American Industrial Hygiene Association Journal

ISSN: 0002-8894 (Print) (Online) Journal homepage: http://www.tandfonline.com/loi/aiha20

Phase Contrast Microscopy Asbestos Fiber Counting Performance in the Proficiency Analytical Testing Program

Paul C. Schlecht & Stanley A. Shulman

PHASE CONTRAST MICROSCOPY ASBESTOS FIBER COUNTING PERFORMANCE IN THE PROFICIENCY ANALYTICAL TESTING PROGRAM

Paul C. Schlecht
Stanley A. Shulman

U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Division of Physical Sciences and Engineering, 4676 Columbia Parkway, Mailstop R8, Cincinnati, OH 45226

This report evaluates 20 years (1972–1992) of asbestos fiber count reporting for the Proficiency Analytical Testing (PAT) program, which is operated by the American Industrial Hygiene Association (AIHA) in cooperation with the National Institute for Occupational Safety and Health (NIOSH). Estimates were obtained for total, intracounter, and intercounter variability. Results show that total variability of counting chrysotile asbestos fibers improved by approximately 35% in recent years when compared with the variability found during 1975–1977, at the lowest filter fiber densities used in the PAT program. Total, intercounter, and intracounter variability for counting amosite and chrysotile asbestos fibers also were compared over a six-year period starting in 1986. PAT program laboratories achieved about one-quarter lower intracounter variability and about one-third lower total and intercounter variability when counting amosite fibers versus chrysotile fibers. In addition, amosite intercounter variability improved by about one-third, with large improvements occurring in the first year that amosite was included in the program. Factors affecting performance, such as changes in phase contrast microscope fiber counting methods, PAT participation, the AIHA Laboratory Accreditation Program, and PAT sample production, are discussed as possible factors affecting variability.

The Proficiency Analytical Testing (PAT) and the Asbestos Analyst Registry (AAR) programs are United States-based proficiency test programs that evaluate the quality of phase contrast microscope (PCM) asbestos fiber counting. Since 1972 the PAT program has measured laboratory performance in several analytical areas, including PCM asbestos fiber counting. Currently over 1200 participating laboratories analyze for asbestos, including over 350 laboratories that have obtained American Industrial Hygiene Association (AIHA) laboratory accreditation. A related program, the Asbestos Analyst Registry (AAR) Program, was started in 1986 to measure the performance of individual fiber counters. The AAR program has over 1800 PCM asbestos counters in 250 organizations participating. It is designed to provide quality audit samples to counters who perform asbestos fiber counts at remote or mobile sites such as at abatement sites or on board ship tenders.

This study identified trends in PCM asbestos fiber counting over a 20-year period and changes in factors such as counting methods, proficiency testing, and laboratory accreditation that may have affected these trends in performance. A similar study of the AAR data will be the topic of a subsequent paper.

Mention of company names or products does not necessarily constitute endorsement by the Centers for Disease Control and Prevention.

PAT PROGRAM OPERATION

The PAT program evaluates laboratories that perform PCM asbestos fiber counting. No fiber identification is involved. It uses audit samples prepared by filtering sonicated suspensions of chrysotile or amosite fibers and aluminum oxide through cellulose ester membrane filters. PAT samples have been generated using chrysotile since 1972, but beginning in 1988 amosite asbestos samples have been generated on alternate rounds.[1] The homogeneity of PAT audit samples is verified before shipment to participants. Manual PCM fiber counting and visual inspection by a single microscopist are used to evaluate filters for excessive fiber fineness or fiber clumping on at least 13 filters from each fiber loading level produced.

The PAT program rates and tracks overall laboratory performance,[2] providing a set of four asbestos fiber loading levels four times a year. The program also provides a blank filter. The set of samples is rotated from round to round among all the counters within the laboratory with one counter officially analyzing all asbestos samples on a round. Only counts from one counter are officially reported on a round, although laboratories are encouraged to fully utilize PAT program samples by using permanent filter mounts and additional filter wedges of PAT program samples to evaluate all counters within the laboratory and to repeat the evaluation of counters.

TABLE I. PAT Asbestos Fiber Counting—Variance Components[A]

Variance Components	Confounding Pattern	Terminology Relative Variance
1. Laboratory to Laboratory, or Between Laboratories $RSD^2 = RSD_B^2$	1. and 2.	intercounter variability $RSD^2 = RSD_{BC}^2 = RSD_B^2 + RSD_C^2$
2. Counter to Counter, Within a Laboratory $RSD^2 = RSD_C^2$		
3. Within Counter $RSD^2 = RSD_W^2$	3. and 4.	intracounter variability $RSD^2 = RSD_{WF}^2 = RSD_W^2 + RSD_F^2$
4. Filter-to-Filter[B] $RSD^2 = RSD_F^2$		
5. Total $RSD^2 = RSD_T^2$		total variability $RSD_T^2 = RSD_{BC}^2 + RSD_{WF}^2$

[A] Unless otherwise indicated the intracounter RSD^2 component will include both within counter and filter-to-filter variability. Total variability, RSD_T^2, is the sum of all four components above, but because of confounding:

$$RSD_T^2 = RSD_{WF}^2 + RSD_{BC}^2.$$

Also, if $S1_T$ and $S2_T$ are unbiased estimates of RSD_T^2, and W1 is an unbiased estimator of RSD_{WF}^2, then:

$$\text{Mean}\left(\frac{S1_T + S2_T}{2} - W1\right) = RSD_{BC}^2.$$

Thus, an unbiased estimator of RSD_{BC}^2 can be obtained when there are unbiased estimators of the other components.

