NEW APPROACHES FOR MONITORING QUALITY CONTROL IN THE CLINICAL LABORATORY

By

JERRY DON DECHERT

Bachelor of Science in Industrial Engineering
Oklahoma State University
Stillwater, Oklahoma
1990

Master of Science in Industrial Engineering
Oklahoma State University
Stillwater, Oklahoma
1992

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 1996

NEW APPROACHES FOR MONITORING

QUALITY CONTROL IN THE

CLINICAL LABORATORY

Thesis Approved:

Kenneth E Case

Thesis Adviser

David B. Pratt

V C C Purdie

Manjunath Kamath

Barry Kurt Moser

Thomas C. Collins

Dean of the Graduate College

ACKNOWLEDGMENTS

Throughout the course of my work on this dissertation, a number of people have done a great deal to both assist and encourage me in my efforts. First, I wish to thank my committee for their work regarding this dissertation. All of the members of the committee were extremely helpful, and their efforts contributed greatly to the quality of the dissertation. Additionally, I would like to single out Dr. Case for his work as the committee chairman. I consider it a privilege to have been able to work with Dr. Case for the last few years, and I will miss working with him in the future.

My heartfelt thanks also goes out to the wonderful people at Abbott Laboratories. I have worked with too many people there to single out particular individuals, but my experiences with Abbott have led to this research. The support I have felt from everyone at Abbott throughout my research has been second to none, and I wish to thank everyone their professionalism and expertise.

I also wish to thank my parents and my sister (Lloyd, Annette, and Renee Dechert) for all their support for all of these years. My family is very special, and I treasure all of them and the fact that they have always been there when I have needed them.

Finally, I want to thank my wife, Becky, for putting up with my status as a graduate student since we became married. She has been there with love and support for all of the ups and downs of the Ph.D. experience, and I cannot thank her enough. In short, I do not know what I would do without her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE RESEARCH PROBLEM

Introduction

One of the most important aspects of diagnostic testing in a clinical laboratory is maintaining the quality of the results being reported by the laboratory. The results reported by such a laboratory are used by physicians and pathologists to diagnose patient conditions and monitor patient progress. Therefore, the quality of the results reported is of utmost importance, and the quality assurance efforts regarding these results must be well planned and effective.

For many years, control materials have been tested and monitored to ensure the continued high quality of the patient results reported. These controls are stable materials of known concentrations tested in the same fashion as patient samples, and they serve the purpose of signaling when the testing system has experienced a change. The assumption is that changes detected by control materials will also impact patient test results in the same fashion.

It is clear that the method for monitoring quality control in the clinical laboratory will have a direct impact on the results reported by the laboratory. This research examines new approaches to monitoring quality control in the clinical setting. These new approaches are compared to one another and current clinical approaches. Comparisons among methods are made from both a statistical and a cost perspective.

The Problem

Clinical laboratories have been using control charts to monitor the stability of control materials (and thereby the measurement system) since the early 1950's. Typical approaches involve monitoring multiple levels of control materials (e.g., high, medium, and low concentrations) over time. However, many clinicians utilize limits on their control charts at +/- 2 standard deviations (SD) of the measurement system. This practice results in almost a 5% probability of false rejection for an individual level when the measurement system is stable, and it is a method long ago outlawed in industrial quality control practice. Tradition and ease of use, however, continue to make +/- 2 SD limits extremely popular in the clinical setting.

Dr. James Westgard has published a great deal in the clinical literature and encouraged the use of runs rules as an alternative to +/- 2 SD limits in the past. Termed the "Westgard Multirule Procedure" (Westgard, Barry, and Hunt 1981), the approach has gained a good deal of support in the literature, but actual implementation in clinical laboratories still does not approach the use of +/- 2 SD limits. Westgard's procedure is discussed in detail in Chapter II of this dissertation, but it is essentially a variation of the AT&T runs rules.

Recent work in the clinical area has surrounded the concept of "clinically significant errors." The underlying premise is that there are actually changes or shifts in the measurement system which are allowable because they are not clinically significant and do not have an impact on the utility of the reported patient results. While there has been

some work towards designing control procedures to detect these clinically significant errors, they have not been widely embraced in clinical laboratories.

It is clear that if clinicians wish to move towards allowing changes to occur in their measurement systems which they deem insignificant from a practical standpoint, new methods must be considered. Using the traditional +/- 2 SD limits approach is definitely not a method which will allow for clinically significant errors. In fact, +/- 2 SD limits will have a high false rejection rate and flag shifts which are non-existent. Part of the popularity of +/- 2 SD limits, however, is that the measurement system is considered either in-control (acceptable) or out-of-control (unacceptable) each time control materials are tested. To the clinician, the use of runs rules can lead to problems. Stopping the process due to a run rule means that the process has been operating with a shift for some time without detection, and stopping for the run rule brings all previous test results into question. For example, if a run rule detects a shift after 4 consecutive points are above or below +/- 1 SD of the measurement system, this means that 3 previous time points also experienced that shift and the lab continued to release patient results. There is a sense of comfort for clinicians in using a monitoring approach that considers their system either acceptable or unacceptable at each time point as opposed to waiting for future data (via runs rules) to detect a problem with the measurement system. Therefore, a method which in some way incorporates clinical significance without the use of runs rules would be very attractive to clinical personnel.

While there is a reasonably large body of literature surrounding clinical quality control approaches, there have been a number of missed opportunities in the field. There

has been a limited treatment of cost models applied to clinical quality control methods, but the field has largely ignored the cost implications of the quality control procedures selected. Some of the current methods used by clinicians can have some very detrimental cost implications (i.e., the high false rejection rate associated with +/- 2 SD limits) and evaluating these methods on a cost basis can be very enlightening.

Another dimension to the problem is the multivariate nature of the quality control monitoring performed in the clinical setting. Clinicians test multiple levels of control material regularly (by law), so multivariate approaches are very applicable to the laboratory situation. The clinical quality control literature, however, has not considered multivariate methods to this point, and these methods need to be investigated for feasibility.

Given these considerations, it is clear that there is a long history of applying quality control methods in the clinical laboratory. However, the methods currently practiced are based primarily on traditional application and ease of use rather than statistical or economic designs. Even though highly automated systems are available enabling the application of sophisticated methods, simple methods are still the choice in the clinical laboratory. In addition, consideration to cost issues has largely been ignored. Therefore, there is a need to capitalize on the automation resources available and provide clinicians with powerful statistical approaches for quality control monitoring.

Research Objective

The objective of this research is to evaluate new approaches to clinical quality control monitoring, comparing these methods with current approaches on both a statistical and a cost basis. The methods used in the comparison include:

Traditional clinical approaches:

1) The strict application of +/- 2 SD limits.

2) The use of +/- 2 SD limits with an immediate retest.

3) The Westgard Multirule Procedure.

Multivariate approaches including:

a) The T^2 chart.

b) The χ^2 chart.

c) Principal component charts.

Research Sub-Objectives

In order to achieve the research objective, the research effort can be broken into two distinct phases or sub-objectives. The first phase is to statistically model the methods identified above. This phase allows for the comparison of the methods on a statistical basis, but it also results in necessary inputs for the second phase. The second phase of the research is the economic modeling of the methods, resulting in method comparisons on a cost basis. The following describes the specific phases considered in the research work.

Phase 1: Compare the alternatives statistically:

The initial phase of the research is to develop models for assessing the statistical performance for each of the methods. The goal is to compare the methods statistically in terms of average run lengths (ARLs) for detecting specified shifts in the measurement system. For purposes of this modeling, sample sizes of two (corresponding to two different levels of control material) and three (three different levels of control material) are used. These are the sample sizes frequently encountered in practice, and larger sample sizes would be quickly rejected by clinicians.

Given that multiple levels are modeled (two or three), the research varies the shifts in centering by level. So, ARL comparisons are performed for the case when levels shift together and also when the levels are shifted independently. Current clinical quality control evaluations always assume that all levels are affected equally by any shifts in centering, so this research explores the impact of shifting levels independently. It should be noted that this independent shifting in levels is perfectly realistic in the clinical setting.

Phase 2: Compare the alternatives economically:

After determining the statistical performance of each of the methods, the focus of the research shifts to comparing the methods on a cost basis. Using a variation of cost models previously utilized in industrial quality control applications, a new model reflective of the clinical setting is developed and used to compare the costs associated with the various alternatives. Results from the first phase of the research (statistical modeling) are used as inputs into the economic modeling portion of the research.

The final stage in the research is to review and analyze the results of the research. A number of interesting results and insights are available as a result of the research, and a synthesis of these findings is included.

Research Contribution

This research contributes to the statistical quality control body of knowledge in a number of respects. The first, the development of new quality control methods for the clinical setting, is most relevant to the clinical laboratory. These methods incorporate the multivariate information available to the application and classify the measurement system as acceptable or unacceptable at each test point. This feature is attractive to clinicians reluctant to employ runs rules to detect changes in their system.

The comparison of the current clinical quality control approaches on a cost basis also constitutes a contribution to research in the field. Only a small amount of cost modeling work exists in the clinical arena, so the results from this research are very powerful in aiding people with the selection of a quality control procedure. Looking at the impact that a quality control method can have on costs is a very useful tool for clinical personnel.

Another contribution is the modeling of clinical quality control where the multiple levels are shifted independently. Previous research efforts assumed the same size shift in each level, but this research explores the effect of shifting the centering of the levels independently. Inducing independent shifts in the different control levels can lead to very different performance of the current approaches than many clinicians believe.

From the traditional quality control perspective, a major research contribution is in the development of an economic model for principal component charts. While the T^2 chart has been modeled from the economic perspective, no such attempt has been made for the principal component chart. Therefore, the economic modeling of the principal component chart constitutes an original contribution to the field of industrial quality control.

It is clear that the research breaks new ground on a number of different fronts. While much of the research is geared towards the clinical setting, the economic modeling of multivariate charts adds an element of research directly impacting industrial quality control applications.

CHAPTER II

LITERATURE REVIEW

Introduction

There is a large and varied body of literature relevant to this research effort. The clinical laboratory has its own body of knowledge, summarized in the first section of this review. Cost modeling of industrial control charts has also long been a topic of interest in the literature, and the second section summarizes the pertinent industrial cost modeling information along with some cost modeling work in the clinical area. The third section reviews multivariate charting techniques and modeling with the final section of the chapter summarizing the literature review.

Quality Control in Clinical Laboratories

Traditional Approaches

The original application of statistical methods for monitoring quality control in the clinical laboratory dates back to the early fifties (Levey, Jennings 1950). In this original application, Levey and Jennings apply a Shewhart chart from industrial quality control and use average and range charts for subgroups of size two. Laboratory personnel today often refer to plots of quality control values as Levey-Jennings charts, but they are typically referring to charts of individual values plotted on a chart with limits set at +/- 2 standard deviations (SD). This evolution from the original Levey-Jennings application of average

charts to charts of individual values with +/- 2 SD limits occurred gradually. Henry and Segalove's paper (1952) discusses the plotting of single replicates against +/- 2 SD limits, and the approach gained popularity in laboratories due to its ease of use.

This evolution of quality control systems began in the fifties and continues in many instances today. Typical approaches for monitoring controls in the clinical laboratory use three different levels of control materials (low, medium, and high) with targeted concentrations across the measurement range of the diagnostic test. Each of these levels is tracked separately with control limits placed at +/- 2 SD limits and applied to individual observations. Linnet's work (1989-b and 1991) advocates averaging replicate observations of control values for monitoring quality control materials and shows the increased power of such an approach, but laboratory personnel have not embraced the practice.

In the late seventies, Westgard, Groth, Aronsson, Falk, and de Verdier (1977) published the first article assessing the statistical performance characteristics of clinical quality control methods. Using computer simulation, their work develops power curves for various combinations of runs rules. The graphs display the probability of detecting specific errors as a function of the number of control observations evaluated. The graphs are based on the probability of detecting a true error (p_{ed}) and the probability of a false rejection (p_{fr}). Additionally, the article discusses the concepts of random error (RE) and systematic error (SE). Random error is the inherent imprecision, or noise, that a testing system will experience in a state of statistical control. A systematic error is a change in the centering of the measurement system. Therefore, a change in the random error is a

change in the measurement system variability, and a systematic error represents a change in the centering of the measurement system.

This work by Westgard et al. is extremely important as it set the direction for future evaluation of quality control approaches. The use of simulation became standard, and future models assumed a common shift to all levels used for monitoring. These assumptions, therefore, set the stage for work continuing into the 1990s.

Westgard, Groth, Aronsson, and de Verdier (1977) used the developments of power curves to evaluate and propose a monitoring method combining a Shewhart control chart with a cumulative sum chart. While the approach performs well statistically, the sophistication of the method makes it unattractive to users and it has not been widely implemented.

Continuing the use of computer simulation to evaluate control methods comprised of combinations of control rules, Westgard and Groth (1979) published what they termed "power function graphs." The graphs are refinements of earlier power curves, displaying probabilities of error detection and false rejection versus shifts in centering (ΔSE) and spread (ΔRE). These power function graphs provide a means for assessing a control monitoring system's performance and determining the statistical acceptability of a given approach.

Clinical Chemistry published the culmination of Westgard's work in the late seventies and early eighties as a "selected method" (Westgard, Barry, and Hunt 1981). Termed the "Multirule Procedure," Westgard describes a combination of control rules to

apply to monitoring control materials. The Westgard Multirule Procedure works as follows:

Reject if:

1 point outside +/- 3 SD limits (1_{3s})

If 1 point outside +/- 2 SD limits, then consider:

a) 2 consecutive points outside +/- 2 SD on the same side of the centerline (2_{2s})

b) Range of 2 points greater than 4 SD (R_{4s})

c) 4 consecutive points outside +/-1 SD limits on the same side of the centerline (4_{1s})

d) 10 consecutive points above or below the mean (10_x)

While many clinicians endorse the procedure, there is still considerable confusion about its application in laboratories. The method can be confusing as there are many comparisons required; many laboratories which claim to be using the Westgard Multirule Procedure are likely using some variation of the approach rather than the procedure as originally published.

Following the appearance of Westgard's Multirule Procedure, there continued to be publication of other combinations of runs rules. Blum (1985) published a method that incorporates the use of 10 different runs rules, justifying the selected rules using computer simulation. Again, the method has proven too sophisticated for application in the laboratory setting.

Westgard continued his work in the area of clinical quality control with the development of selection grids for planning quality control procedures (Westgard, Quam, and Barry 1990). The grids are tools which allow a user to select a set of control rules for a given application. The parameters for selecting the set of rules are the true frequency of actual errors in the laboratory and the critical systematic shift (ΔSE_c) the user wishes to detect. Based on these parameters, the user can determine the number of replicates of controls to run and the control rules to employ.

Clinical Significance

In the traditional quality control approaches described in the previous section, all of the methods attempt to detect any statistically significant change in the measurement system. Westgard's selection grids, however, involve the use of ΔSE_c as a parameter for selecting a method. This marks the appearance of methods designed to detect some specific change identified to be clinically significant or relevant. In other words, changes less than this amount (ΔSE_c) are considered to be of no consequence in the clinical setting while changes this large or larger can have clinical implications. These implications may involve changing a patient's dosage or initiating a change in treatment.

Attempts to define the requirements of clinical laboratory testing systems go back to Skendzel, Barnett, and Platt (1984). Skendzel et al. mailed a questionnaire to physicians to determine the total precision required for a variety of diagnostic tests. Results at that time indicated that almost all the tests considered provided adequate

precision for physicians' requirements. The work was also useful for establishing medically useful guidelines for analytical precision.

Linnet (1989-a) then used Skendzel et al.'s work to calculate what he called "maximum clinically allowable analytical error." Represented by ΔSE_c, this error represents the largest error that can be tolerated according to the requirements outlined through Skendzel et al.'s survey results. Linnet's approach for calculating ΔSE_c is as follows:

$$\Delta SE_c = \Delta_{med} - 1.65 * s_t$$

where Δ_{med} = the median difference of medical importance as reported by the physicians in Skendzel et al.'s work
s_t = total precision including analytical measuring system, sample handling, and patient biological variability

Linnet's paper includes a similar approach for defining a critical change in the spread of the measurement system, ΔRE_c. It is clear that the parameters required to calculate the maximum clinically allowable error are not easy to estimate and can be sources of debate.

Koch, Oryall, Quam, Feldbruegge, Dowd, Barry, and Westgard (1990) describe an application that incorporates the use of the maximum clinically allowable error in the design of the quality control system for a specific analytical system. Their application uses a consensus process at the testing site to determine TE_a, the total allowable error. Then, they calculate ΔSE_c and ΔRE_c using the following formulas:

$$\Delta SE_c = [(TE_a - |\text{bias}|)/s] - 1.65$$

$$\Delta RE_c = (TE_a - |\text{bias}|)/ 1.96s$$

where TE_a = total allowable error for the testing method
bias = known difference between the laboratory's mean and the true mean of the control material
s = analytical testing system precision

Control methods (number of control replicates and control rules) can then be selected for the individual tests with the requirement of a 90% probability of detecting ΔSE_c.

Many other methods are available for determining the maximum clinical allowable error as noted by Fraser (1990) and Westgard and Burnett (1990). Also, Petersen and Fraser (1994) provide an excellent editorial discussing the issues involved with determining this allowable error. Westgard, Seehafer, and Barry (1994-b) further developed their approach for defining maximum clinically allowable errors using criteria from the Clinical Laboratories Improvement Amendments of 1988 (CLIA 1988). This federal regulation outlines total allowable error for many analytical tests. Westgard recommends using the total error indicated from CLIA 88 as follows:

$$TE_a = bias_{meas} + \Delta SE_c \cdot s_{meas} + z \cdot \Delta RE_c \cdot s_{meas}$$

where $bias_{meas}$ = bias in the measurement system
s_{meas} = total precision of the measurement system
z = standard normal value that sets the maximum percentage beyond TE_a

Using the TE_a associated with the test being analyzed, the user can calculate ΔSE_c and ΔRE_c and select an appropriate quality control scheme.

While there does not yet appear to be any general consensus in the literature regarding the best approach for defining changes which are clinically significant, it is clear that quality control approaches in the clinical setting are moving towards some amount of allowable error. This is a result of the technology in the field providing better precision in diagnostic testing. As the variability inherent in these diagnostic tests continues to get smaller and smaller, there will be a larger movement in the clinical field away from methods which define acceptability limits solely on assay variation (i.e., +/- 2 SD limits)

towards methods which incorporate some allowable error or have a total error specification.

Evaluation Methods

In the previous section which discusses traditional methods for monitoring control materials, researchers typically used computer simulation to evaluate the statistical performance of the selected methods. There has been a great deal published in the clinical literature concerning the computer simulations used for this work, and the editorial by Westgard (1992) provides an excellent summary of the assumptions employed in the simulators and lists many of these simulations. Hatjimihail (1992) and Parvin (1991, 1992) provide examples of recent simulation efforts. The literature also contains other approaches including neural networks (Schweiger, Soeregi, Spitzauer, Maenner, and Pohl 1993) and genetic algorithms (Hatjimihail 1993).

Analytical approaches to analyzing control methods utilizing runs rules have been limited in the clinical laboratory literature until recently. Parvin (1993) performed some analytical analysis by looking at data within a run, but his results are not widely applicable. Bishop and Nix (1993) were the first authors in the clinical chemistry literature to obtain analytical results by applying Markov chains to procedures incorporating runs rules. Using Markov modeling to analyze the Westgard Rules as previously published, their work compares their results to previous computer simulation analysis. Bishop and Nix also propose the use of a cumulative sum chart for monitoring control levels and provide results supporting its utility. While the paper is important in that it applies analytical

methods to evaluating clinical quality control procedures, the authors modeled the Westgard Rules for samples of size 5 to 15. These sample sizes are not indicative of what laboratories actually use (sample sizes two or three would be typical), so it is difficult to extrapolate Bishop and Nix's work to actual practice in the clinical laboratory.

Lee (1996) has also done some analytical work regarding the use of Westgard's Multirule Procedure. The sample sizes investigated by Lee were much nearer those actually implemented in practice, and his results provide good insight into the workings of modified approaches to the Westgard Multirule Procedure (i.e., not including the R_{4s} rule or using +/- 2 SD limits as a screening criteria).

Control Chart Cost Modeling

Cost Modeling in Industrial Quality Control

There is a large body of literature regarding the application of cost models to the use of control charts in the industrial quality control literature. Many models have been proposed and used by a number of different researchers with the intent of determining optimal approaches for the application of control charts. Two excellent surveys regarding control chart cost modeling are available with Ho and Case (1994) reviewing the decade spanning 1981-1991 and Montgomery (1980) reviewing earlier work. For the purposes of this research, the literature review only considers cost modeling approaches applied to variables charts.

The original use of cost modeling to design control charts dates back to Duncan (1956). This paper blazed the trail for future researchers and is the foundation of many of the models subsequently developed. In his paper, Duncan identifies the three main control parameters of interest when using control charts: the sample size collected (n), the interval between samples (h), and the control limits used on the chart (k).

The objective, according to Duncan, is to maximize the average net income per time unit through the selection of the parameters n, h, and k. In order to maximize average net income per time unit, Duncan's work uses a net income equation consisting of the sum of in-control income and out-of-control income minus the cost for investigating false alarms, the cost of investigating true problems, and the cost of maintaining the control chart. Duncan converts this income equation into a loss equation, then solves for the optimum values of n, h, and k under assumptions regarding costs.

The basic assumptions in Duncan's model are that the time between process changes is exponentially distributed and that there is a single, assignable cause which affects the process. Additionally, the model assumes that production does not discontinue while investigating a special cause.

Results show that the selection of the sample size is driven by the shift in the process the user wishes to detect. Also, the determination of the control chart limits is tied to the cost of investigating false alarms and the magnitude of true process changes.

Another widely cited paper provides a unified approach to the economic design of control charts (Lorenzen and Vance 1986). Lorenzen and Vance provide a general model that allows the evaluation of attributes charts as well as variables charts. The model

incorporates a general equation for the expected loss per hour of operation. Costs in the equation include the cost of operating both in and out-of-control, the cost for investigating false alarms and true process changes, and fixed sampling costs. An interesting contrast of this model to Duncan's is the use of an indicator variable to allow the modeler to either shut the process down while investigating a signal by the chart or continue production while investigating. Results indicate that a considerable savings can be realized by using an economic approach to designing the control chart parameters.

One of the major concerns regarding the use of cost models for designing control charts is the complexity of the cost models and the difficulties involved with optimizing the sample size, sample interval, and control limits. In an attempt to simplify the design of control charts through economic models, Collani (1986) developed a simpler approach for attaining earlier results published by Montgomery (1982). In his model, Collani assumes a constant production speed and designs the chart to minimize the expected loss per item of production. In order to simplify the model, Collani reduces the number of parameters in the model to two cost parameters. Using his model, Collani was able to produce results very similar to Montgomery's earlier findings without the use of a computer.

An approach which surfaces in a number of cost modeling articles (Collani 1988) is the use of two states to define the status of the process being monitored. One state is considered satisfactory performance and the other state is considered unsatisfactory or unacceptable. The reasoning behind this approach is that there can be subtle shifts in a process which are not causes for alarm, but that larger shifts should result in investigation. Arnold (1989) discusses discrimination between these two states and shows the necessity

of increased sample sizes when there are small differences between states. Tagaras and Lee (1988) took the model further and developed multiple control limits corresponding to minor problems and major problems. Their work shows an improvement from a cost perspective with their model, but with increased complexity of the monitoring approach.

Many of the cost models assume deterministic cost parameters and perfect knowledge of the cost parameters. Pignatiello and Tsai's work (1988) investigates the risk in making these assumptions by using Taguchi's concept of "noise" to model uncertainty in the parameters of the cost model. Using a Taguchi design to set the levels for the sample size, sample interval, and control limits, Pignatiello and Tsai's research shows the danger of assuming no noise in the model parameters when the parameters are actually uncertain.

Additional cost models have been applied to the joint use of average and range charts (Saniga 1989). These models are by necessity more complex than models considering only average charts. Jones and Case (1981) use Duncan's cost model for their work and detail a Markov approach for the state of the process (i.e., out-of-control for range chart, out-of-control for both the range and average chart, etc.,). Results from their modeling provide values for sample size, sample interval, and control limits for given cost assumptions.

All of the models discussed thus far in this literature review assume an exponential time between process changes. Researchers have, however, explored the use of other distributions; primarily the Weibull distribution. Bannerjee and Rahim (1988) use cost models for average charts with the Weibull distribution for the time between process

changes. Additionally, they allow the sample interval to vary over time and show that varying the sample interval can reduce costs. Parkihideh and Case (1989) also use the Weibull distribution and allow the control chart parameters to vary over time with encouraging results. McWilliams (1989) uses the Weibull distribution for the time between process changes and shows that the economic results are not very sensitive to the parameters selected for the Weibull distribution.

From this review, it is evident that there are a large number of models in the literature with varying assumptions. Table 2.1 summarizes some of these models and their assumptions. Included in the table is the objective function used for evaluating the cost model, the distribution used to model the time between changes in the process, whether a single assignable cause or multiple causes can affect the process, and whether or not the process is shut down while searching for a special cause.

Cost Modeling in the Clinical Setting

While there has been a great deal of work published in the industrial quality control literature regarding economic models, there has been relatively little application of economic models in the clinical setting. By and large, the focus of the clinical literature is on the statistical performance of the quality control method. The statistical performance is definitely a priority in selecting a quality control approach, but the cost element must not be ignored. Often, multiple approaches can result in the same statistical power for detecting changes in the measurement process, but the cost associated with the approaches can be drastically different.

Table 2.1 - Cost Model Assumption Summary

Author, Date	Objective Function	Failure Distribution	Single or Multiple Cause	Shut Down?
Duncan, 1956	Max expected net income/ time	Exponential	Single	No
Jones and Case, 1981	Max expected net income/ time	Exponential	Single	No
Arnold, 1989	Min average loss/item	Exponential	Single	No
Pignatiello and Tsai, 1988	Min expected cost/unit time	Exponential	Single	No
Saniga, 1989	Min expected cost/unit time	Exponential	Single	No
Collani, 1986	Min expected loss/item	Exponential	Single	Yes
Collani, 1988	Max profit/ unit	Exponential	Single	Yes
Lorenzen and Vance, 1986	Min expected cost/hour	Exponential	Single	Either
McWilliams, 1989	Min expected cost/unit time	Exponential	Single	Either
Tagaras and Lee, 1988	Min expected cost/unit time	Exponential	Multiple	No
Parkhideh and Case, 1989	Min expected cost/unit time	Weibull	Single	No
Banerjee and Rahim, 1988	Min expected cost/hour	Weibull	Single	Yes

Although the application of economic models in the clinical setting is limited, Westgard and Groth (1983), followed by Westgard, Oryall, and Koch (1990) have done some work in the area. In both articles, the authors use a predictive value model to assign costs to various approaches for monitoring quality control. Basically, their model develops the following four situations:

1) The quality control method detects a true change.
2) The quality control method detects a change when there is no change.
3) The quality control method does not detect a true change.
4) The quality control method does not detect a change when there is no change.

Situations 1 and 4 correspond to a correctly working method while situation 2 corresponds to a Type I error and situation 3 is a Type II error. The authors develop probabilities for being in each of the four situations and costs associated with these situations, comparing methods based on these costs. They can then use their model to predict "quality" (essentially the defect rate) and "productivity" (the test yield).

Multivariate Quality Control Approaches

As described previously, the typical application of quality control in the clinical setting involves running multiple levels of control material and plotting them on a control chart over time. Since these multiple levels of control are tested at the same time, they are in fact correlated. This correlation suggests the potential use of multivariate approaches to monitoring quality control in the clinical setting. Such work has not been expressly proposed in the clinical literature, but there is a long history of the use of multivariate approaches to monitoring quality control in the industrial setting.

Hotelling's T^2 and the χ^2 Chart

Probably the most prevalent multivariate quality control approach is the use of Hotelling's T^2 chart. Hotelling (1931, 1947) derived the T^2 statistic and showed its appropriateness for multivariate applications. The T^2 statistic for a single observation is defined as follows:

$$T^2 = (\mathbf{x} - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}})$$

where $\mathbf{x}$ is a column vector of observations, $\bar{\mathbf{x}}$ is a column vector of means with dimension p, and $\mathbf{S}^{-1}$ is the sample covariance matrix. Hotelling showed that T^2 is related to the F distribution as follows:

$$T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,n-p,\alpha}$$

where n is the number of observations. Through the relationship of T^2 to the F distribution, one can set limits for the T^2 chart and calculate α probability limits using the F distribution. A number of authors develop the T^2 chart in more detail and discuss its application (Alt 1982 and Jackson 1985).

As described in the previous paragraph, the T^2 control chart is useful for monitoring p characteristics when the mean vector and covariance matrix are unknown. For the case when the mean vector and covariance matrix are known (referred to as μ_o and Σ_o respectively), then the χ^2 chart can be used (Alt 1985). Basically, one can plot $n(\bar{\mathbf{x}} - \mu_0)' \Sigma_0^{-1} (\bar{\mathbf{x}} - \mu_0)$ against an upper control limit of $\chi^2_{p,\alpha}$ and a lower control limit of zero where n is the sample size of the $\bar{\mathbf{x}}$ vector. When the calculated statistic exceeds the upper limit of the χ^2 chart, the process is considered out of control.

Furthering the work in multivariate approaches, Tracy, Young, and Mason's research (1992) shows that the Beta distribution can be used to obtain exact control limits for multivariate charts in the start up phase. Their example shows that the use of the Beta distribution can lead to better performance than approximation methods for a small number of subgroups.

Principal Component Charts

Another approach to applying quality control to multivariate situations is the use of principal component charts. Principal components are simply transformations of the original data into new variables which are independent of one another. Therefore, these principal components can be dealt with as independent variables allowing a great range of analysis. Discussed thoroughly by Jackson (1980, 1981-a, 1981-b), principal component analysis is also often used to reduce the number of variables considered in a problem. For example, often a problem involving ten characteristics or variables can be analyzed using two or three principal components without losing much information.

The advantage of principal component charts is that α probability limits can be plotted directly on the chart. Jackson (1956, 1959) shows how principal component analysis can be used for control charting purposes for two or more variables. Jackson discusses how the transformations created through principal component charts can be used for monitoring process stability and provides valuable insight into interpreting the transformed variables. Using principal component charts, one can represent the stability

of the measurement process itself and simply monitor it to see if the measurement system as a whole is in-control.

Another useful characteristic of the principal component chart is the ability to directly plot a specification limit on the chart. Jackson and Bradley (1966) used principal component charts successfully to implement a sequential procedure for evaluating specifications.

Also, Jackson and Mudholkar (1979) developed a nice approach for detecting outliers in the data using principal component charts. Through this approach or a like procedure, new screens can be applied to clinical data to ensure that outliers are not affecting the determination of measurement system status.

Multivariate Cost Modeling

While there is a reasonably large body of literature surrounding multivariate quality control approaches available, very little has been done in the area of economically modeling and designing multivariate charts. Montgomery and Klatt (1972-a and 1972-b) modeled the T^2 chart from an economic perspective, and their economic model closely resembles earlier work by Duncan (1956) with generalization to the multivariate case. The costs incorporated into their model include the expected cost per unit of sampling and testing, the expected cost per unit for investigating out of control signals (both true and false alarms) and the expected cost per unit for producing defective product. Results from the model indicate potential savings may be realized as many actual applications of

multivariate control charts use sample sizes in excess of those determined optimal by Montgomery and Klatt.

There is no evidence of cost modeling applied to principal component charts in the literature. Therefore, any work in this area constitutes an original application in the field.

Summary

From a critical examination of the literature, it is clear that there are a number of research opportunities within these areas. The first opportunity is to apply multivariate approaches for monitoring quality control in the clinical laboratory setting. The use of multivariate methods has long gone ignored in clinical quality control approaches, and there is a tremendous opportunity to use multivariate methods for this application.

Another opportunity in the clinical area is the application of economic models for evaluating quality control approaches. While there has been a small amount of work in the clinical field surrounding the economic impact of the quality control method employed, it is clear that the economic aspects of quality control monitoring have largely gone ignored.

Finally, the literature search reveals that economic modeling of principal component charts has not been pursued in the industrial quality control literature. There has been a limited approach to economically modeling multivariate approaches (T^2 chart), but researchers have not applied economic models to principal component charts to evaluate their performance on a cost basis. Given these gaps in the literature, it is evident that continued research in these areas is warranted.

