring data of the type that had appeared likely to be most useful for organizational epidemiology in energy industry contexts. The observations that follow are similar in many respects to those from the conceptually synergistic strategic work to assess the predictive validity of leading indicators of human performance (EPRI, 2001b).

From the start, it is important to recognize that the primary business of energy facilities is very different from the emphasis of researchers, even though they may share goals such as reducing accidents and increasing productivity. To a large extent, most personnel at energy facilities focus on safe and productive operations on a day-to-day basis, not on carefully and consistently archiving a variety of measures that may have no obvious or immediate pragmatic use. Despite the best intentions of facility management and staff to support strategic research such as that pursued in this project, it is difficult for them to devote sufficient and sustained resources to the research, given the demands of their primary responsibilities.

Since it was recognized in advance that the burden on the participating plants should be as light as possible, a two-stage process was devised in the hope that this would make it relatively straightforward to collect the data required for the project to succeed: (1) conducting interviews to identify potentially useful data, while receiving assurances that both the data existed in the needed form and the providers made some commitment to the effort; and (2) preparing a list of requested data files (including information on what data were sought and who had the data). No novel data were to be demanded; no new observations were to be requested of facility staff. Ideally, it was envisioned that designated data providers would electronically transfer appropriate files (in relatively good order), perhaps to a liaison, who could then pass them along to the research team.

This plan appeared conceptually sound, but the reality proved different. Despite the strong expressed commitment of management at various levels within the three companies that participated in this study, there were barriers to obtaining useful data sets. The initial experience

in this effort was time-consuming and, ultimately, not productive, but the lessons learned in the process facilitated the more successful subsequent work that is emphasized in the present report.

The first efforts made for this research project involved close work with liaisons at two U.S. nuclear power plants owned by two different companies. Visits were made, interviews were conducted, and lists of variables of interest were prepared and provided to the plant staff. It was planned that individual plant and management personnel with access to data files identified as potentially useful would provide their files for the research project, but it proved difficult in practice for the research team to obtain the identified files consistent with the requisite research timeline. Ultimately, a research team member (an information technology specialist with prior experience related to the nuclear power industry) obtained a large number of data files by visiting the plants and working closely with facility information technology specialists, but the files obtained via bulk transfer proved intractable for analysis. It was not possible to decode the fields and identify the entries of interest in order to create a usable data set for organizational epidemiology purposes.

Despite changes in the overall data acquisition procedure implemented to address the problems revealed in initial work, problems—including some that had previously been encountered and some that were novel—were also experienced in the study emphasized in this report, which involved the two fossil plants owned by one non-U.S. company. The data sought often existed only as notations in multi-volume hard-copy files. Some of the data simply did not exist, in spite of impressions from interviews, or the data were located in old legacy software systems no longer used and not compatible with current software. Information existed on different hardware platforms that did not communicate. Other types of data were contained somewhere within vast databases, but extracting the variables of interest in a usable and understandable form would have taken considerable effort on the part of busy programmers.

In the end, a great deal of the work involved in actually obtaining the sought data files and documenting their contents had to be performed by the staff of the two participating fossil plants, aided by corporate personnel. As noted, a year elapsed between when the EPRI-associated research team visited the plants to conduct interviews and when the data sets were finalized for analysis; during that period, considerable work was done to obtain data files, understand their contents, and convert variables of interest into a common monthly format. The primary liaison spent a considerable amount of time (many months) on this project during that year; personnel from the collaborating company are conservatively estimated to have devoted at least 1 to 2 person-months of labor to this effort. Thus, this represented a major resource expenditure on the part of the collaborating company.

Even though a significant percentage of the specified data files was eventually obtained from the two fossil plants, there were problems that compromised the reach of the research. These included missing observations, seemingly inexplicable outliers (values that seemed too high or low to have been valid observations), unavailable files, and inadequate documentation of recording procedures. As one example, mentioned earlier, it became known that the procedures used to produce the thermal index might have changed across time; thus, the large variations in the values (see Figure 6-5) may in part reflect changes in measurement technique rather than changes in plant conditions.

The fact that a predictive model for the thermal index could be found using stepwise regression underscores a key, though not necessarily intuitively obvious, conceptual premise in the primarily empirical approach taken in this study: The presence of a statistical relationship between variables does not imply that a meaningful explanation or understanding can be found—or that one even exists. In the absence of a plausible hypothesis regarding the conceptual nature of an observed predictive relationship, the long-term applicability of such a finding must be viewed with some caution.

