Radioanalytic Counting Instrument Reliability:

An Interim Assessment of A Computer-Assisted Statistical Process Control Approach Under Development by

Robert E. Gladd and James M. Littlefield

International Technology Corporation
OAK RIDGE LABORATORY
Oak Ridge, Tennessee 37830

Calculations of Statistical Process Control (SPC) limits assume representative random sampling of process output for the computation of parameter estimates. Where data accrue rapidly enough for the collection of randomly sampled subsets to be practicable, random sampling is unquestionably the method of choice for the computation of statistical control indices. In a manufacturing environment where output is measured in hundreds or thousands of units per hour or day, randomly drawn Quality Control (QC) data sets are conveniently obtained and SPC limits may be derived in a timely fashion.

In the counting room of the environmental radioanalytical laboratory, however, instrument reliability QC process outputs are typically obtained by the periodic placement of a radioactive standard in each instrument detector chamber and counting the disintegrations over a specified interval. This method produces a single "total counts" QC result per detector per data collection period. Radioactive disintegrations being characteristically stochastic phenomena, one expects a normally distributed variability of source counts over time; the Quality Assurance objective is to insure that counting instruments exhibit an unbiased, random, expected variability that may be taken into account in the error term computations of analytical results, and that any emergent detector performance trends become readily apparent so that corrective adjustments may be applied to the equipment when warranted to assure continuing accuracy of output.

At a data accrual rate of one QC result per detector per collection interval, the practical imperatives of counting instrument control (i.e., procedural, regulatory, and contractual) preclude the use of traditional SPC random sampling techniques; one simply cannot await the amassing of several months worth of data before sampling the database and generating statistical control limits, "flying blind" in the interim. And when an operating detector is subjected to significant electronic and/or mechanical adjustment or repair, the continued use of prior parameter estimates may be improper; new operating conditions properly mandate the collection of a new subset of data for QC evaluation to assure effective instrument control.

Accelerating the data collection process by performing multiple QC counts per detector per day is impractical given the normal production scheduling demands of a laboratory, and, more importantly, would be inappropriate owing to the inevitable resultant biasing of the data by transient environmental variables. Clearly, a routine periodic QC source count procedure with provisions for contingency re-counting is the most practical method of acquiring the necessary instrument reliability QC data.

The Oak Ridge Laboratory of International Technology Corporation (IT/ORL) operates a variety of Alpha, Beta, and Gamma counters, and presently monitors the performance of eight instruments containing a total of seventy-seven detectors. In January of 1989 an in-house developed computerized system of count QC data acquisition and SPC analysis was installed to assist in complying with the requirements of instrument calibration QC ("IQCDATA" and "IQCSTATS"). IT/ORL QC Procedure specifies the use of a minimum of twenty data points for the derivation of SPC parameter estimates, and, of necessity, a sequential sample of the initial twenty source count values for each detector "series" is used to calculate the means and sigmas and construct upper and lower control limits. Individual data points are expressed as normalized deviates ("NDEV") from MEAN=0 and SIGMA=1, with warning limits LWL=-2 and UWL=+2, and control limits LCL=-3 and UCL=+3 (sigmas). Beta and gamma source counts are corrected for temporal source decay before statistics are generated. Source background counts are also entered and plotted as normalized deviates. NDEV results are returned by the program in both tabular and high-resolution control chart scatterplot format. The user may request of the IQCSTATS module tables, scatterplots, or both, for either an individual detector or for all instruments. Additionally, the program contains an option for a daily "Last Entry" report which issues a printout of the input data and statistical indices for the most recent entry in the system for each detector.

The recently completed, fully integrated, interactive prototype of the IQCDATA module provides the analyst with immediate NDEV feedback as data is entered into the system. Input data are rapidly checked for the most common forms of keyboard input errors, such as invalid dates, detector IDs, unauthorized user IDs, exact input duplication, and grossly out-of-range count data stemming from inadvertently mis-keyed or omitted digits. Statistics are computed (for N>=2), and the count input data are converted to NDEVs and returned to the user in a results screen window in a few seconds. NDEVs falling outside +/- 3 sigma require re-counting of the source, and a second consecutive >ABS(3) sigma outlier in the same direction mandates corrective action. NDEV values falling within the "warning" zones [ABS(2<=NDEV<=3)] are monitored for evidence of "runs" indicative of detector trend, or excessive sequential swings denoting inordinate volatility.