CHAPTER III

STATISTICAL MODELING OF THE QUALITY CONTROL APPROACHES

Introduction

In order to compare the performance of clinical quality control monitoring approaches, statistical modeling is required. Each method's statistical performance must be evaluated, then compared against the other methods in order to understand how the methods perform in relation to one another. This chapter takes each of the quality control approaches individually and evaluates its statistical performance. Then, the methods are compared against one another at the close of the chapter in order to draw some conclusions about the performance of the methods with respect to each other.

As previously stated in this dissertation, a number of different instances of measurement system instability are explored. While traditional approaches only assume that the same shift is applied to each of the levels being monitored, this research explores the performance of the quality control methods when the individual quality control levels being monitored shift independently. This potential is very real in the clinical setting. Examples of such instances would include an individual level of control material being improperly stored, resulting in a shift in the actual value of the control material. Another instance would be the case when the calibration curve shifts at one end without shifting at the other end. Since multiple levels of control must be monitored by law, there is definitely a concern that all levels must be periodically evaluated in order to maintain the overall stability of the measurement system.

The sample sizes evaluated in this research include samples of size two (N'=2) and three (N'=3). These numbers were selected as they are the most likely to be encountered in actual laboratories. A minimum of two levels must be run each day, and a maximum of three levels is available (low, medium, and high), therefore sample sizes of N'=2 or N'=3 are justified. Given these sample sizes, there are five scenarios of shifts in the measurement system considered in this research. Using notation N=X/Y where X denotes the number of levels of controls monitored, and Y denotes the number of levels shifted, the scenarios are as follows:

N=2/2: Two levels of control with the same shift in both levels.

N=2/1: Two levels of control with only one level being shifted.

N=3/3: Three levels of control with the same shift in all three levels.

N=3/2: Three levels of control with 2 levels being shifted.

N=3/1: Three levels of control with only 1 level being shifted.

For example, N=3/2 would refer to a situation where three levels of control (low, mid, and high) are being monitored, and two of the levels (say low and mid) shift while the remaining level (high) remains centered on target. After defining these potentials shifts in the measurement process, one can then evaluate the performance of each of the quality control approaches under these five circumstances.

One could consider the case N=1/1 (i.e., only one level of control is monitored and that level is shifted), but the clinical environment will realistically either monitor two or three levels of control. Given that federal law mandates a minimum of two levels of

control material be tested every twenty four hours, only instances of two or three levels of control material are appropriate for this research

Some assumptions made in this research warrant discussion with the first assumption being that all data is normally distributed. This assumption has been shown to be appropriate for clinical quality control data in the past, and facilitates the analyses in the research. Also, subgroups (also referred to as QC timepoints, or moments in time when control values are tested for quality control purposes) are assumed to be independent of one another. The only exception is the instance when shifts are applied to the measurement system. Then, the shift would remain until detected by the quality control approach and remedied.

Throughout this research, statistical comparisons of methods are made on the basis of the approach's average run length (ARL) for a given circumstance. For methods employing runs rules, the ARL is the accepted metric for comparing methods since adding additional control rules will reduce the average time to signal a shift, but not increase the probability of detection of an individual subgroup. For instances where a probability of detection is available (i.e., +/- 2 SD limits), the geometric distribution is used to convert the probability into an ARL by taking the inverse of the probability of detection.

+/- 2 SD Limits

One of the most popular approaches for monitoring measurement system stability in the clinical laboratory is the strict use of +/- 2 SD limits. The approach stipulates that no matter how many levels of control material or samples are run, if any of the replicates

plots outside +/- 2 SD limits as determined from historical data, the measurement system is considered out-of-control and shut down. This approach is very simple in that replicates are compared to a single set of limits and stability determined immediately. There is also a long-standing tradition of use of +/- 2 SD limits, so there is a high degree of comfort with this approach felt by many laboratorians.

While the approach may be very convenient for application, the real question is how the approach performs statistically. In this case, the statistical performance is easily quantifiable. Assuming normally distributed data, one need only calculate the probability of being outside +/- 2 SD limits for a specified shift to evaluate its performance. Then, the inverse of this probability provides the ARL which can be used for method evaluation.

For example, consider the case N=2/1 with a 1.0 SD shift. This means that two levels of control are being monitored, and that one of the levels has experienced a 1.0 SD shift while the other level has remained centered. The probability of detecting this shift can be calculated as follows:

P(detection) = 1-P(neither control value exceeds +/- 2 SD limits)

P(detection) = 1-[(1-0.0456)*(1-(0.1587+0.0013))]

P(detection) = 0.1983

ARL = 1/P(detection) = 1/0.1983 = 5.04

This approach can be used to calculate the ARLs for the various shifts considered in this research.

The ARLs for the +/- 2 SD limits approach are summarized in Table 3.1 and Figure 3.1 respectively. The summary shows the ARLs for varying number of controls

Table 3.1 - ARLs for +/-2 SD limits by Type of Shift

Shift	N=2/2	N=2/1	N=3/3	N=3/2	N=3/1
0	11.22	11.22	7.65	7.65	7.65
0.2	10.3	10.74	7.04	7.23	7.44
0.4	8.19	9.46	5.64	6.17	6.83
0.6	6.11	7.86	4.25	4.95	5.99
0.8	4.51	6.33	3.19	3.89	5.09
1	3.4	5.04	2.46	3.06	4.26
1.2	2.63	4.02	1.95	2.45	3.54
1.4	2.11	3.25	1.62	2.01	2.95
1.6	1.75	2.67	1.39	1.69	2.48
1.8	1.51	2.24	1.24	1.47	2.12
2	1.33	1.91	1.14	1.31	1.84
2.2	1.22	1.67	1.08	1.2	1.62
2.4	1.13	1.49	1.04	1.13	1.46
2.6	1.08	1.35	1.02	1.08	1.33
2.8	1.05	1.25	1.01	1.04	1.24
3	1.03	1.18	1.004	1.02	1.17

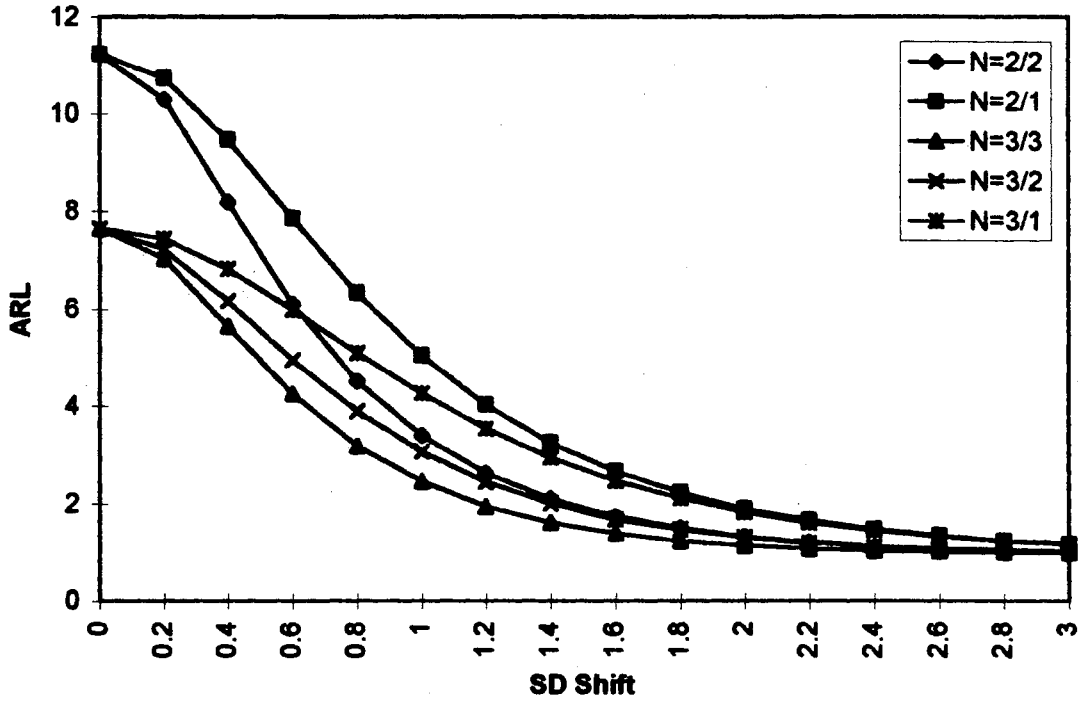

Figure 3.1 - ARLs for +/-2 SD limits by Type of Shift

and for different shifts in the various control levels. From the graph, it is evident that increased subgroup size reduces the ARL for detecting true shifts. However, increased subgroup size also has the detrimental effect of decreasing the average number of quality control evaluations between false rejections resulting in a method with an extremely high rate of false rejections.

Another insight gained from Table 3.1 and Figure 3.1 is that +/- 2 SD limits is more sensitive to a common shift across all levels (either N'=2 or N'=3) than to shifts in individual levels. This is not an unexpected result, but it raises an interesting point. The method is not as sensitive to changes in individual levels as it is to a change across all levels. This means that some clinicians may have a false sense of security in the performance of +/- 2 SD limits for all types of shifts when the method actually has a reduced ability to detect shifts in individual levels of control materials.

+/- 2 SD Limits with a Retest

Another popular approach to monitoring quality control in the clinical laboratory is the use of +/- 2 SD limits verified by a retest. Since some laboratory personnel are aware that +/- 2 SD limits strictly applied will result in a high rate of false rejection, they choose to give the measurement system a "second chance." If the clinician tests controls and a replicate is outside +/- 2 SD limits, then the clinician reruns the controls and classifies the measurement system as unstable only if a replicate exceeds +/- 2 SD limits on the second

run. If all control values are within +/- 2 SD limits on the second run, then the measurement system is considered stable.

Again, this approach is easily evaluated statistically. One need only calculate the probability of at least one replicate of control exceeding +/- 2 SD limits on two consecutive subgroups. Using the assumption of normal data, these probabilities can be readily calculated. Then, using the assumption that the signaling of the quality control system follows the geometric distribution, the ARL for specified shifts in control levels can be determined. These ARLs are summarized in Table 3.2 and Figure 3.2.

For example, the ARL for the case of N=2/1 with a 1.0 SD shift is calculated as follows:

P(detection) = P(at least one replicate exceeds +/- 2 SD limits consecutively)

$P(\text{detection}) = \{1-[(1-0.0456)*(1-(0.1587+0.0013))]\}^2 = 0.039$

ARL = 1/P(detection) = 1/0.029 = 25.43

Using this approach, the ARLs for all the possible shifts considered can be calculated.

As is the case with strict +/- 2 SD limits, increased sample size and the same shift in all the control levels results in decreased ARLs for detecting true changes in the measurement system. The graph also indicates that larger shifts, regardless of sample size or number of levels shifted, tend to have small and similar ARLs. The relatively large ARL when the measurement system is in control gives the approach its attractiveness.

Table 3.2 - ARLs for +/- 2 SD limits with Retest by Type of Shift

Shift	N=2/2	N=2/1	N=3/3	N=3/2	N=3/1
0	125.91	125.91	58.58	58.58	58.58
0.2	106.02	115.3	49.54	52.29	55.29
0.4	67.15	89.46	31.8	38.07	46.59
0.6	37.32	61.8	18.08	24.55	35.86
0.8	20.37	40.09	10.2	15.14	25.94
1	11.54	25.43	6.03	9.38	18.13
1.2	6.93	16.19	3.82	6	12.51
1.4	4.45	10.56	2.62	4.04	8.69
1.6	3.07	7.13	1.94	2.87	6.16
1.8	2.27	5	1.54	2.16	4.48
2	1.78	3.66	1.31	1.72	3.37
2.2	1.48	2.79	1.17	1.45	2.63
2.4	1.29	2.22	1.087	1.27	2.12
2.6	1.17	1.84	1.043	1.16	1.78
2.8	1.1	1.57	1.019	1.091	1.54
3	1.05	1.39	1.008	1.05	1.37

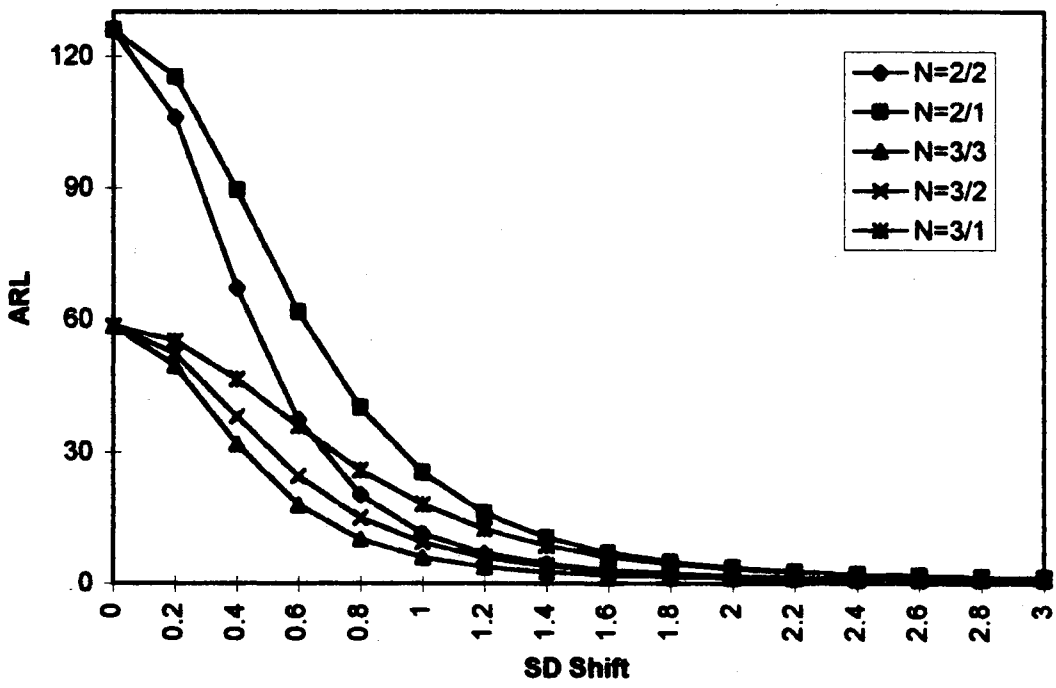

Figure 3.2 - ARLs for +/- 2 SD limits with Retest by Type of Shift

The Westgard Multirule Procedure

As a reaction to the widespread use of +/- 2 SD limits, Westgard (1981) proposed a multirule procedure for detecting changes in clinical analyzer systems. Termed the "Westgard Multirule Procedure", Westgard's approach was warmly received in the clinical literature. Westgard's procedure can be used with either N'=2 or N'=3 samples per subgroup, but it is applied slightly differently for the differing sample sizes. Figures 3.3 and 3.4 show the procedure for N'=2 and N'=3 respectively.

The motivation for Westgard's work was to provide clinicians with a viable alternative to strict +/- 2 SD limits. Due to the high rate of false rejection, alternatives to +/- 2 SD limits needed to be considered. Westgard provided a method that still used +/- 2 SD limits as a "warning" rule, but he added a series of control rules (similar to the AT&T runs rules) to reduce the incidence of false rejection in the clinical laboratory while maintaining a relatively powerful monitoring approach. However, this addition of runs rules adds a great deal of complexity to the monitoring approach which has hindered its widespread application in laboratories.

From a statistical modeling perspective, the Westgard Multirule Procedure can be evaluated in two ways: using a Markov modeling approach or by simulation. The Markov modeling approach would provide exact analytical results for the ARLs of the Westgard Rules, but the procedure is extremely complex (as described in Figures 3.3 and 3.4). Therefore, the transition matrix encountered in using a Markov model would be extraordinarily large. Since the focus of this dissertation is to look at many quality control

Figure 3.3: The Westgard Multirule Procedure Using Two Levels of Control Material (N'=2)

Step 1: Inspect both data points against +/- 2 SD limits.

If both values are within +/- 2 SD limits, then the measurement system is considered stable. Do not continue to Step 2.

If either value is outside +/- 2 SD limits, then consider the additional rules.

Step 2: Inspect control data within the QC timepoint.

The measurement system is unstable if:

(1_{3s}) Either value is outside +/- 3 SD limits.

(2_{2s}) Both values are outside +/- 2 SD limits on the same side of the centerline.

(R_{4s}) The difference between the values is greater than 4 SD. (i.e., The case where one level is at least 2 SD below its centerline and the other level is at least 2 SD above its centerline.)

Step 3: Inspect control data across QC timepoints.

The measurement system is unstable if:

(2_{2s}) Two consecutive values of the same level of control are outside +/2 2 SD limits on the same side of the centerline.

(4_{1s}) All four values (data from this QC timepoint and the previous QC timepoint) exceed +/- 1 SD limits on the same side of the centerline.

(4_{1s}) Four values from the same level of control exceed +/- 1 SD limits on the same side of the centerline.

(10_x) All ten values (data from this QC timepoint and the four previous QC timepoints) are above or below the centerline.

(10_x) Ten consecutive values from the same level of control are above or below the centerline.

Figure 3.4: The Westgard Multirule Procedure Using Three Levels of Control Material (N'=3)

Step 1: Inspect both data points against +/- 2 SD limits.

If all three values are within +/- 2 SD limits, then the measurement system is considered stable. Do not continue to Step 2.

If any value is outside +/- 2 SD limits, then consider the additional rules.

Step 2: Inspect control data within the QC timepoint.

The measurement system is unstable if:

(1_{3s}) Any value is outside +/- 3 SD limits.

(2_{2s}) Two of the three values are outside +/- 2 SD limits on the same side of the centerline.

(3_{1s}) All three values exceed +/- 1 SD limits on the same side of the centerline.

(R_{4s}) The difference between any of the values is greater than 4 SD. (i.e., The case where one level is at least 2 SD below its centerline and either of the other two levels is at least 2 SD above its centerline.)

Step 3: Inspect control data across QC timepoints.

The measurement system is unstable if:

(2_{2s}) Two consecutive values of the same level of control are outside +/- 2 SD limits on the same side of the centerline.

(4_{1s}) Four values from the same level of control exceed +/- 1 SD limits on the same side of the centerline.

(9_x) All nine values (data from this QC timepoint and the two previous QC timepoints) are above or below the centerline.

(9_x) Nine consecutive values from the same level of control are above or below the centerline.

approaches and to focus on their cost impact, simulation is employed for analyzing the Westgard Multirule Procedure. Some analytical results of special cases exist for validating the simulation, and the simulation results generated can be shown quite accurate and adequate for this research.

Given that a simulation approach is employed for analyzing the Westgard Rules, some specific issues regarding the simulation must be addressed. The first is the generation of random numbers. For the generation of Uniform (0,1) variates, the random number generator developed by Marse and Roberts (1983) is used. This generator has been widely tested and shown to be a solid random number generator. Also, the polar method by Marsaglia and Bray (1964) is used for generating Normal variates from the Uniform (0,1) variates. Again, this method has a long standing reputation for providing good random numbers for simulation purposes. The seeds for random number generation are randomly generated between 0 and 1000, then entered as a part of the simulation input in order to allow for the replication of the results reported in this dissertation.

Two simulation programs are used for this research: one program for the Westgard Multirule Procedure applied to two levels of controls and another program for three levels of controls. Written in Turbo Pascal version 6.0, the code for these programs appears in Appendices A and B, respectively. While one program could be written to deal with both sample sizes, the differences in logic between the N'=2 and N'=3 cases is significant enough to warrant separate programs.

As written, the simulation programs allow the user a great deal of freedom. The user can specify which set of control rules to use for evaluation, thus enabling a user to

examine other sets of rules than just the set forming the Westgard Multirule Procedure. Additionally, the user specifies the amount of shift applied to each level of control material. This allows for the evaluation of the set of rules when levels of control material are shifted independently.

In order to validate the simulation and to determine how many realizations are required to obtain reliable results, comparison can be made to some existing results. In his Master's thesis, Lee (1996) provides some analytical results using Markov modeling for some control rule combinations. Lee's work focused on sample sizes N'=2 with the same shift being affected to both levels of control material. The work by Lee does not consider use of the R_{4s} rule or the 2 SD limit as a warning rule, but the simulation can still be compared to special cases considered by Lee. Specifically, two instances considered by Lee are used to validate the developed simulation. The first instance (see summary in Table 3.3) considers the set of control rules $1_{3s}/2_{2s}/4_{1s}/10_{x}$ when no shift has been applied to the measurement system and the second instance (see summary in Table 3.4) considers the same set of control rules but applied to a shift of 2.0 SD in both levels of control material. Each of the ten replications shown in the tables are based on running the simulation for 10,000 realizations (a realization being the eventual signaling of the control system).

From the summaries in Table 3.3 and Table 3.4, it is clear that the simulation provides results that agree with the analytical results. These results also lead to the selected number of realizations employed in the simulation. Each simulation result is

Table 3.3 - Simulation Results for $1_{3s}/2_{2s}/4_{1s}/10_x$ when Shift = 0.0

Replicate	Seed	ARL	ARL SD
1	862	73.00	71.92
2	209	73.27	71.94
3	192	73.96	72.19
4	287	73.64	71.46
5	506	72.78	71.29
6	687	72.20	69.64
7	786	72.80	69.43
8	191	72.06	69.00
9	13	72.22	71.00
10	96	72.44	71.64

Lee's Results:

ARL = 73.21 ARL SD = 70.91

Simulation Results:

ARL 95% Confidence Interval = (72.37,73.29)

Table 3.4 - Simulation Results for $1_{3s}/2_{2s}/4_{1s}/10_x$ when Shift = 2.0 SD

Replicate	Seed	ARL	ARL SD
1	111	1.88	0.99
2	243	1.90	1.00
3	589	1.91	1.00
4	888	1.90	1.00
5	166	1.89	1.00
6	16	1.92	1.00
7	989	1.90	1.00
8	345	1.91	1.00
9	722	1.88	0.98
10	401	1.90	1.00

Lee's Results:

ARL = 1.90 ARL SD = 1.00

Simulation Results:

ARL 95% Confidence Interval = (1.89,1.91)

determined by building a 95% confidence interval around 5 replicates consisting of 5,000 realizations each. In other words, the simulation runs until 5,000 out-of-control signals are observed. It calculates the average of the 5,000 times to signal, and this constitutes a single replicate. Then, the confidence interval is based on 5 replicates of the 5,000 realizations.

Using the simulation just described, the Westgard Multirule Procedure can be evaluated for its statistical performance as a function of sample size (N'=2 or N'=3) and observed shifts in control materials (all control materials, a single material, etc.). Table 3.5 and Figure 3.5 show the ARLs for the Westgard Multirule Procedure for a variety of sample sizes and shifts. As is the case for +/- 2 SD limits and +/- 2 SD limits with a retest, the Westgard Multirule Procedure has smaller ARLs for larger sample sizes and is less able to detect shifts in single levels of control materials as opposed to the same shift across all control materials. Additionally, larger shifts are detected relatively similarly regardless of the sample size or the number of control levels affected.

The χ^2 Chart

The quality control approaches for clinical laboratories discussed thus far in the research are all methods which have some traditional application in the clinical field. The methods covered from this point forward in this chapter are all methods which are new to the clinical environment. While these multivariate approaches have been utilized in industrial applications for some time, they have not been previously applied in the clinical setting.

Table 3.5 - ARLs for the Westgard Multirule Procedure by Type of Shift

Shift	N=2/2	N=2/1	N=3/3	N=3/2	N=3/1
0	83.95	83.95	33.72	33.72	33.72
0.2	68.25	74.3	27.79	29.8	31.72
0.4	42.14	55.07	17.98	21.89	27.17
0.6	23.87	37.5	10.92	14.84	21.1
0.8	13.71	24.57	6.75	9.95	15.68
1	8.59	16.01	4.54	6.88	11.62
1.2	5.81	11.04	3.17	4.92	8.63
1.4	4.08	7.82	2.37	3.68	6.43
1.6	3.09	5.82	1.88	2.87	5
1.8	2.45	4.44	1.54	2.3	3.93
2	1.99	3.53	1.33	1.91	3.19
2.2	1.71	2.87	1.2	1.63	2.66
2.4	1.49	2.42	1.11	1.43	2.27
2.6	1.34	2.09	1.05	1.28	1.97
2.8	1.21	1.84	1.02	1.18	1.74
3	1.14	1.65	1.01	1.12	1.57

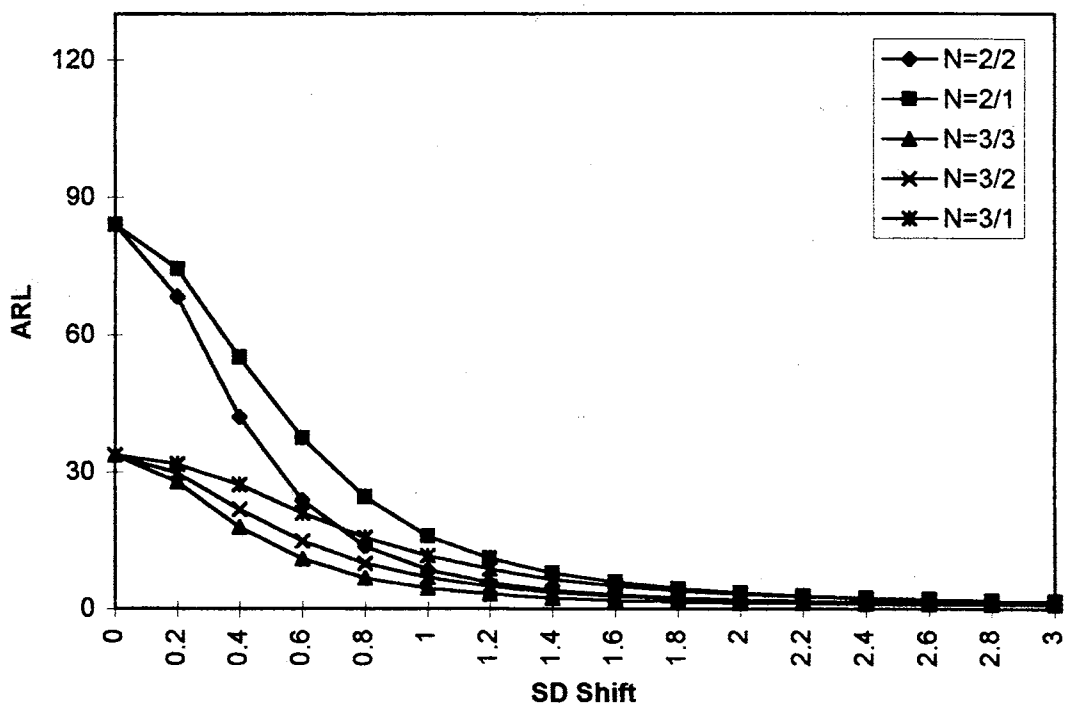

Figure 3.5 - ARLs for the Westgard Multirule Procedure by Type of Shift

The new methods described here are all multivariate in nature. The traditional approaches do not explore the underlying correlation structure of the control levels being monitored, even though it is readily apparent that these levels are in fact correlated. The correlation arises from the fact that all the levels in a given QC timepoint are tested under similar conditions. Testing variation in clinical laboratories is typically largest from run to run with the sources of variability including environmental variability, technician variability, variation within the instrument, and materials variation. When the quality control levels are tested side by side, they are all affected by essentially the same run to run variation. This creates correlation among the levels. The correlation, however, is not perfect as there are some sources of variation that can affect individual levels differently. Therefore, it is evident that there is correlation among levels, and that this correlation is not complete.

While actual testing data can be used to show that there is correlation among levels and provide estimates of correlation values, this research will use three pre-selected correlation structures for investigation. The correlation and covariance matrices are shown subsequently. The three conditions assume the same correlation among all levels of control material with correlations of 0.80, 0.50. and 0.10 respectively. These correlations represent a high degree of correlation ($r = 0.80$), a moderate degree of correlation ($r = 0.50$) and a low degree of correlation ($r = 0.10$). Given these correlation structures, the covariance matrices are determined by using variances from a representative diagnostic test. In this fashion, the correlation matrices can be translated into covariance matrices for investigation in the research.

$$\rho_1 = \begin{bmatrix} 1.0 & 0.80 & 0.80 \\ 0.80 & 1.0 & 0.80 \\ 0.80 & 0.80 & 1.0 \end{bmatrix} \qquad \Sigma_1 = \begin{bmatrix} 0.0625 & 0.108 & 0.152 \\ 0.108 & 0.2916 & 0.328 \\ 0.152 & 0.328 & 0.5776 \end{bmatrix}$$

$$\rho_2 = \begin{bmatrix} 1.0 & 0.50 & 0.50 \\ 0.50 & 1.0 & 0.50 \\ 0.50 & 0.50 & 1.0 \end{bmatrix} \qquad \Sigma_2 = \begin{bmatrix} 0.0625 & 0.0675 & 0.095 \\ 0.0675 & 0.2916 & 0.2052 \\ 0.095 & 0.2052 & 0.5776 \end{bmatrix}$$

$$\rho_3 = \begin{bmatrix} 1.0 & 0.10 & 0.10 \\ 0.10 & 1.0 & 0.10 \\ 0.10 & 0.10 & 1.0 \end{bmatrix} \qquad \Sigma_3 = \begin{bmatrix} 0.0625 & 0.0135 & 0.019 \\ 0.0135 & 0.2916 & 0.041 \\ 0.019 & 0.041 & 0.5776 \end{bmatrix}$$

After determining the covariance structures for investigation, the next step in the research is to analyze the statistical performance of the multivariate approaches, beginning with the χ^2 chart. This multivariate approach allows for the simultaneous control of p quality characteristics. For the clinical application in this research, the number of characteristics, p, will correspond to the number of control levels being monitored, N'. Additional replicates (N'=2 or N'=3) actually increase the number of characteristics for monitoring rather than the sample size from a multivariate perspective. This is because additional replicates are in different levels (i.e., low, mid, or high) rather than replicates of the same level of control material.

For the quality control application to multiple quality characteristics, one can test the hypothesis that $\mu=\mu_o$ where μ_o is a specified vector (the historical mean vector of the control level data). The critical region to test this hypothesis is

$j(\bar{\mathbf{x}} - \mu_0)\Sigma^{-1}(\bar{\mathbf{x}} - \mu_0) > \chi_p^2(\alpha)$ where j is the sample size (in the clinical application, j = 1), p is the number of characteristics being monitored (p = 2 or 3 for the clinical case) and $\bar{\mathbf{x}}$ is the observed sample mean vector.

The assumptions underlying the application of the χ^2 chart (besides that of multivariate normality) are that the specified mean vector (μ_o) and covariance matrix (Σ) are known. While theoretically impractical, these assumptions may be applied in cases where a great deal of data has been collected surrounding the diagnostic measurement system.

After identifying the test statistic used for monitoring the χ^2 chart, the next question is how to statistically model and evaluate the performance of the chart. This is achieved through the use of the non-central χ^2 distribution. Anderson (1958) shows that the power function for evaluating the above hypothesis is the non-central χ^2 distribution, and he provides the density function of the non-central χ^2 with non-centrality parameter $\tau^2 = j(\mu - \mu_0)' \Sigma^{-1}(\mu - \mu_0)$ as:

$$f(v) = \frac{1}{2^{\frac{1}{2}p}} e^{-\frac{1}{2}(\tau^2+v)} v^{\frac{1}{2}p-1} \sum_{\beta=0}^{\infty} \left(\frac{\tau^2}{4}\right)^{\beta} \frac{1}{\beta!\Gamma\left(\frac{1}{2}p+\beta\right)} v^{\beta}$$

The non-central χ^2 distribution, therefore, provides a straightforward approach for calculating the probability of detection for the χ^2 chart. The upper limit of integration on the non-central χ^2 distribution is set by the selected α level of the χ^2 chart, and the non-centrality parameter τ^2 is determined by the amount of shift in the measurement system being evaluated. Given the parameters, integration of the non-central χ^2 distribution will yield the probability of detecting a given shift. As is this case with the traditional quality control approaches, this probability can be inverted to yield an ARL for the χ^2 chart.

For the purposes of this research, Mathcad 6.0, Student Version worksheets are used to perform the integration of the non-central χ^2 distribution. Examples of these worksheets are shown in Appendix C with separate worksheets for the cases where p = 2 (corresponding to N' = 2 from the laboratory perspective) and p = 3 (N' = 3 from the laboratory perspective).