Another problem for the aims of the project was that the analyses had to be conducted using monthly observations. With only 6 years (at most) of data available for many of the variables, the data set was limited to a maximum of 72 observations (months) per variable. That is a fairly severe restriction on the total size of the data set when dealing with a large number of potential predictor variables. A number of additional statistical techniques (especially methods that would depend even more on a larger total number of observations) could be tried if more fine-grained data, such as weekly observations, were appropriate and available.

Nevertheless, the results described in Section 6 suggest the promise of organizational epidemiology, as do the findings from the strategic predictive validity study (EPRI, 2001b). Most of the software tools used in this research study are widely available: Spreadsheet software (Microsoft® Excel) was used for organizing the data sets as well as for generating the graphs shown in Section 6, and the correlation and regression analyses described in Section 7 were performed with MINITAB. Implications and future directions for research and application are discussed in Section 8. Detailed below are some methodological lessons for energy facility analysts and researchers.

Lessons for Energy Companies

The experiences in this project suggest that companies can take a number of steps that would be expected to facilitate analysis of human and facility performance trends and to allow for possible future use of improved analytical tools.

Collect a wider variety of information regarding potential influences on human performance. Numerous committees and reviewers have agreed that human performance in energy facilities reflects a number of organizational and work climate factors, but the majority of these postulated factors seem not to be captured by typical data collection activities. Fortunately, in terms of both performance analysis and the operational goals of energy facilities, it seems that a number of companies are making efforts to address this situation, which would be expected to enhance future analytical capabilities.

Create a central archive for storing data on measures of interest, and update it on a regular basis. If individuals or departments collect data but there is no mechanism for combining what they have, potentially important connections and relations are unlikely to be discovered. Having a wide array of measures available from multiple organizational units of one facility, on one platform, will make it easier to analyze data, examine performance trends, and evaluate the effectiveness of performance improvement programs. In this way, knowledge can be derived from what would otherwise remain only as data or information (EPRI, 2001h; Wildberger *et al.*, 2000).

Maintain data on a per-system, per-unit, or per-sector basis rather than only retaining values that have been aggregated across an entire facility. Systems, units, or sets of units in a facility may have distinct operational, maintenance, and management personnel and organizational attributes. If performance differences exist between systems, units, or sectors, maintaining disaggregated data may allow analysts to identify organizational factors that lead to such differences. In many instances in this study, it appeared that the original data were collected at a finer grain than the form in which they were archived. It is always possible to combine data, but once the original details are lost in the archival process, they cannot be retrieved.

Document the methods and terms for each archived measure. Comparisons across time, or across facilities, can be problematic if there is uncertainty or confusion about what the numbers mean. In fact, erroneous conclusions may be drawn if an observed effect or pattern, such as an increase in a performance measure, is entirely attributable to some definitional change.

Use consistent and common measures across the organization, where applicable. There are many measures that could reasonably be collected in the same way at various facilities within an energy company, as well as for various departments within a facility. To the extent that it is clear that, for example, absenteeism rates, preventive maintenance backlogs, and other measures have the same meaning in each facility or unit, interpretable analyses and comparisons can be conducted.

Use a consistent causal taxonomy for attributing incidents (such as trips) to specific human-related causes or influences. The search for antecedent conditions of performance changes, as well as the development of predictive models, is likely to benefit from having more information about outcomes; it may be possible to specify which antecedent conditions foreshadow which types of errors. For example, trips that are attributed to failure to follow standard procedures may turn out to be statistically associated with absenteeism rates because workers acting as substitutes who have to perform unfamiliar tasks may not know exactly what to do, whereas trips attributed to equipment that has not been adequately serviced may be statistically associated with some measure of preventive maintenance work orders. Ideally, the causal taxonomy to be used would be common across the industry to facilitate comparisons. Further research will be needed to determine what aspects of a causal taxonomy are most useful for development of predictive models, and whether such a taxonomy can be improved with respect to its usefulness (by, for example, increasing the number and the grain of available coding alternatives for human-related causes).

Extend investigations and record-keeping to a wider range of incidents or events. The power and, thus, the usefulness of statistical quantitative analyses depend in part on the quantity (and quality) of data available. Major events or accidents, which are carefully investigated at energy facilities, are very rare. By extending (to the extent feasible) such tools as root cause analysis and human factors causal taxonomies to relatively minor incidents (which are more frequent), companies may improve their abilities to track important changes in the organizational environment, as well as to develop more accurate models for predicting future performance trends.

