NISTIR 5951

Guide for Quality Control on the Qualitative and Quantitative Analysis of Bulk Asbestos Samples: Version 1

Jennifer R. Verkouteren
David L. Duewer*

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Surface and Microanalysis Science Division
Chemical Science and Technology Laboratory
Gaithersburg, MD 20899-0001

*Analytical Chemistry Division

March 1997

U.S. DEPARTMENT OF COMMERCE
William M. Daley, Secretary

TECHNOLOGY ADMINISTRATION
Mary L. Good, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
Arati Prabhakar, Director

PURPOSE

This guide provides a set of approaches for dealing with quality control (QC) data in the bulk asbestos analysis laboratory. This guide was developed in response to numerous requests for detailed information on quality control from the laboratories involved in the National Voluntary Laboratory Accreditation Program (NVLAP) for bulk asbestos.

SCOPE

The guide outlines procedures for characterizing within analyst and within laboratory accuracy for the analysis of bulk asbestos samples. Methods are given for generating control charts using statistics calculated from control data. Real laboratory QC data provided by an anonymous NVLAP laboratory is presented and analyzed.

STATISTICAL APPROACH

A common approach to collecting QC data in the analytical laboratory is to have a small set of samples appropriate for the type of measurements being performed that are analyzed repeatedly. Control charts are used to track analytical response with time for this non-changing set of samples. A real hazard with applying this approach to a measurement in which humans are the detectors is the potential for bias due to familiarity. Any QC system for bulk asbestos analysis must be sensitive to the needs for blind analysis of QC samples, and for limiting the number of times a particular sample is used. To avoid problems with sample familiarity, we recommend using the samples analyzed by the laboratory in the normal course of business as the QC samples. These are "unknowns" in that the amount and type of asbestos are not known by some independent means.

The unknowns are used to determine *within analyst accuracy* and *within laboratory accuracy*. *Within analyst accuracy* is a measure of an individual analyst's ability to repeat an analysis, and *within laboratory accuracy* is a measure of the laboratory's ability to repeat an analysis, given different analysts using different methodologies and equipment. The terms are italicized because we mean them as very specific terms which will be defined for this document but not meant to be used in any general sense. We use the word accuracy in both cases since it combines errors that are both random and systematic in nature, and is therefore a more general term than precision, bias, repeatability, reproducibility, or any number of terms that are used for QC measurements. We would have preferred to use the term precision, rather than accuracy, for the two QC activities described as there is no knowledge of the "true" value of the asbestos concentration for each sample. However, precision is a well-defined term that describes **random** measurement errors. Even in the case of the *within analyst accuracy,* which would seem to be a measurement of precision, we saw evidence of a systematic error, or bias, and therefore could not use the term precision for this activity. **We want to emphasize, however, that the terms *within analyst***

***accuracy* and *within laboratory accuracy*, as defined in this document, describe QC that is internal to the laboratory, and do not say anything about the laboratories performance with respect to other laboratories or with respect to the "true" value. There is a discussion of the analysis of reference samples which will give the laboratory some measure of its ability to obtain the "true" value.**

The QC system is designed to emphasize control near the critical regulatory value of 1% asbestos[1] by analyzing the data with respect to three concentration ranges: 1) trace to less than 1% asbestos, 2) 1% to 10% asbestos, and 3) greater than 10% asbestos. Medians, interquartile ranges (IQRs) encompassing the central 50% of the data, and control limits that approximate 2 times the standard deviation (SD) are calculated for each concentration range. Medians and IQRs are used because they are less sensitive to outliers than means and standard deviations, and therefore are more robust estimators.

DATA MANAGEMENT

Although data management can be done manually, laboratories with computers can choose to enter and manipulate data using commercially available or internally generated spreadsheets or databases. The calculations performed on the data are common to commercially available spreadsheets and can also be performed on hand calculators. Graphical representation of the data is highly recommended, and can be done easily with common spreadsheets or by hand. All the data entry and statistical analysis shown in this guide were performed in Microsoft Excel* and specific information regarding this spreadsheet program is given where applicable.

DATA COLLECTION AND QUALITATIVE ANALYSIS

The NVLAP Bulk Asbestos Handbook[2] stipulates that a minimum of 10% of the analyses performed by the laboratory be QC analyses. The breakdown of the 10% into the various types of analyses and the timing of the QC data collection is left to the discretion of the laboratory. The types of analyses described in this guide are:

1) *within analyst accuracy* : an analyst receives a sample twice and performs each analysis independently. This tests the ability of an analyst to repeat a measurement.

2) *within laboratory accuracy*: two or more analysts receive the same sample and conduct the analyses independently. This tests differences among analysts.

---

*Certain commercial equipment, instruments, or materials are identified to specify adequately the procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

3) reference sample analyses: a material with a known concentration of asbestos is submitted for analysis. This tests the ability of the analyst/laboratory to obtain the "true" value.

A laboratory may choose to have all their analysts do 10% QC for the number of samples they analyze although beginning analysts should perform 100% QC until their proficiency is determined to be within acceptable limits. The relative amounts of the three types of QC analyses are up to the laboratory, and may change with time as needed.

A completely random selection of QC samples is probably not the best approach, given that the concentration range near 1% asbestos is the most critical for QC, and those samples may be relatively uncommon. We recommend selecting all of the samples whose first analysis is between trace and 5% asbestos for inclusion in the QC samples. Samples that are particularly difficult, such as floor tiles, should also be preferentially included in the QC samples. The remaining samples can be selected randomly. The reference materials should cover the concentration range of the samples received for analysis, and should also emphasize the concentration range near 1% asbestos. The proficiency testing materials received by labs in the bulk NVLAP and New York State ELAP programs are appropriate as reference materials and cover a useful range of concentrations. Reference samples containing close to 1% asbestos are relatively uncommon and may need to be synthesized[3].

