Results of Statistical Analysis of Pressure Relief Valve Proof Test Data Designed to Validate a Mechanical Parts Failure Database

by

Julia V. Bukowski, Ph. D.

for

exida.com L.L.C.

September 30, 2007

TABLE OF CONTENTS

Executive Summary	1
1. Introduction	2
2. Notation	3
3. Background	3
4. FMEDA Analysis	4
5. Analysis of Data Set A	4
6. Analysis of Data Set B	9
7. Analysis of Data Set C	11
8. Discussion	13
References	15
Acknowledgements	16
Appendix	16

EXECUTIVE SUMMARY

The purpose of this document is to report on our successful efforts to validate statistically certain random equipment failure rate data used in a mechanical parts failure rate and failure mode database and, by extension, to validate the techniques used to derive the data. To accomplish this, a Failure Modes, Effects, and Diagnostic Analysis (FMEDA) is initially used to predict the useful-life failure rate for the fail-to-open condition of a particular pressure relief valve (PRV) using the failure rates from the mechanical parts database. Next, this prediction is statistically tested against three independent data sets consisting of proof test data for PRV provided by Fortune 500 operating companies. The data sets all meet the intent of the quality assurance of proof test data as documented by the Center for Chemical Process Safety (CCPS) Process Equipment Reliability Database (PERD) initiative. By applying the quantal response method to the results of these PRV proof tests, it is demonstrated that the proof test data are consistent with the predictions of the FMEDA. Specifically, all of the data sets support the FMEDA result at a 95% confidence level. All analyses lead to a useful-life PRV failure rate between 10^{-8} and 10^{-7} failures/hour.

It is very important to note that the results of this study cannot be used to justify extension of proof test intervals beyond the useful life of the PRV. The small value of the failure rate derived from the FMEDA applies only to the useful life of the PRV which depends not only on the equipment's specifications but also on other factors, such as the ambient and process environment in which the PRV is used and the levels and frequency of any on-line maintenance performed. Data analyses place useful life in the range of 4 to 5 years.

Finally, we note that the results of the statistical analyses of the three independent data sets predict an initial failure probability of approximately 1% – 1.6%. This initial failure probability is extremely significant as it accounts for the vast majority of failures observed in proof test. This emphasizes the value of careful installation and thorough commissioning procedures. When commissioning testing cannot be done after installation, as is the case with a PRV, both the initial probability of failing to open, as well as the PFD based upon the random failure rate must be taken into account in the risk analysis. .

1. *INTRODUCTION*

Safety instrumented systems (SIS) are automatic systems designed for the purpose of taking action to avoid danger or to reduce the consequences of a potentially dangerous event. International performance-based standards [1, 2] require that designers of these systems use probabilistic analysis for equipment failures classified as "dangerous" to determine if any given design meets risk reduction goals. This is generally accomplished through an unavailability analysis. The analysis must incorporate all equipment needed for the automation system to protect against pre-identified hazards. Typical equipment includes electronic sensors, electronic signal conditioning modules, microcomputer controllers, relays, solenoids, pneumatic actuators and valves. The unavailability analysis requires, at a minimum, the useful-life failure rates and failure modes data for all subsystems.

For the electrical/electronics equipment, FMEDA techniques [3, 4] have been used to provide failure rates, failure mode distributions, and diagnostic self-test capability measures for subsystems based on extensive component failure rate and failure mode databases [5, 6]. In essence, these techniques compute a subsystem failure rate based on the failure rates and failure modes of the components which comprise the subsystem. These techniques rely on the existence of, and regulatory authorities' acceptance of, these part-level databases of failure rate and failure mode data that have been collected (over many years) from field failure data for a wide variety of electrical/electronics parts.

While it could be argued that a variety of mechanical failure models to predict failure rates exists, it is also true that most work on mechanical part failure models is focused on mechanical failure due to aging or wear-out. As a result, some published useful life failure rates for mechanical equipment are actually based upon data points more reflective of wear-out than random failures. This report deals with failures that principally represent random failures during the useful life of the equipment.

Until recently, a database of useful-life failure rates and failure modes for mechanical components comparable to those for random failures of electrical/electronic components has not existed. In [7] a technique is described for constructing a mechanical component random failure rate and mode database based on a combination of field failure data and expert knowledge. This database [8], if adequately validated, and coupled with end of useful life bounding limit data, would provide the mechanical component counterpart to the electrical/electronic component databases and allow FMEDA techniques to be applied to SIS containing both electrical and mechanical components in order to generate the information required at the subsystem level to comply with the new standards [1, 2].

In this report we describe the results of a FMEDA analysis of a particular PRV to determine the useful-life failure rate of the fail-to-open condition. (The fail-to-open condition occurs when a PRV remains closed when test pressures (TP) reach or exceed 1.5 times the PRV set pressure.) We then use three independent sources of proof test data to validate the predictions made by the FMEDA analysis. While this does not validate the entire mechanical part database, it lends strong support to its validity at least with respect to the component failure modes responsible for the PRV fail-to-open condition and, by extension, to the techniques used to create the database.