The initial sequential sample of N=20 accrues over a period of approximately one month (source checks are not performed on weekends or holidays when the instruments are idle). Tentative statistical control indices are calculated during this data collection interval where 1<N<20 to give the analyst an ongoing indication of instrument performance while awaiting the requisite twenty data values with which to generate the "permanent" detector series parameter estimates. With the addition of each N<20 result the tentative statistics are re-calculated so that the indices converge toward the final N=20 means and sigmas. The final parameter estimates are written to disk from which they are referenced in subsequent operations of the system. These permanent statistics may be cleared from the lookup file upon user request to force re-calculation in the event that data entry errors are found and corrected by a password-authorized editor. A "series" continues from the date a detector is placed in service until such time as significant corrective measures are required, at which point the adjustments and/or repairs undertaken are documented and the next series is assigned an ID for the process to begin anew.

At this writing the IQC system has been in operation for sufficient duration to have acquired detector series datasets numbering from approximately fifty to over one hundred periodic source check results per detector/series, enough data to provide for the assessment of the robustness of the IQCSTATS sequential "convenience" sample-derived control statistics. Detectors were chosen randomly from each instrument, and, with the aid of a random number table, N=20 datasets were drawn from the total available counts for each chamber and compared with the original sequential sets for the respective chamber. The data were entered into a SAS-PC[tm] (Statistical Analysis System) program which performed independent sample T-tests and "Homogeneity of Variance" F-tests on the original datasets vs. the randomly drawn results. The nul hypothesis is that the means and variances of the sequential and random subsets do not differ significantly, i. e., that they are "statistically equivalent," and that, by implication, the method of series sequential sampling of detector data is indeed serviceable for the derivation of instrument SPC indices.

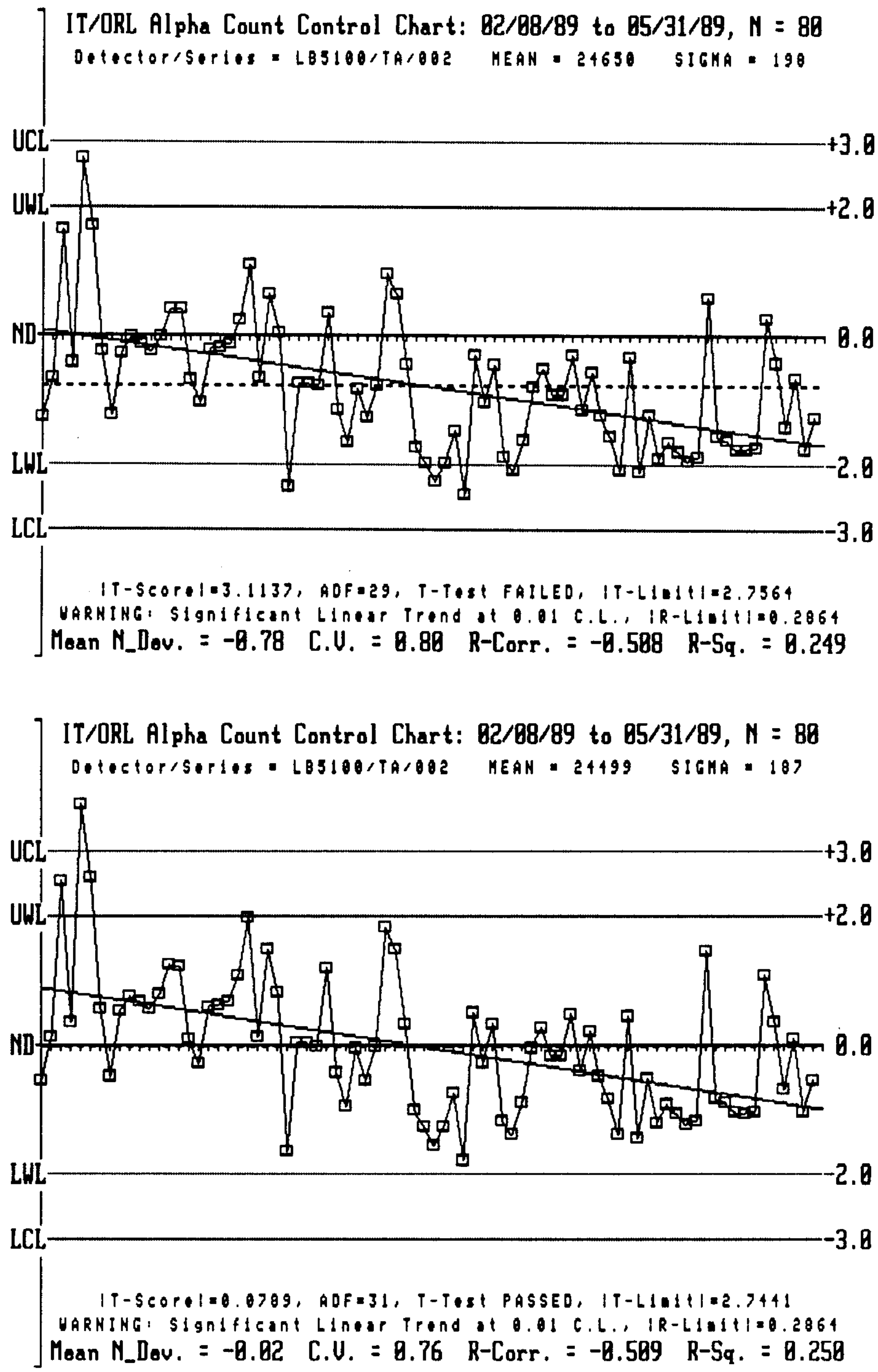
The results, summarized below, are at once reassuring and cautionary. Of the eight instruments, six were subjected to statistical re-examination. One proportional Alpha/Beta counter is a recent acquisition and lacks sufficient operating history to provide a large enough "N" for meaningful random sampling, and a 32 chamber alpha counter is monitored by a pulse count calibration procedure in which the sigma is a fixed percentage of the series mean rather than the usual square root of the adjusted (N-1) mean squared deviation, rendering a conventional T-test inapplicable.