The data entered into the QC records from the laboratory analysis record includes: 1) sample identification number, 2) date, 3) analyst, 4) type(s) of asbestos, and 5) concentration(s) of asbestos. Additional data not included in the following examples could be matrix type and other sample information to characterize errors with respect to sample. This information should be in the laboratory records to be referred to when needed. If entering data in a database, fixed names should be adopted for the asbestos types and analysts in order to aid search and sort functions. An example of data entry is given below:

Sample	Date 1	Analyst 1	Type 1	Conc 1	Date 2	Analyst 2	Type 2	Conc 2	ACM 1	ACM 2	Flag Type	Flag ACM
24	1/5/94	Jill	CHRY	7	1/7/94	Harry	CHRY	15	Y	Y	0	0
36	1/11/94	David	NONE	0	1/12/94	David	CHRY	3	N	Y	F	F
40	1/22/94	Harry	AMOS	13	1/23/94	David	AMOS	5	Y	Y	0	0

Combining the two analyses for each sample can be done at the time of data entry, or afterwards by sorting and merging according to sample number. To handle samples with multiple asbestos types a field (column) can be created that indicates the number of asbestos types in the sample. The information for each asbestos type can be entered as a separate row with the sample number as the common factor. When the data are sorted by sample number, the analyses will be grouped together, and if a 2 has been entered into the field for multiple asbestos types, there should be 2 rows of information per sample per analysis. [In the example above, samples 24 and 40 are tests of *within laboratory accuracy* and sample 36 is a test of *within analyst accuracy*.]

A column is added for each analysis (ACM 1 and ACM 2) to characterize the sample as asbestos-containing based on the concentration reported. A logical statement can be used to determine whether the concentration is 1% or more, in which case a Y for yes is placed in the column,

otherwise N for no is entered. [In Excel, IF column 5 >=1,Y,N.] Two error flags are determined for the QC analyses at this point: 1) whether both analyses have the same asbestos type (s), and 2) whether both analyses agree on asbestos-containing or not. The columns are identified above as Flag Type, and Flag ACM, with an F indicating an error. [In Excel, IF column 4=column 8, 0, F.] The entries can then be sorted by error flags, to pull out all entries that are flagged. Errors in determining whether the material is asbestos-containing, or the type of asbestos, are qualitative errors, and should be considered as very serious. **Every analysis that gets a qualitative flag should be reviewed to determine the source of the problem.** Real data from the volunteer laboratory was analyzed in this manner, and some of the qualitative errors are shown in Table 1. Unfortunately, for these samples, the type of asbestos was not recorded and therefore only the ACM flag was determined.

Sample	Date	Analyst	Analyst 2	Result 1	Result 2	ACM 1	ACM 2
6417453	3/19/94	1	4	34	0	Y	N
6417848	3/22/94	6	4	2	0.1	Y	N
6425033	4/26/94	9	5	15	0	Y	N
6430217	6/6/94	5	4	1.5	0	Y	N
6446608	12/7/94	4	3	6.1	0.1	Y	N
6447423	12/21/94	7	3	0	8.5	N	Y
6447424	12/21/94	7	3	0	6.2	N	Y
6447425	12/21/94	7	3	0	12.2	N	Y
6447567	12/22/94	7	3	0	25.9	N	Y
6447572	12/22/94	7	3	0.1	1.2	N	Y
6447573	12/22/94	7	3	0.1	9.4	N	Y
6448683	12/28/94	4	3	0	47.1	N	Y
6449249	1/3/95	7	3	3	0	Y	N
6503507	2/15/95	7	5	3	0.1	Y	N

Table 1. Qualitative errors from laboratory QC data.

It will generally be impossible to assign an error to an individual analyst for any of the samples until some way of determining the correct answer is found. Some of the differences in the results may be due to measurement uncertainty (e.g., small particle numbers in the point counts) for samples containing close to 1% asbestos, which may be the case for samples 6417848, 6430217, and 6447572. The samples analyzed between 12/21 and 12/28 all have the same analyst pair (7 and 3) and it is possible that the same error occurred for all those samples. All the samples that have qualitative errors such as those shown above must be reviewed by having additional analysts look at them or by using different analytical techniques such as TEM, XRD, or gravimetry, depending on the nature of the discrepancy. Once the nature of the problem is determined, corrective action must be taken.

WITHIN ANALYST ACCURACY

Within analyst accuracy is determined by having an analyst perform 2 independent measurements on a number of samples. The first step in managing the data is to form separate sets of data for each analyst. The data are then analyzed for qualitative errors as described earlier, and any flagged samples are segregated as are samples that do not contain asbestos. [Although it is important to include blank samples in the QC analyses, they are not included in the determination of accuracy.] The data for each analyst can be graphed at this point by plotting the second result *versus* the first result, as shown in Figure 1 for the real data from the volunteer laboratory. This type of graph, called a Youden plot[4], is very useful because it is easy to interpret and provides qualitative information on the overall scatter in the data and any consistent repetition errors. The line drawn on each graph is the y=x line, which is the line that represents exact repetition.

One observation from the graphs in Figure 1 is that analyst 4 is an anomaly in having virtually no repetition error. This data set should be rejected as an indicator of *within analyst accuracy* because it is unrealistic. Analyst 4 is probably not receiving these QC samples in a blind fashion. A second observation from Figure 1 is the consistent repetition error expressed by analyst 5's and 2's data where the data do not scatter uniformly about the y=x line. For both analysts, the second result is systematically higher than the first result, particularly at higher asbestos concentrations. This type of bias is expected for measurements involving two different analysts, but it is peculiar for data from a single analyst. We can speculate that it may be due to fatigue (all the second analyses are conducted in the afternoon) or perhaps to bias during data entry. This should be investigated to discover the cause and eliminate the problem.

The next step in analyzing the data is to calculate the difference between the two results for each sample by subtracting result 2 from result 1 (r1-r2). The average of the two results (r1+r2)/2 is also calculated, and both calculations are shown in Table 2. The data in Table 2 have been sorted to separate the analyses with qualitative errors, as determined by an F in the "Flag ACM" column. Samples with qualitative errors are not included because they represent a large deviation from analytical control. An exception to this would be a sample in which the true amount of asbestos is near 1% and one result is slightly less than 1% and the other result is slightly more than 1%. An analysis of this type may be indicative of true repetition error, whereas a result of 26% *versus* a result of 0 reflects some gross error.

An additional column shown in Table 2 is labeled "category" and represents the separation of the data into the three concentration categories referred to earlier. A value of 1 has been entered for all analyses in which the average of the two results ((r1+r2)/2) is less than 1%, a value of 2 has been entered for (r1+r2)/2 between 1% and 10%, and a value of 3 has been entered for (r1+r2)/2 greater than 10%. [In Excel, IF(column6<1,1,IF(column6>10,3,2)).] All results that contain zeros in both the first result and second result columns have been segregated as blanks. It is not absolutely necessary to create the "category" column, but it helps in sorting and analyzing the data.