2. *NOTATION*

F(t)	cumulative distribution function for time to failure
FIT	1 failure/10^9 operating hours
FMEDA	failure modes, effects, and diagnostics analysis
PRV	pressure relief valve(s)
q_i	estimate for $F(T_i)$ based on failures in i^{th} data interval
R(0)	initial reliability; may be less than 1
R(t)	reliability function for t > 0
SIS	safety instrumented system(s)
SP	set pressure – pressure at which PRV should open
T_i	equivalent failure time associated with q_i
TP	test pressure – pressure required during proof test to cause PRV to open
λ	useful-life failure rate, a constant
$\lambda(t)$	failure rate as a function of time; λ may be a constant

3. *BACKGROUND*

In most SIS, some failure modes cannot be detected while the SIS is in operation. For example, in a PRV, the valve would normally be closed and would open only in the case of an overpressure event. If the valve were stuck in the closed position, this would be undetectable in operation unless an overpressure event occurred and the valve failed to open. While the ensuing injuries/damages would reveal the valve failure, it is preferable to discover the failure before an overpressure event occurs. The only way to uncover these otherwise undetectable failures is through proof tests. The PRV is removed from the process and pressurized on a test bench until the valve opens; the pressure needed to open the valve is the "test pressure" (TP). Each PRV has a "set pressure" (SP), a pressure above which the valve should open in normal operation. The ratio of TP/SP is recorded. If TP/SP $\geq$ 1.5, i.e., if the pressure required to open the PRV during testing is 50% or more above its set pressure, the valve is deemed to have "failed-to-open".

Proof tests are normally conducted during periodic inspection and maintenance. Thus, when an equipment failure is discovered during proof test, the actual time of the failure is not known. All that can be determined for certain is that the failure occurred sometime between the last proof test and the current proof test. As a consequence, the usual methods of time-to-failure analysis used to estimate failure rates cannot be used on proof test data. In the appendix, the reader will find a description of the method of quantal response analysis which is an appropriate analysis method to estimate failure rates from proof test data. We conclude this background section by noting that the proof test data, relevant to this analysis, for each valve consist of the number of in-service hours on the PRV from the last proof test to the current proof test along with an indicator of whether the valve, when tested, opened at a TP/SP < 1.5 or TP/SP $\geq$ 1.5 (in which case the valve failed-to-open).

4. *FMEDA ANALYSIS*

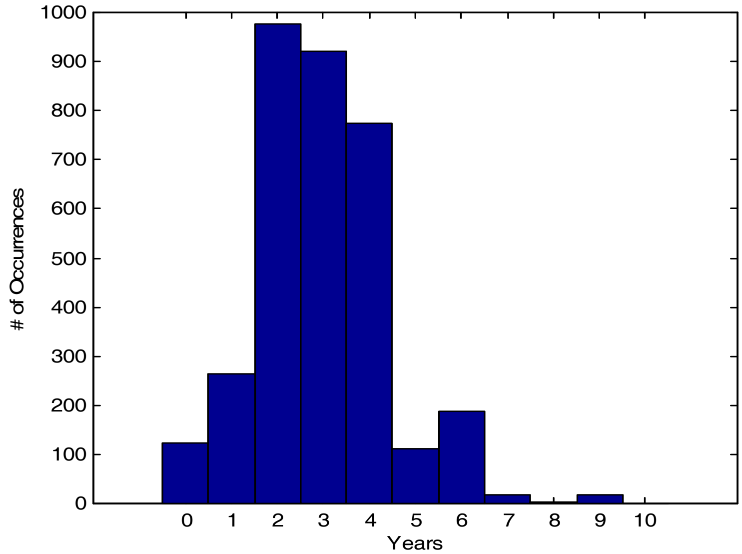
A FMEDA analysis was performed on a particular pressure relief valve containing 29 individual mechanical parts, many with multiple failure modes. Of a total of 43 identified failure modes, 5 lead to the fail-to-open condition. None of the fail-to-open conditions is detectable by means other than a proof test. Based on the failure rates for each mode from the mechanical component database to be validated [8], a useful-life failure rate for the fail-to-open condition was computed as 84 FITS.

5. *ANALYSIS OF DATA SET A*

5.1 *Description of Data Set A*

Data Set A consists of information on 1949 individual PRV that underwent a total of 3403 PRV proof tests which included the number of in-service hours since the previous proof test and an indication of whether the valve did or did not fail-to-open in the proof test. (The previous proof test may or may not be the commissioning proof test before initial installation.) The distribution of the proof test times is shown in the histogram in Figure 1.

Figure 1. Empirical Distribution of Proof Test Times for Data Set A



Among the 3403 proof tests, 48 resulted in fail-to-open. As indicated in the appendix, the quantal response analysis assumes the PRV tested are independently selected. In Data Set A, tests for a given time interval are often clustered in a single plant implying the possibility of similar environmental conditions, testing by the same technicians, and other conditions that might reasonably invalidate the assumption of independence of sampling. The data were carefully reviewed for any indications of possible dependencies. As a result, one data cluster of failures was eliminated from the analysis.

In this case, a single plant tested 66 PRV each with identical in-service hours (equivalent to 2.137 years) since the previous proof test. Of the 66 PRV, a total of 30 failures of all failure types were recorded; ten of these 30 failures were of type fail-to-open. Of these ten, nine were from a single manufacturer and involved three similar model numbers. Because this strong cluster suggested the possibility that these valve failures were not representative of independent data, the ten fail-to-open results were eliminated from the analysis. Thus a total of 38 fail-to-open proof test results were analyzed using the quantal response method. It is worth noting for later discussion that the 38 valves which failed-to-open at proof test had in-service hours since last proof test no greater than 5.01 years.