Eight initial T-tests were performed on the six instruments reviewed (two counters are alpha/beta counters, and the T-tests were applied to both alpha and beta source counts). Of the eight T-tests, seven reported statistical equivalence of parameter estimates, i.e., "non-significant differences at 0.05 probability level." Germanium detector GE1, while non-significant at p=0.05, was a bit close for comfort, owing to three extreme source count values in the initial sequential sample, the most extreme of which, the initial series count, was bypassed during the random sampling while the remaining two, both well below the mean, were included by the random number table, resulting in a negative shift in the mean and a contraction in the sigma from sequential to random sample. The GE1 data were subjected to another round of random sampling, and a second T-test yielded similar marginally "non-significant" test statistics. The GE1 data illustrates vividly the impact extreme values have on process control statistics, particularly the normalized deviates. A few relatively large source count fluctuations in the initial twenty results diminish the magnitudes of the NDEVs for the remaining count values, an important contextual point to keep in mind when interpreting the control charts.

One instrument, coded below in the SAS output summary as "LB51A" ("LB5100/TA/002" on the Alpha Count control chart), showed statistically significant sequential-to-random sample alpha source count shift. This detector series dataset was randomly sampled a second time and T-tested again, and the repeat T-test results confirmed the significant shift in the alpha count parameter estimates, raising a pertinent question; are we to conclude that the SPC parameter estimates derived from the first twenty detector series alpha values are in this instance invalid? If we replace the original mean and sigma in the IQC system lookup file with the second set of random sample-derived indices and re-normalize the daily alpha counts for this series to these revised statistics, the revamped control chart (lower plot in box) exhibits an upward shift in the scatter, with a barely perceptible expansion of the scatter owing to the slight contraction in the sigma. The "Mean Deviation," an index of overall detector bias, rises from -0.78 to -0.02, the new figure indicating that the detector evinces essentially no alpha count reproducibility vertical-axis bias when viewed in the context of the random-sample statistics, an implicit reminder of the inherent methodological strength of the random sample.