As an example, consider the case of N=2/1 where the correlation between the two levels is 0.50 and the single level is shifted 3.0 SD while the other level remains centered. Since the correlation between levels is 0.50, matrices ρ_2 and Σ_2 provide the standard deviation for the single level as 0.25. Since the shift in the single level is 3.0 SD, that means the observed shift in the mean vector will be 0.75. Using this amount of shift, one can calculate the non-centrality parameter for the non-central χ^2 distribution as follows:

$$\tau^2 = j(\mu - \mu_0)' \Sigma^{-1}(\mu - \mu_0) = 1 \cdot \begin{pmatrix} 0.75 & 0 \end{pmatrix} \begin{pmatrix} 0.0625 & 0.0675 \\ 0.0675 & 0.2916 \end{pmatrix}^{-1} \begin{pmatrix} 0.75 \\ 0 \end{pmatrix} = 12.0$$

After determining the non-centrality parameter, the next step is to plug it into the equation for the non-central χ^2 distribution and integrate. The limits of integration, however, must first be determined. The limits of integration come from the upper limit for the χ^2 chart. In this instance, the α level is set at 0.01 and the number of characteristics is 2. So, the upper limit for the chart is determined from a table for the central χ^2 distribution as 9.21. Using this upper limit in the integral, one can calculate the probability of detecting the given shift as follows:

$$P(\text{det.}) = \int_{9.21}^{\infty} \frac{1}{2^{\frac{1}{2}\cdot 2}} e^{-\frac{1}{2}(12+v)} v^{\frac{1}{2}\cdot 2-1} \sum_{\beta=0}^{\infty} \left(\frac{12}{4}\right)^{\beta} \frac{1}{\beta! \Gamma\left(\frac{1}{2}\cdot 2+\beta\right)} v^{\beta} dv = 0.721$$

Turning this probability of detection into an ARL results in:

$$ARL = 1/P(det.) = 1/0.721 = 1.39$$

In other words, integrating the non-central χ^2 distribution from the upper limit of the χ^2 chart to infinity will yield the probability of detecting the given shift. Then, it is a simple matter of translating this probability into an ARL.

Using this statistical model for the χ^2 chart, the performance of the χ^2 chart can be investigated and evaluated. The first area for investigation is the effect of the correlation structure on the performance of the χ^2 chart. Using the three correlation structures described earlier in this section for each of the different types of shifts investigated in this research with α=0.01, one can gain insight into the impact of correlation on the χ^2 chart. Tables 3.6 through 3.10, along with corresponding Figures 3.6 through 3.10, summarize the effect of the correlation structure on the performance of the χ^2 chart.

An interesting insight from analysis of the generated results regards the impact that the correlation structure has on the ability to detect different shifts. For the same shift across all levels being monitored (N=2/2 and N=3/3), the lower degree of correlation case is more sensitive to shifts across all levels. In the situations where the shifts do not occur across all levels (N=2/1, N=3/2, and N=3/1), the highly correlated case is the most sensitive to the defined shifts. This can be attributed to the shape of the in-control area defined by the χ^2 chart . In the case where two levels of control materials are monitored, the in-control area for the χ^2 chart is an ellipse whose axis lies along the regression line between the two levels of control material. See Figure 3.11 for a graphical representation

Table 3.6 - χ^2 Chart Correlation Comparison (N=2/2)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	90.91	90.91	83.33
0.4	71.43	66.67	58.82
0.6	50	43.48	35.71
0.8	33.33	28.57	21.28
1	22.22	18.52	13.16
1.2	15.38	12.35	8.47
1.4	10.64	8.55	5.71
1.6	7.63	6.02	4.02
1.8	5.62	4.42	2.97
2	4.27	3.87	2.29
2.2	3.33	2.65	1.85
2.4	2.67	2.15	1.56
2.6	2.21	1.81	1.36
2.8	2.21	1.56	1.23
3	1.87	1.39	1.14

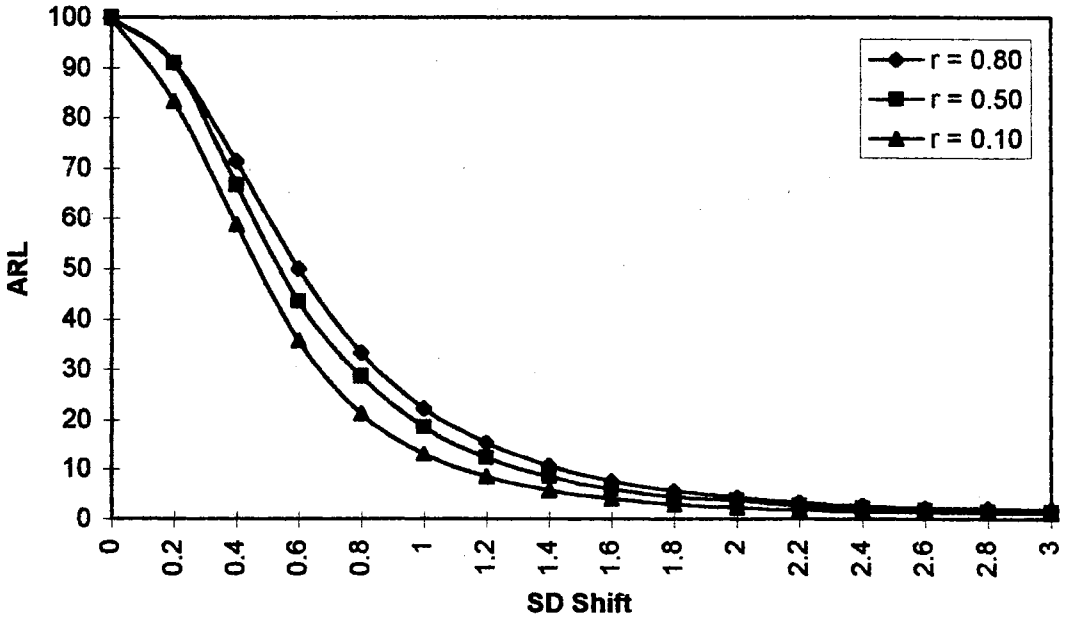

Figure 3.6 - χ^2 Chart Correlation Comparison (N=2/2)

Table 3.7 - χ^2 Chart Correlation Comparison (N=2/1)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	76.92	90.91	90.91
0.4	45.45	66.67	71.43
0.6	25	43.48	52.63
0.8	13.51	28.57	35.71
1	7.87	18.52	24.39
1.2	4.9	12.35	16.95
1.4	3.28	8.55	12.05
1.6	2.35	6.02	8.62
1.8	1.81	4.42	6.37
2	1.48	3.37	4.83
2.2	1.28	2.65	3.76
2.4	1.15	2.15	3.01
2.6	1.08	1.81	2.47
2.8	1.04	1.56	2.07
3	1.02	1.39	1.79

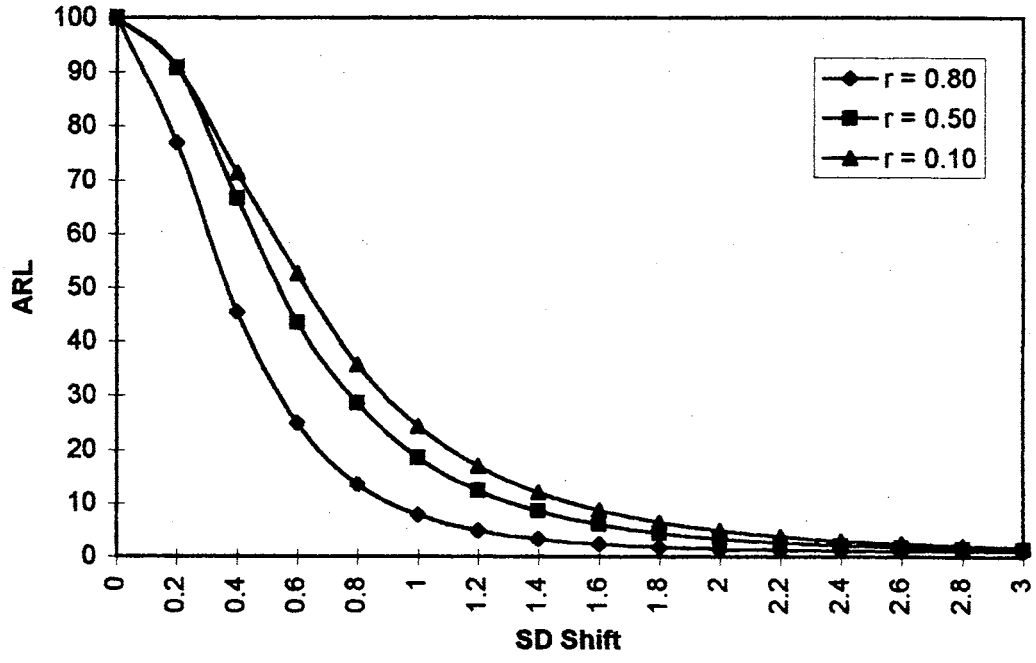

Figure 3.7 - χ^2 Chart Correlation Comparison (N=2/1)

Table 3.8 - χ^2 Chart Correlation Comparison (N=3/3)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	90.91	90.91	83.33
0.4	76.92	66.67	55.56
0.6	55.56	47.62	33.33
0.8	38.46	31.25	19.23
1	26.32	20.41	11.36
1.2	18.52	13.51	7.04
1.4	12.82	9.26	4.61
1.6	9.17	6.45	3.21
1.8	6.71	4.67	2.36
2	5	3.5	1.85
2.2	3.86	2.72	1.52
2.4	3.05	2.18	1.31
2.6	2.48	1.82	1.18
2.8	2.07	1.56	1.09
3	1.77	1.38	1.05

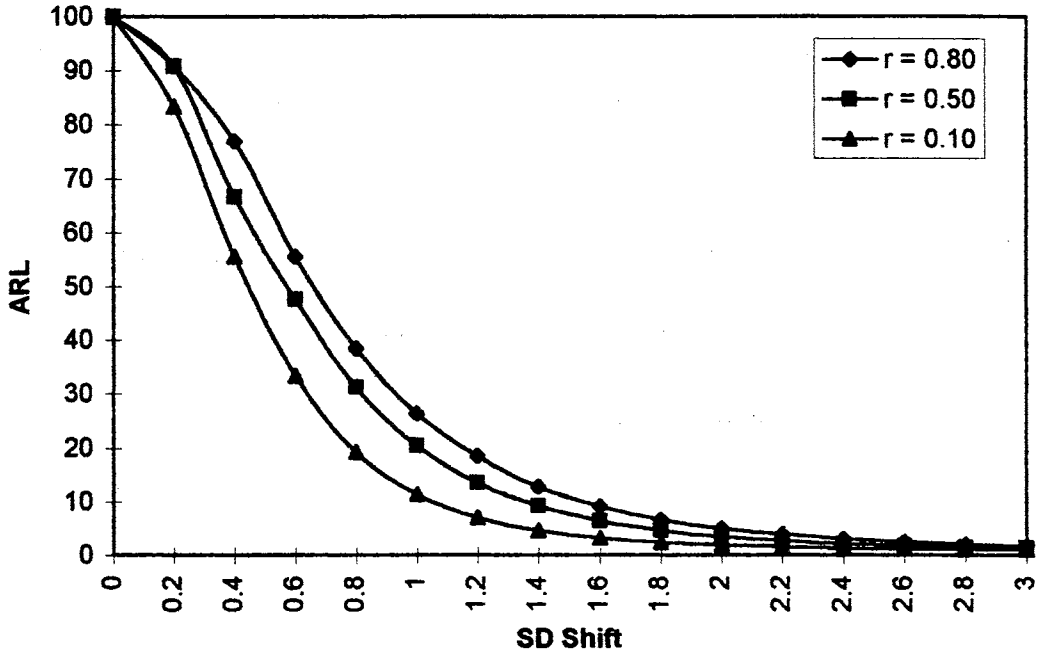

Figure 3.8 - χ^2 Chart Correlation Comparison (N=3/3)

Table 3.9 - χ^2 Chart Correlation Comparison (N=3/2)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	76.92	90.91	90.91
0.4	45.45	62.5	62.5
0.6	22.22	40	41.67
0.8	11.63	23.81	25.64
1	6.45	14.92	16.39
1.2	3.91	9.43	10.42
1.4	2.6	6.29	6.94
1.6	1.88	4.35	4.83
1.8	1.48	3.15	3.5
2	1.26	2.4	2.65
2.2	1.13	1.91	2.09
2.4	1.06	1.59	1.72
2.6	1.03	1.38	1.47
2.8	1.009	1.24	1.31
3	1.003	1.15	1.19

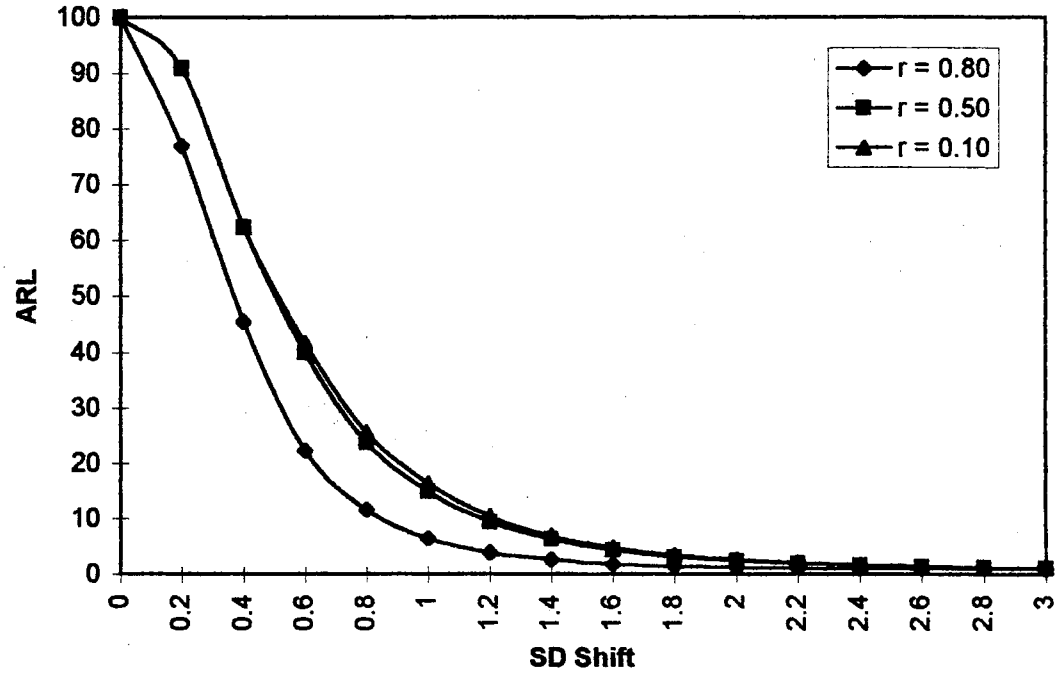

Figure 3.9 - χ^2 Chart Correlation Comparison (N=3/2)

Table 3.10 - χ^2 Chart Correlation Comparison (N=3/1)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	76.92	90.91	90.91
0.4	47.62	66.67	76.92
0.6	24.39	47.62	58.82
0.8	13.16	31.25	41.67
1	7.41	20.41	29.41
1.2	4.5	13.51	20.83
1.4	2.96	9.26	14.93
1.6	2.12	6.45	10.75
1.8	1.63	4.67	7.94
2	1.35	3.5	5.95
2.2	1.19	2.72	4.57
2.4	1.1	2.18	3.61
2.6	1.04	1.82	2.91
2.8	1.02	1.56	2.4
3	1.007	1.38	2.04

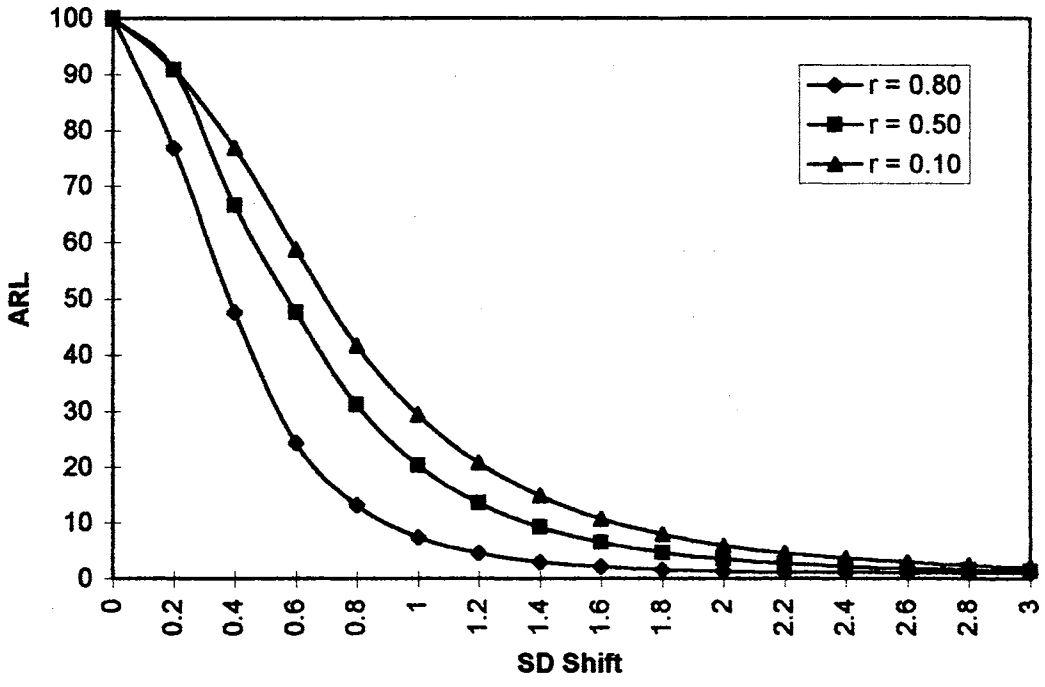

Figure 3.10 - χ^2 Chart Correlation Comparison (N=3/1)

of this in-control ellipse. The figure shows an in-control ellipse for two levels of control material (low and mid) which are highly correlated with the directions for N=2/1 and N=2/2 shifts, respectively. It is clear that a shift in the direction of N=2/1 will move outside the in-control ellipse more rapidly than a shift in the direction of N=2/2. This means that the χ^2 chart will be less sensitive to shifts along this regression line (i.e., the same shift across both levels of control material) than if a single level of control material

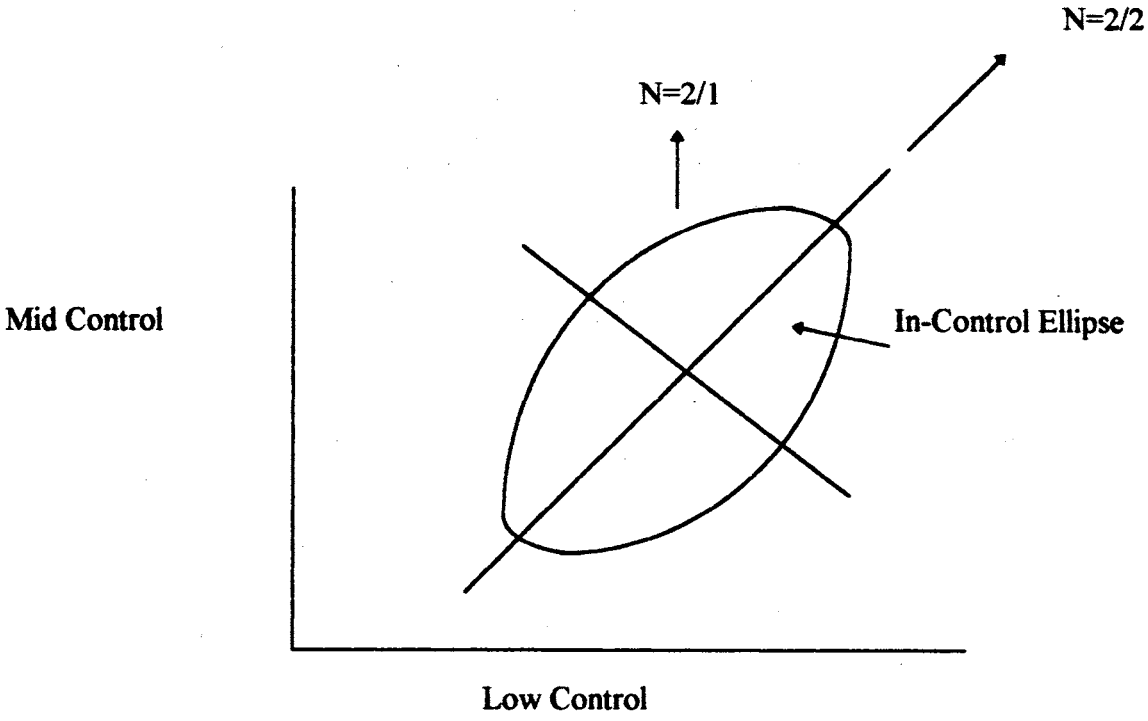

Figure 3.11 - In-Control Ellipse for Two, Highly Correlated Levels of Control Material

shifts. While the difference in performance caused by correlation structure is not dramatic, it is clear that properly determining the correct correlation structure for monitoring is an important issue when applying the χ^2 chart. The correlation structure will dictate the shape of the in-control ellipse, and thereby impact performance of the χ^2 chart.

For further analysis of the χ^2 chart in this research, the correlation structure is fixed as moderate (i.e., $r = 0.50$). Changing this correlation structure will change the comparison of the χ^2 chart to other methods, but the assumption of moderate correlation is appropriate for research. The actual correlation observed by laboratories is bound to fluctuate, but the degree of correlation among control levels is probably moderate at a minimum. Using a single correlation structure of 0.50 allows comparison of a single instance for each of the multivariate approaches against the traditional methods.

After fixing the correlation structure as moderate, the sensitivity of the χ^2 chart to the various shifts considered in the research can be evaluated. Table 3.11 and Figure 3.12 show the performance of the χ^2 chart for the different shifts investigated. The χ^2 chart shows similar performance for all the shifts examined in this research. While due in part to the moderate correlation structure used for analysis, this in an important characteristic for the χ^2 chart. While the traditional methods investigated in this research have very different performance for different shifts, the χ^2 chart shows comparable performance regardless of the type of shift encountered.

Table 3.11 - ARLs for the χ^2 chart (Corr. = 0.50) by Type of Shift

Shift	N=2/2	N=2/1	N=3/3	N=3/2	N=3/1
0	100	100	100	100	100
0.2	90.91	90.91	90.91	90.91	90.91
0.4	66.67	66.67	66.67	62.5	66.67
0.6	43.48	43.48	47.62	40	47.62
0.8	28.57	28.57	31.25	23.81	31.25
1	18.52	18.52	20.41	14.92	20.41
1.2	12.35	12.35	13.51	9.43	13.51
1.4	8.55	8.55	9.26	6.29	9.26
1.6	6.02	6.02	6.45	4.35	6.45
1.8	4.42	4.42	4.67	3.15	4.67
2	3.37	3.37	3.5	2.4	3.5
2.2	2.65	2.65	2.72	1.91	2.72
2.4	2.15	2.15	2.18	1.59	2.18
2.6	1.81	1.81	1.82	1.38	1.82
2.8	1.56	1.56	1.56	1.24	1.56
3	1.39	1.39	1.38	1.15	1.38

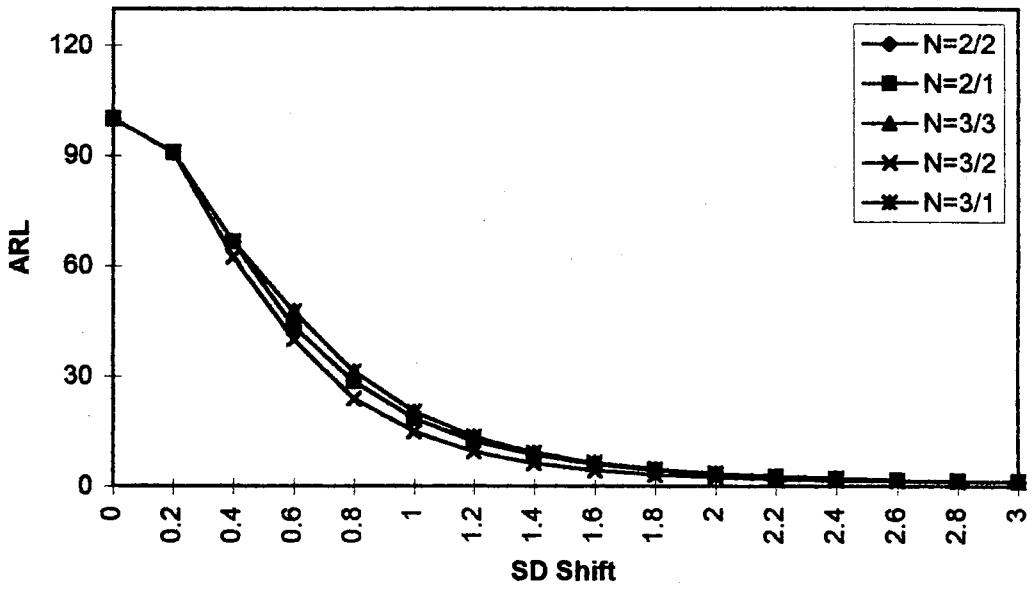

Figure 3.12 - ARLs for the χ^2 chart (Corr. = 0.50) by Type of Shift

The T^2 Chart

As explained in the previous section, the χ^2 chart is a multivariate approach which assumes a known covariance matrix. However, the typical application in practice will involve a situation in which the covariance matrix must be estimated from collected data. This case would then be an instance in which the covariance matrix is unknown, resulting in the requirement of another approach instead of the χ^2 chart. Such an approach is the T^2 chart which assumes an unknown covariance structure and an unknown mean vector. Developed by Hotelling (1931), the T^2 statistic allows one to apply a multivariate control monitoring approach to p characteristics when the covariance matrix is unknown.

In all respects with the exception of the unknown covariance matrix and the mean vector, the T^2 chart is completely analogous to the χ^2 chart. Therefore, this research will evaluate the T^2 chart in the same fashion as the χ^2 chart realizing that the T^2 chart cannot be expected to outperform the χ^2 chart. As the number of subgroups used in establishing the T^2 chart limits increases to infinity, the T^2 chart will approach the χ^2 chart. So, the focus of this section is to identify a workable number of subgroups to collect to set up the T^2 chart limits. In a later section of the dissertation, the T^2 chart's performance will be compared directly to the χ^2 chart along with the rest of the quality control approaches selected for study in this research.

Similarly to the χ^2 chart, Anderson (1958) shows that the T^2 statistic is defined as $T^2 = m(\bar{\mathbf{x}} - \mu_o)' \mathbf{S}^{-1}(\bar{\mathbf{x}} - \mu_o)$ where S is the sample covariance matrix, m is the number of subgroups of size j collected, $\bar{\mathbf{x}}$ is the mean vector of the sample with dimension p, and μ_o

is the historical mean vector. Anderson goes on to show that $(T^2/j)(m-p)/p$ is distributed as a non-central F distribution with non-centrality parameter $\tau^2 = j(\mu - \mu_o)' \Sigma^{-1}(\mu - \mu_o)$. Using this result, Anderson then develops the distribution for T^2 as:

$$f(t) = \frac{e^{-\frac{1}{2}\tau^2}}{(m-1)\Gamma\left(\frac{1}{2}(m-p)\right)} \sum_{\beta=0}^{\infty} \frac{\left(\frac{\tau^2}{2}\right)^{\beta} \left[\frac{t}{(m-1)}\right]^{\frac{1}{2}p+\beta-1} \Gamma\left(\frac{1}{2}m+\beta\right)}{\beta!\Gamma\left(\frac{1}{2}p+\beta\right)\left[1+\frac{t}{(m-1)}\right]^{\frac{1}{2}m+\beta}}$$

Using this expression for the density function of the T^2, one can integrate the function to determine a probability for detecting a shift in the mean vector. Appendix D shows Mathcad 6.0, Student Edition worksheets used for integrating the T^2 distribution to find probabilities of detection which are then converted into ARLs. Two worksheets are included in the appendix: one for the p=2 case (or N'=2 from the clinical perspective) and one for the p=3 case (or N'=3 from the clinical perspective).

For example, consider the case N=2/1 where the single level is shifted by 1.4 SD. Using the same α limit as before (0.01), 20 subgroups, and a correlation of r = 0.50 between the two levels (i.e., using ρ_2 and Σ_2), the probability of detection can be calculated. First, the non-centrality parameter for the non-central T^2 distribution must be calculated. This non-centrality parameter is $\tau^2 = j(\mu - \mu_o)' \Sigma^{-1}(\mu - \mu_o)$, or

$$\tau^2 = 1 \cdot \begin{pmatrix} 0.35 & 0 \end{pmatrix} \begin{pmatrix} 0.0625 & 0.0675 \\ 0.0675 & 0.2916 \end{pmatrix}^{-1} \begin{pmatrix} 0.35 \\ 0 \end{pmatrix} = 2.613$$

Next, the upper limit for the T^2 chart is calculated as:

$$\text{Upper Limit} = \frac{(m-1)p}{(m-p)} F_{p,m-p,\alpha} = \frac{(20-1)2}{(20-2)} \cdot 6.01 = 12.69$$

Using the non-centrality parameter and the upper limit for the T^2 chart, one can then calculate the probability for detecting the shift as follows:

$$P(\text{det.}) = \int_{12.69}^{\infty} \frac{e^{-\frac{1}{2}(2.613)}}{(20-1)\Gamma\left(\frac{1}{2}(20-2)\right)} \sum_{\beta=0}^{\infty} \frac{\left(\frac{2.613}{2}\right)^{\beta}\left[\frac{t}{(20-1)}\right]^{\frac{1}{2}p+\beta-1} \Gamma\left(\frac{1}{2}\cdot 20+\beta\right)}{\beta!\Gamma\left(\frac{1}{2}\cdot 2+\beta\right)\left[1+\frac{t}{(20-1)}\right]^{\frac{1}{2}\cdot 20+\beta}} dt = 0.086$$

ARL = 1/P(det.) = 1/0.086 = 11.63

As was the case with the χ^2 chart, the probability of detecting the shift is simply the area of the non-central T^2 distribution outside the upper limit for the central T^2 distribution.