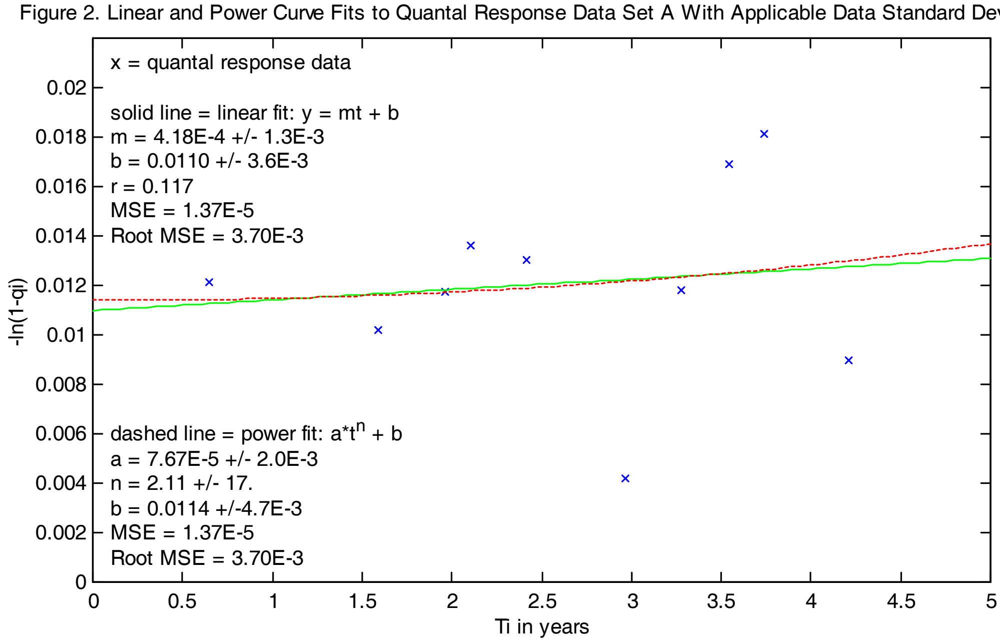
5.2 *Quantal Response Analysis of Data Set A*

After sorting the data in ascending order of in-service times since last proof test, ten non-overlapping time intervals were created usually with four or five failures per interval. (Due to the distribution of proof test times, a few intervals contained only two failures and one interval contained six failures). The cumulative failure probability at time T_i was estimated by q_i which was computed as (number of failures on the interval)/(total number of tests on the interval). T_i was computed as the average in-service hours on the valves found to be failed on the interval. Table 1 gives the results of these computations, i.e., T_i and q_i, i = 1, 2, …, 10, along with the values of $-\ln(1-q_i)$ which are needed for the quantal response analysis as explained in the appendix. Regression techniques were used to fit both a straight line and a power curve to the $-\ln(1-q_i)$ vs T_i data resulting in the plot shown in Figure 2. Note that there is little difference between the power fit and the linear fit. Also note that, as is explained in the appendix, in order for the results to be meaningful, both *a* and *b* must be greater than 0 (which that are). This guarantees that the associated cumulative distribution function for the time to failure, F(t), is monotonically non-decreasing.

Table 1. Quantal Response Data for Data Set A

T_i (years)	q_i	$-\ln(1-q_i)$
0.64658	0.01205	0.01212
1.58904	0.01015	0.01020
1.96301	0.01166	0.01173
2.10753	0.01351	0.01360
2.41699	0.01295	0.01303
2.96301	0.00416	0.00417
3.28082	0.01176	0.01183
3.54795	0.01674	0.01688
3.73863	0.01799	0.01815
4.20913	0.00894	0.00898

Figure 2. Linear and Power Curve Fits to Quantal Response Data Set A With Applicable Data Standard Deviations



5.3 *Interpretation of Quantal Response Analysis Results*

In order to validate the useful-life failure rate resulting from the FMEDA analysis, and, by extension, to validate portions of the mechanical failure database and the techniques used to derive them, it is necessary for the data analysis to test two different hypotheses. The first hypothesis tests whether the data supports the failure rate being constant. If the first hypothesis is accepted, the second hypothesis tests whether the useful-life failure rate could reasonably be 84 FITS.

First Hypothesis Test: Test whether the data $-\ln(1-q_i)$ vs T_i can be fit by a straight line. As is described in the appendix, fitting a curve to $-\ln(1-q_i)$ vs T_i models $\int \lambda(t)\,dt + b$, not $\lambda(t)$ itself. By using nonlinear regression analysis to fit a power curve of the form $at^n + b$ to the data and testing whether $n = 1$, it is possible to determine whether the data can reasonably be fitted by a straight line. This is equivalent to λ being a constant.

Specifically, test H_0: $n = 1$ vs H_1: $n \neq 1$. The point estimate for n, the power, is 2.11 (see Figure 2). The 95% confidence interval is calculated as the point estimate $\pm$ $t_{0.25,7}$ * standard deviation, where t comes from the student t distribution for $\alpha/2 = 0.25$ with 7 degrees of freedom (derived from 10 data points – 3 estimated parameters). Now $t_{0.25,7}$ equals 2.3646. Since $n = 1$ is included within a single standard deviation below the point estimate of 2.11, clearly, the 95% confidence interval must include $n = 1$ and we so state this without further calculation. Therefore the hypothesis H_0 is accepted. (Typically, statisticians avoid the statement that the null hypothesis is accepted; instead they prefer to say that H_0 cannot be rejected. From an engineering standpoint, not rejecting H_0 means that, based on the data, the hypothesis is plausible. Henceforth, when the hypothesis cannot be rejected, we say the hypothesis is accepted recognizing that this differs from standard statistical jargon.)

With $n = 1$, the power curve reduces to the simpler equation of a straight line, $at + b$ or, as more commonly written, $mt + b$. This means that the hypothesis that the data can be fitted by a straight line is accepted at the 95% confidence level. Now, if $\int \lambda(t)\,dt + b$ is a straight line, its derivative, which represents $\lambda(t)$, is equal to the slope of the line and is simply a constant. Therefore, the data does support the failure rate being a constant.