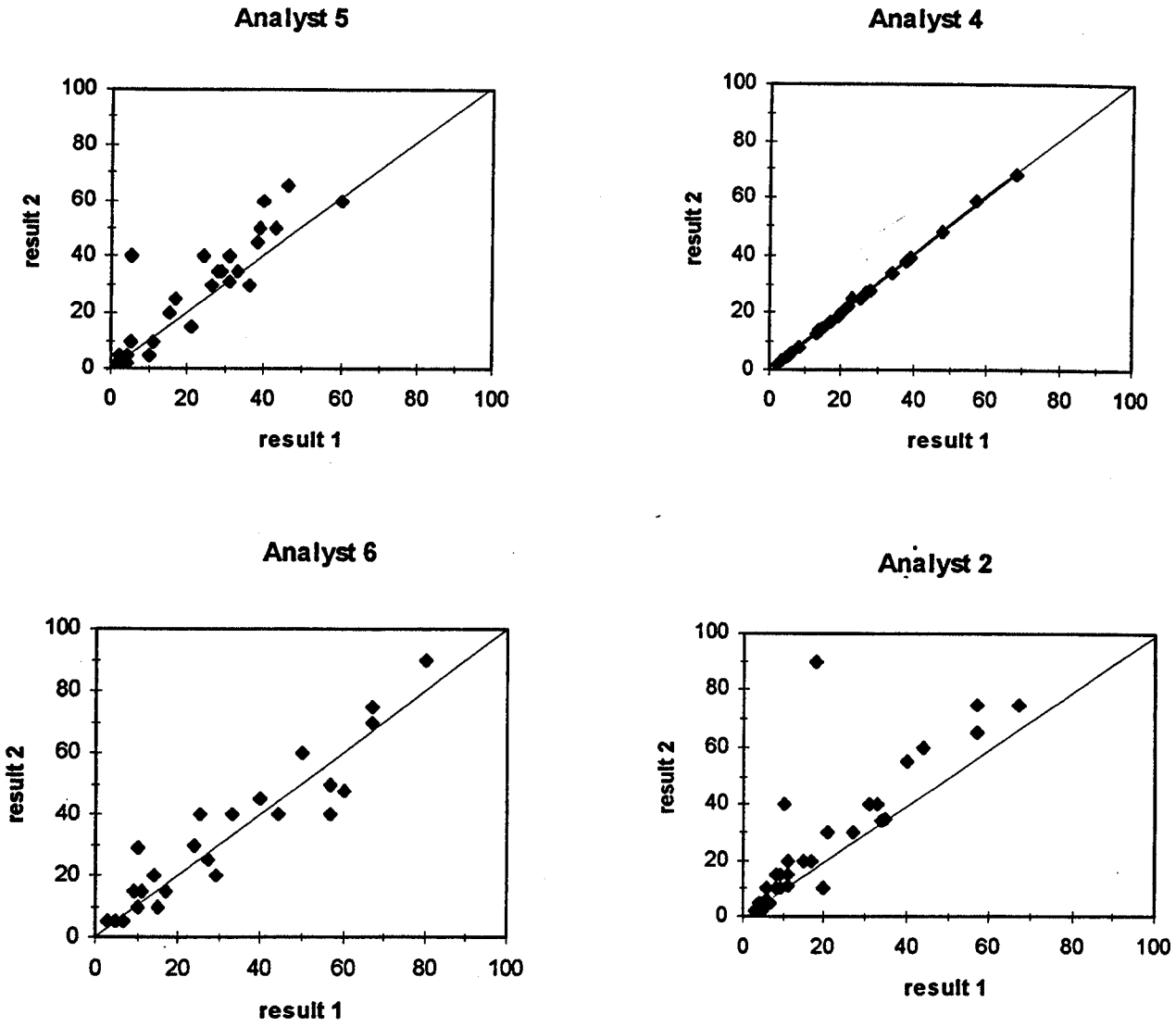

Figure 1. *Within analyst accuracy* data for 4 analysts. The results are in units of % concentration of asbestos.

NOTE: Samples can contain a trace amount of asbestos, defined as asbestos present in the sample but not at a concentration high enough to be counted in a 400-point point count. A numeric value must be assigned for trace concentrations for entry in a spreadsheet, and 0.1 is commonly used for this purpose.

analyst	sample	date	result 1	result 2	(r1+r2)/2	ACM 1	ACM 2	flag ACM	r1-r2	category	rel. diff.
5	6445257	11/21/94	5	0		Y	N	F			
5	6504790	3/1/95	14	0		Y	N	F			
5	6508212	4/27/95	26	0		Y	N	F			
5	6512992	7/18/95	2	0		Y	N	F			
5	6512232	7/17/95	0.1	0.1	0.1	N	N	0	0	1	0
5	6506131	3/31/95	0.1	0.1	0.1	N	N	0	0	1	0
5	6512588	6/21/95	0.1	0.1	0.1	N	N	0	0	1	0
5	6512127	6/16/95	0.3	0.1	0.2	N	N	0	0.2	1	100
5	6513177	7/5/95	0.3	0.1	0.2	N	N	0	0.2	1	100
5	6513268	7/24/95	0.3	0.1	0.2	N	N	0	0.2	1	100
5	6511434	6/28/95	0.5	0.1	0.3	N	N	0	0.4	1	133.33
5	6443432	11/9/94	0.6	0.1	0.35	N	N	0	0.5	1	142.86
5	6503589	2/14/95	5	10	7.5	Y	Y	0	-5	2	-66.667
5	6503601	2/14/95	2	5	3.5	Y	Y	0	-3	2	-85.714
5	6447942	12/19/94	4	5	4.5	Y	Y	0	-1	2	-22.222
5	6519696	8/29/95	2.1	3	2.55	Y	Y	0	-0.9	2	-35.294
5	6500493	1/9/95	2	2	2	Y	Y	0	0	2	0
5	6503611	2/14/95	4	2	3	Y	Y	0	2	2	66.667
5	6503550	2/15/95	10	5	7.5	Y	Y	0	5	2	66.667
5	6441798	9/26/94	5	40	22.5	Y	Y	0	-35	3	-155.56
5	6449321	1/3/95	40	60	50	Y	Y	0	-20	3	-40
5	6508021	4/25/95	46	65	55.5	Y	Y	0	-19	3	-34.234
5	6503678	2/14/95	24	40	32	Y	Y	0	-16	3	-50
5	6507625	4/17/95	39	50	44.5	Y	Y	0	-11	3	-24.719
5	6504026	2/23/95	31	40	35.5	Y	Y	0	-9	3	-25.352
5	6505830	3/20/95	17	25	21	Y	Y	0	-8	3	-38.095
5	6503855	2/15/95	28	35	31.5	Y	Y	0	-7	3	-22.222
5	6431034	6/10/94	38	45	41.5	Y	Y	0	-7	3	-16.867
5	6448615	12/27/94	43	50	46.5	Y	Y	0	-7	3	-15.054
5	6447864	12/19/94	29	35	32	Y	Y	0	-6	3	-18.75
5	6443740	11/9/94	15	20	17.5	Y	Y	0	-5	3	-28.571
5	6503608	2/14/95	15	20	17.5	Y	Y	0	-5	3	-28.571
5	6506336	3/24/95	26	30	28	Y	Y	0	-4	3	-14.286
5	6507121	4/17/95	33	35	34	Y	Y	0	-2	3	-5.8824
5	6506422	3/27/95	31	31	31	Y	Y	0	0	3	0
5	6512484	6/19/95	60	60	60	Y	Y	0	0	3	0
5	6506383	3/24/95	11	10	10.5	Y	Y	0	1	3	9.5238
5	6444178	11/4/94	21	15	18	Y	Y	0	6	3	33.333
5	6429212	6/2/94	36	30	33	Y	Y	0	6	3	18.182