Take note of the measures that have been found to be statistically associated with important facility outcomes. In the present study, statistically predictive models were

developed using measures related to sickness, absenteeism, injuries, and corrective and preventive maintenance; there was also some indication of seasonal patterns at one plant. Indeed, sickness and absenteeism measures, which were available from both plants (unlike the injury and maintenance measures that were available for only one plant each), were included as significant predictors in stepwise regression models for both plants; this consistency lends some confidence in suggesting that such measures are likely to be useful elsewhere. Future research involving a wider range of potential predictors could indicate that there are other measures more useful than the ones that have been identified to date; in the meantime, the findings of the present study can be helpful for planning data archiving and analysis efforts.

Share information about data resources and collection methods, as well as useful analytical approaches and findings, to the extent feasible—within a facility, across facilities owned by the same company, and even across different entities. At the very least, management within a single facility should be aware of data collected and stored by various sectors within the facility; this was not always found to be the case during the interviews conducted for this project. On a wider basis, it may be possible to establish industry-wide information-sharing forums such as communities of practice (e.g., EPRI 2002d) or lessons-learned networks; events such as the annual Human Performance/Root Cause/Trending workshop series (which focuses on topics at nuclear plants but addresses more generic issues), as well as IEEE conferences on Human Factors in Power Plants, help to fill some of this need, but more frequent, extensive, and focused interchanges could prove useful. Given its established role in the industry, EPRI could well play an important role in organizing such sharing, or it could act as a central repository for data that could be analyzed so that general findings could be disseminated to the industry without compromising the confidentiality of data from individual facilities or entities.

Lessons for Researchers

The promising findings of these studies may lead to future research. Attempts to study organizational epidemiology in energy industry settings can be expected to benefit from taking the steps described below.

Obtain formal commitment from data providers, as well as from management of participating organizations. The ultimate data providers are personnel with demanding work schedules who generally are not involved in the original planning discussions and agreements about the research project. Certainly, enthusiastic and articulated support by management is of great value—and is necessary in order to secure a broad range of potentially sensitive data. However, this does not ensure that the data providers will have the time and resources available to devote to extracting, providing, and documenting data files.

Urge management to provide a variety of incentives for participation. If the work of interacting with the research team and providing data files is unrewarded, is not integrated with other duties, and/or impedes more pressing tasks, data providers may not be as helpful as the researchers (or management) would like. Leadership needs to be committed to supporting their staff in this process. In part, this requires that both data providers and their managers understand the magnitude of the task, which is likely to involve many person-days or even person-weeks of work from facility-associated personnel; outside researchers are not suitable replacements for performing some of the required site-based and site-specific tasks. Given these demands, a

reasonable allocation of time, as part of the official workload for the staff involved, would be of value in eliciting the best response from data providers.

Plan on extensive and repeated interaction with data providers. Persistence is required throughout data identification, collection, and preparation stages—from first learning about the existence of certain types of information, to arranging for file extraction and transfer, and on to getting explanations of the file contents (including acronyms). It appears ideal to have a member of the company play a primary role as a project liaison, assuming that person has sufficient time and resources allocated to the project and can interact frequently and extensively with members of the research team.

8

CONCLUSIONS AND RECOMMENDATIONS

EPRI's strategic "Human Performance Management: Database and Analysis" project marks the first major attempt to explore the possibility for organizational epidemiology in energy industry settings. It has involved a literature and experience review (EPRI, 2002a), the organizational epidemiology study (conducted primarily in cooperation with two fossil plants) reported here, and a preliminary evaluation (conducted in cooperation with a nuclear plant) of the predictive validity of leading indicators of organizational health (EPRI, 2001b).

It is widely accepted that the performance of the energy infrastructure—including physical assets, such as generating plants and transmission and distribution systems, as well as energy markets—depends heavily on human performance. Because costly incidents, accidents, outages, and other problems are frequently attributed, at least in part, to human error or inappropriate action, it follows that smooth and productive operations rely on good human performance (in conjunction with well-designed, robust, and error-tolerant systems). It is also widely accepted that human performance deficiencies have a wide variety of causes or contributing factors, and that causal code taxonomies incorporate influences ranging from individual idiosyncrasies through task and workplace conditions up to managerial policies and beyond.

In principle, it should be possible to predict human and facility performance and their variations through careful scrutiny of such contributing factors. The aim of this research project was to assess whether useful predictions are possible in practice, given the data resources currently available at energy facilities. As such, this project could represent an important milestone in the energy industry's longstanding efforts to understand and minimize the contribution of human errors to incidents and accidents, as well as to manage and optimize both human performance and facility performance.

Conclusions

Based on the findings from this strategic evaluation-of-concept project, it appears that useful predictions of human and facility performance are possible based on the analysis of antecedent conditions. Statistically significant models, based on data collected and archived for traditional purposes, emerged for each of the outcome measures that were studied. These results were achieved despite the limitations imposed by the restricted data sets that were ultimately obtained—restricted in the number of years of data that were archived, in the usability of some of these data, and, even more, by the lack of data for potential predictor measures that would more directly tap training, safety efforts, and other factors that had been deemed of substantial interest (based on the background literature review).