[B] Can be estimated from PAT program sample producer quality control of asbestos sample homogeneity for rounds 46 and later

PRECISION ESTIMATES

This study used similar procedures to those used in a 1986 PAT program study to estimate the precision of PCM asbestos fiber counting.[3] Total variability (RSD_T^2) includes four variance components: between-laboratory variance (RSD_B^2), counter-to-counter variance (RDS_C^2), within-counter variance (RSD_W^2), and filter-to-filter variance (RSD_F^2). However, it is possible to estimate only two variance components. Since only one counter from each laboratory counts during a round, the variability between counters within a laboratory (RSD_C^2) is confounded with variability between laboratories (RSD_B^2). This variance is referred to as "intercounter variance" (RSD_{BC}^2). Since only one determination is made on each filter, filter-to-filter variability cannot be separated from within-counter variability (the variability resulting from differences in determinations where a person recounts the sample). This variance is referred to as "intracounter variance" (RSD_{WF}^2). (In the 1986 study this component was called "intralaboratory variance.") The total variance (RSD_T^2) is, then, the sum of the intercounter (RSD_{BC}^2) and intracounter variances (RSD_{WF}^2). (In the 1986 study the total variance was called "interlaboratory variance.") By estimating two of these three components, the third can be obtained. All three variances are expressed as relative variances, the squares of the relative standard deviations (see Table I).

Subtracting estimated RSD_{WF}^2 from estimated RSD_T^2 yields an estimate of the total variance associated with counters and laboratories. (For Rounds 46 and later, RSD_F can be estimated from filters counted by a single laboratory, the producer of PAT samples used to establish sample homogeneity prior to shipment to participants. For any given round from 46 and later, a counter employed by the producer made all counts on that given round—but only a small number of filters were counted, usually 13 filters for each sample.)

Estimates of total variability were obtained from sample standard deviations and means computed for each of the four samples on a round. Asbestos fiber counting was introduced into the PAT program in Round 3. (Round 67 data are not included in this study because of problems with the data.) The estimates were computed after removal of laboratory measurements that failed the Grubbs outlier test at the 1% significance level (two-sided test).[4] Confidence limits for precision estimates were based on the assumption that data are normally distributed after square root transformation, and were computed as shown in the Appendix.

The PAT program does not use duplicate samples. However, estimates of intracounter variability can be obtained by selecting pairs of PAT samples that are near the same filter fiber density (sample medians are within 100 fibers/mm^2). This results in about one-third of PAT program data being used to estimate intracounter variability, since only some rounds have sample pairs with approximately the same filter fiber density.

After estimates for the RSD_T and RSD_{WF} were obtained, estimates for RSD_{BC} could be obtained for those rounds where the pairing was carried out.

SELECTION OF TIME PERIODS

To evaluate changes in performance over time as measured by changes in precision, the data were divided into time periods based on the changes in PCM asbestos fiber counting and the PAT program (see Table II). These included changes in fiber counting methods, laboratory participation, the type of asbestos fiber used in sample generation, changes in PAT program sample producers, and changes in the AIHA Laboratory Accreditation Program. The time periods are similar to the ones used in the 1986 PAT asbestos study, but includes changes in the PCM asbestos fiber counting methods and PAT program asbestos fiber sample production.[3]

Three principal methods of PCM asbestos fiber counting have been used in the United States since the start of the PAT

TABLE II. Changes in the PAT Program—Asbestos

Per	*Round*	*Date*	*Sampling and Analytical Method*[A]	*Labs*[B]	*Sample*[C]	*Producer*[D]	*Accred*[E]
01	01–26	05/72–01/74	USPHS68	g	chry.	NIOSH	None
02	27–29	02/74–06/74	USPHS68	g&p	chry.	NIOSH	Accred.
03	30–32	08/74–02/75	USPHS68	g&p	chry.	CAP	Accred.
04	33–45	03/75–11/77	PCAM239	g&p	chry.	CAP	Accred.
05	46–75	01/78–10/83	PCAM239	g&p	chry.	SRI	Reacc.
06	76–85	01/84–04/86	NIOSH7400	g&p*	chry.	SRI	Reacc.
07	86–end	07/86–end	NIOSH7400	g&p*	+amosite	SRI	Reacc.

[A] USPHS68 = 1968 U.S. Public Health Service method;[4] PCAM239 = NIOSH P&CAM 239 method published in 1979 but distributed to PAT participants starting in 1975;[5] NIOSH7400 = NIOSH 7400 method and subsequent revisions[6]

[B] g = government labs; g&p = government & private labs; g&p* = growth in participation of many labs that perform asbestos but no other industrial hygiene analyses

[C] chry. = PAT samples are chrysotile; +amosite = PAT samples rotated each round among amosite and chrysotile. Starting with Round 86 even rounds are amosite.

[D] NIOSH = NIOSH producer of PAT samples; CAP = College of American Pathology's subcontractor, Hyland Labs, producer of PAT samples; SRI = SRI International producer of PAT samples

[E] Accred. = AIHA accreditation program in operation, and initial site visits of laboratories being performed; Reaccred. = AIHA accreditation program in operation and site revisit and evaluation of laboratory operations being performed

program in 1972. Table III summarizes some of the many significant changes in PCM fiber counting that occurred in terms of counting rules, counter training, microscope standardization, graticules, microscope resolution tests, and quality control.[5–7] Although the NIOSH P&CAM 239 method was published officially in 1979, draft versions of the method were distributed to PAT laboratories and counters attending NIOSH PCM asbestos fiber counting courses starting in 1975. Thus, its use by many participating laboratories pre-dates its official publication.

The type of asbestos laboratory participating in the PAT program also went through several changes: Participants were almost exclusively government laboratories prior to 1974; predominately full-service industrial hygiene laboratories from 1974 to the mid-1980s; and predominately laboratories involved exclusively with asbestos abatement in later years.

The study of trends in PAT program PCM asbestos fiber counting performance is more extensive for chrysotile samples, since amosite samples were only introduced in the PAT program in 1986.

CHANGES IN VARIABILITY ESTIMATES

One way to study the data is to plot total variability estimates over time. Figure 1 is a plot of the estimated RSD_T for chrysotile corresponding to filters with the lowest fiber density filter on each PAT round (lowest loading [L]).