Using the Mathcad worksheets, the effect of the number of subgroups on the performance of the T^2 chart is evaluated. As is the case with the χ^2 chart, the α level for the T^2 chart is set at 0.01 to ensure a low rate of false rejection. Tables 3.12 and 3.13 along with accompanying Figures 3.13 and 3.14 show the impact of the number of subgroups collected on the performance of the T^2 chart. The two sets of tables and figures show the effect of the number of subgroups for a 1.0 SD shift and a 2.0 SD shift respectively when the shifts occur across all three levels (N=3/3) for the three types of correlation structures investigated in this research. These graphics indicate that there is diminishing return in increasing the number of subgroups collected for establishing the T^2 chart limits past 20 and that the correlation structure has little effect on the selection of the number of subgroups collected. Given that 20 subgroups is often a minimum recommended number of subgroups for the establishment of control chart limits in many

Table 3.12 - Number of Subgroups Comparison for the T^2 Chart (N=3/3, 1.0 SD Shift)

m	r = 0.80	r = 0.50	r = 0.10
10	45.45	38.46	25
20	34.48	27.78	16.39
43	30.3	23.26	13.51
63	28.57	22.73	12.82
123	27.78	21.28	12.05

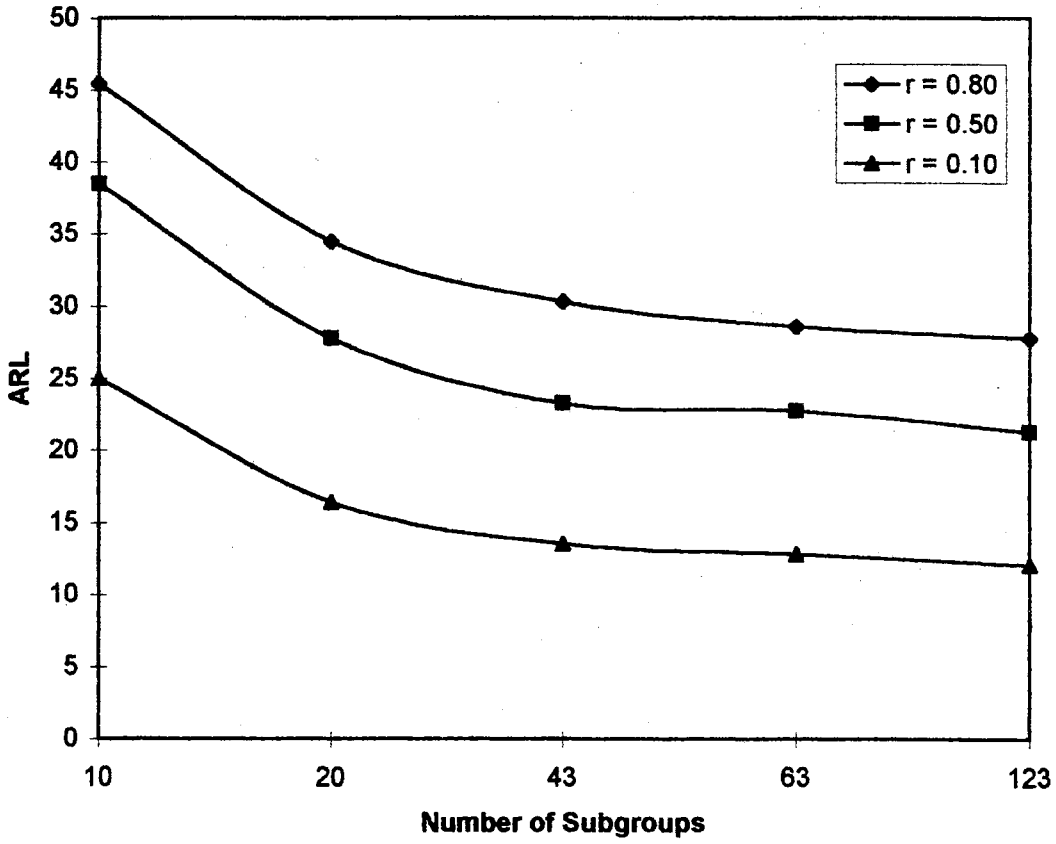

Figure 3.13 - Number of Subgroups Comparison for the T^2 Chart (N=3/3, 1.0 SD Shift)

Table 3.13 - Number of Subgroups Comparison for the T^2 Chart (N=3/3, 2.0 SD Shift)

m	r = 0.80	r = 0.50	r = 0.10
10	13.16	9.62	5
20	7.75	5.43	2.75
43	6.06	4.24	2.17
63	5.71	3.98	2.06
123	5.35	3.73	1.95

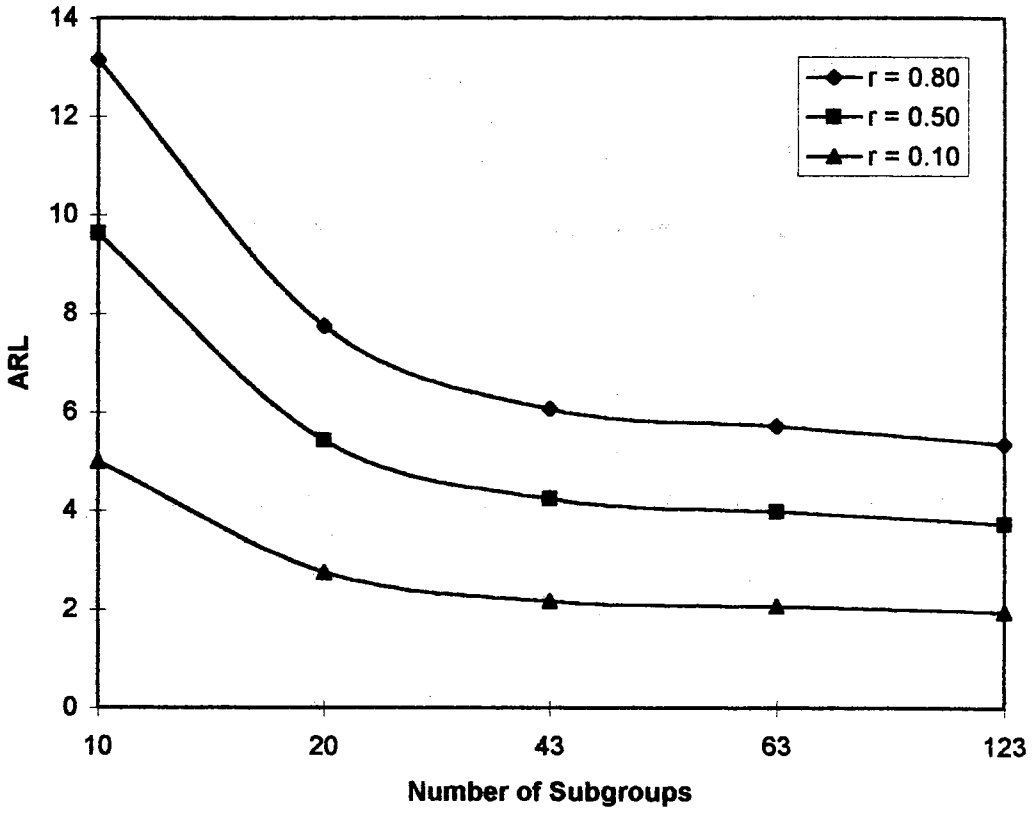

Figure 3.14 - Number of Subgroups Comparison for the T^2 Chart (N=3/3, 2.0 SD Shift)

quality control applications and the fact that this is a realistic number of subgroups to attain for a laboratory, 20 subgroups is used for work with the T^2 chart throughout the remainder of this dissertation.

After fixing the number of subgroups used in the evaluation of the T^2 chart at 20, the next question is how the T^2 chart performs for the variety of shifts in control material considered in the research. This summary is shown in Table 3.14 and Figure 3.15. As should be expected, the T^2 chart mimics the χ^2 chart in its performance. Additionally, it appears that the performance of the T^2 chart is similar for all shifts considered. This is attributable to the correlation structure used for the comparison ($r = 0.50$).

Principal Component Charts

Another multivariate approach which merits consideration for the clinical quality control application is the use of the principal component chart. While this approach is typically used for circumstances where a large number of characteristics are being monitored, the principal component chart is of interest to the clinical application because individual principal components can have some interesting interpretations in the clinical setting. Additionally, a principal components approach may provide some intuitive appeal to clinicians over the other multivariate approaches considered in this research. Once the principal components have been determined, evaluation of the performance of the principal component chart is much more straightforward than the other multivariate approaches (i.e., it does not require integration of the non-central χ^2 distribution), thus making the approach easier for clinicians to accept and understand.

Table 3.14 - ARLs for the T^2 Chart by Type of Shift

Shift	N=2/2	N=2/1	N=3/3	N=3/2	N=3/1
0	100	100	100	100	100
0.2	90.91	90.91	90.91	90.91	90.91
0.4	71.43	71.43	76.92	66.67	76.92
0.6	50	50	55.56	47.62	55.56
0.8	34.48	34.48	40	32.26	40
1	23.81	23.81	27.78	20.83	27.78
1.2	16.39	16.39	19.23	14.08	19.23
1.4	11.63	11.63	13.7	9.62	13.7
1.6	8.4	8.4	9.8	6.76	9.8
1.8	6.17	6.17	7.25	4.9	7.25
2	4.69	4.69	5.43	3.68	5.43
2.2	3.66	3.66	4.2	2.87	4.2
2.4	2.94	2.94	3.32	2.3	3.32
2.6	2.42	2.42	2.7	1.91	2.7
2.8	2.04	2.04	2.25	1.64	2.25
3	1.76	1.76	1.92	1.44	1.92

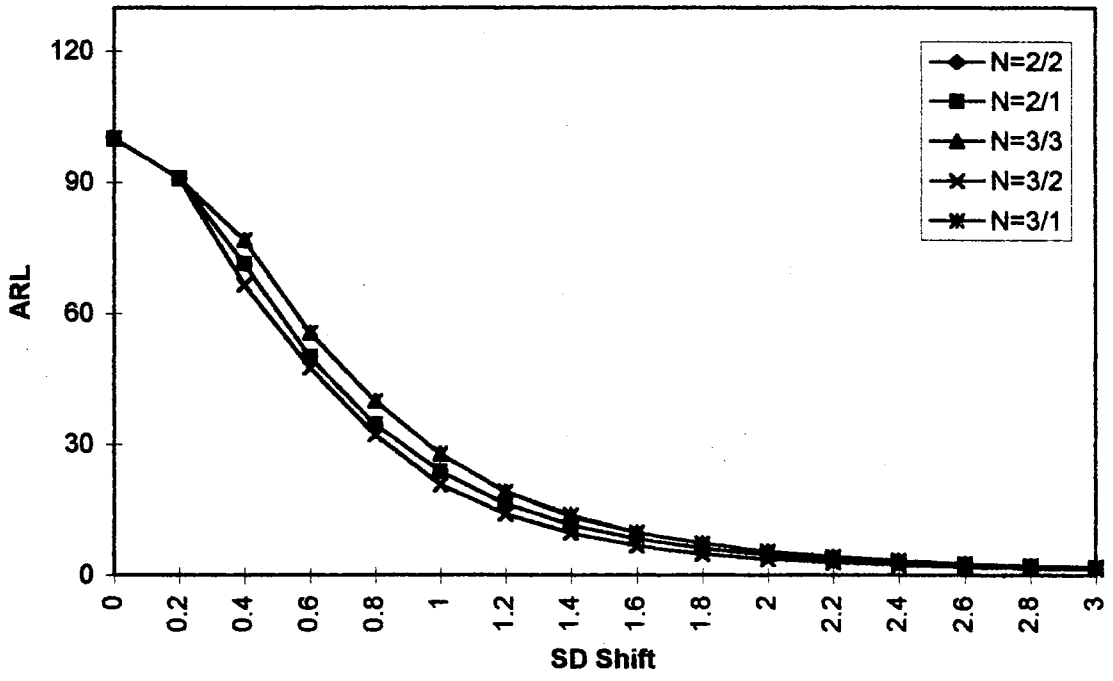

Figure 3.15 - ARLs for the T^2 Chart by Type of Shift

As was the case with the χ^2 chart, analysis of the principal component chart will consider two or three levels of control material with a known mean vector and covariance matrix. Also, it is assumed that the covariance matrix is unaffected by any changes in the mean vector.

Yu (1994) clearly identifies the method for determining probabilities of detection using the principal component chart. He shows that the amount of shift in the ith standardized principal component, Z_i (i.e., $N(0,1^2)$) is:

$$\text{Amount of shift for } Z_i = \frac{\mathbf{u}_i'(\mu - \mu_o)}{\sqrt{l_i}}$$

where $\mathbf{u}_i$ is the ith eigenvector of Σ, l_i is the ith eigenvalue of Σ, μ is the new mean vector upon measurement system shift, and μ_o is the historical, known mean vector. This standardized value can then be used with a standard normal table to determine the probability of detection for the ith principal component.

To facilitate the calculation of the ARLs for the principal component chart, Mathcad 6.0, Student Edition was again employed. In Appendix E, two worksheets (for N'=2 and N'=3, respectively) are displayed. The worksheets calculate the eigenvalues and eigenvectors for the given covariance matrix. Then, the amount of shift in the standardized principal components is calculated as described above. The bottom section of the worksheet calculates a Z value to be read from a standard normal table against an upper limit for each principal component. This upper limit is chosen to yield a combined α level across the control materials of 0.01 in order to remain consistent with the other multivariate approaches considered. These Z values are then read from a standard normal

table and combined into an overall probability of detection for the set of principal components.

As is the case with the other multivariate approaches, a major consideration is the impact of the correlation structure on method performance. Tables 3.15 through 3.19 and corresponding Figures 3.16 through 3.20 display the impact of correlation on the detection of the types of control level shifts identified for this research. The results for the principal component chart are very similar to the results from the χ^2 chart in that high correlation is most effective at detecting shifts in a single level of control material while low correlation is more effective in detecting shifts across levels.

An original intent of the research was to identify a single principal component for monitoring as opposed to monitoring all three principal components. In this fashion, the α level could be applied to one chart rather than being split across two or three (depending on the number of levels being monitored). The problem with that approach is that each principal component detects a given type of shift most effectively, but does not detect others. For example, consider the case of monitoring three levels. If one wished to detect shifts in all three levels simultaneously, one could choose the principal component which is positive for all three levels. In so doing, one creates a chart which is very sensitive to a shift in all three levels. This principal component, however, is extremely insensitive to a shift in a single level. For this reason, it is necessary to monitor all three principal components rather than to select a single principal component for monitoring.

Table 3.15 - Principal Component Chart Correlation Comparison (N=2/2)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	92.84	89.54	87.21
0.4	71.66	68.27	60.86
0.6	50.2	46.51	38.86
0.8	34.64	30.86	24.82
1	23.01	20.45	16
1.2	15.77	13.75	10.63
1.4	10.86	9.4	7.35
1.6	7.8	6.71	5.24
1.8	5.68	4.95	3.89
2	4.32	3.77	2.98
2.2	3.34	2.95	2.37
2.4	2.67	2.39	1.93
2.6	2.21	1.99	1.65
2.8	1.87	1.71	1.44
3	1.63	1.5	1.29

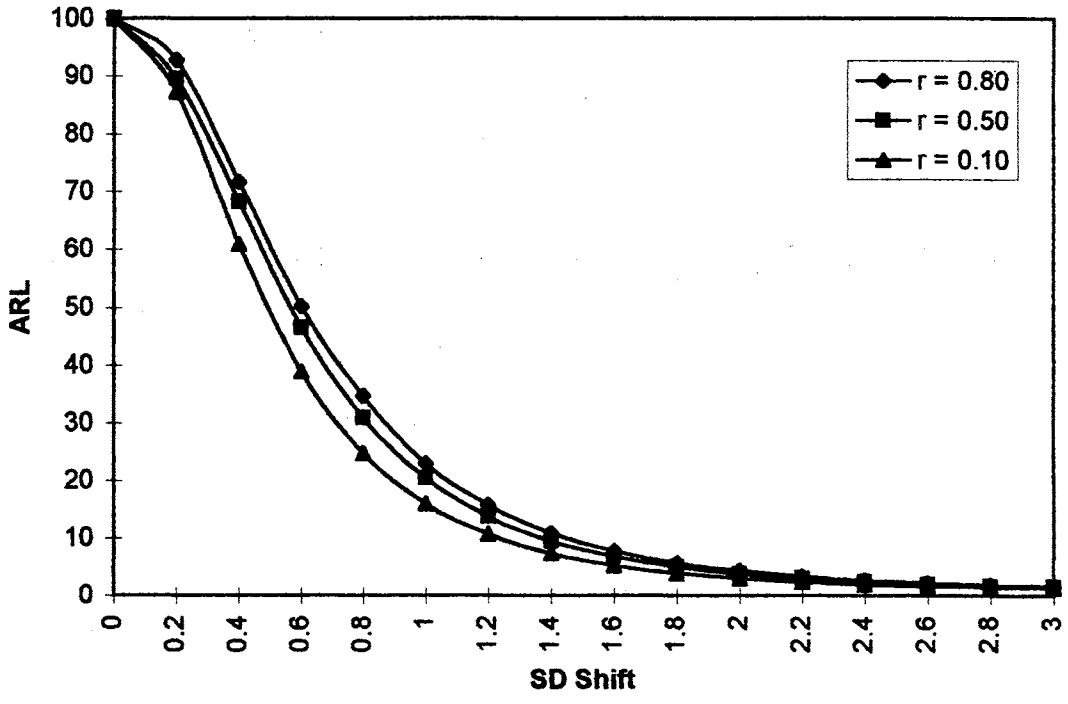

Figure 3.16 - Principal Component Chart Correlation Comparison (N=2/2)

Table 3.16 - Principal Component Chart Correlation Comparison (N=2/1)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	80.89	91.16	92.84
0.4	47.35	67.34	73.23
0.6	25.37	45.43	53.38
0.8	13.45	29.03	37.05
1	7.7	18.68	24.99
1.2	4.75	12.29	16.8
1.4	3.16	8.34	11.71
1.6	2.27	5.86	8.37
1.8	1.74	4.27	6.14
2	1.43	3.23	4.63
2.2	1.25	2.54	3.6
2.4	1.14	2.07	2.87
2.6	1.07	1.74	2.36
2.8	1.03	1.51	1.99
3	1.015	1.35	1.72

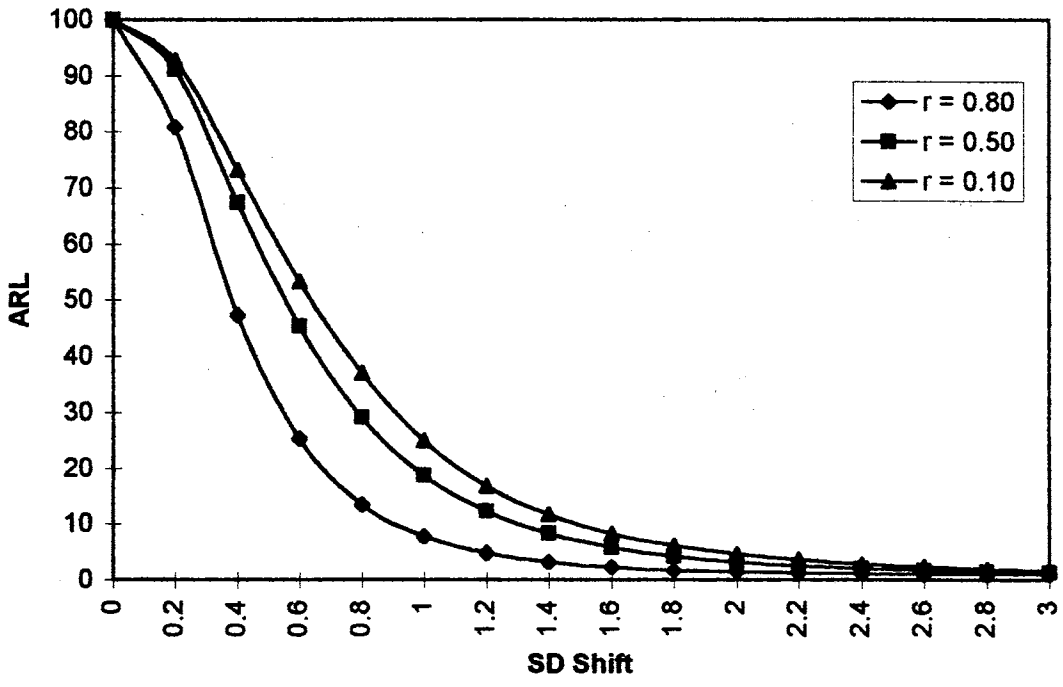

Figure 3.17 - Principal Component Chart Correlation Comparison (N=2/1)

Table 3.17 - Principal Component Chart Correlation Comparison (N=3/3)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	93.79	91.24	86.54
0.4	76.07	70.74	60.21
0.6	57.1	50.03	38.64
0.8	39.44	33.45	24.3
1	26.83	21.88	15.45
1.2	18.24	14.41	10.11
1.4	12.55	9.75	6.9
1.6	8.83	6.87	4.88
1.8	6.38	4.95	3.58
2	4.74	3.69	2.73
2.2	3.64	2.88	2.17
2.4	2.87	2.3	1.78
2.6	2.34	1.9	1.52
2.8	1.95	1.62	1.34
3	2.48	1.44	1.21

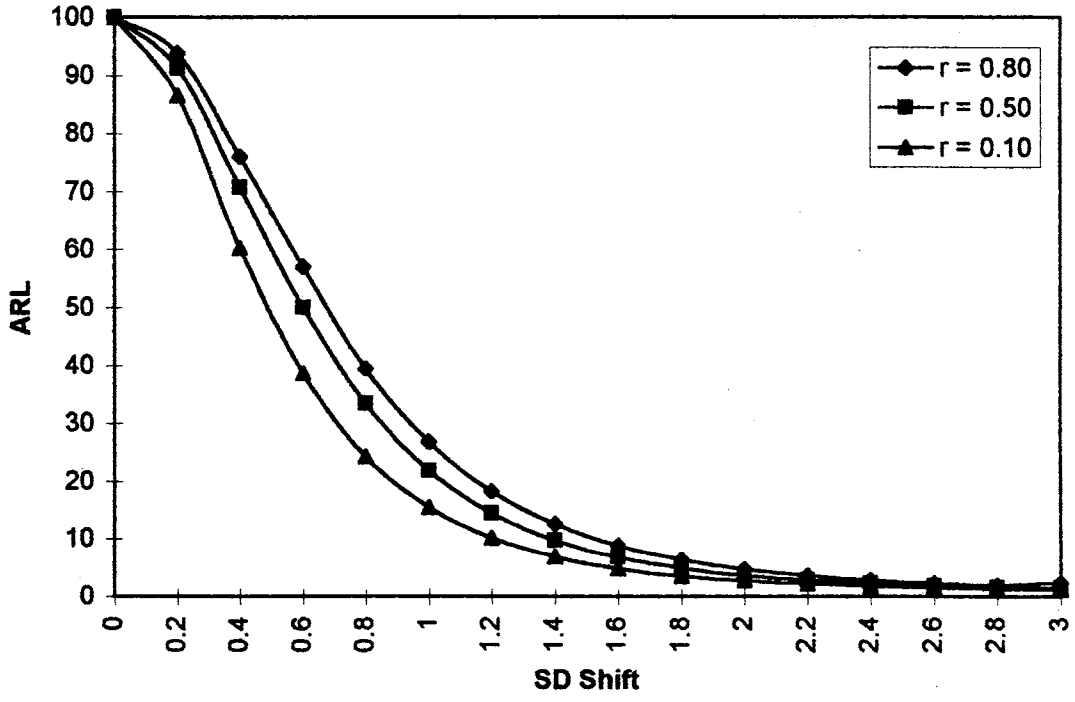

Figure 3.18 - Principal Component Chart Correlation Comparison (N=3/3)

Table 3.18 - Principal Component Chart Correlation Comparison (N=3/2)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	80.33	89.62	90.42
0.4	47.24	64.84	67.89
0.6	25.31	43.78	45.56
0.8	13.52	28.13	29.97
1	7.74	18.02	19.66
1.2	4.705	11.81	12.98
1.4	3.09	8	8.91
1.6	2.21	5.6	6.24
1.8	1.7	4.08	4.58
2	1.39	3.08	3.46
2.2	1.22	2.41	2.68
2.4	1.11	1.96	2.16
2.6	1.05	1.65	1.79
2.8	1.02	1.44	1.54
3	1.009	1.29	1.37

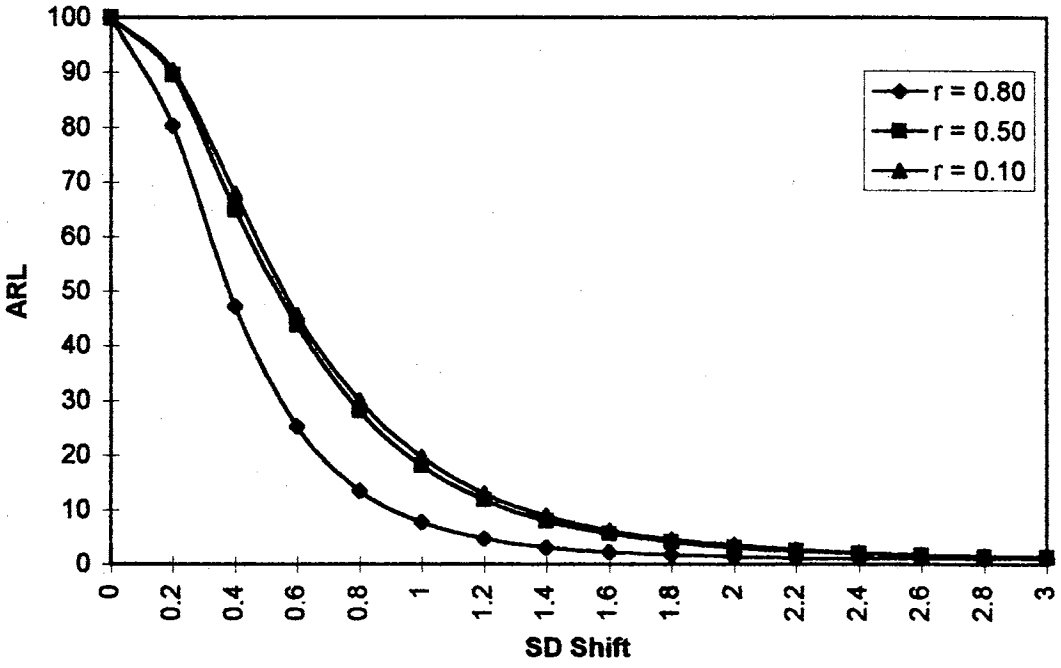

Figure 3.19 - Principal Component Chart Correlation Comparison (N=3/2)

Table 3.19 - Principal Component Chart Correlation Comparison (N=3/1)

Shift	r = 0.80	r = 0.50	r = 0.10
0	100	100	100
0.2	82.29	92.07	94.67
0.4	48.09	71.73	79.06
0.6	24.42	48.56	60.87
0.8	12.53	31.82	43.33
1	6.85	19.96	30.11
1.2	4.1	12.97	20.4
1.4	2.67	8.52	14.14
1.6	1.94	5.91	10.05
1.8	1.51	4.26	7.29
2	1.28	3.15	5.43
2.2	1.14	2.47	4.15
2.4	1.07	1.99	3.23
2.6	1.03	1.67	2.61
2.8	1.012	1.45	2.17
3	1.004	1.3	1.86

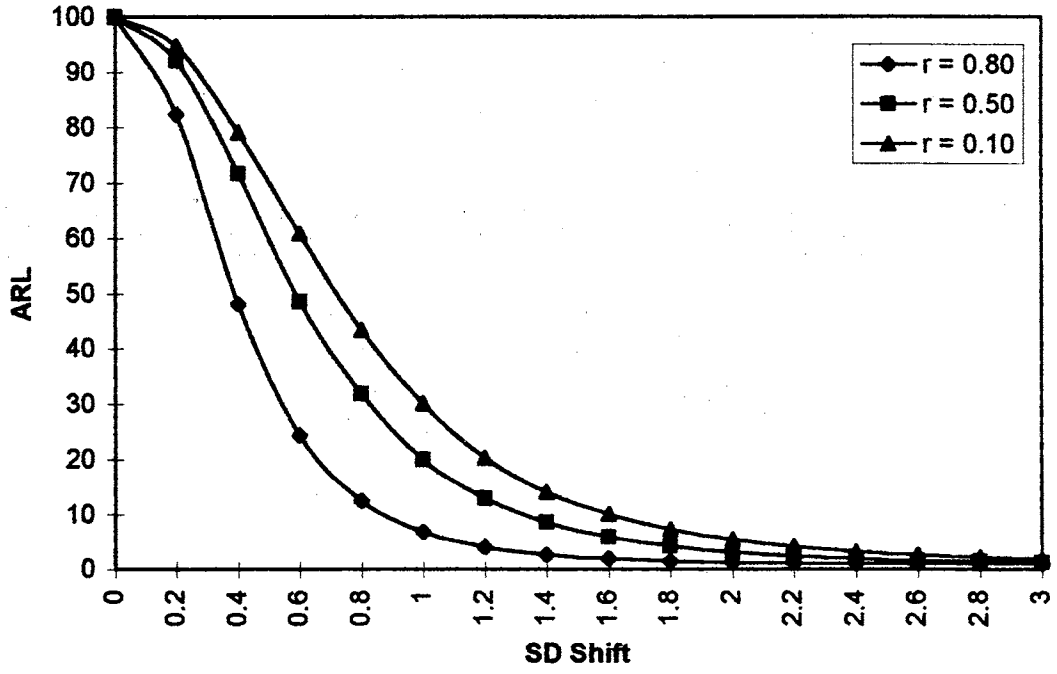

Figure 3.20 - Principal Component Chart Correlation Comparison (N=3/1)

The next comparison (shown in Table 3.20 and Figure 3.21) shows how the principal component chart performs at a correlation of 0.50 for the different varieties of shifts investigated. As with the other multivariate approaches evaluated, the correlation is set at 0.50 to be representative of the clinical application and to also allow comparison of a single instance of the principal component chart to other quality control approaches. It is clear that by monitoring all principal components, the approach provides similar performance for all the shifts considered. This is an important distinction for the principal component chart. While a clinician has a sense of security with the traditional methods for detecting all changes the measurement system might encounter, the traditional methods in fact have very different performance for different kinds of shifts. The principal component chart, however, provides the same level of error protection regardless of the type of shift encountered by the measurement system.

Method Comparisons

After developing the statistical models to analyze and evaluate each of the methods chosen for study, the real use of the modeling is to compare the methods to one another in order to draw some conclusions regarding the relative performance of the quality control methods. Through this statistical comparison, one can truly understand how these methods will perform in the field and make generalities about their use.

Table 3.20 - ARLs for the Principal Component Chart by Type of Shift

Shift	N=2/2	N=2/1	N=3/3	N=3/2	N=3/1
0	100	100	100	100	100
0.2	89.54	91.16	91.24	89.62	92.07
0.4	68.27	67.34	70.74	64.84	71.73
0.6	46.51	45.43	50.03	43.78	48.56
0.8	30.86	29.03	33.45	28.13	31.82
1	20.45	18.68	21.88	18.02	19.96
1.2	13.75	12.29	14.41	11.81	12.97
1.4	9.4	8.34	9.75	8	8.52
1.6	6.71	5.86	6.87	5.6	5.91
1.8	4.95	4.27	4.95	4.08	4.26
2	3.77	3.23	3.69	3.08	3.15
2.2	2.95	2.54	2.88	2.41	2.47
2.4	2.39	2.07	2.3	1.96	1.99
2.6	1.99	1.74	1.9	1.65	1.67
2.8	1.71	1.51	1.62	1.44	1.45
3	1.5	1.35	1.44	1.29	1.3

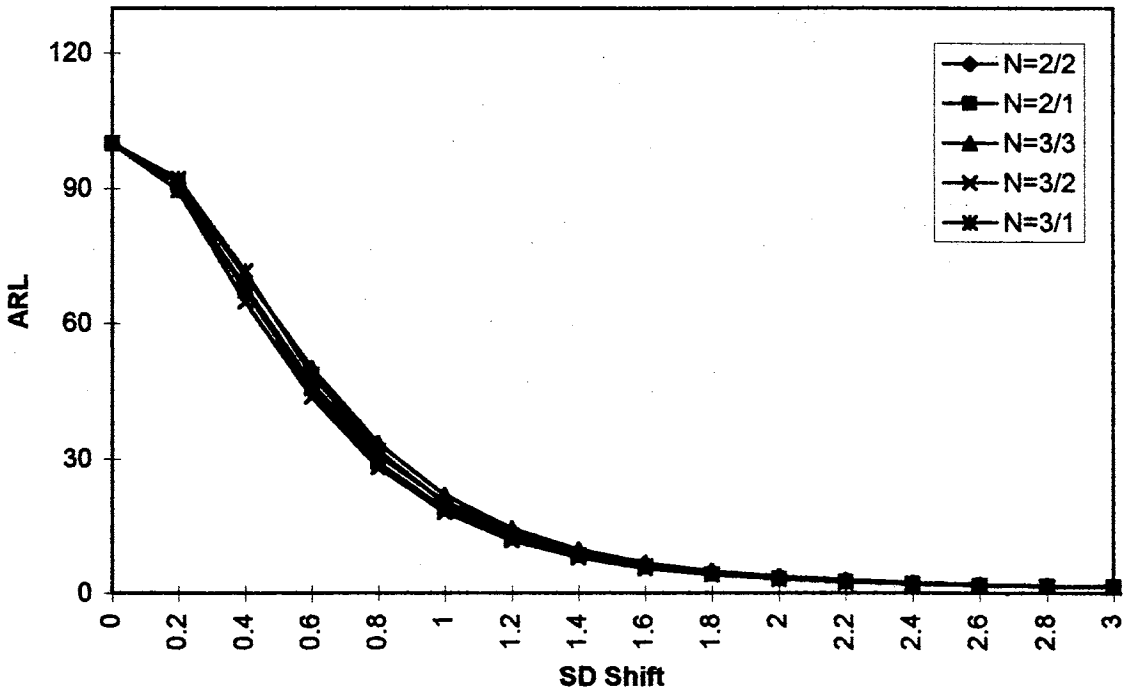

Figure 3.21 - ARLs for the Principal Component Chart by Type of Shift

Tables 3.21 through 3.25 and corresponding Figures 3.22 through 3.26 show how the quality control approaches compare for the shifts in the control levels considered in this research. Using this side by side comparison, one can determine the performance of the approaches in relation to each other.

A few notes are in order to clarify the results shown in the figures. For the χ^2 and T^2 charts, the correlation structure for the covariance matrix utilizes a correlation coefficient of 0.50 for all levels with one another. This would be considered moderate correlation and is suitable for the comparison. Also, the sample size for the T^2 chart is 20 subgroups since that is a reasonable sample size to be used in practice.

A number of conclusions can be drawn from the summary of the research. The first is that +/- 2 SD limits has a much higher rate of false rejection than any of the other methods. The ARL curves for +/- 2 SD limits are extremely flat, regardless of the type of shift observed, resulting in a high frequency of false rejection. This also results in +/- 2 SD limits being the most sensitive monitoring approach, but the high rate of false rejection compromises the positive impact of this sensitivity.