Statistics for r1-r2

Concentration category 2*: Median = -0.9%, IQR = -2% to 1% or 3%, Control limits = ±4.4%

Concentration category 3: Median = -6.5%, IQR = -9.5% to -1.5% or 8%, Control limits ±11.9%

*Only 7 values are available in this concentration category, which is below the recommended minimum of 8 as discussed later in this document.

Table 2. *Within analyst accuracy* data for analyst 5.

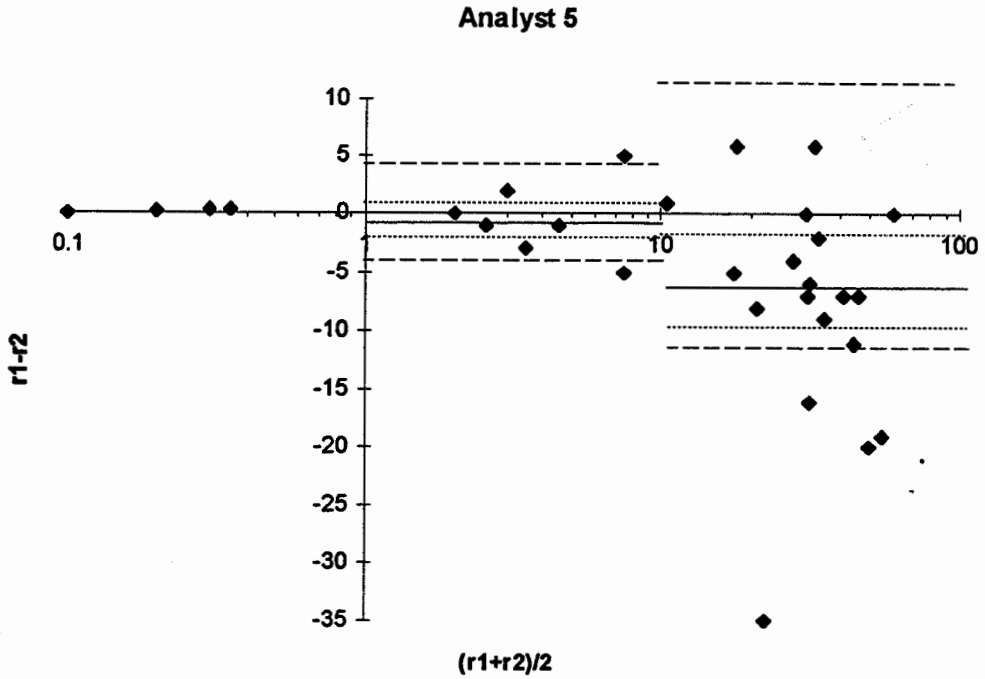

Figure 2. *Within analyst accuracy* data for analyst 5 showing calculated medians (solid lines), IQRs (dotted lines) and control limits (dashed lines) for concentration categories 2 and 3.

The data in Table 2 are shown graphically in Figure 2 by plotting r1-r2 *versus* (r1+r2)/2. The x-axis is plotted using a logarithmic scale to emphasize the concentration categories. The fact that most of the points plot below the zero-difference line reflects the bias discussed earlier in which the second result is consistently higher than the first. The data should scatter uniformly above and below zero. Also shown on the graph are the calculated medians for concentration categories 2 and 3, interquartile ranges (IQRs) that bound the central 50% of the data, and control limits. The calculation of each statistic is discussed below. These statistics were not calculated for concentration category 1, as this is taken to be simply a "trace concentration" category.

Median

The median is the number in the middle of a set of numbers; half of the numbers in the set have a value greater than the median and half of the numbers have a value less than the median. The median of the differences between the first and second results is used to estimate the bias. There is a bias if the median is large in absolute value.

We recommend using the median rather than the mean because it is less sensitive to outliers and

therefore eliminates the necessity of identifying and discarding outliers prior to calculating statistics from the data. (The median is equivalent to the mean for symmetrically distributed data.) For example: there appears to be an obvious outlier in r1-r2 of -35% for concentration category 3 (see Table 2). Most people looking at Figure 2, or the same data displayed differently in Figure 1, would identify that particular point as an outlier. The median of r1-r2 in concentration category 3 is -6.5% but the mean (or average) is -7.4%, due to the presence of the outlier. Although some outlier points are easily discriminated by eye, others are not, and to avoid rigorous tests for outliers we recommend using the median. The median is a standard calculation in spreadsheets [in Excel, =MEDIAN(value 1..value n)] and it can be determined by sorting the data and finding the value at the midpoint. The data in Table 2 were first sorted by concentration category, and then by difference within each category. There are 20 difference values in concentration category 3, and the median can be found at the midpoint between the 10th and 11th value. [See Appendix 1 for manual procedures for calculating medians.]

Interquartile range (IQR)

The IQR encompasses the central 50% of the data and gives an estimate of the spread or scatter of the data. We recommend the use of the IQR as opposed to the standard deviation (SD) because, like the median, it is relatively insensitive to outliers. The calculation of quartiles is a standard feature of spreadsheets and can also be approximated by sorting the data and dividing it into 4 segments. For example, there are 20 entries in concentration category 3 as given in Table 2 that are sorted according to r1-r2. The data separate into 4 segments of 5 entries each, and to get the central 50%, the top five values and the bottom five values are discarded. The first quartile falls between -11% and -9%, and is calculated to be -9.5%. The third quartile falls between -2% and 0%, and is calculated to be -1.5%. [See Appendix 1 for manual calculation procedures.] The IQR is the absolute difference of the first quartile and the third quartile. The IQR of r1-r2 for concentration category 3 is therefore -9.5% to -1.5% or 8%. [In Excel: =QUARTILE(value 1..value n, 1) and =QUARTILE(value 1..value n, 3).]