Second Hypothesis Test: The slope of the line, which is equivalent to the value of a constant failure rate, λ, equals 84 FITS. Test H_0: $\lambda = 84$ FITS vs H_1: $\lambda \neq 84$ FITS. Using least squares regression to fit the data with a straight line, gives a slope $m = 4.18 * 10^{-4}$ failures/year. Converting failures/year to failures/hour by dividing m by 8760 hours/year gives a value for $m = 47.7$ FITS. The upper bound of the 95% confidence interval (as calculated per Theorem 11.3.6 in [9]) is $3.31*10^{-3}$, and, therefore, the 95% confidence interval includes the value $m = 7.36*10^{-4}$ failures/year = 84 FITS. H_0 from the second hypothesis test is accepted at the 95% confidence level.

5.4 *Criticisms and Responses*

It may be noted that, for the linear least squares regression, the sample correlation coefficient, r, is only 0.117. Since r^2 (coefficient of determination) is usually interpreted as "the proportion of the total variation in the [-ln(1-qi)'s] that can be attributed to the linear relationship with [Ti]" [9], and since r is small, does this invalidate the analysis? Let us consider this issue as follows. If the constant failure rate λ is indeed 84 FITS, how many failures would we expect to see in five years among 3404 valve tests? (Recall all the discovered fail-to-open's were found in proof tests that occurred after in-services times of 5.01 years or less.) Let us begin by computing the expected number of random failures assuming that all the failures were uncovered by proof tests.

E(# failures in 5 years) =

3403 valves * [1-exp{-84*10^{-9} failures/hr*8760 hrs/year*5 years) = 12.5 failures (1)

However, any specific failure would be found by the proof test only if the proof test were conducted after the failure occurred. But as is clear in Figure 1, most of the proof tests occurred well before 5 years. Since the distribution of the proof test times is empirical, we used simulation to estimate how many useful-life failures would occur prior to proof test for this proof test distribution. We seeded useful-life failures at a constant failure rate of 100 FITS for 3400 valves with assigned proof test times that were the same as the proof test distribution in our data set. This is equivalent to assuming exponential times to failure with λ = 100 FITS. We found that about 60% of the seeded useful-life failures were discovered by proof tests while the remaining 40% were assigned failure times after the proof test would have occurred. Thus, we would expect that if the data corresponded to a failure rate of 84 FITS, we would have detected in proof tests about 12.5 * 60% = 7.5 fail-to-open conditions. But the data shows 38 failures!

However, as is explained in the appendix, the intercept, b, is related to the model's prediction of initial reliability of the valves when they are re-installed after proof test and this parameter plays an important role in our data fit. The parameter b must be greater than 0 for the results to be valid. When b is greater than zero, the initial reliability is less than 1. According to the model, the probability of initial failure is $1 – e^{-b}$. For our straight line fit, b is estimated to be 0.011 implying a probability of initial failure of 0.0109 with a 95% confidence interval of (0.002730712, 0.019197532) as calculated per Equation (11.3.5) in [9]. Now if about 7-8 failures found in proof test are attributable to useful-life failure, the remaining approximately 30 failures would have to be attributable to initial failure, and all of these failures would be discovered in proof test regardless of when the proof test is performed. This means about 30/3403 = 0.88 % initial failure probability, corresponding to a value of b = 0.008855 – a value which falls in the 95% confidence interval for b.

There is some evidence that commissioning tests on *new* PRV result in an initial failure probability of about 0.4% (see Data Set C below); therefore, the model estimate is reasonable. Given that we are trying to estimate a constant failure rate from data that contains both a very small constant failure rate and a significantly larger probability of initial failure, it is not surprising that the sample correlation coefficient is small. Our useful-life failure data is contained within the "noise" of data for probability of initial failure.

Finally, it may be noted that the FMEDA that produced a value of λ equal to 84 FITS was based on a single PRV model and manufacturer while the data analyzed here consists of PRV from many different manufacturers and includes many different models. Given the small magnitude of the useful-life failure rate, it would be unlikely that we would find sufficient data from a single manufacturer and single model to permit this type of analysis. Further, in [10], which reported a very different analysis of PRV proof test data not directed at estimating λ, it was noted that "no statistical differences [were found] in the average TP/SP [ratios] between manufacturers." Thus, it seems likely that our use of data from many manufacturers is justified.

6. *ANALYSIS OF DATA SET B*

6.1 *Description of Data Set B*

Data Set B consisted of 2578 PRV proof tests of which 57 resulted in fail-to-open. This data set was analyzed in [11] (where the emphasis was on the quantal response method and a data fit to F(t)) and the analysis produced 19 intervals with calculated values for q_i and T_i for each interval. The authors of [11] noted the model's prediction of an initial failure rate of about 1.5%. However, the paper did not estimate λ(t).

6.2 *Quantal Response Analysis of Data Set B and Interpretation of Results*

In [11], based on the data presented in Table 3 of that paper and the power curve fit to that data, it was appropriately concluded that a straight line fit is not supported by the data. The authors noted that "The probability of failure remains relatively constant for the first four years. It then experiences a steep increase." [11] However, recall that the constant failure rate holds only for the useful life of the PRV. Thus, if we seek a useful-life failure rate in the data of [11] we must confine our analysis to those data that represent that useful life region. Based on the observations made in [11], we assume that the first four years of data represent the appropriate data set. In Table 2 below, we repeat from [11] the 14 intervals, with their T_i and q_i, values that are likely based on the first four years of proof test data.