The plot of total variability (RSD_T) at lowest loading levels indicates narrower confidence limits in the later periods than in earlier periods, largely due to the increase in the number of labs

TABLE III. Phase Contrast Microscope Asbestos Fiber Methods

Method	*USPHS 68*[4]	*PCAM 239*[5]	*NIOSH 7400*[6]
Counting Rules	vague	more specific —field boundary —fiber bundles	specific —examples —no hybrid rules
Training	none	recommended 1 counter/lab	required each counter
Microscope Standardization	little	scope differences cited	required resolution test
Graticule	Porton/other	Porton/other	Walton-Beckett
Mounting	dimethyl phthalate 2–30 days	dimethyl phthalate 2 days	acetone/triacetin hot block/video almost permanent
Working Conditions	not specified	not specified	specified —room lighting —breaks —ergonomics
Quality Control	not specified	recommended —10% recounts —workshops —recommended differences	required —field blanks —reference slides —10% blind recounts —proficiency testing —sample exchanges

doing the analyses in later years and also to factors that lowered the RSD estimates themselves. It also shows that the RSD_T estimates were lower and more constant since early in Period 5, but this observation is confounded by the fact that the earlier rounds generally had much bigger spreads from their lowest to their highest loadings. A comparison of the lowest loading plot, Figure 1, with a similar plot of highest loadings, indicates that within each period, there is some dependence on loading levels, with highest loadings having somewhat lower estimated RSD_T than do lowest loadings. However, trends in estimated RSD_T can be affected by differences in fiber loadings from round to round and sample to sample.

The intracounter estimates, RSD_{WF}, are plotted in Figure 2. They indicate narrower confidence limits for intracounter RSD (RSD_{WF}) in the later periods than in the earlier periods, a result of the increased number of laboratories participating in later periods. As with the total RSDs (RSD_T), the estimates from the early part of Period 5 onward were more constant and lower than in the earlier periods. And an examination of median loading levels indicates that the loadings for matched samples were much more variable in rounds before 1976 than after 1976. Also, the data indicate that RSD_{WF} estimates tend to decrease with increasing loading.

Intercounter estimates, RSD_{BC}, also were calculated. A similar plot of RSD_{BC} estimates indicates narrower confidence limits in later periods than in earlier periods, due to the increased number of labs in later periods. For Round 50 and later there was little difference between the RSD_{BC} estimates—all estimates were less than 0.46. In Period 4, Rounds 33 to 45, all estimates exceeded 0.46. Plots of the ratio of estimated RSD_{BC}^2 to estimated RSD_T^2 indicate that this ratio follows no clear time trends and, except for a few low values in Period 1, falls mostly between 0.5 and 0.8. This is not dependent on loading.

Plots of the estimated filter-to-filter variability from PAT sample producer quality-controlled data indicate that, regardless of the loading, this variability drops considerably after Rounds 46 to 60. At 300 fibers/mm^2, many RSD_F estimates exceeded 0.12 during Round 46–60. Afterward, most were less than 0.12. The lowest estimates were in the later rounds of Period 5, before it was necessary to use more than one batch per round in preparing the required number of samples. As more batches were required, estimated RSD_F increased.

However, filter-to-filter variance is a small component of total variance, and elimination of the filter-to-filter variance component resulted in very little change in RSD_T estimates.

TABLE IV. PAT Program Summary Statistics, 1986–1992

Total Variability	*Finding*
Chrysotile	
0.33–0.54 RSD Estimates	About 36% reduction since 1975–1977
0.44–0.47 RSD Model—300 fibers/mm^2 [A]	
0.41–0.43 RSD Model—500 fibers/mm^2 [A]	
0.37–0.39 RSD Model—700 fibers/mm^2 [A]	
Amosite	
0.24–0.59 RSD Estimates	
0.53–0.82 ratio—300 fibers/mm^2 [B]	
0.59–0.71 ratio—500 fibers/mm^2 [B]	About 35% smaller RSDs for amosite than chrysotile
Intracounter Variability	
Chrysotile	
0.18–0.33 RSD Estimates	
0.16–0.24 RSD Model—300 fibers/mm^2 [A]	
0.14–0.22 RSD Model—300 fibers/mm^2 [A]	About 42% reduction since 1975–1977
Amosite	
0.15–0.24 RSD Estimates	
0.62–0.94 ratio—300 & 500 fibers/mm^2 [B]	About 22% smaller RSDs for amosite than chrysotile
Intercounter Variability	
Chrysotile	
0.33–0.44 RSD Estimates	
0.31–0.40 RSD Model—300 fibers/mm^2 [A]	
0.30–0.39 RSD Model—500 fibers/mm^2 [A]	About 37% reduction since 1975–1977
Amosite	
0.21–0.55 RSD Estimates	
0.47–0.85 Ratio—300 fibers/mm^2 [B]	At 300 fibers/mm^2 amosite RSDs about
0.61–1.00 Ratio—500 fibers/mm^2 [B]	34% less than chrysotile

[A] Simultaneous 90% confidence limits
[B] Simultaneous 90% confidence limits on ratio of amosite to chrysotile RSD

MODELS FOR VARIABILITY OVER TIME

Identifying trends solely from plots of variability estimates (either RSD_T or one of the variability components) has a major limitation. PAT fiber loading levels have varied over a wide range from under 100 fibers/mm^2 to over 2500 fibers/mm^2. Even limiting plots to the highest or lowest fiber loading levels such as those shown in Figure 1 still includes samples with a wide range of levels (median range from 23 to 1510 fibers/mm^2 for data plotted in Figure 1). Since some components of variability and total variability may be affected by loading level, confirming observations made from simple plots of variability estimates using models that take into account changes in fiber loading levels are useful.