As described in Section 6, several of the models accounted for more than 50% of the variance in the outcome measures. Energy companies could create real practical value by taking advantage of their data resources to develop such models. It appears, based on analyses of the obtained data sets, that measures related to injuries, sickness, absenteeism, corrective and preventive maintenance activity, and, possibly, seasonality hold particular promise for indicating or reflecting the antecedent conditions associated with changes in human performance and consequent facility performance. Although these findings were developed based on analyses of data from fossil plants, the methodology employed here is likely to be useful in other energy facilities, as well as in additional industrial and organizational settings.

Of perhaps equal importance as the preceding conclusion is the finding that the precise models derived in this work could not be specified in advance. It should not come as a great surprise, for example, that a performance outcome such as the number of trips per month might be higher when sickness and absenteeism are up—or even that sickness and absenteeism levels might be related to trip rates at a later time. However, there was no basis for expecting that the best predictor of trips, within this study, would involve sick days in one staff sector with a 1-month lag, as well as the absentee rate in a different sector at an 8-month lag. These findings do not merely confirm long-held hunches about important performance-shaping factors; they also show that analytical tools are needed to quantify such relationships and optimize predictive models.

A third key finding is that predictive models can be developed by applying conventional techniques to analyze data that are already collected at energy facilities. It was not necessary to institute new data collection procedures for novel observations, nor were esoteric analytical techniques needed (the existence of such requirements could present formidable barriers to the practical implementation of organizational epidemiology). Although novel observations and experimental analyses may ultimately provide improved predictive capabilities and deeper insights into performance-shaping factors, this project has demonstrated that much can be done with widely available statistical software and the data that are currently archived at energy facilities.

Recommendations

The specific models that emerged here probably depend in part on the somewhat arbitrary collection of variables that became available for study. For example, it was anticipated, based on review of the research literature (EPRI, 2002a) and on the predictive validity of leading indicators study (EPRI, 2001b), that measures related to maintenance activity might be useful as predictors of important outcomes. Although that proved to be the case in this study, other useful predictors may well exist among the measures for which useful data were not available. For example, the predictive validity study found a strong correlation between the number of deficiencies in defenses at the participating nuclear plant that were identified first by outsiders rather than by facility personnel (predictor), on the one hand, and estimated going-forward costs (outcome), on the other; there was no indication that such a potential predictor variable was available at the fossil plants that participated in this study. Additional work would be needed to explore a larger variety of potential predictors, as well as to study the usefulness of predictors that might be appropriate in specific contexts. Based on the experiences in this project, such research would not be either easy or inexpensive to conduct.

The quantitative analytical techniques applied in this study are fairly straightforward, when considering the full range of statistical techniques, and they can be streamlined and customized for specific purposes. The three-phase process of pairwise correlations (at various lags), stepwise regressions, and statistical evaluation, as used here, may be automated through the use of simple decision algorithms (such as the one used in this study: taking the predictor and lag with the highest pairwise correlation, then adding predictors one by one until the percent of variance accounted for does not increase substantially). Alternatively, an analyst may intervene in the process in order to emphasize particular predictors, such as those that are thought to be more reliably collected or more readily interpreted in terms of the predictive relationship.

As an example of a possible analyst intervention in predictor selection, note that the number of sick days (for a particular staff sector and lag) turned up as a predictor in several of the regression models. It could be argued that this is an inappropriate or confusing measure, since the number of sick days per month is a function of the total staff size as well as of the typical health of the staff. An analyst might prefer to try to reformulate the model using one measure of sickness rate (illnesses per 100 staff members) plus a count of total staff, on the grounds that this would support a more meaningful interpretation. Such a substitution might help managers avoid making erroneous decisions based on attributing a causal relationship where only a correlation has been observed. As an extreme, an analysis showing that trips are greater in the months following high sick-day counts would not be a sound basis for laying off half of the staff; this action would quite possibly produce a substantial drop in the number of sick days per month (given the reduction in the total number of employees remaining), but would almost certainly not lead to reduced trips. If sickness rate were used as a predictor, the resulting model might not account for quite as much variance in the outcome measure, but the dubious intervention of ordering layoffs would be unlikely to be contemplated, even in some purely hypothetical manner.