Table 2. Quantal Response Data for the First Four Years of Data Set B

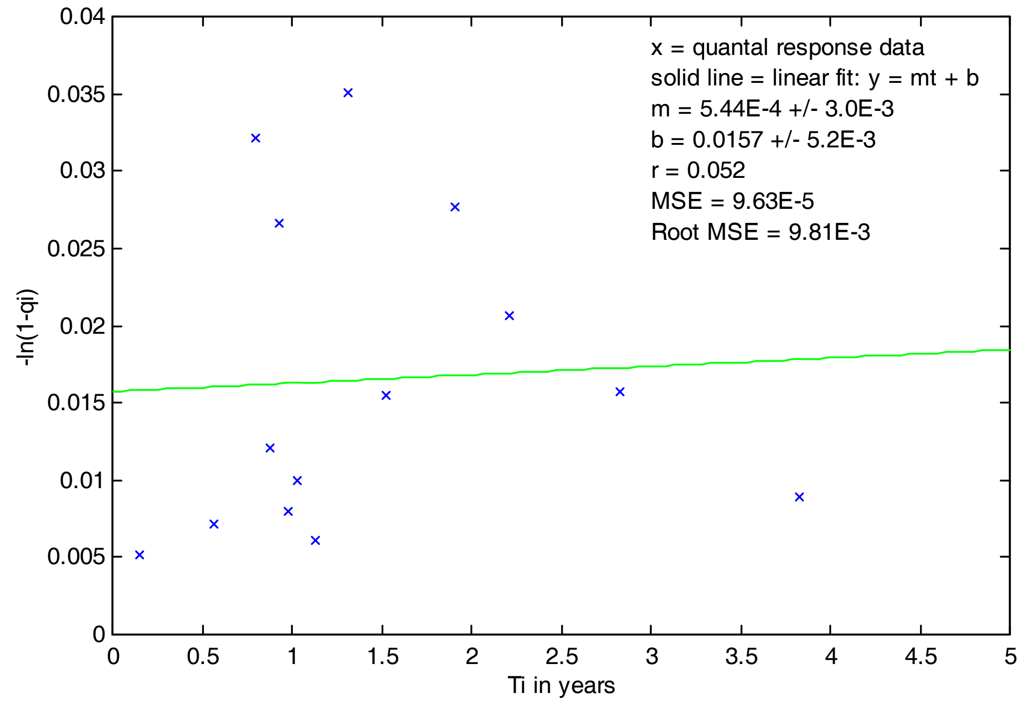
T_i (years)	q_i	$-\ln(1-q_i)$
0.15640	0.00518	0.005193
0.56599	0.00717	0.007196
0.79353	0.03165	0.032162
0.88086	0.01196	0.012032
0.93164	0.02632	0.026673
0.98177	0.00797	0.008002
1.02545	0.00996	0.010010
1.12899	0.00613	0.006149
1.31216	0.03448	0.035088
1.52560	0.01538	0.015499
1.91408	0.02732	0.027700
2.21384	0.02049	0.020703
2.82593	0.01563	0.015753
3.83031	0.00893	0.008970

In Figure 3, we plot $-\ln(1-q_i)$ vs T_i based on the 14 sets of values of T_i and q_i presented in Table 2. The power curve fitted to the first four years of data resulted in parameter values that were invalid given the theoretical constraints on $\lambda(t)$. Therefore, we were not able to confirm statistically the appropriateness of a straight line fit to the data as we had been able to do with Data Set A. Nevertheless, we did use least squares regression to fit a straight line to the first four years of data resulting in the line shown in Figure 3.

Test H_0: λ = 84 FITS vs H_1: $\lambda \neq$ 84 FITS. The fitted straight line has a slope m = $5.44*10^{-4}$ failures/year. Converting failures/year to failures/hour by dividing m by 8760 hours/year gives a value for m = 62 FITS. The upper bound of the 95% confidence interval (as calculated per Theorem 11.3.6 in [9]) is $7.11*10^{-3}$, and, therefore, the 95% confidence interval includes the value m = $7.36*10^{-4}$ failures/year = 84 FITS. H_0 is accepted at the 95% confidence level.

Because we did not have access to the raw data and, consequently, could not determine the number of failures lost in the elimination of data points beyond useful life, we could not more fully analyze the estimate for initial probability of failure. We do note that the estimate of $1 - e^{-1.57}$ = 1.56% for Data Set B is comparable to the estimates for initial probability of failure obtained from Data Sets A and C.

Figure 3. Least Squares Fit to Quantal Response Data Set B With Applicable Data Standard Deviations



7. ANALYSIS OF DATA SET C

7.1 *Description of Data Set C*

Data Set C consisted of 3282 PRV proof tests of which 24 resulted in fail-to-open [12]. Of these 3282 proof tests, 2377 were performed on *new* PRV prior to first time installation. Of these 2377 proof tests, 10 or 0.42% resulted in fail-to-open. The remaining 905 PRV proof tests consisted of 853 proof tests (accounting for 12 fail-to-open) of valves with in-service times less than 6 years and 52 proof tests (accounting for 2 fail-to-open) of valves with in-service times of greater than 6 years. In the analysis below, only those proof tests of new valves or valves with in-service time less than 6 years are included in the analysis because, beyond the six year time frame, the data likely represents service times beyond useful-life. Thus, the analysis below is based on 3230 proof tests of which 22 resulted in fail-to-open. Note that in [12], the original data set included an interval with no failures. This is not permitted in quantal response analysis so the interval containing no failures was combined with the next interval to produce the data in Table 3 at T_i = 0.82 years.