Control limits

Control limits are calculated from the IQR to approximate a 2xSD limit and serve to identify points as in or out of control. The IQR is related to the SD through the relation 0.741xIQR = SD for normally distributed data[5]. Again, we recommend using the IQR rather than the SD because of the relative insensitivity of the IQR to outliers. To calculate the control limit for concentration range 3, the IQR (absolute value of -9.5% minus -1.5% = 8%) is multiplied by 0.741 to get the SD, which is 5.93%, and then by 2 to get a 2xSD limit of ±11.9%. By comparison, calculating the SD directly for this concentration range yields a 2xSD limit of ±19.1%, which is larger due to the outlier. (Removing the outlier point and calculating the SD directly yields a 2xSD limit of ±14.4%.) The control limits of ±11.9% are placed about zero, not about a possibly biased median.

Discussion

A problem with using r1-r2 as an indicator of measurement performance is that the absolute difference should not be independent of concentration, i.e. low concentration samples should have

much smaller values of r1-r2 than high concentration samples. For example, an r1-r2 of 5% for a sample containing 50% asbestos is not as significant as for a sample containing 5% asbestos. There are at least 2 ways to address this problem: 1)calculate measurement performance (i.e. medians, IQRs, control limits) as a function of concentration, or 2)reference the difference to the average value of the concentration to determine a relative difference. We have chosen to use the first approach and have selected the concentration categories to treat low concentration samples separately from the high concentration samples. We saw no significance in further segregating the samples containing greater than 10% asbestos, since the primary goal is to insure that samples containing close to 1% asbestos are closely controlled.

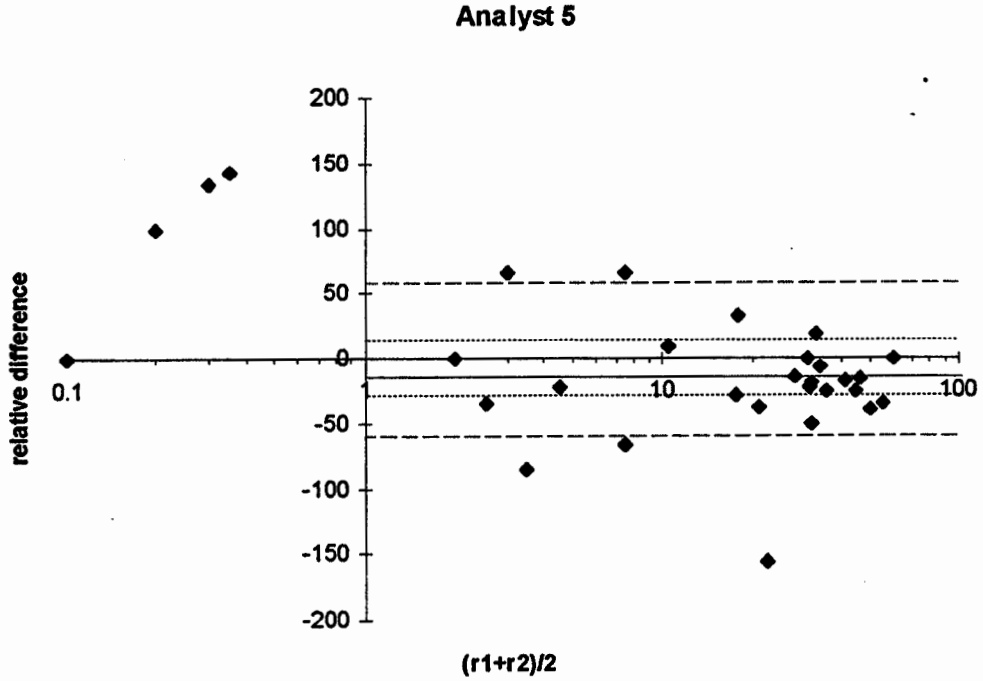

Figure 3. *Within analyst accuracy* for analyst 5 using relative differences (the difference divided by the average and multiplied by 100) rather than differences. Medians (solid lines), IQRs (dotted lines) and control limits (dashed lines) are calculated for the entire data set, with the exception of values from concentration category 1.

We also tested the usefulness of using relative differences, but found that with the data set available to us, calculating relative differences *alone* was insufficient to achieve the goal of appropriately characterizing measurement performance. (Using relative differences in conjunction with concentration categories is perfectly acceptable.) Relative differences were calculated by dividing r1-r2 by (r1+r2)/2 and multiplying by 100, as shown in Table 2. The data were then graphed as shown in Figure 3 and medians, IQRs and control limits were calculated for the entire

set of data, without respect to concentration category (excluding data from concentration category 1). As expected, the relative differences decrease with concentration, just as the differences (r1-r2) increase with concentration. Therefore, control limits calculated for the entire set of data are probably too restrictive for the low concentration samples (samples with results of 2% and 4%, and 2% and 5% are out of control) and too generous for the high concentration samples (anything within ±60% relative is in control, which is ±30% asbestos for a sample containing 50% asbestos). We simply do not have enough data at this point to say which is the best approach, or even what are acceptable performance levels. The laboratories using this document will need to decide which approach best demonstrates acceptable results for low concentration samples for their data.

WITHIN LABORATORY ACCURACY

Within laboratory accuracy is determined by performing multiple measurements of the same sample using different analysts. The scatter in the data for *within laboratory accuracy* should be larger than the scatter in the data for *within analyst accuracy*, as there may be systematic differences among analysts. *Within laboratory accuracy* data can be collected in two ways: 1) by randomly forming pairs of analysts to produce 2 analyses per sample, or 2) by having all analysts measure a set of samples in random order. The first approach allows for a larger number of QC samples to be analyzed and provides an estimate of the scatter in the data expected for the entire laboratory, whereas the second approach is a systematic way of testing for relative differences among analysts. If it is a two-person laboratory, the two approaches are, of course, the same.