Given the negative implications of such a high false rejection rate as is the case for +/- 2 SD limits, the results show that the performance of the +/- 2 SD limits with a retest reduces the incidence of false rejection while showing competitive sensitivity. While not as sensitive as the strict application of +/- 2 SD limits, +/- 2 SD limits with a retest has the steepest slope of the approaches evaluated, starting with a longer in-control ARL than most methods and resulting in shorter ARLs for larger sized shifts. These are encouraging

Table 3.21 - Method Comparison for N=2/2

SD Shift	2 SD	2 SD/RT	W. Rules	Chi-sq	T-sq	P-C
0	11.22	125.91	83.95	100	100	100
0.2	10.3	106.02	68.25	90.91	90.91	89.54
0.4	8.19	67.15	42.14	66.67	71.43	68.27
0.6	6.11	37.32	23.87	43.48	50	46.51
0.8	4.51	20.37	13.71	28.57	34.48	30.86
1	3.4	11.54	8.59	18.52	23.81	20.45
1.2	2.63	6.93	5.81	12.35	16.39	13.75
1.4	2.11	4.45	4.08	8.55	11.63	9.4
1.6	1.75	3.07	3.09	6.02	8.4	6.71
1.8	1.51	2.27	2.45	4.42	6.17	4.95
2	1.33	1.78	1.99	3.37	4.69	3.77
2.2	1.22	1.48	1.71	2.65	3.66	2.95
2.4	1.13	1.29	1.49	2.15	2.94	2.39
2.6	1.08	1.17	1.34	1.81	2.42	1.99
2.8	1.05	1.1	1.21	1.56	2.04	1.71
3	1.03	1.05	1.14	1.39	1.76	1.5

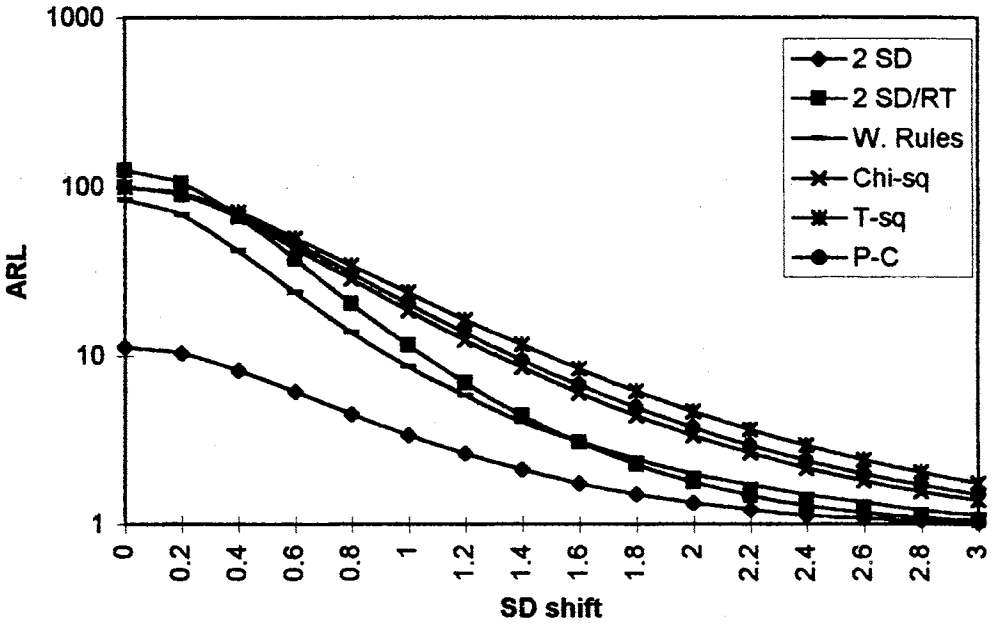

Figure 3.22 - Method Comparison for N=2/2

Table 3.22 - Method Comparison for N=2/1

SD Shift	2 SD	2 SD/RT	W. Rules	Chi-sq	T-sq	P-C
0	11.22	125.91	83.95	100	100	100
0.2	10.74	115.3	74.3	90.91	90.91	91.16
0.4	9.46	89.46	55.07	66.67	71.43	67.34
0.6	7.86	61.8	37.5	43.48	50	45.43
0.8	6.33	40.09	24.57	28.57	34.48	29.03
1	5.04	25.43	16.01	18.52	23.81	18.68
1.2	4.02	16.19	11.04	12.35	16.39	12.29
1.4	3.25	10.56	7.82	8.55	11.63	8.34
1.6	2.67	7.13	5.82	6.02	8.4	5.86
1.8	2.24	5	4.44	4.42	6.17	4.27
2	1.91	3.66	3.53	3.37	4.69	3.23
2.2	1.67	2.79	2.87	2.65	3.66	2.54
2.4	1.49	2.22	2.42	2.15	2.94	2.07
2.6	1.35	1.84	2.09	1.81	2.42	1.74
2.8	1.25	1.57	1.84	1.56	2.04	1.51
3	1.18	1.39	1.65	1.39	1.76	1.35

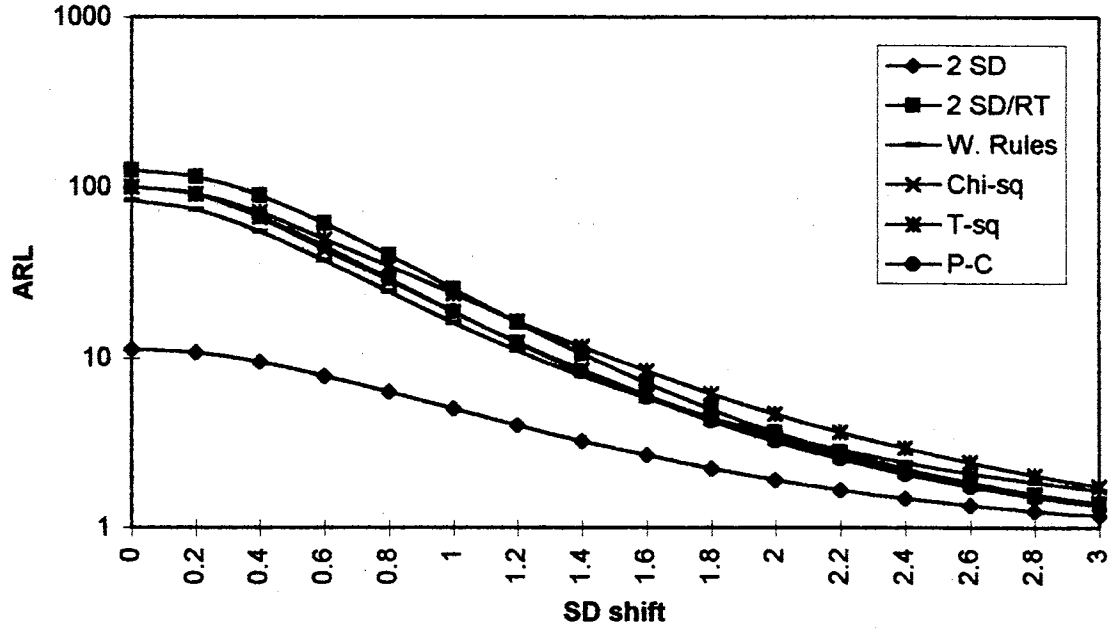

Figure 3.23 - Method Comparison for N=2/1

Table 3.23 - Method Comparison for N=3/3

SD Shift	2 SD	2 SD/RT	W. Rules	Chi-sq	T-sq	P-C
0	7.65	58.58	33.72	100	100	100
0.2	7.04	49.54	27.79	90.91	90.91	91.24
0.4	5.64	31.8	17.98	66.67	76.92	70.74
0.6	4.25	18.08	10.92	47.62	55.56	50.03
0.8	3.19	10.2	6.75	31.25	40	33.45
1	2.46	6.03	4.54	20.41	27.78	21.88
1.2	1.95	3.82	3.17	13.51	19.23	14.41
1.4	1.62	2.62	2.37	9.26	13.4	9.75
1.6	1.39	1.94	1.88	6.45	9.8	6.87
1.8	1.24	1.54	1.54	4.67	7.25	4.95
2	1.14	1.31	1.33	3.5	5.43	3.69
2.2	1.08	1.17	1.2	2.72	4.2	2.88
2.4	1.04	1.087	1.11	2.18	3.32	2.3
2.6	1.02	1.043	1.05	1.82	2.7	1.9
2.8	1.01	1.019	1.02	1.56	2.25	1.62
3	1.004	1.008	1.01	1.38	1.92	1.44

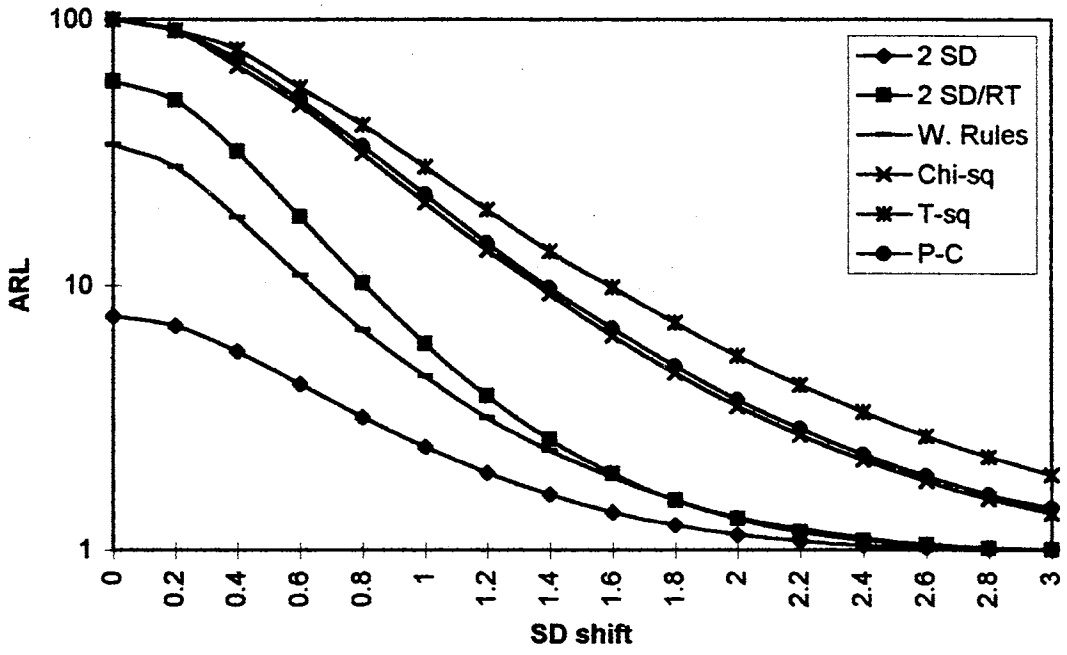

Figure 3.24 - Method Comparison for N=3/3

Table 3.24 - Method Comparison for N=3/2

SD Shift	2 SD	2 SD/RT	W. Rules	Chi-sq	T-sq	P-C
0	7.65	58.58	33.72	100	100	100
0.2	7.23	52.29	29.8	90.91	90.91	89.62
0.4	6.17	38.07	21.89	62.5	66.67	64.84
0.6	4.95	24.55	14.84	40	47.62	43.78
0.8	3.89	15.14	9.95	23.81	32.26	28.13
1	3.06	9.38	6.88	14.93	20.83	18.02
1.2	2.45	6	4.92	9.43	14.08	11.81
1.4	2.01	4.04	3.68	6.29	9.62	8
1.6	1.69	2.87	2.87	4.35	6.76	5.6
1.8	1.47	2.16	2.3	3.15	4.9	4.08
2	1.31	1.72	1.91	2.4	3.68	3.08
2.2	1.2	1.45	1.63	1.91	2.87	2.41
2.4	1.13	1.27	1.43	1.59	2.3	1.96
2.6	1.08	1.16	1.28	1.38	1.91	1.65
2.8	1.04	1.09	1.18	1.24	1.64	1.44
3	1.02	1.05	1.12	1.15	1.44	1.29

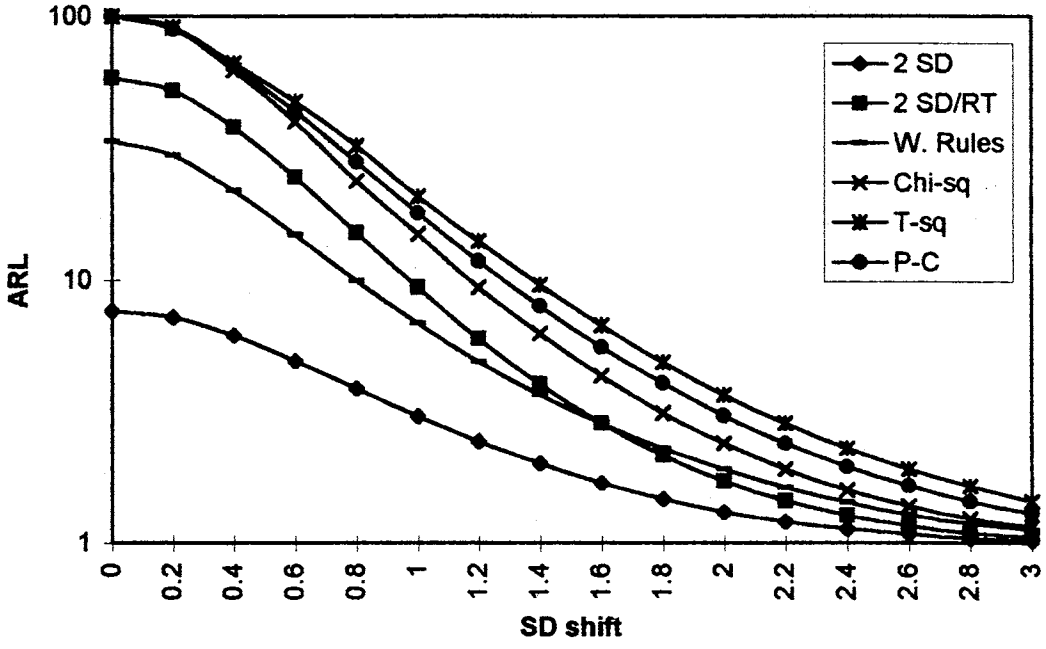

Figure 3.25 - Method Comparison for N=3/2

Table 3.25 - Method Comparison for N=3/1

SD Shift	2 SD	2 SD/RT	W. Rules	Chi-sq	T-sq	P-C
0	7.65	58.58	33.72	100	100	100
0.2	7.44	55.29	31.72	90.91	90.91	92.07
0.4	6.83	46.59	27.17	66.67	76.92	71.73
0.6	5.99	35.86	21.1	47.62	55.56	48.56
0.8	5.09	25.94	15.68	31.25	40	31.82
1	4.26	18.13	11.62	20.41	27.78	19.96
1.2	3.54	12.51	8.63	13.51	19.23	12.97
1.4	2.95	8.69	6.43	9.26	13.7	8.52
1.6	2.48	6.16	5	6.45	9.8	5.91
1.8	2.12	4.48	3.93	4.67	7.25	4.26
2	1.84	3.37	3.19	3.5	5.43	3.15
2.2	1.62	2.63	2.66	2.72	4.2	2.47
2.4	1.46	2.12	2.27	2.18	3.32	1.99
2.6	1.33	1.78	1.97	1.82	2.7	1.67
2.8	1.24	1.54	1.74	1.56	2.25	1.45
3	1.17	1.37	1.57	1.38	1.92	1.3

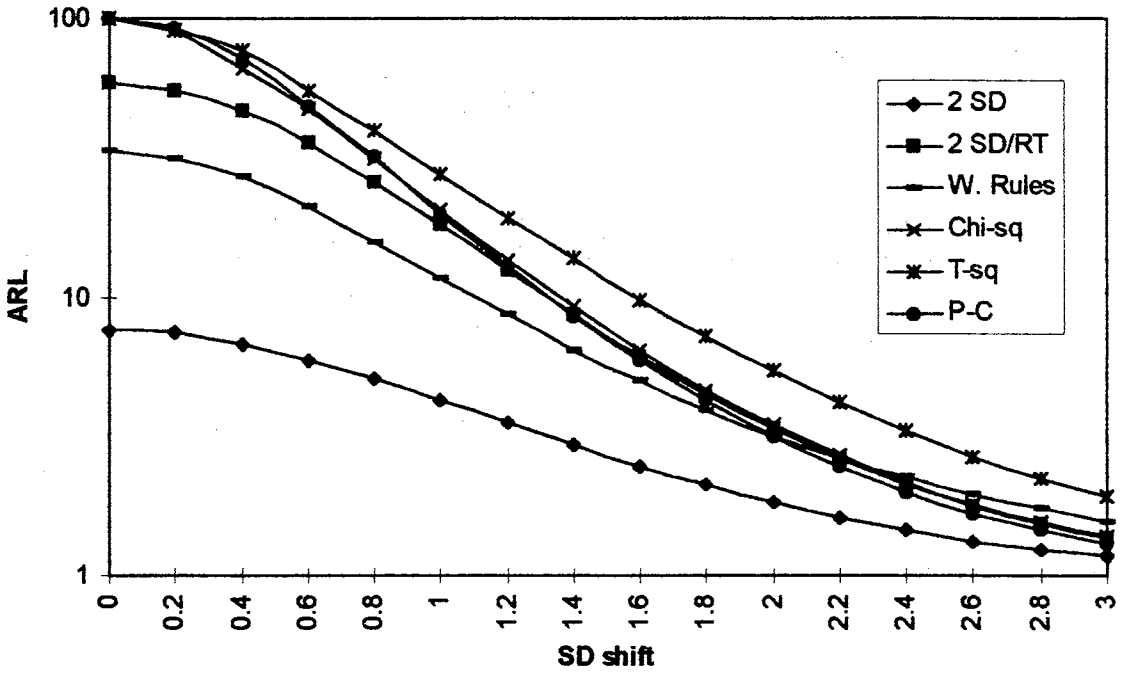

Figure 3.26 - Method Comparison for N=3/1

results from the clinical perspective as +/- 2 SD limits with a retest is a simple and straightforward method for monitoring quality control.

The Westgard Multirule Procedure shows performance similar to +/- 2 SD limits with a retest, but the in-control ARL for the Westgard Multirule Procedure is generally lower than for +/- 2 SD limits with a retest. Given the similar sensitivity to detecting larger shifts along with the fact that it is much easier to implement, many clinicians would prefer the use of +/- 2 SD limits with a retest as opposed to the implementation of the Westgard Multirule procedure.

As for the multivariate approaches, they appear to perform very similarly to one another. As expected, the T^2 chart is the least sensitive for the same in-control ARL since it assumes an unknown covariance matrix and uses only 20 subgroups. Its performance, however, is not drastically worse than the χ^2 chart, the best case given that it assumes a known covariance. The principal component chart and χ^2 chart have very similar performance across the different shifts considered. All of the multivariate approaches, though, are hampered from a sensitivity perspective by their in-control ARL of 100. By increasing the α level used in the design of the multivariate approaches, improved sensitivity to true shifts can be improved. The +/- 2 SD limits with a retest approach, however, has a longer in-control ARL than the multivariate methods with better sensitivity for the cases where two levels of control material are monitored. This does not, however, hold for three levels of control material as the in-control ARL for +/- 2 SD limits with a retest drops to 58.58 while the multivariate approaches retain in-control ARLs of 100.

CHAPTER IV

AN ECONOMIC MODEL FOR THE CLINICAL LABORATORY

Economic Model Description

While relatively little work has been pursued in the clinical literature regarding cost models for monitoring quality control, there is a long history of cost modeling in the industrial quality control literature. The model used in this research capitalizes on this history, building primarily on the work by Duncan (1956) and Lorenzen and Vance (1986). Basically, the model considers a full cycle to consist of starting in a state of statistical control (SOSC), eventually experiencing a shift in the measurement process becoming out-of-control (OOC), and then returning to SOSC to complete the cycle. The assumptions are that the process starts in control, finishes in control, and each repair attempt for a true shift is always successful.

The economic model employed for this research determines the total cost per unit time. The costs considered in the model include costs for sampling and testing control materials, costs for operating in-control, costs for operating out-of-control, and costs for downtime when the testing system is not in operation. The model allows for three different assignable causes to affect the measurement system with Exponential failure rates used for the assignable causes. These three assignable causes allow the model to tie different costs to different sizes of shift. The model assumes that these three shifts are applied to all levels affected by shifts. For example, if shifts of 1.0 SD, 2.0 SD, and 3.0 SD are applied to the N=2/1 case, this means that one level of control material is affected

by these three shifts in the economic model while the other level remains centered and unaffected. Another model assumption is that the testing process shuts down to look for causes when the quality control method signals an out-of-control condition. Figure 4.1 summarizes the parameters required in the model.

Figure 4.1 - Cost Model Parameters

λ = Vector of Exponential failure rates ($1/\lambda_i$ is the average time between shifts in the measurement process for the ith assignable cause).

δ = Vector of possible shifts in the measurement process.

$\mathbf{ARL}_{OOC}$ = A vector of average run lengths corresponding to δ for the quality control method when the process is out-of-control.

$\mathbf{SRT}_{OOC}$ = A vector of the average search and repair times in hours for shifts in the measurement process defined by δ.

$\mathbf{p}$ = A vector of the probabilities for the shifts in δ.

ARL_{SOSC} = The average run length of the quality control method when the process is in-control (i.e., experiencing no shift).

SRT_{SOSC} = The average search and repair time in hours for a false alarm.

h = sampling interval (the time between quality control samples in hours).

N = The number of control samples tested at each QC timepoint.

C_1 = The cost per unit time for operating in-control.

$\mathbf{C_2}$ = A vector of costs per unit time for operating out-of-control according to δ.

C_3 = The cost per unit time for downtime or not generating patient results.

C_4 = The cost per control value observation (including labor and materials).

Before continuing with the description of the cost model, some further discussion of these parameters is warranted. The first notable point concerns the use of a vector (δ) of possible shifts in the measurement process. A vector is utilized here to allow for varying shifts in the measurement process ranging from a very small shift up to a large shift which would be considered clinically significant. It should be noted that the search and repair times for given shifts ($\mathbf{SRT_{OOC}}$) and the cost to the process after experiencing a shift ($\mathbf{C_2}$) are also vectors which are functions of the size of the shifts in δ.

For this multiple assignable cause model, the overall failure rate can be taken as the sum of the failure rates. Therefore, $\lambda=\lambda_1+\lambda_2+\lambda_3$ and is the overall failure rate for the model. The individual failure rates define **p** as $p_i=\lambda_i/\lambda$. Again, these failure rates are applied only to levels which are being shifted in the economic model. For example, the case of N=3/3 would mean that all three levels of control material are affected by shifts, and that the overall failure rate for each level is λ.

Also, additional detail about the cost factors (C_1 through C_3) is necessary in order to fully understand the model. Cost C_1 is essentially a function of testing volume. It includes costs for labor, testing consumables, equipment depreciation, overhead, etc. It can be considered the cost of doing business when things are performing as expected. The cost factor $\mathbf{C_2}$ includes the basic costs related to C_1, but there is an additional cost component attached. This additional cost relates to reporting patient results which are actually shifted from their true mean. Failing to detect large shifts and operating out-of-control can result in governing bodies shutting the laboratory down, so this additional cost is very large for large shifts in the measurement process. Finally, C_3 is a cost for not

generating results when they are needed. This cost reflects responsiveness of the laboratory, and this cost may be extremely high for situations when a physician needs rapid turn around of patient results but the laboratory is down or not operating. The cost assigned for C_3 will include the basic overhead costs incurred in C_1, but it will include additional costs for having idle resources on hand and may include overtime operations for making up lost throughput.

Given these model parameters, a total cost equation can be determined. Figure 4.2 graphically illustrates a complete cycle for the economic model which leads to the total cost equation.

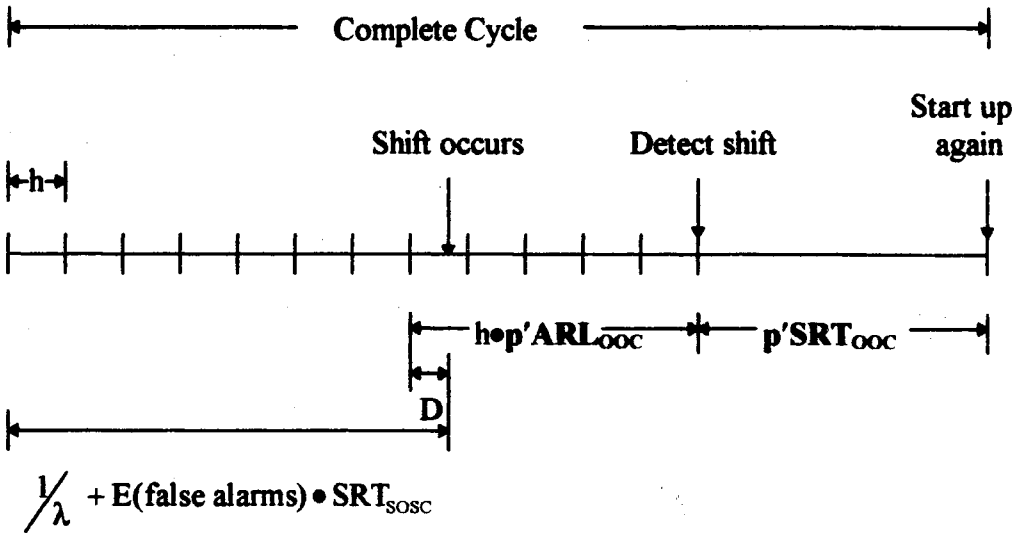

Figure 4.2 - Illustration of a Complete Cycle for the Cost Model

In the preceding graphic, reference is made to the expected number of false alarms while the process is actually in-control, referred to as E(false alarms). Lorenzen and Vance

(1986) showed that the following expression can be used for the expected number of false alarms when the economic model assumes the process shuts down to look for changes in the measurement process.

$$\text{E(false alarms)} = \frac{\dfrac{e^{-\lambda h}}{\left(1-e^{-\lambda h}\right)}}{ARL_{SOSC}}$$

Figure 4.2 also refers to D, the amount of time between the last testing point which is in-control to the point in the sampling interval when the process actually shifts. Duncan (1956) developed an expression for the expected time for the shift to occur in an interval which can be used to find the following equation for D.

$$D = \left(\frac{1-(1+\lambda h)e^{-\lambda h}}{\lambda\left(1-e^{-\lambda h}\right)}\right)$$

Next, the pieces of the cycle can be put together to determine the average length of a total cycle (TCT). This is the average amount of time starting from an in-control condition, experiencing a shift to an out-of-control condition, and being restored to an in-control state again.

$$TCT = \frac{1}{\lambda} + \left(\frac{\dfrac{e^{-\lambda h}}{\left(1-e^{-\lambda h}\right)}}{ARL_{SOSC}}\right) \bullet SRT_{SOSC} + h \bullet \mathbf{p'ARL_{OOC}} - \left(\frac{1-(1+\lambda h)e^{-\lambda h}}{\lambda\left(1-e^{-\lambda h}\right)}\right) + \mathbf{p'SRT_{OOC}}$$

After finding an expression for the total cycle time, the percentage of total cycle time operating in a state of statistical control (%SOSC) and the percentage of total cycle time operating out-of-control (%OOC) can both be determined as ratios to the total cycle time.

$$\%SOSC = \frac{1/\lambda}{TCT}$$

$$\%OOC = \frac{h \bullet p'ARL_{ooc} - \frac{1-(1+\lambda h)e^{-\lambda h}}{\lambda(1-e^{-\lambda h})}}{TCT}$$

In the same manner, the percentage of total cycle time looking for false alarms (%FR) and the percentage of total cycle time looking for true shifts (%TR) in the measurement process can also be determined. Since both result in shut down of the system, these two percentages can be combined to determine the total percentage of downtime (%DT) for the testing system.

$$\%FR = \frac{\left(\frac{e^{-\lambda h}}{(1-e^{-\lambda h})}\right)}{\frac{ARL_{SOSC}}{TCT}} \bullet SRT_{SOSC}$$

$$\%TR = \frac{p'SRT_{ooc}}{TCT}$$

$$\%DT = \frac{\left(\frac{e^{-\lambda h}}{(1-e^{-\lambda h})}\right)}{\frac{ARL_{SOSC}}{TCT}} \bullet SRT_{OOC} + p'SRT_{ooc}$$

The final piece to be determined for the cost model is the cost for sampling and testing control materials (SC). This cost can be determined as follows:

$$SC = \frac{N' \bullet C_4}{h}$$

Now, all of these costs can be combined into a total cost equation which considers all the costs incurred during the operation of the testing system.

Total Cost Per Unit Time = $\%SOSC \bullet C_1 + \%OOC \bullet p'C_2 + \%DT \bullet C_3 + SC$

At this point, the entire model is a single cost equation which can be used to compare the various quality control approaches.

The above model must be slightly altered for the +/- 2 SD limit with a retest method. The problem arises in the determination of the sampling cost. For all of the other models, the amount of testing is the same for each QC timepoint, but this is not the case for +/- 2 SD limits with a retest. For +/- 2 SD limits with a retest, the number of observations per QC timepoint will be doubled when a false or true rejection occurs. Therefore, the value used for N' in the equation for SC must be altered to reflect the fact that the expected sample size will be greater than the assigned sample size. If one assumes that the amount of time to do a retest is negligible (the same as the current assumption regarding the amount of time to get an initial result), then the only parameter from the cost model which must be modified is N' in order to estimate the correct sampling cost. The following develops the approach for calculating N' for the +/- 2 SD limits with a retest case.

Let N' = the expected number of samples per QC timepoint

n = the original sample size (either 2 or 3 for this application)

2n = the sample size upon retest

λ_i = the failure rate for the ith shift (i =1,2 or 3)

$\lambda_4 = \dfrac{1}{(ARL_{SOSC})(h)}$ where ARL_{SOSC} is the ARL for strict +/- 2 SD limits

$\lambda_{tot} = \sum_{i=1}^{4} \lambda_i$

$P(in_i)$ = the probability of shift i being within +/- 2 SD limits

$P(out_i)$ = the probability of shift i being outside +/- 2 SD limits

$$\text{Then, } N' = \sum_{i=1}^{4} \frac{\lambda_i}{\lambda_{tot}} \left[n \bullet P(in_i) + 2n \bullet P(out_i) \right]$$

Using this expression for the expected number of samples per QC timepoint, the same model as described above can be used for assessing +/- 2 SD limits with a retest.

The cost model developed above contains a number of attractive features. The first is that there is a separation between the statistical modeling of a quality control approach and the cost model. The required inputs to the economic model from the statistical model are only the ARLs for being in-control and the ARLs for being out-of-control. While these ARLs will vary from method to method and the approach for calculating these ARLs will also vary from method to method, the only statistical requirements of the economic model are these ARLs.

Another attractive aspect of this model is that it easily breaks down the proportion of time that the system is in the various operating states. The amount of time spent looking for non-existent problems can be broken out by method, allowing clinicians to better understand the implications of their selection of a quality control approach.

Cost Model Validation

Following the development of the theoretical cost model, it is still of interest to validate the model through another approach. The model makes a number of assumptions

in its development, and the use of another approach for validation provides an additional degree of comfort with the theoretical model.

The first aspect to explore is the memoryless property of the Exponential failure rate employed in the model. The model assumes that the testing process is shut down during false rejections, so the theoretical model defines the amount of time until a true shift (tts) as being:

$$tts = 1/\lambda + E(\text{false alarms}) \bullet SRT_{SOSC}$$

One way to think of this development is that at time 0, an unknown time to a true failure is generated. Before this time to failure is reached, false rejections may occur. The model assumes that the time to failure takes up where it left off once a false rejection has been completed. An alternative approach to the failure mechanism is that a completely new failure is generated each time a false rejection occurs. In this instance, one would be assuming that action taken during a false rejection can affect the underlying failure rate, and this is undoubtedly a valid viewpoint.

With this in mind (i.e., that failures can regenerate themselves during false rejections), one can take another approach to the development of the expected time until a true shift without considering time for false rejection. If one assumes the time to false rejection is also Exponentially distributed, then one can define the failure rate due to false rejections as (λ_{fr}).

$$\lambda_{fr} = \frac{1}{(ARL_{SOSC})(h)}$$

Then, one can view the situation as four competing failure rates: three true failures competing with a failure due to a false rejection. This would make the overall failure rate $\lambda_{tot}=\lambda+\lambda_{fr}$. This result can be used to find the probabilities for true and false rejections, then to determine the expected time until a true shift not considering shutdown time for false rejections.