Another analytical decision could be to limit predictors to longer lags. In the present study, predictor variables with 0-month lag (i.e., not really “predictions” but simply correlations between two simultaneous measures) were excluded from the stepwise regressions in order to look for models that would predict at least 1 month ahead. In practical applications, it might make sense to limit models to, for example, predictors with at least a 3-month lag or more—long enough to allow time to gather data, develop statistical predictive models, and consider and implement interventions to improve performance, but not so long as to reduce the likelihood of finding such models.

With regard to the statistical analytical methods used in this study, it is very encouraging to find that widely used tools are adequate to uncover predictive relationships in energy facility data. Even though it might not be possible to anticipate the exact terms (the variables, the lags, or the model coefficients) of the regression models, the relationships that emerge invite interpretation and may lead to increased knowledge of the factors that influence performance in a facility. Such relatively straightforward understanding would be much less likely if more exotic analytical methods were required, such as neural nets or very complicated software models that have been developed to make predictions of financial markets.

Thus, there is reason for optimism about the practical application of organizational epidemiology in energy industry facilities, and further research and development work would seem warranted

by this conclusion. Before rushing ahead, however, several important caveats are to be considered.

One caveat is that a statistically predictive model does not necessarily help analysts *understand* the bases for performance changes. As discussed earlier, the search for antecedent conditions includes but is not limited to the search for causes of human performance problems. A variable that helps predict outcomes might be only indirectly related to an underlying causal factor. For example, increased absenteeism might be a direct cause of deteriorating facility performance, but it may also be that increased absenteeism reflects poor morale, poor labor relations, or uncomfortable work conditions that lead to performance problems. The distinction between prediction and understanding may not be crucial if the aim is to predict performance, but it can have critical pragmatic implications for selection of appropriate interventions or corrective action programs in a real-world setting. Future research and sharing of results, as feasible, across many and diverse settings could lead to recommendations for specific performance improvement efforts tied to certain patterns of predictor variables and predicted outcomes. A software package, such as the Corrective Action Research and Evaluation (CARE) Tool developed by the EPRI Nuclear Sector (EPRI, 2001i), could be used to store and retrieve information on intervention programs. An enhanced version of this concept might make it relatively straightforward for an analyst to go from a statistical predictive model to one or more selection algorithms or (potentially) to recommended interventions.

Another caveat to bear in mind is that the models produced by stepwise regression have been referred to as *predictive* models, insofar as they find statistically significant relationships between various measures and a later outcome. However, the models derived to date are only predictive after the fact—due to the structure of this study, as constrained by pragmatic considerations, the models describe relationships only found in historical data. It remains an unanswered (and unaddressed) research question as to whether models such as these can make useful predictions going forward.

It may not be possible to test the predictive usefulness of such models under real operational conditions. As noted previously in this report and elsewhere (EPRI, 2001b), the business of energy facility managers and personnel is to ensure safe and productive operations, not to conduct scientific research for its own sake. Given a prediction of performance problems ahead, managers and staff would be expected to take action to prevent those problems, even though such action would interfere with the research aim of finding out whether the problems would arise as predicted. As soon as a predictive model is developed or a suggestion is made that a particular factor is believed to influence human or facility performance, management is likely to implement changes (preventive rather than corrective action) in an effort to improve future performance. Furthermore, when everything is going well, outcome measures such as trips are likely to have very little variability (close to zero trips per month), making it impossible to find statistical predictive models. This is a positive state of affairs for the industry setting—but it is an impediment to this type of research (or a challenge to discover a more appropriate outcome measure).

One way around this apparent quandary regarding testing the predictive usefulness of models under operational conditions may be to develop predictive models from a subset of historical data and then to see how well their predictions hold up for more recent data. That approach is commonly used in situations where the total set of available data is large enough to permit

developing a model based on a subset and then evaluating how well the model extrapolates to the rest of the data. With respect to the present study, the time window for observations was a 6-year period ending in December 2000—too short to allow subdivision into a subset for model development and a subset for model evaluation; the entire 6 years of monthly data were, necessarily, used for model development.

Now that those data have been obtained, processed, and analyzed, however, an additional 18 months have elapsed (as of the time of preparation of this report). As a useful follow-up study, an effort could, theoretically, be made to obtain observations of key variables (the predictors that were included in the regression models, along with the outcome measures) for these past 18 months, and then to determine whether the existing models developed from the first 6 years of data yield moderately accurate predictions into the following 18 months. In effect, the research (and practical) question would be whether, hypothetically, these models would have been useful to the plants if development of the models had been completed in early January 2001. That evaluation is outside the scope of the present study, but it could be very valuable for assessing how far into the future such regression models might be useful.