The volunteer laboratory submitted data using the first approach with 2 analyses per sample by 2 different analysts. Their data contains 28 different analyst pairs. The difference (r1-r2) between the two results for each sample was determined, and plotted with respect to (r1+r2)/2 as shown in Figure 4. The medians, IQRs, and control limits were calculated for r1-r2 for the two concentration categories, and are:

1%-10%: Median = -1%, IQR = -3% to 1.3% or 4.3%, Control limit = ±6.4%
>10%: Median = -1%, IQR = -10% to 7.5% or 17.5%, Control limit = ±25.9%

The *within laboratory accuracy* data show more scatter than the *within analyst accuracy* data for both concentration regions with a 40% increase for the 1%-10% IQR and more than a doubling of the >10% IQR. The medians for both concentration categories are close to zero, implying either random variation among the analysts, or that there are just as many analysts that report high values as report low values. [It is important when collecting data in this fashion that the analyst pairs are randomly selected and that the position of each analyst in the pair is also random to prevent an apparent bias due to weighting of one analyst's data.] To determine whether the variation among the analysts is random or systematic, the second approach where

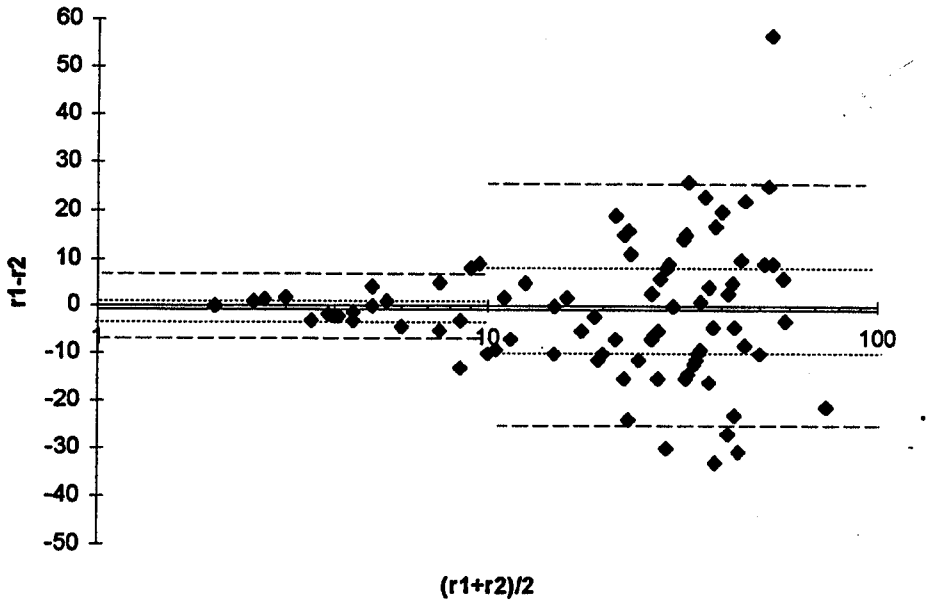

Figure 4. *Within laboratory accuracy* data in which each point represents the difference in results between 2 analysts for a given sample. Medians (solid lines), IQRs (dotted lines) and control limits (dashed lines) are shown for concentration categories 2 and 3.

all analysts measure the same set of samples should be used. The volunteer laboratory did not submit data of this type, and so an example set of data was created, containing results for 4 samples from 4 analysts. The results for each analyst were plotted against the average for each sample, as shown in Figure 5. It is easy to see from this graph that analyst 2 is consistently lower than the other analysts. There are more rigorous ways to investigate data of this type, but we suggest using this approach in a qualitative sense to investigate gross differences among analysts and try to correct large systematic differences. This type of approach does not indicate which analyst is correct, only the relative differences among analysts.

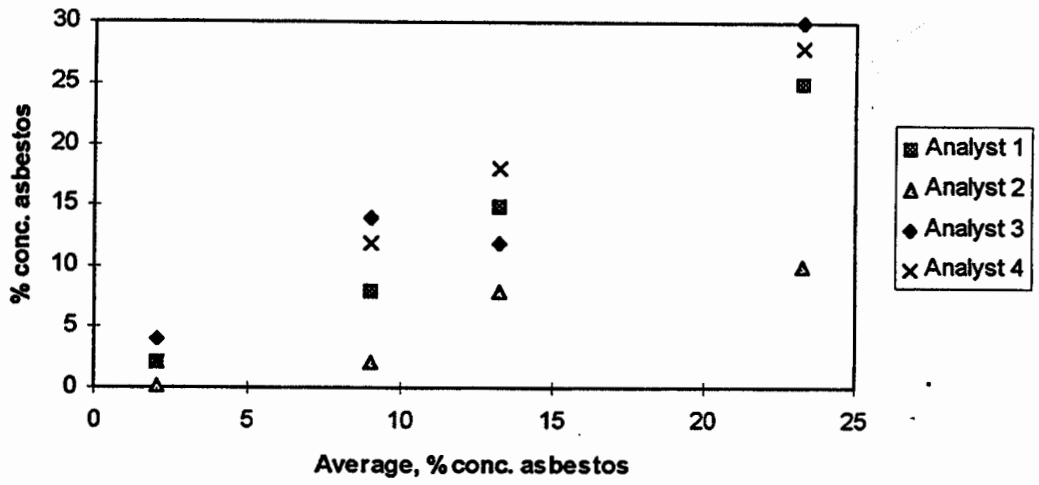

Figure 5. An example of *within laboratory accuracy* data to test for systematic differences among analysts. The results for each of 4 analysts for 4 samples are plotted with respect to the average for each sample.

REFERENCE ANALYSES

Reference materials are samples for which the concentration of asbestos is known, either through an independent, accurate, technique such as XRD or gravimetry, or because the materials are synthesized to contain known concentrations of asbestos. Analysis of reference materials is the only way to determine the ability of an analyst to obtain the "true" value. The difference between the analyst's result and the reference value is determined, but it is called an error, not a difference, because the true value is known. The data are handled in exactly the same way as described for *within analyst* and *within laboratory accuracy* by separating the blanks and the samples with qualitative errors, and then determining medians, IQRs, and control limits on the errors according to concentration category.

The reference data for one of the analysts from the volunteer laboratory are shown in Table 3 and in Figure 6. The "error" column in Table 3 was determined by subtracting the reference value from the analyst's result. The reference sample data for an individual analyst should show scatter, as indicated by the IQR, comparable to the *within analyst accuracy* for that analyst. The medians determined for reference data indicate a bias, or systematic error, on the part of the analyst. The

reference data should be looked at quite seriously, since these are the data that will demonstrate the laboratory's accuracy in determining whether a sample is asbestos-containing or not.