Some of the differences across periods and by loading can be quantified by making statistical models for variability consider differences in loading from round to round. Several different

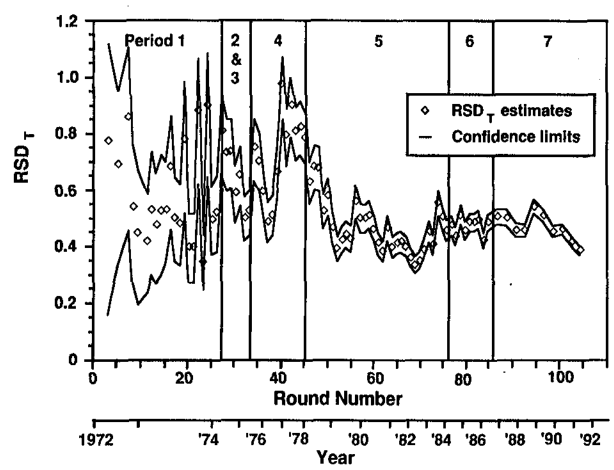

FIGURE 1. RSD_T estimates and 95% confidence limits for lowest loadings—chrysotile

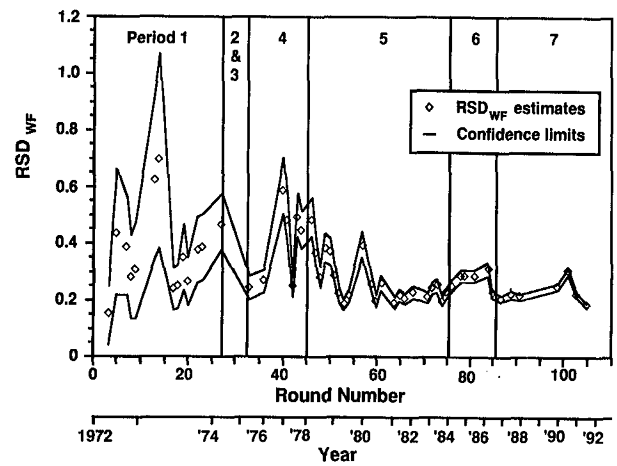

FIGURE 2. RSD_{WF} estimates and 95% confidence limits for paired loadings—chrysotile

models at these fiber loading levels (300, 500, and 700 fibers/square) have been constructed, resulting in similar conclusions about performance trends. These are discussed in the Appendix. Two periods are said to differ if their confidence intervals do not overlap.

The amount of intracounter data was limited, because no more than one sample pair resulting in a single predicted value was available for each round. Therefore, matched samples were not separated by loading. As for the intercounter RSDs, two periods were considered statistically different if the lower confidence limits on one period's RSDs exceeded the upper confidence limits on the other period's RSDs.

Models for intercounter variability based on the differences of estimated intracounter RSD were determined by procedures analogous to intracounter models and were obtained at 300 fibers/mm^2.

CHANGES IN TOTAL VARIABILITY

All statements of statistically significant differences control the Type I error at the 10% level simultaneously, over all comparisons of periods at the three loadings. The plots (total RSD by round) show that data from some periods were modeled by horizontal lines, and from other periods by straight lines with nonzero slope (indicating change within period) or even by quadratics (also indicating change within period). The choices of which forms to fit by loading and period were based on use of Mallow's C_p statistic. The models chosen were not the only reasonable models that might have been chosen, but the alternative choices would have yielded similar conclusions. Although models were fitted for data at 300, 500, and 700 fibers/mm^2, only results of data at 300 fibers/mm^2 are discussed.

At 300 fibers/mm^2, the lowest filter fiber density modeled, the fitted model for RSD_T indicates that both Periods 6 and 7 intercounter RSDs were lower than the total RSDs of Period 4. In other words, intercounter variability after 1984 was consistently lower than the intercounter variability found from 1975 to 1977—about 35% lower. According to the preferred models indicated by the C_p statistic, there is no trend (changes in intercounter RSD) within period for Periods 4, 6, and 7. Thus, the fitted values and the associated confidence limits are all horizontal lines (see Figure 3). The lower confidence limits for Period 4 were higher than the upper confidence limits for the Periods 6 and 7 model-based predictions, indicating that the differences in intercounter RSDs were statistically significant for all rounds within these time periods. (It is not clear why rounds in some periods differ. It is certainly easiest to draw conclusions when the differences between periods are consistent over all rounds within periods, as shown by horizontal lines.)

At 300 fibers/mm^2 Periods 1 and 5 showed changes in intercounter variability that occurred within these time periods.

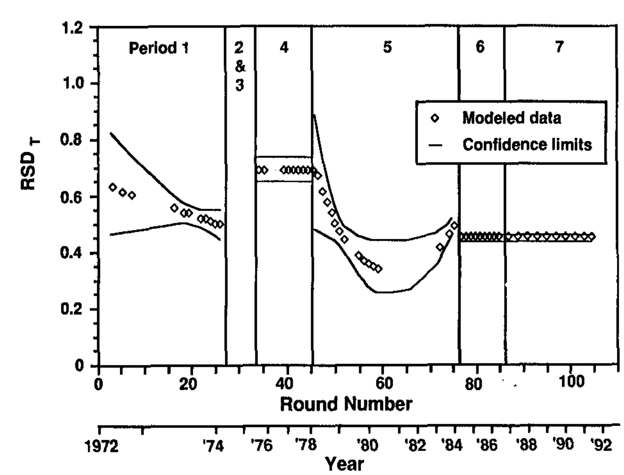

FIGURE 3. RSD_T model, 300 fibers/mm^2, chrysotile; 90% confidence limits for all comparisons

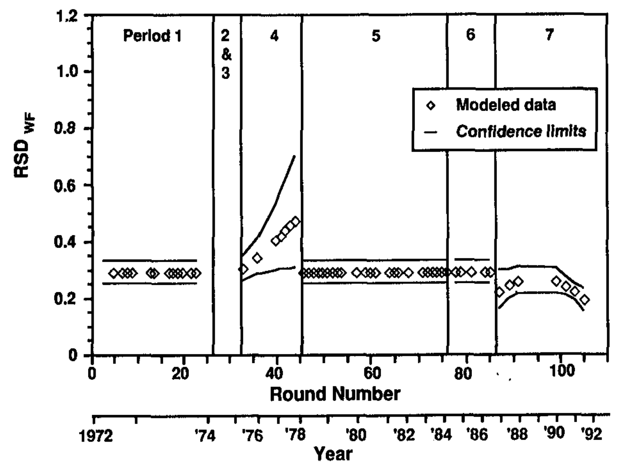

FIGURE 4. RSD_{WF} model, 300 fibers/mm², chrysotile; 90% confidence limits for all comparisons

The early rounds of Period 5 (from early 1978) were indistinguishable from Period 4 (1975 to 1977). However, by the fourth round of Period 4 (Round 49; mid-1978), Period 5 intercounter variability was lower than that in Period 4. Thus, the 1975 to 1977 intercounter variability exceeded later intercounter variability, with the exception of counting conducted in early 1978 (Figure 3).