Under the revised SPC limits, one early alpha data point (fifth in the series, on 02/14/89, originally at NDEV=+2.74) now becomes, post-hoc, an "outlier" at NDEV=+3.71, and a comparision of the original vs. the revised plot reveals a reduction, from seven to two, in normalized deviation points lying between the warning and control limits (i.e., ABS(2<NDEV<3), one of which is now our retroactive outlier). The "warning zone" points on the original plot were, with the exception of the 02/14 result, all <LWL, and a declining count trend is visibly evident on the alpha plot, one that merits testing for significance of slope.

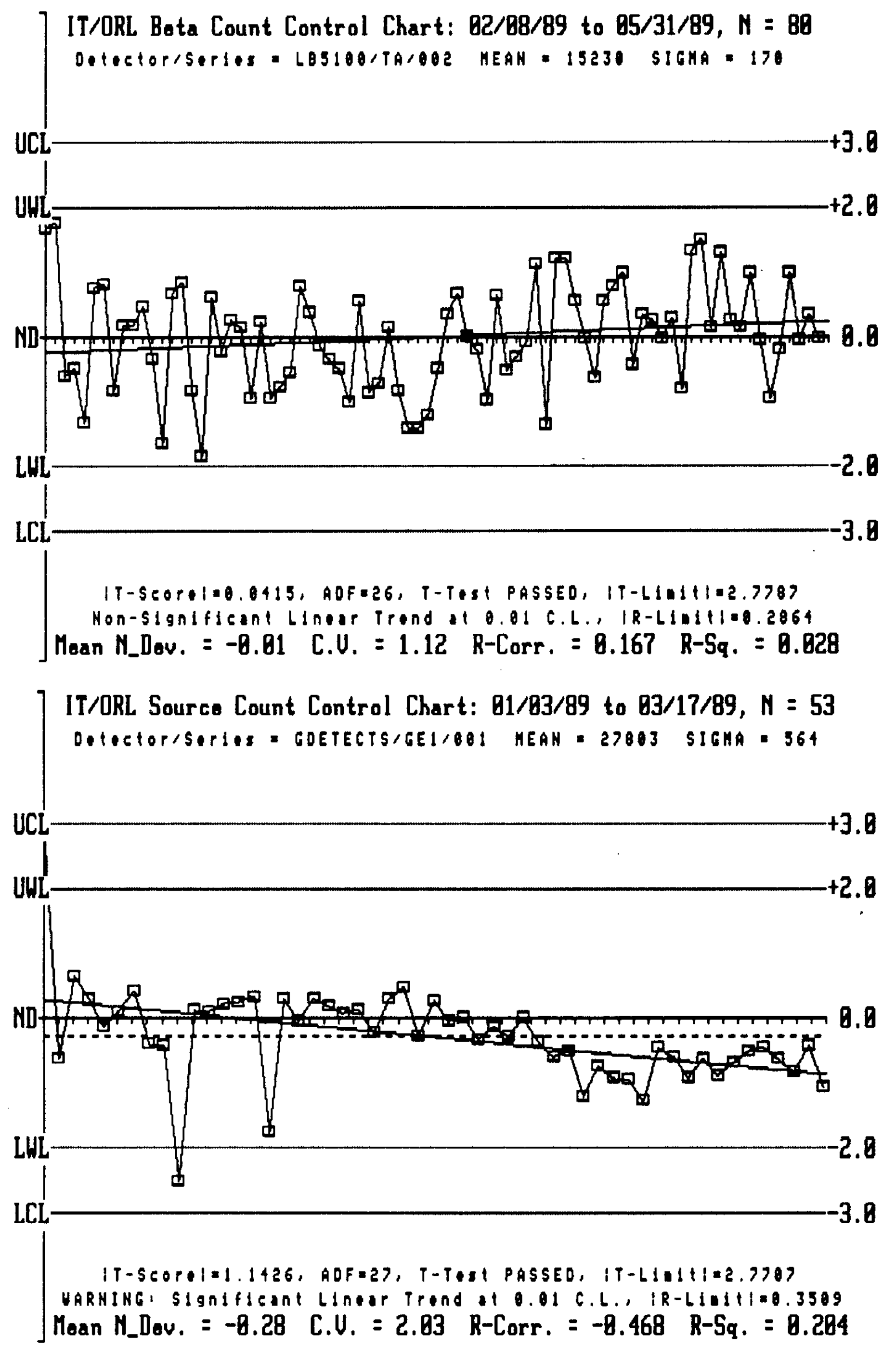
Using the SAS regression procedure PROC REG, a test of the simple linear regression model, "Y = a*X + b," for the LB51A/002 dataset yields a statistically significant slope of -4.453262 (Prob > |T|, 0.0001) with a model Adjusted R-Squared of 0.2489. Formal statistical analysis confirms what is readily apparent to the eye; a declining alpha source count trend is occuring for this detector, one that merits closer scrutiny. Confirmation of the statistical significance of the negative trend is obtained by deleting three early series results (02/10, 02/14, 02/15), all with disproportionately high NDEV values, and testing the remaining data via the same SAS regression model. The revised model, while diluted somewhat in strength, remains