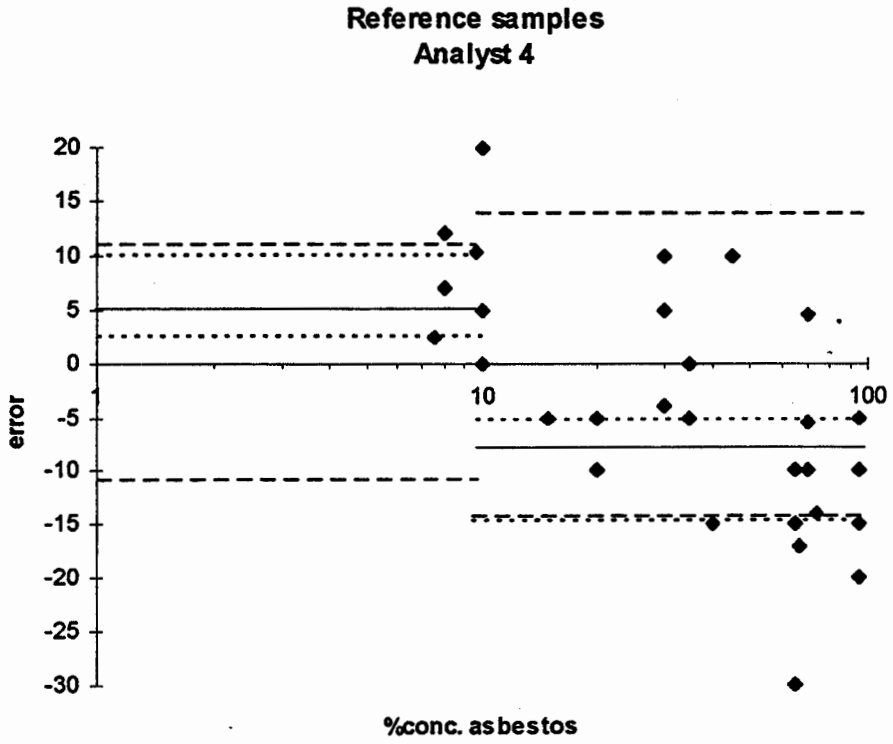

Figure 6. Reference data for analyst showing errors, calculated as result minus reference value, *versus* concentration. Medians (solid lines), IQRs (dotted lines) and control limits (dashed lines) are shown for concentration categories 2 and 3.

Analyst	Sample	Date	Reference	Result	ACM Ref	ACM Res	ACM flag	error	category
4	7217	34584	0.1	20	N	Y	F		
4	4118	34414	10	10	Y	Y	0	0	2
4	9718	34677	10	10	Y	Y	0	0	2
4	7628	34456	7.5	10	Y	Y	0	2.5	2
4	7628	34677	7.5	10	Y	Y	0	2.5	2
4	9718	34456	10	15	Y	Y	0	5	2
4	8119	34486	10	15	Y	Y	0	5	2
4	3836	34584	8	15	Y	Y	0	7	2
4	1929	34584	9.7	20	Y	Y	0	10.3	2
4	1929	34366	9.7	20	Y	Y	0	10.3	2
4	3836	34366	8	20	Y	Y	0	12	2
4	7045	34696	10	30	Y	Y	0	20	2
4	5824	34414	65	35	Y	Y	0	-30	3
4	770	34388	95	75	Y	Y	0	-20	3
4	3131	34612	67	50	Y	Y	0	-17	3
4	2819	34346	40	25	Y	Y	0	-15	3
4	1223	34661	65	50	Y	Y	0	-15	3
4	1024	34359	65	50	Y	Y	0	-15	3
4	636	34397	95	80	Y	Y	0	-15	3
4	535	34346	74	60	Y	Y	0	-14	3
4	6436	34470	20	10	Y	Y	0	-10	3
4	1223	34351	65	55	Y	Y	0	-10	3
4	2720	34612	70	60	Y	Y	0	-10	3
4	2138	34379	70	60	Y	Y	0	-10	3
4	770	34547	95	85	Y	Y	0	-10	3
4	8625	34470	70.5	65	Y	Y	0	-5.5	3
4	114	34612	15	10	Y	Y	0	-5	3
4	1730	34486	20	15	Y	Y	0	-5	3
4	6436	34605	20	15	Y	Y	0	-5	3
4	5818	34470	35	30	Y	Y	0	-5	3
4	4116	34536	95	90	Y	Y	0	-5	3
4	770	34388	95	90	Y	Y	0	-5	3
4	5346	34470	30	26	Y	Y	0	-4	3
4	5818	34605	35	35	Y	Y	0	0	3
4	8625	34605	70.5	75	Y	Y	0	4.5	3
4	5346	34605	30	35	Y	Y	0	5	3
4	9120	34346	30	40	Y	Y	0	10	3
4	9549	34346	45	55	Y	Y	0	10	3

Statistics for r1-r2

Concentration category 2: Median = 5%, IQR = 2.5% to 10.3%, Control limits = ±11.6%

Concentration category 3: Median = -7.75%, IQR = -14.75% to -5%, Control limits = ±14.4%

Table 3. Reference data for analyst 4.

The first entry in Table 3 represents a serious problem in that a result of 20% was returned for a reference sample containing 0.1% asbestos, generating an ACM flag. Unfortunately, there are no other reference samples close in concentration to this one, and so the single data point that would demonstrate proficiency near 1% shows the absence of proficiency. All the samples in concentration category 2 (1%-10% asbestos) are in the 7.5%-10% range, which is not satisfactory for demonstrating proficiency near 1% or for generating statistics that are appropriate for the full concentration range. The median, IQR, and control limit calculated for this range are given at the bottom of Table 3, and indicate a large positive bias (+5%) and a larger scatter than was seen even in the *within laboratory accuracy* data. The analyst has a negative bias at higher concentrations (concentration category 3), so much so that the lower limit of the IQR falls outside of the control limit, which is positioned about zero.

The reference sample data should be used initially to demonstrate the ability of the analyst to perform an adequate analysis; in particular, to discriminate between samples that contain asbestos and those that do not. Only after general proficiency has been demonstrated do "control limits" have any meaning. After a level of competence has been demonstrated, the reference sample data can be used to characterize the accuracy of the analyst by demonstration of minimum bias and acceptable scatter.

IMPLEMENTATION

A sufficient amount of QC data must be collected prior to determining any statistics to have a meaningful result. We recommend a minimum of 8 points per concentration category because of the symmetry needed for a useful interquartile range (an equal number of high points to low points). Once the minimum number of points has been collected for each concentration category for each type of measurement (*within analyst accuracy, within laboratory accuracy*, and reference) the graphs can be generated and the statistics calculated.