Also, Periods 5 (after the fourth round), 6, and 7 RSDs were not distinguishable from each other. There appears to have been little improvement in intercounter variability since Round 46 (1978) at the 300 fibers/mm² level.

Comparisons also could be made with Period 1. (Periods 2 and 3 have only three rounds each.) However, few laboratories participated in early PAT rounds, resulting in wide confidence intervals for Period 1. Period 1 results therefore were not significantly different from the later periods.

Thus, the main observation is that the four RSDs for the most recent period were statistically lower than those of Period 4 at the lowest filter fiber density modeled (300 fibers/mm² loading), and the reduction was about 35%. This is of special interest, because the performance of the method at low loadings is important for fiber counting near the present standards.

CHANGES IN INTRACOUNTER VARIABILITY

The results of the intracounter model are displayed in Figure 4. All statistically significant differences are at the 10% significance level.

At the 300 fiber/mm² level most estimates for intracounter variability from Period 7 were lower than estimates for intracounter variability from Period 4, 1975 to 1977. Except for the first few rounds of the period, Period 4 RSDs were larger than Period 7 RSDs in tests done at the 10% significance level. Because the estimates for the first few rounds of Period 4 correspond to loadings in excess of 600 fibers/mm², the predictions at 300 fibers/mm² for those early rounds seem uncertain. Thus, it is reasonable to call the difference between Periods 4 and 7 RSDs statistically significant; Period 7 RSDs were over 40% less, on average.

All other comparisons of intracounter variability among time periods were not conclusive, since confidence intervals overlap considerably.

CHANGES IN INTERCOUNTER VARIABILITY

The results of the intercounter RSD model at 300 fibers/mm² are shown in Figure 5. The models fitted to the intercounter variability data at 300 fibers/mm² indicated that the differences between the Periods 4 and 6 RSDs and between the Periods 4 and 7 RSDs were significant at the 10% level at both loadings except for comparisons involving the first few rounds of Periods 6 and 7, which do not yield RSD_{BC}s statistically distinguishable from the first few rounds of Period 4. On average, the Period 7 RSD_{BC}s were about 63% of the Period 4 RSD_{BC}s. One reason to minimize the importance of the comparisons involving the first few rounds of Periods 4 and 7 is that for those rounds, the available data corresponds to loadings higher than 600 fibers/mm². Therefore, the quality of the predictions at 300 fibers/mm² for those early rounds is uncertain.

The reduction in intercounter RSDs can be related to the changes observed in total RSDs. As was previously indicated, the ratio of estimated RSD_{BC} to RSD_T did not change in a systematic way over time periods. When the ratio was modeled, there was no statistically significant difference between the ratio for Period 4 compared with that for Period 7, although the ratio varied slightly by loading—about 0.84 at 300 fibers/mm² and a little higher as the loading increased.

These estimated values were used to obtain the following result, at 300 fibers/mm²:

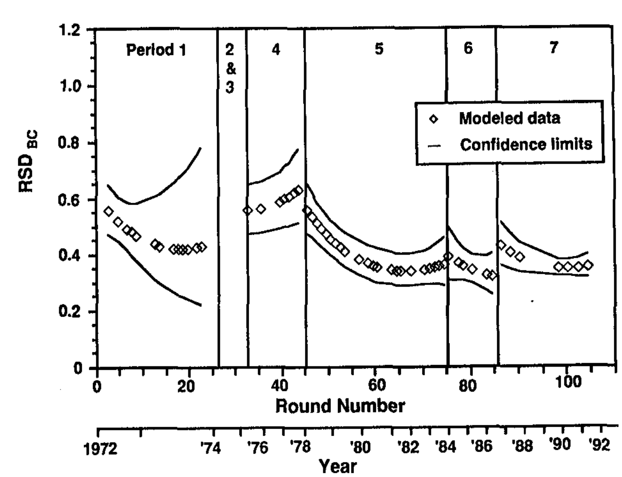

FIGURE 5. RSD_{BC} model, 300 fibers/mm², chrysotile; 90% confidence limits for all comparisons

$$RSD_T \text{ (Period 7)} \sim (1/0.84)\, RSD_{BC} \text{ (Period 7)}$$
$$\sim (1/0.84)(0.63)\, RSD_{BC} \text{ (Period 4)}$$
$$\sim (1/0.84)(0.63)(0.84)\, RSD_T \text{ (Period 4)}$$
$$\sim 0.63\, RSD_T \text{ (Period 4)},$$

where "~" means "approximately."

Thus,

$$RSD_T \text{ (Period 7)} \sim 0.63\, RSD_T \text{ (Period 4)}.$$

This approximation only makes sense because RSD_{BC} is a much larger part of RSD_T than RSD_{WF}.

From the model for RSD_T at 300 fibers/mm^2, whose predicted values are shown in Figure 3, the ratio of the Period 7 to Period 4 RSD_Ts is 0.64. The results here indicate that a main reason for the improvement in total variability at 300 fibers/mm^2 is that intercounter variability has improved.