P(true reject) = $f = \lambda/\lambda_{tot}$

P(false reject) = $1-f = \lambda_{fr}/\lambda_{tot}$

T=time until a true shift (not considering shutdown time for false rejections)

$$E(T) = \frac{1}{\lambda_{tot}}f + \left(\frac{1}{\lambda_{tot}}+\frac{1}{\lambda_{tot}}\right)(1-f)f + \left(\frac{1}{\lambda_{tot}}+\frac{1}{\lambda_{tot}}+\frac{1}{\lambda_{tot}}\right)(1-f)^2 f + \cdots$$

$$E(T) = \frac{1}{\lambda_{tot}}\sum_{n=1}^{\infty} n(1-f)^{n-1} f = \frac{1}{\lambda_{tot}} \bullet \frac{1}{f} = \frac{1}{\lambda}$$

Therefore, it appears that the memoryless property of the Exponential acts such that it does not matter how one views the generation of the failure.

To explore this issue further and to validate the theoretical model, two computer simulation programs of the cost model are employed. The first program (found in Appendix F and referred to as CostA) simulates the cost model by generating a single time to failure at the beginning of the cycle. Then, it suspends the failure during a false rejection, but then continues the time to this failure following the false rejection. The second simulation program (found in Appendix F and referred to as CostB) regenerates a failure each time a false rejection is realized. In this fashion, both assumptions can be tested against the theoretical model for validation purposes.

Both programs are coded in Turbo Pascal version 6.0, and the same uniform random number generator from Marse and Roberts (1983) is used as in the simulation programs concerning the Westgard Multirule Procedure. In addition to generating Exponential times to failures and geometric times to false rejections, the simulations also generate random search and repair times for all true rejections as well as false rejections. The repair times are modeled as uniform random variables in the simulation programs.

Figure 4.3 summarizes the input parameters used to compare the theoretical results with the simulation results and Table 4.1 shows the results of the comparisons. Table 4.1 shows the 95% confidence intervals for 5 simulation runs, each consisting of 5,000 realizations (where a realization is a completion of the cost cycle). The measures chosen for comparison do not include cost figures as the true validation of the model concerns the proportion of time the system is in a given state. It is clear from these results that the memoryless property is in effect such that the model is not affected by the assumption regarding failure generation (i.e., initially generated or regenerated with each false rejection). Additionally, the theoretical results match those of the simulation models, indicating that the theoretical model is valid. The only exception is with the %OOC measure for program CostB, but this can be considered as an α error since 95% limits were chosen and 18 confidence intervals were constructed.

Given the additional simulation and validation work, it is clear that the theoretical cost model is indeed appropriate. Therefore, the research effort can use this theoretical model for making comparisons among the various quality control approaches.

Figure 4.3 - Simulation and Theoretical Results Comparison (Input Parameters)

$$\lambda = \begin{bmatrix} 0.001 \\ 0.003 \\ 0.005 \end{bmatrix} \qquad ARL_{OOC} = \begin{bmatrix} 30 \\ 40 \\ 50 \end{bmatrix} \qquad SRT_{OOC} = \begin{bmatrix} (8,12) \\ (20,30) \\ (45,55) \end{bmatrix}$$

$SRT_{SOSC} = (25,35)$ $\qquad ARL_{SOSC} = 2$ $\qquad h=5$

Table 4.1 - Simulation and Theoretical Results Comparison

Measure	Theoretical	CostA	CostB
Avg. time to shift	437	(429.9, 445.7)	(436.56, 451.93)
E(false rejects)	10.86	(10.68, 11.08)	(10.85, 11.23)
D	2.48	(2.45, 2.48)	(2.47, 2.49)
TCT	693.96	(691.71, 704.98)	(691.00, 711.95)
%SOSC	0.160	(0.158, 0.161)	(0.160, 0.161)
%OOC	0.317	(0.313, 0.326)	(0.310, 0.316)
%FR	0.47	(0.462, 0.472)	(0.470, 0.475)
%TR	0.054	(0.053, 0.054)	(0.053, 0.054)
%DT	0.523	(0.516, 0.525)	(0.523, 0.528)

Method Comparisons Using the Economic Model

After formulating and validating the theoretical cost model, the next step is to use the model to compare quality control approaches on a cost basis. While the statistical performance of the methods cannot be ignored, the ultimate question regarding their relative performance must be answered on a cost basis.

A major consideration when comparing the methods from a cost perspective, however, is the selection of the cost parameters chosen for the comparisons. The parameter values used in this research are shown in Figure 4.4. Note that the sampling

interval (h) and the sample size (N') are not mentioned in Figure 4.4 as the research will vary these values in order to see their impact on the cost model.

The selection of the cost parameters in Figure 4.4 incorporates input from people with clinical backgrounds, but there would undoubtedly be disputes among other clinicians regarding the values selected for the parameters. Even so, the parameter settings used in this research are very representative of real world costs. The purpose of this modeling is to determine the relative costs of the examined methods for this particular set of parameters, so any conclusions drawn will apply specifically to this set of parameters.

Some discussion about these parameters is, however, warranted. Three different shifts are considered with respective magnitudes of 1 SD, 2 SD and 3 SD. These shifts represent small, intermediate, and large shifts respectively. The size of the shifts then dictate the selection of λ, $\mathbf{SRT}_{OOC}$, and $\mathbf{C}_2$ in turn. The values for λ used for the model translate into the frequency of a 1 SD shift being every 500 hours, the frequency of a 2 SD shift is every 1000 hours, and 3 SD shifts occur every 2000 hours. This means that the measurement system is relatively stable, and that small shifts occur more frequently than

Figure 4.4 - Cost Parameters for Method Comparisons

$$\delta = \begin{bmatrix} 1\ SD \\ 2\ SD \\ 3\ SD \end{bmatrix} \qquad \lambda = \begin{bmatrix} 0.002 \\ 0.001 \\ 0.0005 \end{bmatrix} \qquad \mathbf{SRT}_{OOC} = \begin{bmatrix} 4\ hrs \\ 3\ hrs \\ 2\ hrs \end{bmatrix} \qquad SRT_{SOSC}=5\ hrs$$

$$C_1=\$200/hr \qquad \mathbf{C}_2 = \begin{bmatrix} \$220/hr \\ \$800/hr \\ \$2000/hr \end{bmatrix} \qquad C_3=\$1600/hr \qquad C_4=\$10/sample$$

large shifts (which is the case in the clinical laboratory). Also tied to λ, the SRT times reflect the fact that smaller shifts will be more difficult to find than larger shifts, and that the amount of time to repair true shifts is relatively small compared to the time between shifts occurring. Additionally, the search and repair time for a false rejection is larger than the time for true rejections as the search for a phantom problem can be very time consuming. The costs assigned to each shift (C_2) also reflect the magnitude of the shift. The standard cost of doing business (C_1) is set at \$200/hr, and the costs for C_2 are relative to this baseline. The cost for a 1 SD shift is \$220/hr, or a 10% increase in cost while the costs for shifting to 2 SD and 3 SD are fourfold and tenfold increases, respectively. These costs reflect the fact that small shifts will have less impact on patient results than will larger shifts. Again, the emphasis is on the costs relative to one another for the identified shifts. The same thought process went into the selection of C_3, the cost for downtime, as being eight times the cost of regular operation. The remaining cost parameter, C_4 ,is a standard cost per sample that is reflective of many diagnostic tests.

In order to facilitate the cost comparisons made in this research, a Turbo Pascal program is used to perform the calculations employed in the theoretical cost model. The code for this program is included in Appendix G, and all cost model results reported here are generated using this program.

Given the selection of the cost model parameters (shown in Figure 4.4), methods can be compared from a cost perspective for these parameters. Comparisons for this research are made for the six quality control approaches considered. In the spirit of traditional economic modeling of control charts, the parameters varied in this research are

the traditional n (i.e., the cases N=2/2, N=2/1, etc.,), h (the sampling interval which is varied between 8 and 24 hours in this research), and k (typically a control chart limit, but dictated by the quality control approach in this research). The cost model results are shown in Tables 4.2 through 4.11. The tables show the proportion of the time the system is in the various states including the percent time in-control (%SOSC), the percent time operating out-of-control (%OOC), the percent time investigating false rejections (%FR), the percent time investigating true rejections (%TR), the total percent of down time (%DT), and finally the total cost per unit time for the system (Cost).

To better understand the interpretation of the shifts for the cost modeling, N=3/2 (for example) means that two of three control levels monitored are shifted. They are both shifted according to δ, so the two levels being shifted are each shifted by δ in the cost model. The third level, however, remains centered on target.

The results exhibited in Tables 4.2 through 4.11 lead to some interesting conclusions. First, note that each pair of tables (i.e., Table 4.2 and 4.3, Table 4.4 and 4.5, etc.,) has the same shift and that the only difference between the pairs is the sampling interval, h. Since the same shift is common to the pairs of tables, one can make comparisons across pairs of tables to select the least cost alternative. When comparing across sampling interval for the same type of shift, the Westgard Multirule Procedure with a sampling interval of 8 hours in most instances has the least cost per unit time of the methods considered. For N=2/1, the Westgard Multirule Procedure is not the least cost alternative, but it is very close to the least cost alternative, +/- 2 SD limits with a sampling interval of 8 hours. For N=3/3, +/- 2 SD limits with a retest narrowly edges out

Table 4.2 - Cost Model Method Comparison for N=2/2, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	89.1	4.9	4.9	1.1	6.0	307.65
2 SD RT	82.9	15.7	0.4	1.0	1.4	291.61
W.Rules	86.0	12.4	0.6	1.0	1.7	280.24
Chi-Sq.	75.0	23.6	0.5	0.9	1.4	325.51
T^2	69.9	28.8	0.4	0.8	1.3	346.95
P.C.	73.1	25.6	0.5	0.9	1.3	333.52

Table 4.3 - Cost Model Method Comparison for N=2/2, h=24

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	83.6	13.9	1.5	1.0	2.5	296.82
2 SD RT	63.2	35.9	0.1	0.8	0.9	371.10
W.Rules	69.1	29.9	0.2	0.8	1.0	346.24
Chi-Sq.	51.0	48.2	0.1	0.6	0.7	423.10
T^2	44.5	54.9	0.1	0.5	0.6	451.18
P.C.	48.5	50.8	0.1	0.6	0.7	434.03

Table 4.4 - Cost Model Method Comparison for N=2/1, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	86.7	7.5	4.8	1.0	5.8	316.79
2 SD RT	69.2	29.6	0.3	0.8	1.2	349.56
W.Rules	77.1	21.4	0.6	0.9	1.5	317.42
Chi-Sq.	75.0	23.6	0.5	0.9	1.4	325.51
T^2	69.9	28.8	0.4	0.8	1.3	346.95
P.C.	74.9	23.7	0.5	0.9	1.4	325.82

Table 4.5 - Cost Model Method Comparison for N=2/1, h=24

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	77.5	20.2	1.4	0.9	2.3	322.00
2 SD RT	43.5	55.9	0.1	0.5	0.6	455.18
W.Rules	54.2	45.0	0.1	0.7	0.8	409.95
Chi-Sq.	51.0	48.2	0.1	0.6	0.7	423.10
T^2	44.5	54.9	0.1	0.5	0.6	451.18
P.C.	50.9	48.3	0.1	0.6	0.7	423.53

Table 4.6 - Cost Model Method Comparison for N=3/3, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	88.4	3.4	7.1	1.1	8.2	333.31
2 SD RT	89.3	8.7	0.9	1.1	2.0	270.87
W.Rules	90.6	6.6	1.7	1.1	2.7	271.41
Chi-Sq.	73.3	25.4	0.5	0.9	1.3	334.02
T^2	66.7	32.1	0.4	0.8	1.2	362.01
P.C.	72.0	26.7	0.4	0.9	1.3	339.66

Table 4.7 - Cost Model Method Comparison for N=3/3, h=24

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	86.6	10.1	2.3	1.0	3.3	291.72
2 SD RT	76.5	22.3	0.3	0.9	1.2	316.32
W.Rules	80.7	17.8	0.5	1.0	1.4	299.87
Chi-Sq.	48.7	50.6	0.1	0.6	0.7	433.46
T^2	40.7	58.8	0.1	0.5	0.6	467.81
P.C.	47.0	52.4	0.1	0.6	0.7	440.86

Table 4.8 - Cost Model Method Comparison for N=3/2, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	87.6	4.3	7.1	1.1	8.1	336.31
2 SD RT	85.0	13.1	0.9	1.0	1.9	288.94
W.Rules	87.2	10.1	1.6	1.0	2.6	285.23
Chi-Sq.	78.9	19.6	0.5	0.9	1.4	310.21
T^2	72.8	25.8	0.4	0.9	1.3	335.92
P.C.	75.6	23.0	0.5	0.9	1.4	324.20

Table 4.9 - Cost Model Method Comparison for N=3/2, h=24

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	84.2	12.6	2.2	1.0	3.2	301.46
2 SD RT	67.7	31.3	0.2	0.8	2.0	353.95
W.Rules	73.2	25.5	0.4	0.9	1.3	331.74
Chi-Sq.	56.8	42.4	0.1	0.7	0.8	398.97
T^2	48.1	51.2	0.1	0.6	0.7	435.98
P.C.	51.9	47.4	0.1	0.6	0.7	419.88

Table 4.10 - Cost Model Method Comparison for N=3/1, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	85.7	6.3	6.9	1.0	7.9	342.65
2 SD RT	75.1	23.2	0.8	0.9	1.7	330.12
W.Rules	81.0	16.5	1.5	1.0	2.5	310.76
Chi-Sq.	73.3	25.4	0.5	0.9	1.3	334.02
T^2	66.7	32.1	0.4	0.8	1.2	362.01
P.C.	73.9	24.8	0.5	0.9	1.3	331.66

Table 4.11 - Cost Model Method Comparison for N=3/1, h=24

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	79.4	17.6	2.1	1.0	3.0	320.92
2 SD RT	51.5	47.7	0.2	0.6	0.8	422.55
W.Rules	61.4	37.5	0.4	0.7	1.1	381.81
Chi-Sq.	48.7	50.6	0.1	0.6	0.7	433.46
T^2	40.7	58.8	0.1	0.5	0.6	467.81
P.C.	49.5	49.8	0.1	0.6	0.7	430.28

the Westgard Multirule Procedure as the least cost alternative when the sampling interval is 8 hours. Closely following the Westgard Multirule Procedure in overall performance is the use of +/- 2 SD limits with a retest. Given that the retest procedure is much easier to apply in practice, it may be a more attractive alternative than the Westgard Multirule Procedure to many clinicians.

With a sampling interval of 8 hours, the performance of the strict +/- 2 SD limits is behind the other traditional approaches selected here on a cost basis. While +/- 2 SD limits is definitely the most sensitive approach considered and the percentage of time in a state of statistical control (%SOSC) is generally higher for the +/- 2 SD limits approach, the costs and downtime associated with the false rejection rate of the method make it unattractive economically. It is interesting to note that the performance of +/- 2 SD limits actually improves by going from h=8 hours to 24 hours relative to the other approaches. This is because using a sampling interval of 24 hours will lower the impact of the false rejection rate, yet +/- 2 SD limits still has a good ability to detect real shifts as evidenced in the %SOSC measure. For all the other methods, the larger interval increases costs as their relative insensitivity to true changes begins to increase costs by operating out of control (%OOC) for a greater period of time.

The multivariate methods in general did not prove to be as cost effective as the traditional methods of +/- 2 SD limits with a retest or the Westgard Multirule Procedure. For the instances where not all of the levels shifted together (N=2/1, N=3/2, and N=3/1), the χ^2 chart is competitive with the other methods. The multivariate approaches overall, however, do not appear sensitive enough to operate cost effectively for the given set of

cost model parameter values. The percent of time operating in-control (%SOSC) is consistently smaller for the multivariate approaches. This would indicate that the lower sensitivity of the multivariate approaches is a handicap for this particular set of model parameter settings. As stated earlier in this paper, the lower α rate may be inhibiting the performance of the multivariate approaches relative to the traditional clinical approaches.

Another interesting insight from the cost model results is that for the instances where all levels shift together (N=2/2 and N=3/3), the use of three levels of controls is preferable. Again, this goes to the fact that more replicates increase the sensitivity of the quality control monitoring approach and that the cost per control observation is very small compared to the other operating costs. Given this result, clinicians may wish to consider increasing their control testing volume from two levels per subgroup to three levels per subgroup. They would also be best advised to test control values every 8 hours as opposed to every 24 hours if the cost parameters described here apply to their operation (unless they choose to use strict +/- 2 SD limits).

Overall, the cost modeling shows the importance of being sensitive to real changes if there are large costs associated with operating shifted away from the target. Since the cost of testing control samples is relatively small, increasing testing volume makes economical sense for this situation. Additionally, the amount of time required for fixing problems (both real and imaginary) is small enough compared to the expected frequency of true problems that they have limited impact on the results of the economic model.

Cost Model Sensitivity Analysis

In the preceding section, a single set of cost model parameters settings are used for comparing the quality control approaches in this research. While these settings are probably the most appropriate settings in terms of reflecting the actual costs incurred by laboratories, insight into the performance of the methods can be gained by evaluating other parameter settings. This sensitivity analysis enables the evaluation of other cost scenarios which could potentially be encountered. Therefore, this section considers two alternative sets of cost parameters from an economic perspective.

The first additional set of cost parameters makes only one change to the original set of cost parameters. The change is in the cost parameter, $\mathbf{C}_2$. While leaving the magnitudes of the three assignable causes the same at 1.0 SD, 2.0 SD, and 3.0 SD, the cost parameter $\mathbf{C}_2$ is altered such that the cost for operating with either a 1.0 SD shift or a 2.0 SD shift is the same as regular operation ($200/hour). This change effectively means that operating under these smaller shifts does not add any cost to the system, but a large shift (3.0 SD) still results in a large cost penalty. The cost parameters for the first sensitivity analysis (referred to as "Sensitivity Analysis A") are shown in Figure 4.5. From the cost modeling work in the previous section of this paper, it is clear that a sampling interval of 8 hours provides the best performance in general, so the sampling interval is fixed at 8 hours for the sensitivity analysis. Results of the first sensitivity analysis are shown in Tables 4.12 through 4.16 with each table summarizing the results for a different type of shift (i.e., N=2/2, N=2/1, etc.,).

Figure 4.5 - Cost Parameters for Sensitivity Analysis A

$$\delta = \begin{bmatrix} 1\ SD \\ 2\ SD \\ 3\ SD \end{bmatrix} \qquad \lambda = \begin{bmatrix} 0.002 \\ 0.001 \\ 0.0005 \end{bmatrix} \qquad \mathbf{SRT}_{OOC} = \begin{bmatrix} 4\ hrs \\ 3\ hrs \\ 2\ hrs \end{bmatrix} \qquad SRT_{SOSC}=5\ hrs$$

$$C_1=\$200/hr \qquad \mathbf{C}_2 = \begin{bmatrix} \$200/hr \\ \$200/hr \\ \$2000/hr \end{bmatrix} \qquad C_3=\$1600/hr \qquad C_4=\$10/sample$$

Table 4.12 - Sensitivity Analysis A for N=2/2, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	89.1	4.9	4.9	1.1	6.0	298.65
2 SD RT	82.9	15.7	0.4	1.0	1.4	262.93
W.Rules	86.0	12.4	0.6	1.0	1.7	257.60
Chi-Sq.	75.0	23.6	0.5	0.9	1.4	282.32
T^2	69.9	28.8	0.4	0.8	1.3	294.31
P.C.	73.1	25.6	0.5	0.9	1.3	286.80

Table 4.13 - Sensitivity Analysis A for N=2/1, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	86.7	7.5	4.8	1.0	5.8	303.05
2 SD RT	69.2	29.6	0.3	0.8	1.2	295.42
W.Rules	77.1	21.4	0.6	0.9	1.5	278.34
Chi-Sq.	75.0	23.6	0.5	0.9	1.4	282.32
T^2	69.9	28.8	0.4	0.8	1.3	294.31
P.C.	74.9	23.7	0.5	0.9	1.4	282.49

Table 4.14 - Sensitivity Analysis A for N=3/3, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	88.4	3.4	7.1	1.1	8.2	327.08
2 SD RT	89.3	8.7	0.9	1.1	2.0	255.02
W.Rules	90.6	6.6	1.7	1.1	2.7	259.25
Chi-Sq.	73.3	25.4	0.5	0.9	1.3	287.63
T^2	66.7	32.1	0.4	0.8	1.2	303.29
P.C.	72.0	26.7	0.4	0.9	1.3	290.78

Table 4.15 - Sensitivity Analysis A for N=3/2, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	87.6	4.3	7.1	1.1	8.1	328.37
2 SD RT	85.0	13.1	0.9	1.0	1.9	264.99
W.Rules	87.2	10.1	1.6	1.0	2.6	266.73
Chi-Sq.	78.9	19.6	0.5	0.9	1.4	274.31
T^2	72.8	25.8	0.4	0.9	1.3	288.69
P.C.	75.6	23.0	0.5	0.9	1.4	282.13

Table 4.16 - Sensitivity Analysis A for N=3/1, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	85.7	6.3	6.9	1.0	7.9	331.10
2 SD RT	75.1	23.2	0.8	0.9	1.7	287.72
W.Rules	81.0	16.5	1.5	1.0	2.5	280.56
Chi-Sq.	73.3	25.4	0.5	0.9	1.3	287.63
T^2	66.7	32.1	0.4	0.8	1.2	303.29
P.C.	73.9	24.8	0.5	0.9	1.3	286.31

The results from this first sensitivity analysis do not appear to depart substantially from the results for the original cost parameter settings. Again, the Westgard Multirule Procedure shows very good performance with +/- 2 SD limits with a retest showing comparable performance. The multivariate methods, however, show comparatively better performance for this set of cost parameters. This is attributable to the fact that the sensitivity of the multivariate approaches is less than the other methods, but higher sensitivity is not as necessary in this instance since small shifts do not incur any additional cost. Coupled with the low false rejection rate of the multivariate approaches, their cost performance is nearer the other methods for Sensitivity Analysis A than for the original set of cost parameter settings.

Another set of cost parameter settings is used for Sensitivity Analysis B (see Figure 4.6). The same settings from Sensitivity Analysis B are used with the exceptions of SRT_{SOSC} which is increased to 15 hours and the cost of down time, C_4, which is increased to $3000 per hour. Costs for both 1.0 SD shifts and 2.0 SD shifts remain at $200/hour meaning that small shifts do not add additional cost to the system. By using these parameter settings, one is basically saying that small shifts do not increase costs, it takes a substantial amount of time to resolve false rejections, and down time is very expensive. The results for Sensitivity Analysis B are shown in Tables 4.17 through 4.21.

Figure 4.6 - Cost Parameters for Sensitivity Analysis B

$$\delta = \begin{bmatrix} 1\ SD \\ 2\ SD \\ 3\ SD \end{bmatrix} \qquad \lambda = \begin{bmatrix} 0.002 \\ 0.001 \\ 0.0005 \end{bmatrix} \qquad \mathbf{SRT}_{OOC} = \begin{bmatrix} 4\ hrs \\ 3\ hrs \\ 2\ hrs \end{bmatrix} \qquad SRT_{SOSC}=15\ hrs$$

$$C_1=\$200/hr \qquad \mathbf{C}_2 = \begin{bmatrix} \$200/hr \\ \$200/hr \\ \$2000/hr \end{bmatrix} \qquad C_3=\$3000/hr \qquad C_4=\$10/sample$$

Table 4.17 - Sensitivity Analysis B for N=2/2, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	81.2	4.5	13.4	1.0	14.3	615.81
2 SD RT	82.2	15.6	1.2	1.0	2.2	304.44
W.Rules	84.9	12.2	1.9	1.0	2.9	314.80
Chi-Sq.	74.3	23.4	1.4	0.9	2.3	326.14
T^2	69.3	28.5	1.3	0.8	2.1	335.09
P.C.	72.5	25.3	1.3	0.9	2.2	329.48

Table 4.18 - Sensitivity Analysis B for N=2/1, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	79.1	6.9	13.0	0.9	14.0	611.91
2 SD RT	68.8	29.4	1.0	0.8	1.8	329.90
W.Rules	76.3	21.1	1.7	0.9	2.6	329.50
Chi-Sq.	74.3	23.4	1.4	0.9	2.3	326.14
T^2	69.3	28.5	1.3	0.8	2.1	335.09
P.C.	74.3	23.5	1.4	0.9	2.3	326.27

Table 4.19 - Sensitivity Analysis B for N=3/3, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	77.4	3.0	18.7	0.9	19.6	761.10
2 SD RT	87.7	8.5	2.8	1.1	3.8	333.38
W.Rules	87.7	6.4	4.8	1.1	5.9	384.41
Chi-Sq.	72.6	25.1	1.3	0.9	2.2	330.42
T^2	66.1	31.9	1.2	0.8	2.0	342.11
P.C.	71.3	26.5	1.3	0.9	2.2	332.77

Table 4.20 - Sensitivity Analysis B for N=3/2, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	76.7	3.8	18.5	0.9	19.5	758.56
2 SD RT	83.5	12.9	2.6	1.0	3.6	339.44
W.Rules	84.5	9.8	4.6	1.0	5.7	387.16
Chi-Sq.	78.2	19.4	1.4	0.9	2.4	320.49
T^2	72.2	25.6	1.3	0.9	2.2	331.21
P.C.	74.9	22.8	1.4	0.9	2.3	326.32

Table 4.21 - Sensitivity Analysis B for N=3/1, h=8

Method	%SOSC	%OOC	%FR	%TR	%DT	Cost
2 SD	75.3	5.5	18.2	0.9	19.1	753.17
2 SD RT	74.0	22.8	2.3	0.9	3.2	353.32
W.Rules	78.7	16.0	4.3	0.9	5.3	392.26
Chi-Sq.	72.6	25.1	1.3	0.9	2.2	330.42
T^2	66.1	31.9	1.2	0.8	2.0	342.11
P.C.	73.2	24.6	1.4	0.9	2.2	329.43

The first observation one can make based on Sensitivity Analysis B is that +/- 2 SD limits does not perform very well at all from an economic perspective for the selected parameter settings. This is not surprising since Sensitivity Analysis B includes a large amount of time for investigating false rejections and has a large cost for down time. Coupled with its high rate of false rejection, +/- 2 SD limits shows very poor performance based on costs for this set of parameters.

Interestingly, the multivariate approaches perform very well for the parameter settings used for Sensitivity Analysis B. The χ^2 chart is the least cost alternative for three of the five types of shifts investigated with the principal component chart being the least cost alternative for one of the other two types of shifts. In this instance, the low false rejection rates of the multivariate approaches pay dividends by reducing unnecessary down time associated with investigating false rejections. Since only large shifts result in major cost penalties for this set of parameters, the relative lack of sensitivity of the multivariate approaches is not as damaging as for the original set of cost parameters investigated in this research.

While the cost parameter settings used in the sensitivity analysis are probably not the most reflective of real world laboratory situations, they do provide insight into the performance of the alternatives under investigation. Particularly, the potential harm of the high false rejection rate for strict +/- 2 SD limits is highlighted. Additionally, it is clear that multivariate approaches can be extremely attractive under the right circumstances.

CHAPTER V

SUMMARY

Conclusions and Recommendations

As a result of this research, straightforward approaches are available for evaluating both the statistical and economic performance of quality control approaches used in the clinical laboratory setting. Six different approaches are analyzed in this research, but other methods could be evaluated in the same fashion. The important aspects of the evaluations are how well the method detects shifts as determined through its respective ARLs and the total cost per unit time for the method. By evaluating the method's performance on these two fronts, one can make intelligent decisions regarding the appropriateness of the quality control approach for his or her situation.

From the evaluation of the six methods explored in this research, a few overall conclusions may be drawn. From a statistical perspective, the traditional methods used by laboratorians tend to have good sensitivity to true changes in the measurement system. For the strict application of +/- 2 SD limits, this sensitivity is in fact excessive as it results in a high rate of false rejection. Both +/- 2 SD limits with a retest and the Westgard Multirule Procedure temper this high false rejection rate while maintaining good sensitivity to true changes in the measurement system. The +/- 2 SD limits with a retest approach may be favored over the Westgard Multirule Procedure as it is easier to implement in a laboratory testing situation. Additionally, +/- 2 SD limits with a retest tends to have a

longer in-control ARL than the Westgard Multirule Procedure while having comparable sensitivity to true shifts.

For the traditional approaches, the methods all have increased sensitivity with higher numbers of replicates. However, this increased sensitivity comes with a price as the false rejection rates for the traditional approaches also increase with higher numbers of controls. Also, the historical approach of evaluating these methods only for shifts across all levels can be very deceiving. As shown in this research, the traditional methods have very different performance for shifts in individual levels while the other levels remain centered. Clinicians may have a false sense of security in their quality control approach's ability to detect any kind of shift when in fact their method has been designed around detecting shifts across all levels of control materials.

The multivariate approaches considered show excellent promise, but they do not quite have the sensitivities of the traditional approaches. One of the advantages of the multivariate approaches is that a user can fix the false rejection rate and maintain the same false rejection rate regardless of the number of replicates being monitored. For this research, the in-control ARL is fixed at 100 across the board whereas traditional approaches have false rejection rates which fluctuate with the number of replicates being monitored.

One aspect of using the multivariate approaches which must be considered is the correlation structure of the measurement system. This research shows that the ability of the multivariate approaches to detect different types of shifts will be impacted by the correlation structure of the data being monitored. Data which is highly correlated among

the three levels will result in a multivariate approach which will be less sensitive to shifts across all levels of control materials, but sensitive to shifts in individual levels. Low correlation results in better sensitivity to shifts across all levels. While a user cannot manipulate the correlation of the control levels being monitored, one must be aware of the impact that the correlation structure can have on the performance of the multivariate approaches.

For the multivariate approaches examined in this research (i.e., the χ^2 chart, the T^2 chart, and the principal component chart), the performance of the three approaches is surprisingly similar. While the χ^2 chart will always outperform the T^2 chart, it is interesting to note that the T^2 chart with limits based on 20 subgroups is not dramatically outperformed by the χ^2 chart. Also, the χ^2 chart and principal component chart show very similar performance.

An advantage of the multivariate approaches is that they have more consistent performance regardless of the type of shift encountered while the traditional approaches can have very different performance depending on the shift. This can change with a changing correlation structure, but the work in this research indicates that the multivariate approaches have similar sensitivities to all the shifts investigated. This is an advantage for the multivariate approaches over the traditional methods as it provides a clinician with a given level of error protection, regardless of the nature of the shift. For traditional methods, the degree of error protection provided to a clinician will vary by the type of shift encountered.

From a cost perspective, this research also leads to a number of interesting insights. The first is that the Westgard Multirule Procedure is generally the most economic approach for a set of cost parameter settings representative of laboratory operation. However, +/- 2 SD limits with a retest shows performance very near to the Westgard Multirule Procedure. Given that +/- 2 SD limits does not require past testing data and the ease with which it can be implemented, +/- 2 SD limits should be given consideration as a valid alternative to the Westgard Multirule Procedure.

For the cost parameter settings typical of a functioning laboratory, the multivariate approaches did not prove to be as cost effective as either the Westgard Multirule Procedure or +/- 2 SD limits with a retest. For the instances where not all of the levels shifted together (N=2/1, N=3/2, and N=3/1), the χ^2 chart is competitive with the other methods. The multivariate approaches overall, however, do not appear sensitive enough to operate cost effectively for the given set of cost model parameter values.

As for setting the testing interval from an economic perspective, a testing interval of eight hours shows better performance over a 24 hour interval. For +/- 2 SD limits, though, increasing the testing interval actually reduces costs since it lowers the impact of the high false rejection rate for the method. In general, however, the testing interval of eight hours is preferred.

Another interesting insight from the cost model results is that for the instances where all levels shift together (N=2/2 and N=3/3), the use of three levels of control materials is preferable. Again, this goes to the fact that more replicates increase the sensitivity of the quality control monitoring approach and that the cost per control

observation is very small compared to the other operating costs. Given this result, clinicians may wish to consider increasing their control testing volume from two levels per subgroup to three levels per subgroup.

The cost model parameter settings typical of laboratory operation are the most appropriate for evaluating how the methods will perform in the field, but sensitivity analysis in this research also provides some valuable insights into the methods' performance. In one sensitivity analysis scenario, no additional cost is added for 1.0 SD or 2.0 SD shifts. The results for this scenario do not depart substantially from the results for the original parameter settings, but the multivariate approaches are more competitive with the traditional methods. In another sensitivity analysis scenario, the amount of time to investigate a false rejection is tripled and the cost per hour for down time is increased from $2000 to $3000 per hour. The +/- 2 SD limits approach performs horribly for this scenario as its high rate of false rejection proves to be a true handicap. The multivariate methods show the best performance for this scenario as they have a low rate of false rejection, and there is not an additional cost for small shifts (i.e., 1.0 SD or 2.0 SD).