Once the minimum amount of data has been collected, additional data are appended to the file as collected. The results of the QC data should be analyzed and summarized monthly. The tabulation should include any qualitative errors and their resolution, the *within analyst accuracy* and reference sample statistics of each analyst including graphs, and the *within laboratory accuracy* statistics, including graphs. The number of points that fall outside the control limits, and the actions taken to remedy the problems should be identified. Any improvements seen in the data that imply a reduction of the bias or decrease in the scatter should be discussed. At some point, older data can be segregated by simply excluding some data from the analysis according to date (but without deleting the data). If the data are improving with time, the statistics will better reflect this if the older data are gradually removed. The implications of the QC data on the ability of the laboratory to perform quality analyses should be stated, in particular the data that support the laboratory's ability to discriminate asbestos-containing samples from samples containing <1% asbestos. As the data for the low concentration samples increases, it may be advisable to further segregate concentration category 2 (1-10% asbestos) into smaller categories such as 1-5% and 5-

10%. This would provide more meaningful limits for the critical 1% samples.

REFERENCES

1. AHERA - 40 CFR Part 763, Vol. 52, No. 210: "Asbestos Containing Materials in Schools; Final Rule and Notice.

2. Steel, E.B., Verkouteren, J.R., Alderman, D.F., "Bulk Asbestos Analysis", NIST Handbook 150-3, U.S. Dept. Of Commerce, Technology Administration, National Institute of Standards and Technology.

3. Perkins, R.L. and Harvey, B.W., Test Method, Method for the Determination of Asbestos in Bulk Building Materials, EPA/600/R-93/116, July 1993.

4. Youden, W.J. "Graphical Diagnosis of Interlaboratory Test Results", Industrial Quality Control 1959, 15.

5. Stuart, A., and Ord, J.K., Kendall's Advanced Theory of Statistics, Vol. 1, 5th Ed., Oxford University Press, New York, 1987, pp. 332-334.

APPENDIX 1
METHOD FOR CALCULATING QUARTILES AND MEDIANS

Preliminary steps

1a) Sort data into ascending order
1b) Define the first datum as "D_1", the second as "D_2", and so on.

2a) Count the number of data.
2b) Define "N" to be the number of data: N = COUNT(data)
2c) Define "R" to be the remainder of the integer division of N/4: R = MOD(N,4). R will be equal to: 1, 2, 3, or 0. (R is the numerator of the fractional remainder of the division, e.g. 9/4 = 2 1/4, R=1.)
2d) Define "N_1" to be the integer part of (3+N)/4: N_1 = FLOOR((3+N)/4,1). This is not a rounding procedure, e.g. the integer part of 3.75 is 3.
2e) Define "N_2" to be the integer part of N/2: N_2 = FLOOR(N/2,1).
2f) Define "N_3" to be the integer part of (1+3*N)/4: N_3 = FLOOR((1+3*N)/4,1).

To determine the first Quartile (Q_1)

3a) If R is 1, the first quartile is the N_1th data value: $Q_1 = D_{N1}$
(D_{N1} is the value found at the position in the sorted data given by N_1, not the value calculated for N_1)
3b) If R is 2, the first quartile is 1/4th the sum of 3*(the N_1th) and the (N_1+1)th data values: $Q_1 = (3*D_{N1} + D_{N1+1})/4$
3c) If R is 3, the first quartile is 1/2th the sum of the N_1th and (N_1+1)th data values: $Q_1 = (D_{N1} + D_{N1+1})/2$
3d) If R is 0, the first quartile is 1/4th the sum of the N_1th and 3*(the (N_1+1)th) data values: $Q_1 = (D_{N1} + 3*D_{N1+1})/4$

To determine the Median (Q_2)

4a) If R is either 1 or 3 (N is odd), the Median is the (N_2+1)th data value: Median = D_{N2+1}
4b) If R is either 0 or 2 (N is even), the Median is 1/2th the sum of the N_2th and (N_2+1)th data values: Median = $(D_{N2} + D_{N2+1})/2$

To determine the third Quartile (Q_3)

5a) If R is 1, the third quartile is the N_3th data value: $Q_3 = D_{N3}$
5b) If R is 2, the third quartile is 1/4th the sum of the N_3th and 3*(the (N_3+1)th) data values: $Q_3 = (D_{N3} + 3*D_{N3+1})/4$
5c) If R is 3, the third quartile is 1/2th the sum of the N_3th and the (N_3+1)th data values: $Q_3 = (D_{N3} + D_{N3+1})/2$
5d) If R is 0, the third quartile is 1/4th the sum of 3*(the N_3th) and the (N_3+1)th data values: $Q_3 = (3*D_{N3} + D_{N3+1})/4$

To determine the InterQuartile Range (IQR)

6) The InterQuartile Range is the difference between Q_3 and Q_1: $IQR = Q_3 - Q_1$

Example Calculations

Data

Datum	Set 1	Set 2	Set 3	Set 4
D_1	1.0	1.0	1.0	1.0
D_2	2.0	2.0	2.0	2.0
D_3	3.0	3.0	3.0	3.0
D_4	4.0	4.0	4.0	4.0
D_5	5.0	5.0	5.0	5.0
D_6	6.0	6.0	6.0	6.0
D_7	7.0	7.0	7.0	7.0
D_8	8.0	8.0	8.0	8.0
D_9	9.0	9.0	9.0	9.0
D_{10}		10.0	10.0	10.0
D_{11}			11.0	11.0
D_{12}				12.0

Intermediate terms

N	9	10	11	12
R	1 (9/4)	2 (10/4)	3 (11/4)	0 (12/4)
N_1	3 (12/4)	3 (13/4)	3 (14/4)	3 (15/4)
N_2	4 (9/2)	5 (10/2)	5 (11/2)	6 (12/2)
N_3	7 (28/4)	7 (31/4)	8 (34/4)	9 (37/4)

Statistics

Q_1	3.00 (D_3)	3.25 ($3*D_3+D_4$)/4	3.50 (D_3+D_4)/2	3.75 (D_3+3*D_4)/4
Median (Q_2)	5.00 (D_5)	5.50 (D_5+D_6)/2	6.00 (D_6)	6.50 (D_6+D_7)/2
Q_3	7.00 (D_7)	7.75 (D_7+3*D_8)/4	8.50 (D_8+D_9)/2	9.25 ($3*D_9+D_{10}$)/4
IQR	4.0	4.5	5.0	5.5