LB-5100/TA/002 alpha, original (top) and re-normalized (bottom) control plots.

statistically significant, with a slope of -3.281061 (Prob > |T|, still 0.0001) and a model Adjusted R-Squared of 0.1739.

Although LB51A/002 indisputably manifests a negative trend, a concern with "statistical significance" should be tempered with an awareness of "practical significance." The Adjusted R-Square statistic, a correlative measure of model predictive strength, indicates that no more than 24.89% of the alpha source count variability is accounted for by the relationship between the source count results and their chronological positions. Experimental removal of the possibly anomalous early extreme values lowers the predictive power of the linear trend model to 17.39%. This "trend," while not to be casually dismissed, is in fact rather loosely coupled. Leaning too heavily upon textbook statistical trend-line analysis has its perils, as many a former Wall Street broker would unhappily attest. Continued evidence of detector trend does, however, warrant inspection of the equipment for contamination or damage, as well as inspection of the source material. Several of our germanium detectors recently exhibited a declining source count trend (see the GE1 example plot on the left); the cause was found to be a nearly invisible crack in the source casing.



LB-5100/TA/002 beta plot (top) and GE1/001Bi-207 source plot (bottom).

Effective radioanalytic counting instrument control is accomplished through a blend of adaptive SPC empirical methodology and expert judgement. While traditional statistical procedures are useful analytical tools, a proper perspective is essential; even in the "worst" case found in this study, that of the LB5100 alpha series, the difference in the sequential vs. the random-sample means, while "significantly different" in terms of a formal statistical T-test, is on the order of 0.6 percent, and this relative 0.6% difference in the means is far less than the customary 2.5 to 5.0% "expected laboratory precision parameters" specified for analytical results. When we renormalize the data to the random-sample-derived mean and sigma, the scatter is shifted upward and all but one of the individual points remain within the control bounds. Further, our revisionist "outlier" must be viewed in the context of a fundamental characteristic of the normal distribution; 99.7% of normally distributed data fall within +/- 3 standard deviations of a mean value. We would therefore expect, on average, three out of 1000 points to randomly fall outside of 3 sigma. Only when a detector exhibits repeated extreme deviations from the mean in the same direction may we infer the presence of a systematic problem. And, with respect to the longer-term trend indicated by a predictively weak but "statistically significant" linear slope, similar interpretive caution is warranted; just as transient environmental variables may impact a single count, so too might seasonal environmental influences be substantive factors in observed detector drift, and such trends, while worthy of closer scrutiny, may tend to slowly flatten out and reverse in response to such subtle long-term background influences.

The IQC developmental system now in place at the IT Oak Ridge Laboratory provides analysts and management with timely and, as indicated by this study, generally robust SPC indicators of detector performance. System enhancements installed since our initial empirical inquiry provide the counting room staff with a comprehensive set of statistical tools for the evaluation of instrument reliability. The IQC control chart routine now fits a least-squares trend line through the scatter and reports the Regression Correlation Coefficient and R-Square values on the plot. The regression coefficient is tested for significance using the standard 0.01 Confidence Limit Critical Values. Correlation indices falling outside the critical values are noted on the plot with a warning statement. In a similar manner, the parameter estimates are now continually re-assessed with the addition of each new set of source and background check data; the "Mean Normalized Deviation" and its standard deviation are calculated for the entire series dataset and a T-Test is performed to reveal whether or not the overall series Normalized Mean differs significantly from zero (zero being, you will recall, the normalized mean calculated from the initial N=20 set). Again, a 0.01 Confidence Level is used, and the T-Test is computed under the conservative assumption of "unequal variances," a method requiring the computation of "approximate degrees of freedom" (ADF on the plot). T-scores falling outside the critical values are noted on the plot with a "T-Test FAILED" warning. The Mean Normalized Deviation is indicated on the plot by a broken line. Where a detector exhibits neither trend nor "Mean N_Dev" bias, the least squares trend line and the population normalized mean "bias line" collapse to the X-axis.