COMPARISON OF AMOSITE AND CHRYSOTILE PERFORMANCE

Since 1986, amosite samples have been produced in the PAT program on even rounds, chrysotile samples on odd rounds. Plots of Period 7 RSD_Ts, by lowest loadings, are given in Figure 6, which shows a downward trend in both amosite and chrysotile intercounter RSDs, with considerable improvement in amosite occurring on the first three rounds after amosite was introduced into the program. As a result amosite total RSDs were consistently lower than chrysotile total RSDs.

The simultaneous 90% confidence limits on the ratio of the amosite total RSD to the chrysotile RSD were (0.534, 0.823) at 300 fibers/mm^2.

As with the comparison of total RSDs for chrysotile, the RSDs for amosite were modeled by a quadratic in each round. The predicted values at the concentration used above were produced from these fitted models, from which the following observations can be made. The data at 300 fibers/mm^2 indicate that whereas the amosite RSDs consistently declined from Round 86 to 104, the chrysotile RSDs increased slightly from Round 87 to 95, and then decreased to Round 105. Also, amosite RSD_T estimates at 300 fibers/mm^2 in Round 90 were one-third to one-half lower than in the first amosite round, Round 86.

The analysis of the intracounter RSDs indicates that the difference between the amosite and chrysotile RSDs varied with the loading, but the amosite RSD estimates tended to be lower. The estimated 90% confidence limits on the ratio of amosite intracounter RSD to chrysotile intracounter RSD were 0.616, and 0.939, respectively.

The estimated 90% confidence limits on the ratio of amosite intercounter RSDs to chrysotile intercounter RSDs were (0.471, 0.847) at 300 fibers/mm^2.

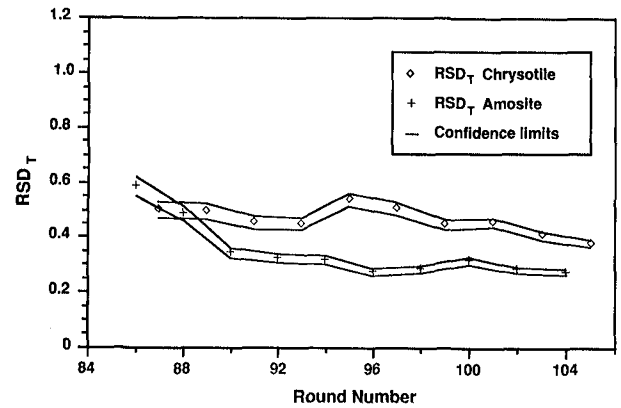

FIGURE 6. RSD_T estimates and 95% confidence limits for lowest loadings, Period 7, chrysotile and amosite

DISCUSSION

Total Variability

In recent years laboratories have demonstrated total precision of 0.33 to 0.54 for chrysotile and 0.24 to 0.59 for amosite in the PAT program. Models based on these estimates, which take into account the dependence of total variability on fiber loading, yielded simultaneous 90% confidence limits for chrysotile intercounter variability of 0.44 to 0.47 at 300 fibers/mm^2 and somewhat lower variability at higher loading levels for recent rounds.

The variability of asbestos fiber counting when compared to other analytes in the PAT program was quite large. For example, for the 4 samples in PAT Round 116, the 66 reference laboratories in the PAT program yielded RSD_Ts from 0.179 to 0.220. For all other analytes in the PAT program only silica had higher RSDs (0.225–0.264), about a 20–25% higher variability than asbestos fiber counting. For all other PAT analytes (various metals and organics) in Round 116, the estimated RSD_Ts on Round 116 were no higher than 0.068, or 70% or more lower variability than PAT asbestos fiber counting.

Initially when amosite was introduced to the PAT program in 1986, amosite total variability was similar to chrysotile total variability. But after one year (two amosite PAT program rounds) amosite total RSD estimates improved by one-third to one-half. Chrysotile PAT samples generally contain curved fibers longer than the 5 μ length rule for fiber counting. Although one expects amosite samples to be easier to count than chrysotile fibers, since amosite fibers are straight, the process of producing homogeneous amosite fibers results in a large number of fibers near the 5 μ cut-off for length. Therefore, PAT amosite samples are a much better challenge than either chrysotile fibers or typical field samples of amosite to measure a counter's ability to accurately estimate if fibers are longer than 5 μ.

Comparisons across time periods of individual total RSD estimates and total variability model predictions demonstrated that total variability has improved since the mid-1970s at the lowest fiber density level modeled (300 fibers/mm^2) by approximately 35%, and that improvement was statistically significant at the 90% confidence level. Improvements at higher fiber densities also may have occurred, but comparisons were not found to be statistically significant for all PAT program rounds within the time periods. The most important component of this improvement at 300 fibers/mm^2 appears to be reduction in variability between laboratories and between counters within laboratories.

Of special interest is the performance of PAT laboratories in 1978, since most of the improvement in total variability at 300 fibers/mm^2 appears to have taken place in the second half of 1978.

The PAT program contains one of the two largest databases on PCM asbestos fiber counting in the United States. Of the factors that are identified to describe time periods—analytical methods, type of participating laboratory, changes in PAT program sample production, and changes in AIHA laboratory accreditation requirements—three of these factors could have changed in 1978: (1) the widespread adoption of NIOSH P&CAM 239 or its predecessor draft methods by participating laboratories; (2) the effect on laboratories of the AIHA Laboratory Accreditation Program, which began the process of reaccrediting laboratories; and (3) the changes in the production of PAT program asbestos samples resulting in improved sample-to-sample homogeneity.

The changes in the NIOSH P&CAM 239 PCM fiber counting method over the U.S. Public Health Service method of 1968 are significant. These include an emphasis on training of counters, the identification of bias problems among different microscopes, more specific fiber counting rules, precautions concerning fiber migration over time with the dimethyl phthalate mounting technique, and an emphasis on frequent recounting of slides.[5]

Drafts of NIOSH P&CAM 239 were distributed to counters in NIOSH training courses and to PAT program laboratories starting in 1975, although the method was not officially published until 1979. Unfortunately, it is not possible from PAT program rounds to ascertain when during this four-year period most participating laboratories adopted the various improvements recommended by the NIOSH method. It is possible that the improvement in PAT program total variability was a result of adoption of P&CAM 239 improvements, if the majority of laboratories adopted P&CAM 239 in 1978, just prior to its official publication.