7.2 *Quantal Response Analysis of Data Set C and Interpretation of Results*

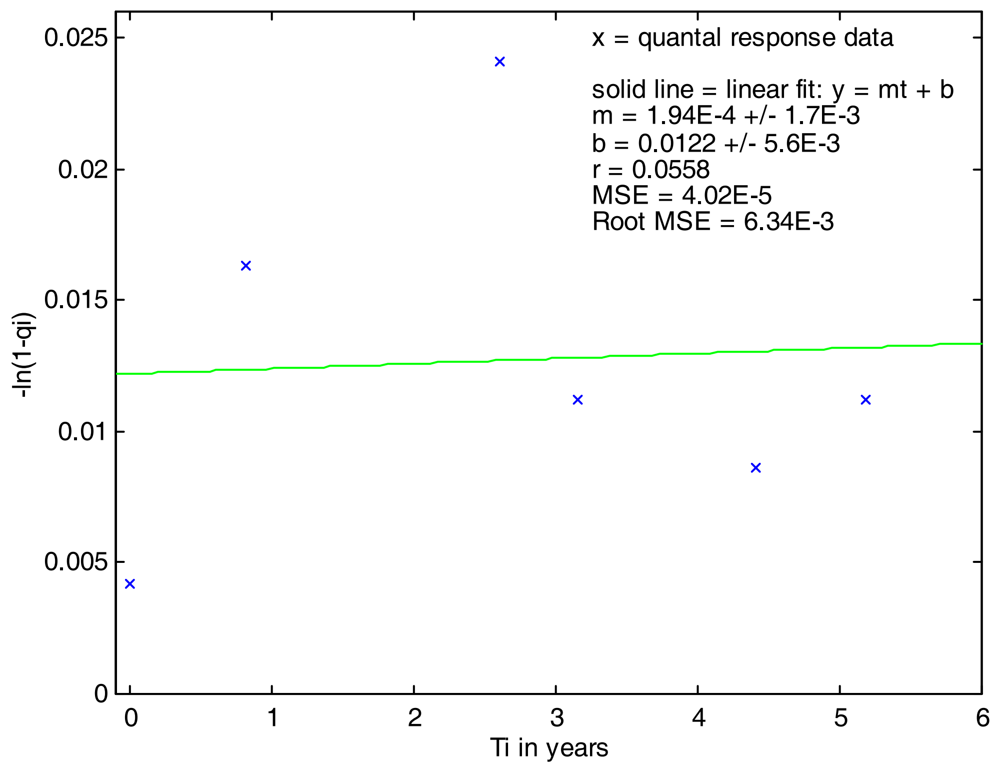
The quantal response analysis data for Data Set C is presented in Table 3.

Table 3. Quantal Analysis Data for Data Set C

T_i (years)	q_i	$-\ln(1-q_i)$
0	0.00421	0.0042
0.82	0.01613	0.0163
2.60	0.02410	0.02439
3.15	0.01119	0.01126
4.41	0.00862	0.00866
5.18	0.01117	0.01124

In Figure 4, we plot $-\ln(1-q_i)$ vs T_i based on the data presented in Table 3. The power curve fit of the data resulted in parameter values that were invalid given the theoretical constraints on $\lambda(t)$. Therefore, we were not able to confirm statistically the appropriateness of a straight line fit to the data as we had been able to do with Data Set A. Nevertheless, we did use least squares regression to fit a straight line to the data resulting in the line shown in Figure 4.

Figure 4. Linear Fit to Quantal Response Data Set C With Applicable Data Standard Deviation



Test H_0: λ = 84 FITS vs H_1: $\lambda \neq$ 84 FITS. The fitted straight line has a slope m = 1.94×10^{-4} failures/year. Converting failures/year to failures/hour by dividing m by 8760 hours/year gives a value for m = 22 FITS. The upper bound of the 95% confidence interval (as calculated per Theorem 11.3.6 in [9]) is $5.00*10^{-3}$, and, therefore, the 95% confidence interval includes the value m = $7.36*10^{-4}$ failures/year = 84 FITS. H_0 is accepted at the 95% confidence level.

Our data shows 12 failures on the 853 PRV proof tests subject to in-service time less than 6 years. (The 10 failures on new PRV are eliminated from this group because these PRV did not have any in-service time at the time of the proof test. Therefore, those failures are not part of the random failures during useful life.) According to our FMEDA, with λ = 84 FITS, among the 853 valves with in-service time to we would expect to see

E(# failures in 6 years) =

853 valves * [1-exp{-84*10^{-9} failures/hr*8760 hrs/year*6 years) = 3.77 failures (2)

due to random failure during useful life. For this data set, we do not know the distribution of proof test times and, therefore, are unable to accurately estimate the number of random failures that would be detected in proof test. Certainly it would not be all. If we use the 60% figure determined by simulation for Data Set A, then this would translate into about 2 observed failures attributable to random failure and the remaining 10 failures due to initial failure. This means about 10/853 = 1.17% initial failure probability, corresponding to a value for *b* of 0.0179. The 95% confidence interval is calculated as the point estimate $\pm$ $t_{0.25,4}$ * standard deviation, where t comes from the student t distribution for α/2 = 0.25 with 4 degrees of freedom (derived from 6 data points – 2 estimated parameters). Now $t_{0.25,4}$ equals 2.7764. Since *b* = 0.0179 is included within a single standard deviation below the point estimate of 0.0122, clearly, the 95% confidence interval must include our calculated value for *b* with λ = 84 FITS and we so state this without further calculation.

8. *DISCUSSION*

Recall that the purpose of this study was to validate certain useful-life failure rates contained in [8] and, by extension, to validate the techniques used to derive those failure rates. While the analysis above does not *prove* the correctness of the values in the database, it clearly lends very strong credibility to the database and the techniques used to derive it for installed lives of 4 to 5 years. Given three independent data sets for PRV proof test results, all data sets estimated useful-life failure rates close to and below that of the useful-life failure rate predicted by the FMEDA analysis of the single PRV using the mechanical component database being validated. As reported in [7], the techniques used to construct that database intentionally attempted to overestimate the useful-life failure rate so as to build an added margin of safety into the analysis. Thus it seems appropriate that the mechanical FMEDA method in conjunction with the mechanical component database [8] is currently being used for analysis of safety protection functions and has been accepted by regulatory authorities in various parts of the world.