Finally, a "Coefficient of Variation" (C.V.) statistic is returned to provide an index of relative variability. Sometimes referred to as the "Percentage Standard Deviation," this parameter estimate is simply the ratio of Sigma to Mean expressed as a percentage, enabling the analyst to quickly assess both day-to-day and overall fluctuations of the detector series data. For example, where C.V. = 0.8, the interval from the LCL (-3) to the UCL (+3), a range of 6 Sigma, is equivalent to 4.8% of the mean. Use of the C.V. permits the user to easily ascertain both the relative volatility of the series scatter and the proportional magnitudes of individual sequential fluctuations.

Our empirical review of the accruing IQC database and analytical system finds that the sequential sampling method employed by the software is indeed a conceptually sound one and a practical procedure yielding SPC estimates of a generally robust nature with which to monitor, analyze, and manage instrument performance. The study further demonstrates that, even under eventualities where original detector series parameter estimates are found to be statistically unrepresentative of the detector series count population, adverse impact upon the accuracy of analytical results is highly unlikely. The IQC system provides the IT/ORL counting room staff with up-to-date and statistically comprehensive decision data for optimal equipment management.

Appendix a:
T-TEST & REGRESSION MODEL RESULTS

ITAS OAK RIDGE LABORATORY: IQC SYSTEM SAMPLING TEST
T-TEST SEQUENTIAL vs. RANDOM SOURCE CHECK COUNTS
SAS 12:28 Wednesday, May 31, 1989

****************************** DETECTOR=C2A ******************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	10323.10000000	92.50541947	20.68484062
S	20	10337.15000000	106.91032888	23.90587629

Variances	T	DF	Prob>|T|
Unequal	-0.4444	37.2	0.6593
Equal	-0.4444	38.0	0.6592

For H0: Variances are equal, F'= 1.34 with 19 and 19 DF
Prob > F'= 0.5343

****************************** DETECTOR=C2B ******************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	12432.40000000	155.70126592	34.81586148
S	20	12430.95000000	190.52792999	42.60334031

Variances	T	DF	Prob>|T|
Unequal	0.0264	36.6	0.9791
Equal	0.0264	38.0	0.9791

For H0: Variances are equal, F'= 1.50 with 19 and 19 DF
Prob > F'= 0.3868

****************************** DETECTOR=GE1 ******************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	27553.75000000	425.46258107	95.13632532
S	20	27802.65000000	563.73008426	126.05387894

Variances	T	DF	Prob>|T|
Unequal	-1.5761	35.3	0.1239
Equal	-1.5761	38.0	0.1233

For H0: Variances are equal, F'= 1.76 with 19 and 19 DF
Prob > F'= 0.2290

****************************** DETECTOR=GE1 ******************************

*T-TEST #2

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	27554.75000000	381.98620463	85.41471200
S	20	27802.65000000	563.73008426	126.05387894

Variances	T	DF	Prob>|T|
Unequal	-1.6281	33.4	0.1129
Equal	-1.6281	38.0	0.1118

For H0: Variances are equal, F'= 2.18 with 19 and 19 DF
Prob > F'= 0.0981

******************************** DETECTOR=L6 ********************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	17627.65000000	127.47394987	28.50404173
S	20	17624.70000000	117.21464791	26.20999207

Variances	T	DF	Prob>|T|
Unequal	0.0762	37.7	0.9397
Equal	0.0762	38.0	0.9397

For H0: Variances are equal, F'= 1.18 with 19 and 19 DF
Prob > F'= 0.7183

******************************** DETECTOR=LB51A ********************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	24510.25000000	174.42774858	39.00323030
S	20	24650.10000000	197.83483276	44.23721344