Also, the first large-scale reaccreditation of laboratories by AIHA was undertaken in 1978. As a result, considerable changes were made in the accreditation program to ensure timely implementation of corrective actions and development of a system to follow up quickly on PAT program performance.

Although the procedures used to generate PAT asbestos samples have remained essentially unchanged, the samples have been produced and quality controlled for sample-to-sample homogeneity by different groups over the years. Sample-to-sample homogeneity is a component of variation of both the intracounter and total variabilities in this study, but not of interlaboratory variability. Examination of Figures 3 through 5 indicates that the improvement in total variability at 300 fibers/mm^2 in 1978 is related to the improvement in intercounter variability, not to intracounter variability or sample-to-sample homogeneity.

Since 1978, PAT program total and intercounter variability have been fairly stable. This means that a noticeable improvement in intercounter variability with the introduction of NIOSH 7400 could not be demonstrated. This is somewhat surprising given the heavy emphasis of NIOSH 7400 on counter training, microscope standardization, standardization to the Walton-Beckett graticule, introduction of the HSE/NPL microscope resolution test, standardization of counting rules, the adoption of the acetone/triacetin mounting technique, and extensive quality control required by the method. These included daily use of reference slides, 10% blind recounts, and exchange of field samples among laboratories—all elements designed to improve intercounter variability.

Perhaps the number of laboratories outside an accreditation program in the 1980s and 1990s, and the fact that much of the asbestos fiber counting conducted today takes place at abatement sites, may be significant confounding factors in this study, since field counts may have different variability than laboratory counts.

As with the 1986 study, very little could be concluded about asbestos fiber counting prior to 1975, because few laboratories participated in PAT, and there were significant program changes in 1974 and early 1975.

Intracounter Variability

In recent years laboratories have demonstrated an intracounter RSD of 0.18 to 0.33 for chrysotile and 0.15 to 0.24 for amosite in the PAT program. Models based on these estimates predicted a 90% intracounter RSD for chrysotile in recent rounds of 0.16 to 0.24 and did not change very much by fiber loading. Comparisons among time periods indicated that except for the early rounds in Period 4, the Period 7 RSDs were lower than those for Period 4 for tests at the 10% significance level, at 300 fibers/mm^2.

The largest consistent difference in intracounter variability occurred between Period 4, 1975 to 1977, and the later time periods. A large reduction in intracounter variability of over 40% was found.

As with amosite total variability, some improvement in amosite intracounter variability occurred after the first year that amosite was introduced into the program.

Intercounter Variability

In recent years the estimated intercounter RSDs have been between 0.33 and 0.44 for chrysotile, and between 0.21 and 0.55 for amosite. For chrysotile, models based on these estimates predicted a 90% confidence interval for the intercounter RSD of 0.31 to 0.40 at 300 fibers/mm^2, and slightly lower limits at higher loading levels, 500 fibers/mm^2. Comparisons across time periods indicate that except for early in Period 4, the Period 7 RSDs were lower than Period 4 for 300 fibers/mm^2. The reduction in variability was about 37% for chrysotile.

CONCLUSIONS

This study used data from almost 20 years of the PAT program to evaluate the performance of participating laboratories and the changes in asbestos fiber counting that have occurred.

Chrysotile total variability and intercounter variability in the PAT program improved by approximately 35% since 1978. Statistically significant improvement in total variability occurred at the lowest fiber loading levels studied—300 fibers/mm^2. Much of the improvement occurred during the latter half of 1978. Two factors may have contributed to this performance improvement: (1) fiber counting methods such as NIOSH P&CAM 239 improved the standardization of counting rules and microscopes and recommended counter training and internal quality control involving frequent recounting, and (2) AIHA's industrial hygiene laboratory accreditation program began its first large-scale reaccreditation involving site visits to laboratories. Much of the improvement in total variability in 1978 was related to a reduction in intercounter variability, a possible consequence of these two factors.

A similar improvement in chrysotile total and intercounter performance could not be demonstrated when NIOSH 7400 was published in 1984. NIOSH 7400 incorporated improvements in fiber counting rules, counter training requirements, microscope and graticule standardization, mounting techniques, working conditions for counters, and extensive internal quality control including the exchange of field samples. However, other factors such as an increase in the number of PAT program participating laboratories not participating in a laboratory accreditation program, an increase in the amount of fiber counting being conducted in the field at abatement sites, and the rapid growth in counters may be significant confounding factors that mask improvements in analytical methods.

Amosite total RSD estimates at the 300 fibers/mm^2 loading improved by one-third to one-half within one year, two amosite PAT program rounds after amosite was introduced into the program in 1986. This is consistent with the finding of the 1986 PAT program study that laboratories improve chrysotile performance as they gain proficiency test experience, and suggests that it is useful to challenge laboratories with a variety of fiber types in proficiency testing to both obtain a good estimate of a laboratory's ability to count a variety of fiber types and to improve performance. For 300, 500, and 700 fibers/mm^2, amosite RSD_T^S were about one-third lower than chrysotile.