Aside from doing such additional retrospective evaluations, a field demonstration might be initiated, in collaboration with one or more energy companies, to set up procedures for implementing the organizational epidemiology process as described in this report. This would first involve gathering data for appropriate measures—such as those that were included in the predictive models in the present study, as well as ones related to training, safety efforts, and other factors that had been deemed of interest in this study but were not available from the two participating fossil plants. Data streams would then be collected on a central platform, which would merge them into a common data set on a regular and timely basis. Existing statistical analytical software tools would be used to develop and revise predictive models on a monthly basis. The primary barriers to such research would involve ensuring that data are provided reasonably quickly and thoroughly; the methodology for the work has already been developed, as described in this report.

Such a project would be designed to yield true predictions of performance—in advance, rather than in hindsight, as was done in this proof-of-concept study. These predictions would provide useful insights for the design of proactive interventions to improve both human and facility performance. Ultimately, experiences with organizational epidemiology in various industry settings could be shared through a consortium or community of practice. Lessons learned about performance prediction and about the effectiveness of specific interventions for optimizing human performance in specific workplace environments could help improve the reliability, safety, security, productivity, and environmental performance of energy infrastructure.

9

REFERENCES

American Society for Quality. (1999) *Recommendations for the Implementation of Selected Leading Indicators of Performance at Nuclear Power Production Plants*. Rev. 0 Report EED-99-01, Nuclear Power Production Committee, Energy and Environmental Division.

Ayres, T., Gross, M., Kalinowski, A., Ramachandran, K., Moloi, D. & McGhee, D. (2001) "Organizational epidemiology in fossil electric power generation facilities." Presented at the 6th Annual International Conference on Industrial Engineering Theory, Applications and Practice, San Francisco.

Ayres, T., Gross, M., & McCarthy, R. (1993) "A Retrospective on Attempts to Reduce Vehicular Risk Through Operator Training." In F. A. Elia, Jr. & D. W. Pyatt (Eds.), *SERA-Vol. 1, Safety Engineering and Risk Analysis*. ASME Book No. H00894-1993.

Chernoff, H. (1973) "The use of faces to represent points in k-dimensional space graphically." *Journal of the American Statistical Association*, 68 (342), 361-368.

Committee on the Safety of Nuclear Installations. (2000) *Identification and assessment of organisational factors related to the safety of NPPs: State-of-the-art report*. Nuclear Energy Agency, NEA/CSNI/R(99)21/Vol 1. (<http://www.oecdnea.org/html/nsd/docs/1999/csni-r99-21-vol1.pdf>)

Connelly, E. M., Van Hemel, S. B. & Haas, P. M. (1990) *Industry based performance indicators for nuclear power plants*. NUREG/CR-5568, Vol. 1. McLean, VA: Communications Technology Applications, Inc.

Cooper, S.E., Bley, D.C. & Parry, G.W. (1996) "Knowledge-Base for the New Human Reliability Analysis Method, 'A Technique for Human Error Analysis' (ATHEANA)." *Proceedings from PSA '96, International Topical Meeting on Probabilistic Safety Assessment: Moving Toward Risk-Based Regulation*.

EPRI. (1994) *Wrong Unit, Train, and Component Events at U. S. Nuclear Power Plants: Joint EPRI-CRIEPI Human Factors Studies*. Palo Alto, CA. TR-103954.

EPRI. (1999a) *Guidelines for Leading Indicators of Human Performance: Preliminary Guidance for Use of Workplace and Analytical Indicators of Human Performance*. Palo Alto, CA. TR-107315.

EPRI. (1999b) *Pilot Study: Occupational Health and Safety Surveillance Database*. Palo Alto, CA. TR-113884.