Areas for Future Research

The biggest opportunity for future research in this area is with regards to the multivariate approaches investigated. Future research could explore the impact of increasing the power of the multivariate approaches to more closely match the traditional methods. While this will increase the false rejection rate for the multivariate approaches, they will more closely match the sensitivity of the traditional methods. Given that the

difference in cost performance of the multivariate approaches from the traditional approaches appears to be driven by sensitivity, improving the power of the multivariate methods and increasing the false rejection rate could result in cost performance very similar to the traditional methods.

Another potential area for research is the determination of the cost model parameter settings used in this research. While the settings considered here are representative, further input from a large sample of clinicians could result in better inputs to the cost model parameters. As evidenced in the sensitivity analysis of this research, changing the model parameter settings can impact the relative performance of the methods. Therefore, further information about these model parameter settings would be useful.

As stated in the literature review of this research, the principal component chart has not been previously modeled from a cost perspective. This research develops that model and analyzes it for the clinical application. This cost model could be adapted and used, however, to optimize the principal component chart in general circumstances.

REFERENCES

Alt, F.B. (1982). Multivariate quality control: State of the art. *ASQC Quality Congress Transactions*, 886-893.

Alt, F.B. (1985). Multivariate quality control. *Encyclopedia of Statistical Sciences, 6* (S. Kotz and N. Johnson, eds.) John Wiley & Sons, New York, NY, 110-122.

Arnold, B. F. (1989). Optimal control charts and discrimination between 'Acceptable' and 'Unacceptable' states. *Sankhya, 51*, Series B, 375-389.

Anderson, T.W. (1958) *An introduction to multivariate statistical analysis*. John Wiley and Sons Inc., New York, NY.

Banerjee, P.K. and Rahim, M.A. (1988). Economic design of $\overline{X}$ control charts under Weibull shock models. *Technometrics, 30*, 407-414.

Bishop, J. and Nix, A.B.J. (1993). Comparison of quality-control rules used in clinical chemistry laboratories. *Clinical Chemistry, 39*, 1638-1649.

Blum, A.S. (1985). Computer evaluation of statistical procedures and a new quality-control statistical procedure. *Clinical Chemistry, 31*, 206-212.

Collani, E.v. (1986). A simple procedure to determine the economic design of an $\overline{X}$ control chart. *Journal of Quality Technology, 18*, 145-151.

Collani, E.v. (1988). A unified approach to optimal process control. *Metrika, 35*, 145-159.

Duncan, A.J. (1956). The economic design of $\overline{X}$ charts used to maintain current control of a process. *Journal of the American Statistical Association, 51*, 228-242.

Fraser, C.G., Petersen, H., and Larsen, M.L. (1990). Setting analytical goals for random analytical error in specific clinical monitoring situations. *Clinical Chemistry, 36*, 1625-1628.

Hatjimihail, A.T. (1992). A tool for the design and evaluation of alternative quality-control procedures. *Clinical Chemistry, 38*, 204-210.

Hatjimihail, A.T. (1993). Genetic algorithms-based design and optimization of statistical quality-control procedures. *Clinical Chemistry, 39*, 1972-1978.

Henry, R.J. and Segalove, M. (1952). The running of standards in clinical chemistry and the use of the control chart. *Journal of Clinical Pathology, 5*, 305-311.

Ho, C. and Case, K.E. (1994). Economic design of control charts: A literature review for 1981 - 1991. *Journal of Quality Technology, 26*, 39-53.

Hotelling, H. (1931). The generalization of Student's ratio. *Annals of Mathematical Statistics, 2*, 360-378.

Hotelling, H. (1947). Multivariate quality control. *Techniques of Statistical Analysis*, (C. Eisenhart, M. Hastay, and W.A. Wallis, eds.) McGraw Hill, New York, NY, 111-184.

Jackson, J.E. (1956). Quality control methods for two related variables. *Industrial Quality Control, 12* (7), 4-8.

Jackson, J.E. (1959). Quality control methods for several related variables. *Technometrics, 1* (4), 359-377.

Jackson, J.E. (1980). Principal components and factor analysis: Part I - Principal components. *Journal of Quality Technology, 12*, 201-213.

Jackson, J.E. (1981-a). Principal components and factor analysis: Part II - Additional topics related to principal components. *Journal of Quality Technology, 13*, 46-58.

Jackson, J.E. (1981-b). Principal components and factor analysis: Part III - What is factor analysis? *Journal of Quality Technology, 13*, 125-130.

Jackson, J.E. (1985). Multivariate quality control. *Communications in Statistics, 14* (11), 2657-2688.

Jackson, J.E. and Bradley, R.A. (1966). Sequential multivariate procedures with quality control applications. *Multivariate Analysis I*, (P.R. Krishnaiah, ed.) Academic Press, New York, NY, 507-519.

Jackson, J.E. and Hearne, F.T. (1973). Relationships among coefficients of vectors used in principal components. *Technometrics, 15* (3), 601-610.

Jackson, J.E. and Mudholkar G.S. (1979). Control procedures for residuals associated with principal components analysis. *Technometrics, 21* (3), 341-349.

Jones, L.L. and Case, K.E. (1981). Economic design of a joint $\overline{X}$ and R control chart. *AIIE Transactions, 13*, 182-195.

Koch, D.D., Oryall, J.J., Quam, E.F., Feldbruegge, D.H., Dowd, D.E., Barry, P.L., and Westgard, J.O. (1990). Selection of medically useful quality-control procedures for individual tests done in a multi-test analytical system. *Clinical Chemistry, 36*, 230-233.

Lee, W.K. (1996). *Performance characteristics of multi-rule Shewhart control charts used in clinical chemistry laboratories*. Master's Report, Oklahoma State University, Stillwater, Oklahoma.

Levey, S. and Jennings, E.R. (1950). The use of control charts in the clinical laboratory. *American Journal of Clinical Pathology, 20*, 1059-1066.

Linnet, K. (1989-a). Choosing quality-control systems to detect maximum clinically allowable analytical errors. *Clinical Chemistry, 35*, 284-288.

Linnet, K. (1989-b). The between-run component of variation in internal quality control. *Clinical Chemistry, 35*, 1416-1422.

Linnet, K. (1991). Mean and variance rules are more powerful or selective than quality control rules based on individual values. *European Journal of Clinical Chemistry and Clinical Biochemistry, 29*, 417-424.

Lorenzen, T.J. and Vance, L.C. (1986). The economic design of control charts: A unified approach. *Technometrics, 28*, 3-10.

Marsaglia, G. and Bray, T.A. (1964) A convenient method for generating normal variables. *SIAM Rev., 6*, 260-264.

Marse, K. and Roberts, S.D. (1983) Implementing a portable FORTRAN uniform (0,1) generator. *Simulation, 41*, 135-139.

Mathcad User's Guide, Version 6.0 (1995). Mathsoft, Inc., Cambridge, MA.

McWilliams, T.P. (1989). Economic control chart designs and the in-control time distribution: A sensitivity analysis. *Journal of Quality Technology, 21*, 103-110.

Montgomery, D.C. (1980). The economic design of control charts: A review and literature survey. *Journal of Quality Technology, 21*, 75-87.

Montgomery, D.C. (1982). Economic design of an $\overline{X}$ control chart. *Journal of Quality Technology, 14*, 40-43.

Montgomery, D.C. and Klatt, P.J. (1972-a). Economic design of T^2 control charts to maintain current control of a process. *Management Science, 19* (1), 76-89.

Montgomery, D.C. and Klatt, P.J. (1972-b). Minimum cost multivariate quality control tests. *AIEE Transactions, 4* (2), 103-110.

Parkhideh, B. and Case, K.E. (1989). The economic design of a dynamic $\overline{X}$ control chart. *IIE Transactions, 21*, 313-323.

Parvin, C.A. (1991). Estimating the performance characteristics of quality-control procedures when error persists until detection. *Clinical Chemistry, 37*, 1720-1724.

Parvin, C.A. (1992). Comparing the power of quality-control rules to detect persistent systematic error. *Clinical Chemistry, 38*, 358-363.

Parvin, C.A. (1993). New insight into the comparative power of quality-control rules that use control observations within a single analytical run. *Clinical Chemistry, 39*, 440-447.

Petersen, P.H., and Fraser, C.G. (1994). Setting quality standards in clinical chemistry: Can competing based on analytical, biological, and clinical outcomes be harmonized? *Clinical Chemistry, 40*, 1865-1868.

Pignatiello, J.J. and Tsai, A. (1988). Optimal economic design of $\overline{X}$ control charts when cost model parameters are not precisely known. *IIE Transactions, 20*, 103-110.

Saniga, E.M. (1989). Economic statistical control-chart designs with an application to $\overline{X}$ and R charts. *Technometrics, 31*, 258-264.

Schweiger, C.R., Soeregi, G., Spitzauer, S., Maenner, G., and Pohl, A.L. (1993). Evaluation of laboratory data by conventional statistics and by three types of neural networks. *Clinical Chemistry, 39*, 1966-1971.

Skendzel, L.P., Barnett, R.N., and Platt, R. (1985). Medically useful criteria for analytical performance of laboratory tests. *American Journal of Clinical Pathology, 83*, 200-205.

Tagaras, G. and Lee, H.L. (1988). Economic design of control charts with different control limits for different assignable causes. *Management Science, 34*, 1347-1366.

Tracy, N.D., Young, J.C., and Mason, R.L. (1992). Multivariate control charts for individual observations. *Journal of Quality Technology, 24* (2), 88-95.

Turbo Pascal User's Guide, Version 6.0 (1990). Borland, International, Scotts Valley, CA.

Westgard, J.O. (1992). Simulation and modeling for optimizing quality control and improving analytical quality assurance. *Clinical Chemistry, 38*, 175-178.

Westgard, J.O., Barry, P.L., and Hunt, M.R. (1981). A multi-rule Shewhart chart for quality control in clinical chemistry. *Clinical Chemistry, 27*, 493-501.

Westgard, J.O. and Burnett, R.W. (1990). Precision requirements for cost-effective operation of analytical processes. *Clinical Chemistry, 36*, 1629-1632.

Westgard, J.O. and Groth, T. (1979). Power functions for statistical control rules. *Clinical Chemistry, 25*, 863-869.

Westgard, J.O. and Groth, T. (1983) A predictive value model for quality control: effects of the prevalence of errors on performance of control procedures. *American Journal of Clinical Pathology, 80*, 49-56.

Westgard, J.O., Groth, T., Aronsson, T. and deVerdier, C.H. (1977). Combined Shewhart-cusum control chart for improved quality control in clinical chemistry. *Clinical Chemistry, 23*, 1881-1887.

Westgard, J.O., Groth, T., Aronsson, T., Falk, H., and de Verdier, C.H. (1977). Performance characteristics of rules for internal quality control: Probabilities for false rejection and error detection. *Clinical Chemistry, 23*, 1857-1867.

Westgard, J.O., Oryall, J.J., and Koch, D.D. (1990) "Predicting effects of qc practices on the cost-effective operation of a multi-test analytical system. *Clinical Chemistry, 36*, 1760-1764.

Westgard, J.O., Quam, E.F., and Barry, P.L. (1990). Selection grids for planning quality control procedures. *Clinical Laboratory Science, 3*, 271-278.

Westgard, J.O., Seehafer, J.J., and Barry, P.L. (1994-b). Allowable imprecision for laboratory tests based on clinical and analytical test outcome criteria. *Clinical Chemistry, 40*, 1909-1914.

Yu, G. (1994) *The evaluation and design of two multivariate quality control charts using the method of principal components*. Doctoral Dissertation, Oklahoma State University, Stillwater, Oklahoma.

APPENDIX A

Computer Code for Westgard Multirule Procedure (N'=2)

```
program multirule;
uses crt;

{This program will determine ARLs for subgroup size 2 using a variety of control rules}

var
  warn,rult,rultw,rulf,rulxb,rulrfs,m,n,l,i : integer;
  tsdw, blta,alta,bmta,amta,s,flag,tresd,rfs : integer;
  bfw,afw,blfa,alfa,bmfa,amfa: integer;
  btenw,atenw,bltena,altena,bmtena,amtena: integer;
  low,mid,arl,rl,lshft,mshft,diff,sd,uci,lci,x : real;
  sumsq,sum,y,z,vone,vtwo,w,yy,nsum,nsumsq,farl,fas :real;
  seed : longint;

function RandUnif : real;

{Function RandUnif generates Uniform 0,1 variates using the Marse-Roberts code}

const
 B2E15 =  32768;
 B2E16 =  65536;
 Modulus = 2147473647;
 Mult1 =  24112;
 Mult2 =  26143;

var  Hi15, Hi31, Low15, Lowprd, Ovflow,Zi : longint;

begin
 Zi := Seed;
 Hi15 := Zi DIV B2E16;
 Lowprd := (Zi - Hi15 * B2E16) * Mult1;
 Low15 := Lowprd DIV B2E16;
 Hi31 := Hi15 * Mult1 + Low15;
 Ovflow := Hi31 DIV B2E15;
 Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
      (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
 IF Zi < 0 THEN Zi := Zi + Modulus;
 Hi15 := Zi DIV B2E16;
 Lowprd := (Zi - Hi15 * B2E16) * Mult2;
 Low15 := Lowprd DIV B2E16;
 Hi31 := Hi15 * Mult2 + Low15;
 Ovflow := Hi31 DIV B2E15;
 Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
      (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
 IF Zi < 0 THEN Zi := Zi + Modulus;
```

```
  Seed := Zi;
  RandUnif := (2* (Zi DIV 256) + 1) / 16777216.0;
end;

{The next section allows the user to input the control rules for the simulation}

begin
  clrscr;
  writeln('This program is specifically for N=2');
  writeln;
  writeln('2 sd as screening? (1=yes, 2=no)');
  readln(warn);
  writeln('1-3 sd rule? (1=yes, 2=no)');
  readln(rult);
  writeln('2-2 sd rule? (1=yes, 2=no)');
  readln(rultw);
  writeln('4-1 sd rule? (1=yes, 2=no)');
  readln(rulf);
  writeln('10 xbar rule? (1=yes, 2=no)');
  readln(rulxb);
  writeln('R - 4 sd rule? (1=yes, 2=no)');
  readln(rulrfs);
  writeln;
  writeln('How many realizations? ');
  readln(m);
  writeln('Enter the seed for the random number generator:');
  readln(seed);
  writeln('Enter the low control shift in sd: ');
  readln(lshft);
  writeln('Enter the mid control shift in sd: ');
  readln(mshft);

  rl:=0;  flag:=0;   tresd:=0;
  rfs:=0;  tsdw:=0;  blta:=0;  alta:=0;
  bmta:=0; amta:=0;  bfw:=0;  afw:=0;  blfa:=0;
  alfa:=0; bmfa:=0;  amfa:=0;  btenw:=0;  atenw:=0;
  bltena:=0;  altena:=0; bmtena:=0;amtena:=0;
  nsum:=0;  nsumsq:=0;
  clrscr;
  for s:=1 to 5 do
   begin
   sum:=0; sumsq:=0;
  for l:= 1 to m do
   begin
    repeat
```

```
begin
 repeat

{Here the program generates two normal variates}

  begin
   y:=RandUnif;
   z:=RandUnif;
   vone:=2*y-1;
   vtwo:=2*z-1;
   w:=vone*vone+vtwo*vtwo;
  end;
until w<1;
yy:=sqrt((-2*ln(w))/w);
low:=lshft+vone*yy;
mid:=mshft+vtwo*yy;

{3 sd checking logic}
if (low<-3.0) then tresd:=1;
if (low>3.0) then tresd:=1;
if (mid<-3.0) then tresd:=1;
if (mid>3.0) then tresd:=1;

{R 4sd checking logic}
diff:=mid-low;
if (diff>4.0) or (diff<-4.0) then rfs:=1;

{2-2 sd checking logic within a run}
if (low<-2.0) and (mid<-2.0) then tsdw:=1;
if (low>2.0) and (mid>2.0) then tsdw:=1;

{2-2 sd checking logic across runs}
if (low<-2.0) then blta:=blta+1
  else blta:=0;
if (low>2.0) then alta:=alta+1
  else alta:=0;
if (mid<-2.0) then bmta:=bmta+1
  else bmta:=0;
if (mid>2.0) then amta:=amta+1
  else amta:=0;

{4 1-sd checking logic within a run}
if (low<-1.0) and (mid<-1.0) then bfw:=bfw+2
  else bfw:=0;
if (low>1.0) and (mid>1.0) then afw:=afw+2
```

```
  else afw:=0;

{4 1-sd checking logic across runs}
if (low<-1.0) then blfa:=blfa+1
  else blfa:=0;
if (low>1.0) then alfa:=alfa+1
  else alfa:=0;
if (mid<-1.0) then bmfa:=bmfa+1
  else bmfa:=0;
If (mid>1.0) then amfa:=amfa+1
  else amfa:=0;

{10 xbar within}
if (low<0.0) and (mid<0.0) then btenw:=btenw+2
  else btenw:=0;
if (low>0.0) and (mid>0.0) then atenw:=atenw+2
  else atenw:=0;

{10 xbar across runs}
if (low<=0.0) then bltena:=bltena+1
  else bltena:=0;
if (low>0.0) then altena:=altena+1
  else altena:=0;
if (mid<=0.0) then bmtena:=bmtena+1
  else bmtena:=0;
if (mid>0.0) then amtena:=amtena+1
  else amtena:=0;

{Summary logic}
{If no 2 sd screen}
if (warn=2) then
 begin
   if (rult=1) and (tresd=1) then flag:=1;
   if (rulrfs=1) and (rfs=1) then flag:=1;
   if (rultw=1) and (tsdw=1) then flag:=1;
   if (rultw=1) and (blta>=2) then flag:=1;
   if (rultw=1) and (alta>=2) then flag:=1;
   if (rultw=1) and (bmta>=2) then flag:=1;
   if (rultw=1) and (amta>=2) then flag:=1;
   if (rulf=1) and (bfw=4) then flag:=1;
   if (rulf=1) and (afw=4) then flag:=1;
   if (rulf=1) and (blfa=4) then flag:=1;
   if (rulf=1) and (alfa=4) then flag:=1;
   if (rulf=1) and (bmfa=4) then flag:=1;
   if (rulf=1) and (amfa=4) then flag:=1;
```

```
      if (rulxb=1) and (btenw=10) then flag:=1;
      if (rulxb=1) and (atenw=10) then flag:=1;
      if (rulxb=1) and (bltena=10) then flag:=1;
      if (rulxb=1) and (altena=10) then flag:=1;
      if (rulxb=1) and (bmtena=10) then flag:=1;
      if (rulxb=1) and (amtena=10) then flag:=1;
    end;

  {Using 2 sd screen}
  if (warn=1) then
  begin
   if(low<-2.0) or (low>2.0) or (mid<-2.0) or (mid>2.0) then
   begin
    if (rult=1) and (tresd=1) then flag:=1;
    if (rulrfs=1) and (rfs=1) then flag:=1;
    if (rultw=1) and (tsdw=1) then flag:=1;
    if (rultw=1) and (blta>=2) then flag:=1;
    if (rultw=1) and (alta>=2) then flag:=1;
    if (rultw=1) and (bmta>=2) then flag:=1;
    if (rultw=1) and (amta>=2) then flag:=1;
    if (rulf=1) and (bfw=4) then flag:=1;
    if (rulf=1) and (afw=4) then flag:=1;
    if (rulf=1) and (blfa=4) then flag:=1;
    if (rulf=1) and (alfa=4) then flag:=1;
    if (rulf=1) and (bmfa=4) then flag:=1;
    if (rulf=1) and (amfa=4) then flag:=1;
    if (rulxb=1) and (btenw=10) then flag:=1;
    if (rulxb=1) and (atenw=10) then flag:=1;
    if (rulxb=1) and (bltena=10) then flag:=1;
    if (rulxb=1) and (altena=10) then flag:=1;
    if (rulxb=1) and (bmtena=10) then flag:=1;
    if (rulxb=1) and (amtena=10) then flag:=1;
   end;
  end;
  {Summary statistics}
  rl:=rl+1;
 end;
until flag=1;

sum:=sum+rl;
sumsq:=sumsq+(rl*rl);
rl:=0;
flag:=0;
tresd:=0; rfs:=0;  tsdw:=0;  blta:=0;  alta:=0;
bmta:=0; amta:=0;  bfw:=0;  afw:=0;  blfa:=0;
```

```
    alfa:=0; bmfa:=0;  amfa:=0;  btenw:=0;  atenw:=0;
    bltena:=0;  altena:=0; bmtena:=0;amtena:=0;
   end;
    arl:=sum/m;
    sd:=sqrt((sumsq-(m*arl*arl))/(m-1));
    writeln('The ARL is ',arl:5:2);
    writeln('The ARLSD is ',sd:5:2);
    nsum:=nsum+arl;
    nsumsq:=nsumsq+(arl*arl);
   end;

{Final Output}
   writeln('Low Control Shift = ',lshft:5:2);
   writeln('Mid Control Shift = ',mshft:5:2);
   writeln;
   farl:=nsum/5;
   fas:=sqrt((nsumsq-(5*farl*farl))/4);
   uci:=farl+2.776*(fas/(sqrt(5)));
   lci:=farl-2.776*(fas/(sqrt(5)));
   writeln('The ARL estimate is ',farl:5:2);
   writeln('The upper 95% CI is ',uci:5:2);
   writeln('The lower 95% CI is ',lci:5:2);

end.
```

APPENDIX B

Computer Code for Westgard Multirule Procedure (N'=3)

```
program multirule;
uses crt;

{The program will determine ARLs for subroup size 3 using a variety of control rules}

var
  warn,rult,rultw,rulf,rulxb,rulrfs,m,n,l,i : integer;
  tsdw, blta,alta,bmta,amta,flag,tresd,rfs,s : integer;
  bfw,afw,blfa,alfa,bmfa,amfa: integer;
  btenw,atenw,bltena,altena,bmtena,amtena: integer;
  bhta,ahta,bhfa,ahfa,bhtena,ahtena : integer;
  low,mid,high,arl,rl,lshft,mshft,hshft,diff,sd : real;
  uci,lci,x,nsum,nsumsq,farl,fas : real;
  sumsq,sum,y,z,vone,vtwo,w,yy,diffa,diffb :real;
  seed : longint;

function RandUnif : real;

{Function RandUnif generates Uniform 0,1 variates using the Marse-Roberts code}

const
 B2E15 =  32768;
 B2E16 =  65536;
 Modulus = 2147473647;
 Mult1 =  24112;
 Mult2 =  26143;

var  Hi15, Hi31, Low15, Lowprd, Ovflow,Zi : longint;

begin
 Zi := Seed;
 Hi15 := Zi DIV B2E16;
 Lowprd := (Zi - Hi15 * B2E16) * Mult1;
 Low15 := Lowprd DIV B2E16;
 Hi31 := Hi15 * Mult1 + Low15;
 Ovflow := Hi31 DIV B2E15;
 Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
      (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
 IF Zi < 0 THEN Zi := Zi + Modulus;
 Hi15 := Zi DIV B2E16;
 Lowprd := (Zi - Hi15 * B2E16) * Mult2;
 Low15 := Lowprd DIV B2E16;
 Hi31 := Hi15 * Mult2 + Low15;
 Ovflow := Hi31 DIV B2E15;
 Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
```

```
      (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
  IF Zi < 0 THEN Zi := Zi + Modulus;
  Seed := Zi;
  RandUnif := (2* (Zi DIV 256) + 1) / 16777216.0;
end;

{The next section allows the user to input the control rules for the simulation}

begin
  clrscr;
  writeln('This program is specifically for N=3');
  writeln;
  writeln('2 sd as screening? (1=yes, 2=no)');
  readln(warn);
  writeln('1-3 sd rule? (1=yes, 2=no)');
  readln(rult);
  writeln('2-2 sd rule? (1=yes, 2=no)');
  readln(rultw);
  writeln('4-1 sd rule? (1=yes, 2=no)');
  readln(rulf);
  writeln('10 xbar rule? (1=yes, 2=no)');
  readln(rulxb);
  writeln('R - 4 sd rule? (1=yes, 2=no)');
  readln(rulrfs);
  writeln;
  writeln('How many realizations? ');
  readln(m);
  writeln('Enter the seed for the random number generator:');
  readln(seed);
  writeln('Enter the low control shift in sd: ');
  readln(lshft);
  writeln('Enter the mid control shift in sd: ');
  readln(mshft);
  writeln('Enter the high control shift in sd: ');
  readln(hshft);

  nsum:=0;  nsumsq:=0;  rl:=0;  flag:=0;  tresd:=0;
  rfs:=0;  tsdw:=0;  blta:=0;  alta:=0;
  bmta:=0; amta:=0;  bfw:=0;  afw:=0;  blfa:=0;
  alfa:=0; bmfa:=0;  amfa:=0;  btenw:=0;  atenw:=0;
  bltena:=0;  altena:=0; bmtena:=0;amtena:=0;
  bhta:=0; ahta:=0; bhfa:=0; ahfa:=0; bhtena:=0; ahtena:=0;

  clrscr;
  for s:=1 to 5 do
```

```
  begin
  sum:=0; sumsq:=0;
 for l:= 1 to m do
  begin
   repeat
    begin
     repeat
      begin

{This section generates 3 normal variates for the simulation}

        y:=RandUnif;
        z:=RandUnif;
        vone:=2*y-1;
        vtwo:=2*z-1;
        w:=vone*vone+vtwo*vtwo;
       end;
      until w<1;
      yy:=sqrt((-2*ln(w))/w);
      low:=lshft+vone*yy;
      mid:=mshft+vtwo*yy;

      repeat
       begin
        y:=RandUnif;
        z:=RandUnif;
        vone:=2*y-1;
        vtwo:=2*z-1;
        w:=vone*vone+vtwo*vtwo;
       end;
      until w<1;
      yy:=sqrt((-2*ln(w))/w);
      high:=hshft+vone*yy;

      {3 sd checking logic}
      if (low<-3.0) then tresd:=1;
      if (low>3.0) then tresd:=1;
      if (mid<-3.0) then tresd:=1;
      if (mid>3.0) then tresd:=1;
      if (high<-3.0) then tresd:=1;
      if (high>3.0) then tresd:=1;

      {R 4sd checking logic}
      diff:=mid-low;
      diffa:=high-low;
```

```
diffb:=high-mid;
if (diff>4.0) or (diff<-4.0) then rfs:=1;
if (diffa>4.0) or (diffa<-4.0) then rfs:=1;
if (diffb>4.0) or (diffb<-4.0) then rfs:=1;

{2-2 sd checking logic within a run}
if (low<-2.0) and (mid<-2.0) then tsdw:=1;
if (low>2.0) and (mid>2.0) then tsdw:=1;
if (low<-2.0) and (high<-2.0) then tsdw:=1;
if (low>2.0) and (high>2.0) then tsdw:=1;
if (high<-2.0) and (mid<-2.0) then tsdw:=1;
if (high>2.0) and (mid>2.0) then tsdw:=1;

{2-2 sd checking logic across runs}
if (low<-2.0) then blta:=blta+1
  else blta:=0;
if (low>2.0) then alta:=alta+1
  else alta:=0;
if (mid<-2.0) then bmta:=bmta+1
  else bmta:=0;
if (mid>2.0) then amta:=amta+1
  else amta:=0;
if (high<-2.0) then bhta:=bhta+1
  else bhta:=0;
if (high>2.0) then ahta:=ahta+1
  else ahta:=0;

{3 1-sd checking logic within a run}
if (low<-1.0) and (mid<-1.0) and (high<-1.0) then bfw:=1;
if (low>1.0) and (mid>1.0) and (high>1.0) then bfw:=1;

{4 1-sd checking logic across runs}
if (low<-1.0) then blfa:=blfa+1
  else blfa:=0;
if (low>1.0) then alfa:=alfa+1
  else alfa:=0;
if (mid<-1.0) then bmfa:=bmfa+1
  else bmfa:=0;
If (mid>1.0) then amfa:=amfa+1
  else amfa:=0;
if (high<-1.0) then bhfa:=bhfa+1
  else bhfa:=0;
if (high>1.0) then ahfa:=ahfa+1
  else ahfa:=0;
```

```
{9 xbar within}
if (low<0.0) and (mid<0.0) and (high<0.0) then btenw:=btenw+3
  else btenw:=0;
if (low>0.0) and (mid>0.0) and (high>0.0) then atenw:=atenw+3
  else atenw:=0;

{9 xbar across runs}
if (low<=0.0) then bltena:=bltena+1
  else bltena:=0;
if (low>0.0) then altena:=altena+1
  else altena:=0;
if (mid<=0.0) then bmtena:=bmtena+1
  else bmtena:=0;
if (mid>0.0) then amtena:=amtena+1
  else amtena:=0;
if (high<=0.0) then bhtena:=bhtena+1
  else bhtena:=0;
if (high>0.0) then ahtena:=ahtena+1
  else ahtena:=0;

{Summary logic}
{If no 2 sd screen}
if (warn=2) then
 begin
   if (rult=1) and (tresd=1) then flag:=1;
   if (rulrfs=1) and (rfs=1) then flag:=1;
   if (rultw=1) and (tsdw=1) then flag:=1;
   if (rultw=1) and (blta>=2) then flag:=1;
   if (rultw=1) and (alta>=2) then flag:=1;
   if (rultw=1) and (bmta>=2) then flag:=1;
   if (rultw=1) and (amta>=2) then flag:=1;
   if (rultw=1) and (bhta>=2) then flag:=1;
   if (rultw=1) and (ahta>=2) then flag:=1;
   if (rulf=1) and (bfw=1) then flag:=1;
   if (rulf=1) and (blfa=4) then flag:=1;
   if (rulf=1) and (alfa=4) then flag:=1;
   if (rulf=1) and (bmfa=4) then flag:=1;
   if (rulf=1) and (amfa=4) then flag:=1;
   if (rulf=1) and (bhfa=4) then flag:=1;
   if (rulf=1) and (ahfa=4) then flag:=1;
   if (rulxb=1) and (btenw=9) then flag:=1;
   if (rulxb=1) and (atenw=9) then flag:=1;
   if (rulxb=1) and (bltena=9) then flag:=1;
   if (rulxb=1) and (altena=9) then flag:=1;
   if (rulxb=1) and (bmtena=9) then flag:=1;
```

```
      if (rulxb=1) and (amtena=9) then flag:=1;
      if (rulxb=1) and (bhtena=9) then flag:=1;
      if (rulxb=1) and (ahtena=9) then flag:=1;
     end;

   {Using 2 sd screen}
   if (warn=1) then
   begin
   if(low<-2) or (low>2) or (mid<-2) or (mid>2) or (high<-2) or (high>2) then
    begin
      if (rult=1) and (tresd=1) then flag:=1;
      if (rulrfs=1) and (rfs=1) then flag:=1;
      if (rultw=1) and (tsdw=1) then flag:=1;
      if (rultw=1) and (blta>=2) then flag:=1;
      if (rultw=1) and (alta>=2) then flag:=1;
      if (rultw=1) and (bmta>=2) then flag:=1;
      if (rultw=1) and (amta>=2) then flag:=1;
      if (rultw=1) and (bhta>=2) then flag:=1;
      if (rultw=1) and (ahta>=2) then flag:=1;
      if (rulf=1) and (bfw=1) then flag:=1;
      if (rulf=1) and (blfa=4) then flag:=1;
      if (rulf=1) and (alfa=4) then flag:=1;
      if (rulf=1) and (bmfa=4) then flag:=1;
      if (rulf=1) and (amfa=4) then flag:=1;
      if (rulf=1) and (bhfa=4) then flag:=1;
      if (rulf=1) and (ahfa=4) then flag:=1;
      if (rulxb=1) and (btenw=9) then flag:=1;
      if (rulxb=1) and (atenw=9) then flag:=1;
      if (rulxb=1) and (bltena=9) then flag:=1;
      if (rulxb=1) and (altena=9) then flag:=1;
      if (rulxb=1) and (bmtena=9) then flag:=1;
      if (rulxb=1) and (amtena=9) then flag:=1;
      if (rulxb=1) and (bhtena=9) then flag:=1;
      if (rulxb=1) and (ahtena=9) then flag:=1;
     end;
   end;
   {Summary statistics}
   rl:=rl+1;
  end;
until flag=1;

sum:=sum+rl;
sumsq:=sumsq+(rl*rl);
rl:=0;
flag:=0;
```

```
    tresd:=0; rfs:=0;  tsdw:=0;  blta:=0;  alta:=0;
    bmta:=0; amta:=0;  bfw:=0;  afw:=0;  blfa:=0;
    alfa:=0; bmfa:=0;  amfa:=0;  btenw:=0;  atenw:=0;
    bltena:=0;  altena:=0; bmtena:=0;amtena:=0;
    bhta:=0; ahta:=0; bhfa:=0; ahfa:=0; bhtena:=0; ahtena:=0;

  end;
  arl:=sum/m;
  sd:=sqrt((sumsq-(m*arl*arl))/(m-1));
  writeln('The ARL is ',arl:5:2);
  writeln('The ARLSD is ',sd:5:2);
  nsum:=nsum+arl;
  nsumsq:=nsumsq+(arl*arl);
  end;

{Final Output}
  writeln('Low Control Shift = ',lshft:5:2);
  writeln('Mid Control Shift = ',mshft:5:2);
  writeln('High Control Shift = ',hshft:5:2);
  writeln;
  farl:=nsum/5;
  fas:=sqrt((nsumsq-(5*farl*farl))/4);
  uci:=farl+2.776*(fas/(sqrt(5)));
  lci:=farl-2.776*(fas/(sqrt(5)));
  writeln('The ARL estimate is ',farl:5:2);
  writeln('The upper 95% CI is ',uci:5:2);
  writeln('The lower 95% CI is ',lci:5:2);
end.
```