It is also critically important to remind the reader that the results of this study **cannot** be used to conclude that, because the predicted useful-life failure rate of the analyzed PRV is very small, proof tests are unnecessary or at least can be extended to significantly longer intervals. The small value of the failure rate derived from the FMEDA applies **only** to the useful life of the PRV which depends not only on the equipment's specifications but also on other factors, such as the process and ambient environment in which the PRV is used and the levels and frequency of any on-line maintenance performed. It should also be noted that the analysis did not look at partial plugging that could reduce the capacity of the valve. Such a failure mode could indeed be dangerous if upon a pressure relief demand the resulting process pressure became excessive.

Given that the validated database applies only for the useful life, for a given PRV, how can the useful life be determined? If this question could be satisfactorily answered, it might establish clear guidelines for when proof test intervals can and cannot be extended. While the data analysis presented in [10] is a promising beginning, this question offers fertile ground for further research. Until this question is resolved, it is imperative that conservative guidelines for end of useful life be employed, meaning that proof test intervals should not be permitted to exceed 4 to 5 years unless reasonably justified by quantitative data from a strict quality assurance program.

Another point raised by the current analyses revolves around the estimates of probability of initial failure. The data analyses suggest that initial failures are responsible for the vast majority of fail-to-open conditions identified in proof test. What are their underlying causes and how can these causes be addressed to improve overall safety and reliability? Irrespective, when taking credit for a PRV in a risk assessment, both the initial probability of failing to open, as well as the PFD based upon the random failure rate must be taken into account. Even then the results are only defensible when it can be demonstrated that the proof test interval occurs before wear-out, i.e., end of useful life, begins.

REFERENCES

1. IEC 61508, *Functional safety of electrical/electronic/ programmable electronic safety-related systems*, Geneva, Switzerland, 2000.
2. ANSI/ISA SP84.00.02 – 2004 (IEC 61511 Mod.), *Application of Safety Instrumented Systems for the Process Industries*, Raleigh, NC, ISA, 2004.
3. W. M. Goble, J. V. Bukowski, and A. C. Brombacher, "How diagnostic coverage improves safety in programmable electronic systems," *ISA Transactions*, Vol. 36, No. 4, The Netherlands: Amsterdam, Elsevier Science B. V. , 1998.
4. W. M. Goble and A. C. Brombacher, "Using a Failure Modes, Effects and Diagnostic Analysis (FMEDA) to Measure Diagnostic Coverage in Programmable Elecronic Systems," *Reliability Engineering and System Safety*, Vol. 66, No. 2, November 1999.
5. H.A. Amer and E. J. McCluskey, "Weighted Coverage in Fault-Tolerant Systems," *1987 Proceedings of the Annual Reliability and Maintainabiltiy Symposium*, NY: NY, IEEE, 1987.
6. W. M. Goble and Harry Cheddie, *Safety Instrumented Systems Verification: Practical Probabilistic Calculations*, Appendix E, ISA, NC: Research Triangle Park, 2005.
7. W. M. Goble and J. V. Bukowski, "Development of a Mechanical Component Failure Database," *Proceedings of the 2007 Annual Reliability and Maintainability Symposium*, Orlando, FL, January 2007.
8. Electrical & Mechanical Component Reliability Handbook, exida, PA: Sellersville, 2006.
9. R. J. Larsen and M. L. Marx, An Introduction to Mathematical Statistics and Its Applications, 3rd Ed, Prentice Hall, Upper Saddle River, NJ, 2001.
10. R. Gross and S. P. Harris, "Extending Pressure Relief Valve Inspection Intervals by Using Statistical Analysis of Proof Test Data," *Proceedings of ASME PVP 2006/ICPVT-11 Conference*, Vancouver, Canada, July 2006.
11. J. H. Sheesley, H. W. Thomas, and C. A. Valenzuela, "Quantal Response Analysis of Relief Valve Test Data," *ASQC 49th Annual Quality Congress Proceedings*, pp. 741-748.
12. R. Gross and S. Harris, "Analysis of Safety Relief Valve Proof Test Data to Optimize Lifecylce Maintenance Costs", (Draft H), *Proceedings of the 2008 Annual Reliability and Maintainability Symposium*, Los Vegas, NV, January 2008.
13. W. B. Nelson, *Appplied Life Data Analysis*, John Wiley & Sons, New York, 1982.

ACKNOWLEDGEMENTS

The author wishes to acknowledge L. Bredemeyer (exida) for the FMEDA analysis, and the following individuals for their contributions and helpful suggestions: Harold Thomas (Air Products & Chemicals, Inc.), Carlos Valenzuela (Air Products & Chemicals, Inc.), Robert Gross (Savannah River Site) and Stephen Harris (Savannah River National Laboratory).

APPENDIX

A.1. *Introduction*

A general description of the quantal response method may be found in [13]. An excellent description of the method as applied to the analysis of proof test data is available in [11]. In order for this report to be relatively self-contained, a brief description of the application of the quantal response method to proof test data is given here. The reader is referred to [11] and [13] for greater detail.