References

- EPRI. (2000) *Guidelines for Trial Use of Leading Indicators of Human Performance: The Human Performance Assistance Package*. Palo Alto, CA. 1000647.
- EPRI. (2001a) *Human Performance Optimization: Emerging Management Issues and Artificial Intelligence Methods. Volume 2: Forecasting Individual Human Behavior in a Constrained Environment*. Palo Alto, CA. 1004672.
- EPRI. (2001b) *Predictive Validity of Leading Indicators: Human Performance Measures and Organizational Health*. Palo Alto, CA. 1004670.
- EPRI. (2001c) *The Real Challenge of Safe Behavior: Transitioning from Being Accountable to Feeling Responsible*. Palo Alto, CA. 1004667.
- EPRI. (2001d) *Final Report on Leading Indicators of Human Performance*. Palo Alto, CA. 1003033.
- EPRI. (2001e) *Occupational Health and Safety Annual Report 2001: Illness and Injury Among the Electric Energy Workforce, 1995-2000*. Palo Alto, CA. 1005198.
- EPRI. (2001f) *Collecting and Using Near-Miss Information: Enhancing Switching Safety and Reliability*. Palo Alto, CA. 1001956.
- EPRI. (2001g) *PAOWF Users' Guide: Users' Guide to the PAOWF System Version 1.0*. Palo Alto, CA. 1006317.
- EPRI. (2001h) *Human Performance Optimization: Emerging Management Issues and Artificial Intelligence Methods. Volume 3: Finding and Building Expertise*. Palo Alto, CA. 1004673.
- EPRI. (2001i) *Corrective Action Research and Evaluation (CARE) Tool Users' Guide: User's Guide to the CARE Tool Version 1.1*. Palo Alto, CA. 1006319.
- EPRI. (2002a) *Organizational Epidemiology and Energy Facilities: Review of Antecedent Conditions for Human Performance Optimization and Error Prevention*. Palo Alto, CA. 1004668.
- EPRI. (2002b) *Organizational Epidemiology: Analytical Approaches for Predicting Human and Energy Facility Performance*. Palo Alto, CA. Available from www.epri.com.
- EPRI. (2002c) *Human Performance Management and Optimization: Strategic R&D Directions for Human-Centered Technologies, Tools, and Methods*. Palo Alto, CA. 1004665.
- EPRI. (2002d) *Guidelines for Capturing Valuable Undocumented Knowledge from Energy Industry Personnel*. Palo Alto, CA. 1004663.
- Fujimoto, H., Fukuda, M. & Tabata, H. (1994) "Sensitivity study of human errors as a basis for human error reductions on new safety system design." *Reliability Engineering and System Safety*, 45 (1-2), 215-221.

- Gross, M. M., Ayres, T. J. & Murray, J. (2000) "Analysis of human error at electric utilities." *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society*, 3, 173-176.
- Gross, M., Ayres, T., Wreathall, J., Merritt, A. & Moloi, D. (2001) "Predicting human performance trends." *Proceedings of the 7th Annual Human Performance/Root Cause/Trending Workshop*, Baltimore, MD.
- Hale, A., Wilpert, B. & Freitag, M. (1997) *After the Event: From Accident to Organisational Learning*. New York: Pergamon.
- Haugset, K. (1997) "Overview of Human-System Research at the OECD Halden Reactor Project." IEEE Sixth Annual Human Factors Meeting.
- Heyes, A. G. (1995) "PRA in the nuclear sector - Quantifying human error and human malice." *Energy Policy*, 23 (12), 1027-1034.
- Kerlinger, F. N. (1979) *Behavioral Research: A Conceptual Approach*. New York: Holt.
- Lanoie, P. & Trottier, L. (1998) "Costs and benefits of preventing workplace accidents: Going from a mechanical to a manual handling system." *Journal of Safety Research*, 29 (2), 65-75.
- Lehto, M. R. (1991) "A proposed conceptual model of human behavior and its implications for design of warnings." *Perceptual and Motor Skills*, 73, 595-611.
- Moray, N. (1992) "Toward an agenda for error research." *Proceedings of the Human Factors Society 36th Annual Meeting*, 640-643.
- Morris, M. W., Moore, P. C., & Sim, D. L. H. (1999) "Choosing remedies after accidents: Counterfactual thoughts and the focus on fixing 'human error.'" *Psychonomic Bulletin & Review*, 694, 579-585.
- Murray, J., Gross, M. M. & Ayres, T. J. (1999) "Human error in power plants: A search for pattern and context." *Proceedings of the Silicon Valley Ergonomics Conference & Exposition*, 187-191.
- Muschara, T. (1997) "Eliminating plant events by reducing the number of shots on goal." *Proceedings of the 1997 IEEE Sixth Conference on Human Factors and Power Plants*, 12-1 to 12-6.
- Park, K. S. (1997) "Human error." In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics*. John Wiley: New York.
- Reason, J. (1990) *Human Error*. New York: Cambridge University Press.
- Reason, J. (1997) *Managing Organizational Risks*. Ashcroft: Brookfield.