Variances	T	DF	Prob>|T|
Unequal	-2.3713	37.4	0.0230
Equal	-2.3713	38.0	0.0229

For H0: Variances are equal, F'= 1.29 with 19 and 19 DF
Prob > F'= 0.5885

******************************** DETECTOR=LB51A ********************************

*T-TEST #2

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	24498.70000000	186.72865199	41.75379592
S	20	24650.10000000	197.83483276	44.23721344

Variances	T	DF	Prob>|T|
Unequal	-2.4889	37.9	0.0173
Equal	-2.4889	38.0	0.0173

For H0: Variances are equal, F'= 1.12 with 19 and 19 DF
Prob > F'= 0.8038

******************************** DETECTOR=LB51B ********************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	15204.00000000	127.22421153	28.44819854
S	20	15229.65000000	169.71191535	37.94873793

Variances	T	DF	Prob>|T|
Unequal	-0.5408	35.2	0.5920
Equal	-0.5408	38.0	0.5918

For H0: Variances are equal, F'= 1.78 with 19 and 19 DF
Prob > F'= 0.2182

******************************** DETECTOR=LSC1 ********************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	116905.55000000	460.74167611	103.02497079
S	20	116907.85000000	385.97631165	86.30692706

Variances	T	DF	Prob>|T|
Unequal	-0.0171	36.9	0.9864
Equal	-0.0171	38.0	0.9864

For H0: Variances are equal, F'= 1.42 with 19 and 19 DF
Prob > F'= 0.4474

******************************** DETECTOR=LSC1 ********************************

Variable: COUNTS

SAMPLED	N	Mean	Std Dev	Std Error
R	20	116741.75000000	543.53481223	121.53807883
S	20	116592.90000000	463.73789103	103.69494481

Variances	T	DF	Prob>|T|
Unequal	0.9317	37.1	0.3575
Equal	0.9317	38.0	0.3574

For H0: Variances are equal, F'= 1.37 with 19 and 19 DF
Prob > F'= 0.4955

LB5100/TA/002: SAS PROC REG OUTPUT, LINEAR TREND REGRESSION ANALYSIS
SAS 10:24 Tuesday, June 6, 1989 1

Model: MODEL1
Dep Variable: SOURCE

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	814679.73710	814679.73710	26.843	0.0001
Error	77	2336946.9464	30349.96034		
C Total	78	3151626.6835			

Root MSE	174.21240	R-Square	0.2585
Dep Mean	24496.06329	Adj R-Sq	0.2489
C.V.	0.71119		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	24674	39.57598903	623.464	0.0001
DAYS	1	-4.453262	0.85953556	-5.181	0.0001

LB5100/TA/002: SAS PROC REG OUTPUT, LINEAR TREND REGRESSION ANALYSIS
(TRIAL #2)

Model: MODEL1
Dep Variable: SOURCE

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	400279.31645	400279.31645	16.791	0.0001
Error	74	1764128.7230	23839.57734		
C Total	75	2164408.0395			

Root MSE	154.40070	R-Square	0.1849
Dep Mean	24474.19737	Adj R-Sq	0.1739
C.V.	0.63087		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	24610	37.58078801	654.856	0.0001
DAYS	1	-3.281061	0.80072240	-4.098	0.0001

REFERENCES

Goldin, Abraham S. (1983), *Internal Quality Control for Radioassay*, USEPA Office of Radiation Protection Programs, Wash., DC.

Kanipe, Larry G. (1977), *Handbook for Analytical Quality Control in Radioanalytical Laboratories*, TVA/USEPA Interagency Energy-Environment Research & Development Program (EPA 600/7-77-088), Wash., DC.

Freund, R.J. and Little, R.C. (1986), *SAS System For Linear Models*, SAS Institute, Cary, NC.

Steel, R.G.B. and Torrie, J.H. (1980), *Principles and Procedures of Statistics, Second Edition*, New York, McGraw-Hill Book Co.

Diem, K. and Lentner, C., Ed. (1970), *Scientific Tables, Seventh Edition*, Ciba-Geigy, Ltd., Basle, Switzerland.

ROBERT E. GLADD is Vice President of Corporate and Academic Media Associates, Inc. (CAM3) of Knoxville, TN. He has served as a microcomputer applications consultant to the IT Oak Ridge Laboratory since 1986.

JAMES M. LITTLEFIELD is the Radioanalytical Counting Room Supervisor of the IT Oak Ridge Laboratory.