First, consider a theoretically ideal situation in which a very large number of PRV are in service. At time T_i, a group of n_i independently chosen valves is tested and k_i of these valves fail to open during the test. The ratio $q_i = k_i/n_i$ forms a binary or binomial point estimate of the total percentage of in-service valves that are in this undetected mode of failure at T_i. If at some later time, say, T_j, another independently chosen group of n_j valves is tested and k_j are found to fail to open, then at T_j, the ratio $q_j = k_j/n_j$ is a point estimate for the percentage of in-service valves that are in this undetected mode of failure at T_j. By plotting the q_i's vs T_i's and fitting an appropriate curve to the data, F(t), the cumulative distribution of the time to failure, can be estimated. From F(t), information about λ(t), the failure rate, can be derived.

In a real setting, some of the ideal conditions will not be met. It is unlikely that there will be a sufficiently large group of valves tested at precisely the same time, T_i. A more likely scenario is that a group of valves of size n_i will be tested between times t_i and t_j. A value for T_i, $t_i \le T_i \le t_j$, representative of this group must then be computed. In [11], two different methods for computing T_i are studied via simulation and compared. In the first method, T_i is computed as the average of all service hours (from prior proof test to current proof test) on all valves tested on that interval. In the second method, T_i is computed as the average of all in-service hours on only those valves in that interval found to be failed as a result of the proof test. In [11] it is noted that in estimating F(t) there is usually little difference between the two methods of computing T_i. When more substantial differences existed, using T_i computed from the average in-service hours for the failed valves alone was generally better.

Another deviation from ideal conditions is that, in the real world, valves that have undergone proof test are then maintained and re-installed. At the time of reinstallation, it is generally assumed that the valve is "as good as new". This assumption is made in [11] and is equivalent to assuming that the initial reliability of any valve is 1 even after testing and re-installation. It is not actually necessary to make this assumption as is discussed below. Allowing for the possibility of initial failure is well within the capability of the quantal response method and leads to interesting results.

The steps of the quantal response method may be summarized as follows:

- Arrange the valve data in ascending order of in-service hours since last proof test without regard to whether the valve passed or failed the proof test.
- Divide the data into m non-overlapping intervals each containing some suitable nonzero number of failures.
- For each interval i, i = 1, 2, …, m, let n_i = number of valves tested and k_i = number of valves failed and t_p = in-service hours since last proof test, p = 1, 2, …, n_i.
- Form the ratio $q_i = k_i/n_i$, i = 1, 2, …, m.
- Compute T_i for i = 1, 2, …, n_i as

$$T_i = (1/n_i) \sum_{p=1}^{n_i} t_p \qquad \text{(A.1)}$$

or

$$T_i = (1/k_i) \sum t_f \qquad \text{(A.2)}$$

where t_f are the in-service hours since last proof test for only those valves which tested as fail-to-open on the interval.

- Plot q_i vs T_i and estimate F(t) and λ(t).

A.2. *Some Additional Details About Quantal Response Analysis*

The last step of our summary above actually consists of a number of steps and they are worth exploring here to give a better understanding of our data analysis presented in the body of the report. During their lifetimes, most components and systems display three distinctly different reliability behaviors – a region of decreasing failure rate due to infant mortality, a region of constant failure rate during useful life due to random failure, and a region of increasing failure rate due to some form of ageing. These three distinct failure rate behaviors are easily recognized in plots of λ(t) vs t and, when pieced together, give rise to the familiar bathtub curve. On the other hand, it is generally much more difficult if not impossible to recognize one or more of these behaviors from a graph of the cumulative distribution function, F(t). Thus we seek a connection between the q_i's that estimate F(t) and the more familiar and easily interpreted λ(t). With R(t) representing the reliability function, recall that

$$F(t) = 1 - R(t) = 1 - \exp\{-\int_0^t \lambda(s)\, ds\} \qquad t \geq 0 \qquad \text{(A.3)}$$

In (A.3), when t = 0, F(0) = 0 and R(0) = 1 corresponding to the assumption that all units are fully operational at time t = 0. (Time t = 0 represents either initial installation after commissioning proof testing or re-installation after subsequent proof testing.) If this is not the case, we must revise (A.3) to include the possibility that some units are initially failed by adding a constant to the exponent of exp. Specifically, assuming b > 0 and letting R(0) = exp{-b},

$$F(t) = 1 - R(0)R(t) = 1 - \exp\{-[\int_0^t \lambda(s)\,ds + b]\} \quad t > 0 \qquad (A.4)$$

a form which explicitly indicates the possibility of initial failures. Now, since q_i estimates F(t) at t = T_i, we may write

$$q_i \sim F(T_i) = 1 - \exp\{-[\int_0^{T_i} \lambda(s)\,ds + b]\} \qquad (A.5)$$

from which we can solve for $\int \lambda(s)\,ds + b$ as

$$\int_0^{T_i} \lambda(s)\,ds + b = -\ln(1-q_i). \qquad (A.6)$$

Fitting a power function of the form $at^n + b$ to the points $-\ln(1-q_i)$ vs T_i is equivalent to estimating the left side of (A.6) by the fitted curve $at^n + b$. If the fit is to be valid, a, n, and b must all be greater than 0. Then $1-e^{-b}$ gives an estimate of the initial percentage of failures and at^n gives an estimate for $\int \lambda(s)\,ds$. Differentiating at^n provides an estimate of λ(t) as

$$\lambda(t) = n\,a\,t^{n-1}. \qquad (A.7)$$

Now if we test the hypothesis H_0: n = 1 vs H_1: n ≠ 1 and find that the 95% confidence interval for n includes the value 1, we can accept H_0 thereby concluding that n = 1 is plausible, reducing the power curve to a straight line and λ(t) in (A.7) to a constant value for useful-life failure rate equal to λ. If we then fit a straight line to the data, the slope of the straight line will provide a point estimate for the constant useful-life failure rate λ.