References

- Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York: North-Holland.
- Rosenthal, I. *Organizational Analysis in High-hazard Production systems: An Academy/industry Dialogue*. Wharton Risk Management and Decision Process Center.
- Smith, T. J. & Larson, T. L. (1991) "Integrating quality management and hazard management: A behavioral cybernetic perspective." *Proceedings of the Human Factors Society 35th Annual Meeting*, 903-907.
- Tamuz, M. (1994) "Developing Organizational Safety Information Systems for Monitoring Potential Dangers." *Proceedings of PSAM - II*, 71-7 to 71-12.
- Thompson, R. C., Hilton, T. F. & Witt, L. A. (1998) "Where the safety rubber meets the shop floor: A confirmatory model of management influence on workplace safety." *Journal of Safety Research*, 29 (1), 15-24.
- Van Hemel, S. B., Connelly, E. M. & Haas, P. M. (1991) "Management and organizational indicators of process safety." *Proceedings of the Human Factors Society 35th Annual Meeting*, 908-912.
- Wildberger, A. M., Gross, M. M. & Ayres, T. J. (2000) "Learning lessons intelligently in the electric power industry." Presented at AAAI-00 Workshop on Intelligent Lessons Learned Systems. July 31, 2000, Austin TX.
- Wilpert, B. & Miller, R. (1999). *Organizational factors: Their definition and influence on nuclear safety (ORFA)*. Report on needs and methods. Commission of the European Communities, Fourth Framework Programme on Nuclear Fission Safety, AMM-ORFA(99)-R03.
- Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL: Wolfram Media.
- Wreathall, J. & Jones, L. (2000) "Leading indicators of human performance - The story so far." Presented at the 6th Annual Human Performance/Root Cause/Trending Conference, Philadelphia.
- Wreathall, J., Schurman, D.L. & Anderson, N. (1991) "An Observation on Human Performance and Safety: the Onion Model of Human Performance Influence Factors." *Probabilistic Safety Assessment & Management*, 1, 25-30.
- Yu, C. (1995). *Visualization techniques of different dimensions*. Arizona State University, <http://seamonkey.ed.asu.edu/~behrens/asu/reports/compre/comp1.html>.

A

FACIAL DISPLAYS

This appendix details results from trial use of a multivariate visualization technique to analyze data from Plant A. The technique is briefly introduced in Section 5, and results are described at the end of Section 6.

For this trial examination of the technique, variables selected for display using Chernoff faces and the associated features were as follows:

- Area of Face: Thermal Index
- Shape of Face: Gross Sickness and Absentee Rate
- Length of Nose: Absentee Frequency Rate
- Location of Mouth: Preventive Maintenance Orders, Total
- Curve of Smile: Preventive Maintenance Orders, Initiated
- Width of Mouth: Preventive Maintenance Orders, Outstanding
- Location of Eyes: Ratio of Corrective to Preventive Maintenance Orders, Outstanding
- Separation of Eyes: Ratio of Corrective to Preventive Maintenance Orders, Initiated
- Angle of Eyes: Ratio of Corrective to Preventive Maintenance Orders, Total

The faces are shown on the following pages. Missing facial features (e.g., no facial outline for Face 14) indicate that the corresponding value was missing or had not yet been obtained at the time these plots were prepared. The faces are ordered by number of trips per month, with the fewest trips for the first faces in the set and the greatest number of trips for the last faces. (The number under each face is an identifier, not the number of trips that occurred in the corresponding month.)



1



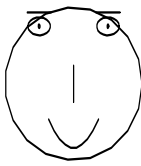
2



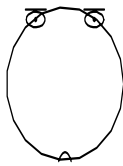
3



4



5



6



7



8



9



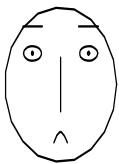
10



11



12



13



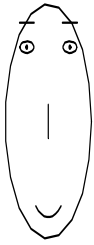
14



15



16



17



18



19



20



21



22



23



24



25



26



27



28



29



30



31



32



33



34



35



36



37



38



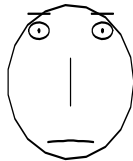
39



40



41



42



43



44



45



46



47



48



49



50



51



52



53



54



55



56



57



58



59



60



61



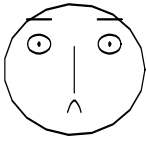
62



63



64



65



66



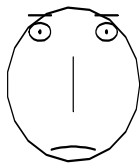
67



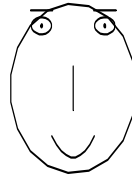
68



69



70



71



72

About EPRI

EPRI creates science and technology solutions for the global energy and energy services industry. U.S. electric utilities established the Electric Power Research Institute in 1973 as a nonprofit research consortium for the benefit of utility members, their customers, and society. Now known simply as EPRI, the company provides a wide range of innovative products and services to more than 1000 energy-related organizations in 40 countries. EPRI's multidisciplinary team of scientists and engineers draws on a worldwide network of technical and business expertise to help solve today's toughest energy and environmental problems.

EPRI. Electrify the World

© 2002 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute and EPRI are registered service marks of the Electric Power Research Institute, Inc. EPRI. ELECTRIFY THE WORLD is a service mark of the Electric Power Research Institute, Inc.

 Printed on recycled paper in the United States of America

I004669