APPENDIX C

Mathcad Worksheets for the χ^2 Chart

Integration of the Non-Central Chi-Sq for N'=2

$p := 2 \qquad c := 9.21 \qquad t := 12$

$$\Sigma := \begin{pmatrix} .0625 & .0675 \\ .0675 & .2916 \end{pmatrix} \qquad \mu := \begin{pmatrix} 0.75 \\ 0.0 \end{pmatrix} \qquad \mu^{T} \cdot \Sigma^{-1} \cdot \mu = 12$$

$$\int_{c}^{75} \left[\frac{1}{2^{\left(\frac{1}{2}\right)\cdot p}} \right] \cdot e^{\left(-\frac{1}{2}\right)\cdot(t+v)} \cdot v^{\left(\frac{1}{2}\right)\cdot p - 1} \cdot \sum_{\beta=0}^{25} \left(\frac{t}{4}\right)^{\beta} \cdot \left(\frac{1}{\beta! \cdot \Gamma\left(\frac{1}{2}\cdot p + \beta\right)} \right) \cdot v^{\beta} \, dv = 0.721$$

Integration of the Non-Central Chi-Sq for N'=3

$p := 3 \qquad c := 11.344 \qquad t := .074$

$$\Sigma := \begin{pmatrix} 0.0625 & 0.0135 & 0.019 \\ 0.0135 & 0.2916 & .041 \\ 0.019 & .041 & 0.5776 \end{pmatrix} \qquad \mu := \begin{pmatrix} 0.05 \\ .108 \\ 0 \end{pmatrix} \qquad \mu^{T} \cdot \Sigma^{-1} \cdot \mu = 0.074$$

$$\int_{c}^{75} \left[\frac{1}{2^{\left(\frac{1}{2}\right) \cdot p}} \right] \cdot e^{\left(-\frac{1}{2}\right) \cdot (t+v)} \cdot v^{\left(\frac{1}{2}\right) \cdot p - 1} \cdot \sum_{\beta = 0}^{25} \left(\frac{t}{4} \right)^{\beta} \cdot \left(\frac{1}{\beta! \cdot \Gamma\left(\frac{1}{2} \cdot p + \beta\right)} \right) \cdot v^{\beta} \, dv = 0.011$$

APPENDIX D

Mathcad Worksheets for the T^2 Chart

Integration of the Non-Central T for N'=2

$p := 2 \qquad N := 20 \qquad L := 0.0 \qquad U := 12.69 \qquad t := 2.613$

$$\Sigma := \begin{pmatrix} 0.0625 & 0.0675 \\ 0.0675 & 0.2916 \end{pmatrix} \qquad \mu := \begin{pmatrix} 0.35 \\ 0.0 \end{pmatrix} \qquad \mu^{T} \cdot \Sigma^{-1} \cdot \mu = 2.613$$

$$\int_{L}^{12.69} \frac{e^{\left(-\frac{1}{2}\right)\cdot t}}{(N-1)\cdot\Gamma\left[\left(\frac{1}{2}\right)\cdot(N-p)\right]} \cdot \sum_{\beta=0}^{75} \frac{\left[\left[\left(\frac{t}{2}\right)^{\beta}\right]\cdot\left[\left[\frac{x}{(N-1)}\right]^{\left(\frac{1}{2}\right)\cdot p+\beta-1}\right]\cdot\Gamma\left[\left(\frac{1}{2}\right)\cdot N+\beta\right]\right]}{\beta!\cdot\Gamma\left[\left(\frac{1}{2}\right)\cdot p+\beta\right]\cdot\left[1+\left[\frac{x}{(N-1)}\right]\right]^{\left(\frac{1}{2}\right)\cdot N+\beta}} \, dx = 0.914$$

Integration of the Non-Central T for N'=3

$p := 3 \qquad N := 20 \qquad L := 0.0 \qquad U := 17.36 \qquad t := 13.5$

$$\Sigma := \begin{pmatrix} 0.0625 & 0.0675 & 0.095 \\ 0.0675 & 0.2916 & .2052 \\ 0.095 & .2052 & 0.5776 \end{pmatrix} \qquad \mu := \begin{pmatrix} 0.75 \\ 0.0 \\ 0.0 \end{pmatrix} \qquad \mu^{T} \cdot \Sigma^{-1} \cdot \mu = 13.5$$

$$\int_{L}^{17.36} \frac{e^{\left(-\frac{1}{2}\right)\cdot t}}{(N-1)\cdot\Gamma\left[\left(\frac{1}{2}\right)\cdot(N-p)\right]} \cdot \sum_{\beta=0}^{75} \frac{\left[\left[\left(\frac{t}{2}\right)^{\beta}\right]\cdot\left[\left[\frac{x}{(N-1)}\right]^{\left(\frac{1}{2}\right)\cdot p+\beta-1}\right]\cdot\Gamma\left[\left(\frac{1}{2}\right)\cdot N+\beta\right]\right]}{\beta!\cdot\Gamma\left[\left(\frac{1}{2}\right)\cdot p+\beta\right]\cdot\left[1+\left[\frac{x}{(N-1)}\right]\right]^{\left(\frac{1}{2}\right)\cdot N+\beta}} \, dx = 0.478$$

APPENDIX E

Mathcad Worksheets for the Principal Components Chart

Principal Component Calculations for N'=2

$$\Sigma := \begin{pmatrix} 0.0625 & 0.0135 \\ 0.0135 & .2916 \end{pmatrix}$$

$$\text{eigenvals}(\Sigma) = \begin{pmatrix} 0.06171 \\ 0.29239 \end{pmatrix}$$

$$\text{eigenvec}(\Sigma, 0.06171) = \begin{pmatrix} 0.99828 \\ -0.05862 \end{pmatrix} \qquad \text{eigenvec}(\Sigma, 0.29239) = \begin{pmatrix} 0.05862 \\ 0.99828 \end{pmatrix}$$

$$u := \begin{pmatrix} 0.99828 \\ -0.05862 \end{pmatrix} \qquad v := \begin{pmatrix} 0.05862 \\ 0.99828 \end{pmatrix} \qquad \mu := \begin{pmatrix} 0.75 \\ 0.0 \end{pmatrix}$$

$l := 0.06171$

$m := 0.29239$

$$\frac{u^T \cdot \mu}{\sqrt{l}} = 3.01395$$

$$\frac{v^T \cdot \mu}{\sqrt{m}} = 0.08131$$

$z := 3.01395$

$x := .08131$

$2.81 - z = -0.2$

$2.81 - x = 2.73$

$-2.81 - z = -5.82$

$-2.81 - x = -2.89$

Principal Component Calculations for N'=3

$$\Sigma := \begin{pmatrix} 0.0625 & 0.0135 & 0.019 \\ 0.0135 & 0.2916 & 0.041 \\ 0.019 & 0.041 & 0.5776 \end{pmatrix}$$

$$\text{eigenvals}(\Sigma) = \begin{pmatrix} 0.06117 \\ 0.58418 \\ 0.28635 \end{pmatrix}$$

$$\text{eigenvec}(\Sigma, 0.06117) = \begin{pmatrix} 0.99808 \\ -0.05268 \\ -0.03254 \end{pmatrix} \quad \text{eigenvec}(\Sigma, 0.58418) = \begin{pmatrix} 0.03967 \\ 0.14046 \\ 0.98929 \end{pmatrix} \quad \text{eigenvec}(\Sigma, 0.28635) = \begin{pmatrix} -0.04755 \\ -0.98868 \\ 0.14228 \end{pmatrix}$$

$$u := \begin{pmatrix} 0.99808 \\ -0.05268 \\ -0.03254 \end{pmatrix} \qquad x := \begin{pmatrix} 0.03967 \\ 0.14046 \\ 0.98929 \end{pmatrix} \qquad y := \begin{pmatrix} -0.04755 \\ -0.98868 \\ 0.14228 \end{pmatrix} \qquad \mu := \begin{pmatrix} 0.75 \\ 0.0 \\ 0.0 \end{pmatrix}$$

$l := 0.06117$ $\qquad m := 0.58418$ $\qquad n := 0.28635$

$$\frac{u^T \cdot \mu}{\sqrt{l}} = 3.02662 \qquad \frac{x^T \cdot \mu}{\sqrt{m}} = 0.03893 \qquad \frac{y^T \cdot \mu}{\sqrt{n}} = -0.06664$$

$a := 3.02662$ $\qquad b := .03893$ $\qquad c := .06664$

$2.935 - a = -0.09$ $\qquad 2.935 - b = 2.9$ $\qquad 2.935 - c = 2.87$

$-2.935 - a = -5.96$ $\qquad -2.935 - b = -2.97$ $\qquad -2.935 - c = -3$

APPENDIX F

Computer Code for programs CostA and CostB

```
program costa;
uses crt,printer;
var  i,n,r,j,cause,flag,num : integer;
     seed : longint;
     efr,ttfr,tottfr,sfr,ttns : real;
     x,h,time,fr,fix,min,tfr : real;
     total, over, shift, blamb, inter, tshift : real;
     totns, ens, tns, d, td, dtot, dfin, tct : real;
     dumb, gen, flagi, catch, tfix, ttds, totfin : real;
     mint, mintot,tic,tsfr,tcsh : real;
     ffin, fin, parl, iparl, psrt, ipsrt, tottds, fttds,thetds : real;
     psosc,pooc,ptlfr,ptltr,pdt,sc,tcost,cone,ctre : real;
     gtic,gtsfr,gtcsh,gfix,gfin,spsosc,spooc,stlfr,stltr,spdt : real;
     lambda,arl,srt,p,a,b,pfact : array[1..4] of real;
     fail,ctwo : array[1..3] of real;

     {This version of the cost model generates a single minimum failure
     from the three failure rates}

function RandUnif : real;

    {Function RandUnif uses the Marse-Roberts code for generating
    random uniform 0,1 variates}

const
  B2E15 =  32768;
  B2E16 =  65536;
  Modulus = 2147473647;
  Mult1 =  24112;
  Mult2 =  26143;

var  Hi15, Hi31, Low15, Lowprd, Ovflow,Zi : longint;

begin
  Zi := Seed;
  Hi15 := Zi DIV B2E16;
  Lowprd := (Zi - Hi15 * B2E16) * Mult1;
  Low15 := Lowprd DIV B2E16;
  Hi31 := Hi15 * Mult1 + Low15;
  Ovflow := Hi31 DIV B2E15;
  Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
       (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
  IF Zi < 0 THEN Zi := Zi + Modulus;
  Hi15 := Zi DIV B2E16;
  Lowprd := (Zi - Hi15 * B2E16) * Mult2;
```

```
  Low15 := Lowprd DIV B2E16;
  Hi31 := Hi15 * Mult2 + Low15;
  Ovflow := Hi31 DIV B2E15;
  Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
       (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
  IF Zi < 0 THEN Zi := Zi + Modulus;
  Seed := Zi;
  RandUnif := (2* (Zi DIV 256) + 1) / 16777216.0;

end;

begin
  clrscr;
  total:=0;  tottfr:=0;   totns:=0;    dtot:=0; totfin:=0; tottds:=0;
  gtic:=0; gtcsh:=0; gtsfr:=0; gfix:=0; gfin:=0;

  {Upfront information entry}

  write('Enter the 3 failure rates: ');
  readln( lambda[1], lambda[2], lambda[3]);
  writeln;
  write('Enter the 3 arls for the corresponding failure rates: ');
  readln( arl[1], arl[2], arl[3]);
  writeln;
  write('Enter SRT a and b for the three failures respectively: ');
  for i:= 1 to 3 do
    readln(a[i],b[i]);
  writeln;
  write('Enter SRT a and b for a false alarm: ');
  readln(a[4],b[4]);
  writeln;
  write('Enter the in-control arl: ');
  readln(arl[4]);
  writeln;
  write('Enter the sampling interval (h): ');
  readln(h);
  writeln;
  write('Enter the seed for the random number generator: ');
  readln(Seed);
  writeln;
  write('Enter the number of realizations for the simulation: ');
  readln(r);
  writeln;

  {Simulation Section of the program}
```

```
for j:= 1 to r do
 begin
  time:=0;  flag:=0;  shift:=0; min:=0; ttfr:=0; ttns:=0; tfr:=0;
  tic:=0; tsfr:=0;

  {This portion of code generates the minimum failure time of the
  three failure rates}

  for i:= 1 to 4 do
    pfact[i]:=ln(1-(1/arl[i]));
  for i:= 1 to 3 do
    fail[i]:= -1*(1/lambda[i])*ln(Randunif);
  if (fail[1]<fail[2]) and (fail[1]<fail[3]) then
    begin
     min:=fail[1];
     cause:=1;
    end;
  if (fail[2]<fail[1]) and (fail[2]<fail[3]) then
    begin
      min:=fail[2];
      cause:=2;
    end;
  if (fail[3]<fail[1]) and (fail[3]<fail[2]) then
    begin
      min:=fail[3];
      cause:=3;
    end;

  repeat

   {This portion of code compares the failure time with a generated
   time to false rejection.  The comparison continues until the
   failure time is reached.}

   begin
    fr:=(trunc((ln(randunif))/pfact[4])+1);
    tfr:=time+((fr)*h);   {ttns taken out}
    if tfr<min then
      begin
       time:=tfr;
       tic:=tic+time;
       fix:=a[4]+(b[4]-a[4])*Randunif;
       tsfr:=tsfr+fix;
       ttfr:=ttfr+1;
```

```
      time:=time+fix;
      min:=min+fix;
      ttns:=h-(h*((fix/h)-trunc((fix/h))));
     end
   else
    begin
     shift:=min;
     flag:=1;
    end;
end;
until flag=1;

d:=(min-time)-(trunc((min-time)/h)*h);


{This portion of code determines the time to fix the failure
given which of 3 failures has occurred}

if cause=1 then
 begin
  catch:=(trunc((ln(randunif))/pfact[1])+1);
  tfix:=a[1]+(b[1]-a[1])*randunif;
 end;
if cause=2 then
 begin
  catch:=(trunc((ln(randunif))/pfact[2])+1);
  tfix:=a[2]+(b[2]-a[2])*randunif;
 end;
if cause=3 then
 begin
  catch:=(trunc((ln(randunif))/pfact[3])+1);
  tfix:=a[3]+(b[3]-a[3])*randunif;
 end;

ttds:=(min-d) + (catch*h);
tcsh:=ttds-min;
fin:=ttds+tfix;

{This portion of code collects information on the amount
of time the system is in a given state}

total:=total+shift;
tottfr:=tottfr+ttfr;
dtot:=dtot+d;
totfin:=totfin+fin;
```

```
    gtic:=gtic+(min-tsfr);
    gtsfr:=gtsfr+tsfr;
    gtcsh:=gtcsh+tcsh;
    gfix:=gfix+tfix;
    gfin:=gfin+fin;

  end;

  over:=total/r;
  sfr:=tottfr/r;
  dfin:=dtot/r;
  ffin:=totfin/r;
  spsosc:=gtic/gfin;
  spooc:=gtcsh/gfin;
  stlfr:=gtsfr/gfin;
  stltr:=gfix/gfin;
  spdt:=stlfr+stltr;

 {theory section }

 iparl:=0;  ipsrt:=0;
 blamb:=lambda[1]+lambda[2]+lambda[3];
 inter:=exp((-1*blamb)*h);
 efr:= (inter/(1-inter))/arl[4];
 tshift:=(1/blamb)+((inter/(1-inter))/arl[4])*((a[4]+b[4])/2);
 tns:=(inter/(1-inter));
 td:=(1-(1+blamb*h)*inter)/(blamb*(1-inter));
 for i:= 1 to 3 do
   srt[i]:=(a[i]+b[i])/2;
 for i:= 1 to 3 do
   p[i]:=lambda[i]/blamb;
 parl:=p[1]*arl[1]+p[2]*arl[2]+p[3]*arl[3];
 psrt:=p[1]*srt[1]+p[2]*srt[2]+p[3]*srt[3];
 thetds:=h*parl;
 tct:=tshift+(h*parl)-td+psrt;
 psosc:=(1/blamb)/tct;
 pooc:=(h*parl-td)/tct;
 ptlfr:=(efr*((a[4]+b[4])/2))/tct;
 ptltr:=psrt/tct;
 pdt:=ptlfr+ptltr;
{  sc:=(num*cfour)/h;
 tcost:=psosc*cone+pooc*(p[1]*ctwo[1]+p[2]*ctwo[2]+p[3]*ctwo[3])+pdt*ctre+sc;}

 {The following section prints out the comparison between the simulation
 and theoretical resuls}
```

```
    writeln('The simulation average time to shift is ',over:4:2);
    writeln('The theoretical average time to shift is ',tshift:4:2);
    writeln('The simulation expected num fr is ',sfr:4:2);
    writeln('The theoretical exp. num fr is ',efr:4:2);
    writeln('The simulation d is ',dfin:4:2);
    writeln('The theoretical d is ',td:4:2);
    writeln('The simulation tct is ',ffin:4:2);
    writeln('The theoretical tct is ',tct:4:2);
    writeln('The simulation percent sosc is ',spsosc:4:3);
    writeln('The theoretical percent sosc is ',psosc:4:3);
    writeln('The simulation percent ooc is ',spooc:4:3);
    writeln('The theoretical percent ooc is ',pooc:4:3);
    writeln('The simulation percent time looking for fr is ',stlfr:4:3);
    writeln('The theoretical percent time looking for fr is ',ptlfr:4:3);
    writeln('The simulation percent time looking for tr is ',stltr:4:3);
    writeln('The theoretical percent time looking for tr is ',ptltr:4:3);
    writeln('The simulation percent down time is ',spdt:4:3);
    writeln('The theoretical percent down time is ',pdt:4:3);

end.
```

```
program costb;
uses crt,printer;
var  i,n,r,j,cause,flag,num : integer;
     seed : longint;
     efr,ttfr,tottfr,sfr,ttns : real;
     x,h,time,fr,fix,min,tfr : real;
     total, over, shift, blamb, inter, tshift : real;
     totns, ens, tns, d, td, dtot, dfin, tct : real;
     dumb, gen, flagi, catch, tfix, ttds, totfin : real;
     mint, mintot,tic,tsfr,tcsh : real;
     ffin, fin, parl, iparl, psrt, ipsrt, tottds, fttds,thetds : real;
     psosc,pooc,ptlfr,ptltr,pdt,sc,tcost,cone,ctre : real;
     gtic,gtsfr,gtcsh,gfix,gfin,spsosc,spooc,stlfr,stltr,spdt : real;
     lambda,arl,srt,p,a,b,pfact : array[1..4] of real;
     fail,ctwo : array[1..3] of real;

{This version of the cost model regenerates a failure following
each false rejection}

function RandUnif : real;

{Function RandUnif uses the Marse-Roberts code for generating a
uniform 0,1 random variate}

const
  B2E15 =   32768;
  B2E16 =   65536;
  Modulus = 2147473647;
  Mult1 =   24112;
  Mult2 =   26143;

var   Hi15, Hi31, Low15, Lowprd, Ovflow,Zi : longint;

begin
  Zi := Seed;
  Hi15 := Zi DIV B2E16;
  Lowprd := (Zi - Hi15 * B2E16) * Mult1;
  Low15 := Lowprd DIV B2E16;
  Hi31 := Hi15 * Mult1 + Low15;
  Ovflow := Hi31 DIV B2E15;
  Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
       (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
  IF Zi < 0 THEN Zi := Zi + Modulus;
  Hi15 := Zi DIV B2E16;
  Lowprd := (Zi - Hi15 * B2E16) * Mult2;
```

```
  Low15 := Lowprd DIV B2E16;
  Hi31 := Hi15 * Mult2 + Low15;
  Ovflow := Hi31 DIV B2E15;
  Zi := (((Lowprd - Low15 * B2E16) - Modulus) +
       (Hi31 - Ovflow * B2E15) * B2E16) + Ovflow;
  IF Zi < 0 THEN Zi := Zi + Modulus;
  Seed := Zi;
  RandUnif := (2* (Zi DIV 256) + 1) / 16777216.0;

end;

begin
  clrscr;
  total:=0;  tottfr:=0;   totns:=0;    dtot:=0; totfin:=0; tottds:=0;
  gtic:=0; gtcsh:=0; gtsfr:=0; gfix:=0; gfin:=0;

  {Upfront information entry}

  write('Enter the 3 failure rates: ');
  readln( lambda[1], lambda[2], lambda[3]);
  writeln;
  write('Enter the 3 arls for the corresponding failure rates: ');
  readln(arl[1], arl[2], arl[3]);
  writeln;
  write('Enter SRT a and b for the three failures respectively: ');
  for i:= 1 to 3 do
    readln(a[i],b[i]);
  writeln;
  write('Enter SRT a and b for a false alarm: ');
  readln(a[4],b[4]);
  writeln;
  write('Enter the in-control arl: ');
  readln(arl[4]);
  writeln;
  write('Enter the sampling interval (h): ');
  readln(h);
  writeln;
  write('Enter the seed for the random number generator: ');
  readln(Seed);
  writeln;
  write('Enter the number of realizations for the simulation: ');
  readln(r);
  writeln;

  {Simulation Section of the program}
```

```
for j:= 1 to r do
 begin
  time:=0;  flag:=0;  shift:=0; min:=0; ttfr:=0; tic:=0; tsfr:=0;

  {This portion of code generates 3 failures times plus a time to
  false rejection. It takes the minimum of the four and continues
  generating failures untial a failure occurs before a false rejection.}

  for i:= 1 to 4 do
    pfact[i]:=ln(1-(1/arl[i]));
  repeat
   begin
    for i:= 1 to 3 do
      fail[i]:= -1*(1/lambda[i])*ln(Randunif);
    fr:=(trunc((ln(randunif))/pfact[4])+1)*h;

  if (fail[1]<fail[2]) and (fail[1]<fail[3])  and (fail[1]<fr) then
    begin
     shift:=time+fail[1];
     cause:=1;
     tic:=tic+fail[1];
     flag:=1;
    end;
  if (fail[2]<fail[1]) and (fail[2]<fail[3]) and (fail[2]<fr) then
    begin
      shift:=time+fail[2];
      cause:=2;
      tic:=tic+fail[2];
      flag:=1;
    end;
  if (fail[3]<fail[1]) and (fail[3]<fail[2]) and (fail[3]<fr) then
    begin
      shift:=time + fail[3];
      cause:=3;
      tic:=tic+fail[3];
      flag:=1;
    end;
  if(fr<fail[1]) and (fr<fail[2]) and (fr<fail[3]) then
    begin
     fix:=a[4]+(b[4]-a[4])*Randunif;
     tsfr:=tsfr+fix;
     time:=time+fix+fr;
     tic:=tic+fr;
     ttfr:=ttfr+1;
```

```
    end;

  end;
  until flag=1;

   d:=(shift-time)-(trunc((shift-time)/h)*h);

   {This portion of code determines the time to fix the
   failure given which of the 3 failures has occurred.}

   if cause=1 then
    begin
     catch:=(trunc((ln(randunif))/pfact[1])+1);
     tfix:=a[1]+(b[1]-a[1])*randunif;
    end;
   if cause=2 then
    begin
     catch:=(trunc((ln(randunif))/pfact[2])+1);
     tfix:=a[2]+(b[2]-a[2])*randunif;
    end;
   if cause=3 then
    begin
     catch:=(trunc((ln(randunif))/pfact[3])+1);
     tfix:=a[3]+(b[3]-a[3])*randunif;
    end;

   ttds:=(shift-d) + (catch*h);
   tcsh:=ttds-shift;
   fin:=ttds+tfix;

  {This portion of code collects information on the amount of
  time the system is in a given state.}

  total:=total+shift;
  tottfr:=tottfr+ttfr;
  dtot:=dtot+d;
  totfin:=totfin+fin;
  gtic:=gtic+tic;
  gtsfr:=gtsfr+tsfr;
  gtcsh:=gtcsh+tcsh;
  gfix:=gfix+tfix;
  gfin:=gfin+fin;

end;
```

```
  over:=total/r;
  sfr:=tottfr/r;
  dfin:=dtot/r;
  ffin:=totfin/r;
  spsosc:=gtic/gfin;
  spooc:=gtcsh/gfin;
  stlfr:=gtsfr/gfin;
  stltr:=gfix/gfin;
  spdt:=stlfr+stltr;


{Theory section of the program}

iparl:=0;  ipsrt:=0;
blamb:=lambda[1]+lambda[2]+lambda[3];
inter:=exp((-1*blamb)*h);
efr:= (inter/(1-inter))/arl[4];
tshift:=(1/blamb)+((inter/(1-inter))/arl[4])*((a[4]+b[4])/2);
tns:=(inter/(1-inter));
td:=(1-(1+blamb*h)*inter)/(blamb*(1-inter));
for i:= 1 to 3 do
   srt[i]:=(a[i]+b[i])/2;
for i:= 1 to 3 do
   p[i]:=lambda[i]/blamb;
parl:=p[1]*arl[1]+p[2]*arl[2]+p[3]*arl[3];
psrt:=p[1]*srt[1]+p[2]*srt[2]+p[3]*srt[3];
thetds:=h*parl;
tct:=tshift+(h*parl)-td+psrt;
psosc:=(1/blamb)/tct;
pooc:=(h*parl-td)/tct;
ptlfr:=(efr*((a[4]+b[4])/2))/tct;
ptltr:=psrt/tct;
pdt:=ptlfr+ptltr;
{  sc:=(num*cfour)/h;
tcost:=psosc*cone+pooc*(p[1]*ctwo[1]+p[2]*ctwo[2]+p[3]*ctwo[3])+pdt*ctre+sc;}

{The following section prints out the comparison between the
simulation and theoretical results.}

  writeln('The simulation average time to shift is ',over:4:2);
  writeln('The theoretical average time to shift is ',tshift:4:2);
  writeln('The simulation expected number of false rejects is ',sfr:4:2);
  writeln('The theoretical expected number of false rejects is ',efr:4:2);
  writeln('The simulation d is ',dfin:4:2);
  writeln('The theoretical d is ',td:4:2);
```

```
    writeln('The simulation total cycle time (tct) is ',ffin:4:2);
    writeln('The theoretical total cycle time (tct) is ',tct:4:2);
    writeln('The simulation percent sosc is ',spsosc:4:3);
    writeln('The theoretical percent sosc is ',psosc:4:3);
    writeln('The simulation percent ooc is ',spooc:4:3);
    writeln('The theoretical percent ooc is ',pooc:4:3);
    writeln('The simulation percent time looking for false rejects is ',stlfr:4:3);
    writeln('The theoretical percent time looking for false rejects is ',ptlfr:4:3);
    writeln('The simulation percent time looking for true rejects is ',stltr:4:3);
    writeln('The theoretical percent time looking for true rejects is ',ptltr:4:3);
    writeln('The simulation percent down time is ',spdt:4:3);
    writeln('The theoretical percent down time is ',pdt:4:3);

end.
```

APPENDIX G

Computer Code for Theoretical Cost Model

```
program costc;
uses crt,printer;
var  i,n,r,j : integer;
    efr,ttfr,tottfr,sfr,ttns,thetds,num : real;
    x,hr, h, blamb, inter, tshift, ens, tns,  td, tct : real;
    dumb, gen, flagi, catch, parl, iparl, psrt, ipsrt  : real;
    psosc,pooc,ptlfr,ptltr,pdt,sc,tcost,cone,ctre,cfour : real;
    lambda,arl,srt,p,pfact : array[1..4] of real;
    ctwo : array[1..3] of real;

{This program contains only the theoretical cost model results without any simulation}

begin
 clrscr;

 {Upfront information entry}

 write('Enter the 3 failure rates: ');
 readln( lambda[1], lambda[2], lambda[3]);
 writeln;
 write('Enter the 3 arls for the corresponding failure rates: ');
 readln(arl[1], arl[2], arl[3]);
 writeln;
 write('Enter expected SRTs for the three failures respectively: ');
 for i:= 1 to 3 do
  readln(srt[i]);
 writeln;
 write('Enter the expected SRT for a false alarm: ');
 readln(srt[4]);
 writeln;
 write('Enter the in-control arl: ');
 readln(arl[4]);
 writeln;
 write('Enter the sampling interval (h): ');
 readln(h);
 writeln;
 write('Enter cost factor C1: ');
 readln(cone);
 writeln;
 write('Enter the three respective costs for C2: ');
 for i:= 1 to 3 do
   readln(ctwo[i]);
 writeln;
 write('Enter the cost factor C3: ');
 readln(ctre);
```

```
  writeln;
  write('Enter the cost factor C4: ');
  readln(cfour);
  writeln;
  write('Enter the sample size, N: ');
  readln(num);
  writeln;

  {Theory calculation section of the program}

  iparl:=0;  ipsrt:=0;
  blamb:=lambda[1]+lambda[2]+lambda[3];
  inter:=exp((-1*blamb)*h);
  efr:= (inter/(1-inter))/arl[4];
  tshift:=(1/blamb)+((inter/(1-inter))/arl[4])*(srt[4]);
  tns:=(inter/(1-inter));
  td:=(1-(1+blamb*h)*inter)/(blamb*(1-inter));
  for i:= 1 to 3 do
    p[i]:=lambda[i]/blamb;
  parl:=p[1]*arl[1]+p[2]*arl[2]+p[3]*arl[3];
  psrt:=p[1]*srt[1]+p[2]*srt[2]+p[3]*srt[3];
  thetds:=h*parl;
  tct:=tshift+(h*parl)-td+psrt;
  psosc:=(1/blamb)/tct;
  pooc:=(h*parl-td)/tct;
  ptlfr:=(efr*(srt[4]))/tct;
  ptltr:=psrt/tct;
  pdt:=ptlfr+ptltr;
  sc:=(num*cfour)/h;
  tcost:=psosc*cone+pooc*(p[1]*ctwo[1]+p[2]*ctwo[2]+p[3]*ctwo[3])+pdt*ctre+sc;

  {The following section prints out the results.}

    writeln('The theoretical percent sosc is ',psosc:4:3);
    writeln('The theoretical percent ooc is ',pooc:4:3);
    writeln('The theoretical percent time looking for false rejects is ',ptlfr:4:3);
    writeln('The theoretical percent time looking for true rejects is ',ptltr:4:3);
    writeln('The theoretical percent down time is ',pdt:4:3);
    writeln('The total cost per unit time is ',tcost:4:3);
end.
```

VITA

Jerry Dechert

Candidate for the Degree of

Doctor of Philosophy

Thesis: NEW APPROACHES FOR MONITORING QUALITY CONTROL IN THE CLINICAL LABORATORY

Major Field: Industrial Engineering

Biographical:

Personal Data: Born in Riverton, Wyoming, on July 19, 1968, the son of Lloyd and Annette Dechert.

Education: Graduated from Shoshoni High School, Shoshoni, Wyoming in May 1986. Received a Bachelor of Science degree in Industrial Engineering from Oklahoma State University, Stillwater, Oklahoma in May 1990 and completed a Master of Science in Industrial Engineering at Oklahoma State University, Stillwater, Oklahoma in May 1992. Completed the requirements for the Doctor of Philosophy degree with a major in Industrial Engineering and a minor in Statistics at Oklahoma State University in December, 1996.

Experience: Employed as an Industrial Engineer for Abbott Laboratories at Abbott Park, Illinois, from August 1992 through August 1993. Employed by Oklahoma State University, School of Industrial Engineering and Management as a graduate research assistant, August 1993 to present.

Professional Memberships: Institute of Industrial Engineers, American Society for Quality Control.