REFERENCES

1. **National Institute for Occupational Safety and Health:** "Proficiency Analytical Testing (PAT) Program Standard Reference Sample Generation Procedures," by SRI International, Menlo Park, CA, November 1987.
2. **National Institute for Occupational Safety and Health:** *Laboratory Reports and Rating Criteria for the Proficiency Analytical Testing Program (PAT)*, by J.H. Groff, P.C. Schlecht, and S.A. Shulman. [DHHS/NIOSH Pub. No. 91–102] Washington, D.C.: Government Printing Office, 1991.
3. **Schlecht, P.C. and S.A. Shulman:** Performance of asbestos fiber counting laboratories in the niosh proficiency analytical testing (PAT) program. *Am. Ind. Hyg. Assoc. J. 47*:259–269 (1986).
4. **Barnett, V. and T. Lewis:** *Outliers in Statistical Data*. New York: John Wiley and Sons, 1978. pp. 93–94.
5. **Edwards, G.H. and J.R. Lynch:** The method used by the U.S. Public Health Service for enumeration of asbestos dust on membrane filters. *Ann. Occup. Hyg. 11*:1–6 (1968).
6. **National Institute for Occupational Safety and Health:** *USPHS/NIOSH Membrane Filter Method for Evaluating Airborne Asbestos Fibers*, by N.A. Leidel, S.G, Bayer, R.D. Zumwalde, and K.A. Busch. [DHEW/NIOSH Pub. No. 79–127] Washington, D.C.: Government Printing Office, 1979.
7. **National Institute for Occupational Safety and Health:** *NIOSH Manual of Analytical Methods*, 3rd ed., by P.M. Eller, ed. [DHHS/NIOSH Pub. No. 84–100] Washington, D.C.: Government Printing Office, 1989. Method 7400 (supplement issued 5/15/89).
8. **Schlecht, P.C. and S.A. Schulman:** PCM fiber counting performance in the Proficiency Analytical Testing Program. Springfield, VA: National Technical Information Service. [Submitted]
9. **Hald, A.:** *Statistical Theory with Engineering Applications*. New York: John Wiley and Sons, 1952, p. 301.
10. **Daniel, C. and F. Wood:** *Fitting Equations to Data*, 2nd ed. New York: John Wiley and Sons, 1980. pp. 86–88.

APPENDIX

Estimation of Intercounter RSD

The intracounter standard deviation of the PAT data must be estimated differently from the total variability. The approach used in the previous study[3] selected rounds with paired samples, in which the means of these paired samples differed by no more than 100 fibers/mm^2. The present analyses used the same criterion, except that medians were used instead of means. The estimates were obtained from the residual mean square of an analysis of variance, in which square roots of the individual measurements for the paired samples in each selected round were used as the dependent variable. The explanatory factors in the analysis of variance were the sample number and the laboratory identifier. The residual mean square was used as an estimate of the intracounter variance. Taking the square root of this estimate, dividing by the mean, and converting the estimate to the original scale (since the data were on the square root scale) yielded an estimate of the intracounter RSD. The estimates can be converted from the square root scale to the original scale by using the relationship given later in this Appendix.

Thus, if $S1_T$ and $S2_T$ are unbiased estimates of RSD_T^2 for the two paired samples of Round r, and if W1 is the estimate of RSD_{WF}^2 obtained, as described above, from the analysis of variance, then $((S1_T \text{ and } S2_T)/2 - W1)$ is an unbiased estimate of RSD_{BC}^2. Again, the estimated RSD_{BC} must be transformed from the square root scale to the original scale.

SQUARE ROOT TRANSFORMATION TO ATTAIN NORMALITY

The data were assumed to follow a normal distribution after square root transformation. Normal probability plots of the samples in each round indicated that the square root transformation was usually appropriate to obtain approximate normality of the data. This approach differed from that used in the 1986 PAT program asbestos study, in which the square root was used for all data except eight PAT rounds, generally rounds with very high intercounter variability (0.47 to 1.20 RSD). In that study several samples were considered together for selection of a transformation.[3] In the present study the samples within each round were used separately to estimate the relative standard deviations, and the square root is usually no worse than and often better than the log transformation or cube root transformations previously used. As was demonstrated in the earlier PAT asbestos study, the estimates of precision (RSD) on the original scale are approximately twice those produced on the square root scale for normally distributed data. The exact relationship between results on the square root and the original scales is given in the next section. For approximately normally distributed data for large samples, the RSD_T estimator described above is approximately normally distributed. Except for the very early rounds, the number of laboratories involved should assure approximate normality of RSD_T estimates.

Conversion of Estimated RSD from Square Root to Original Scale

It can be shown that for data that are normally distributed on the square root scale, the true RSD on the original scale (RSDO) is approximately twice the true RSD on the square root scale. This approximation works best for small values of the RSD. An exact formula can be derived as follows: if X is normally distributed on the square root scale, with mean μ and variances σ^2, then the mean of X^2 is $\mu^2 + \sigma^2$, and the variance of X^2 is $2\sigma^2(2\mu^2 + \sigma^2)$.

Thus:

$$\text{RSDO} = \frac{\sqrt{2}\sigma\sqrt{(2\mu^2 + \sigma^2)}}{(\mu^2 + \sigma^2)} = \frac{\sqrt{2}\,\text{RSD}\sqrt{(2 + \text{RSD}^2)}}{(1 + \text{RSD}^2)}$$

Complete tables of RSD estimates, a more complete discussion of the models used, and plots of RSD trends at fiber loading levels above 300 fibers/mm^2 are contained in a National Technical Information Service report.[8]

Computation of Confidence Limits for RSDS

The confidence limits (CL) were computed from the following formula,[9] in which "Est." before a parameter denotes an estimate of the parameter:

$$\text{CL} = \text{Est. RSD} + 1.96 * \text{Est. Std. Err (RSD)},$$

where:

$$\text{Est. Std. Err. (RSD)} = \frac{\text{Est. RSD}}{(2 \times \text{D.F.})^{.5}} \times (1 + 2 \times (\text{Est. RSD})^2)^{.5}.$$

(D.F. is the degrees of freedom of the standard deviation used in the computation of the estimated RSD.)

REGRESSION MODELS FOR RSDS

Total RSD

The models at 300 fibers/mm^2 presented here modeled the total RSD (RSD_T) estimates in each round as a quadratic function of the sample median. These models were then used to obtain values for the total variability corresponding to 300 fibers/mm^2. Similarly, models have been constructed for 500 and 700 fibers/mm^2. The regression models were fitted over the seven time periods. The natural log of the predicted value of RSD was regressed on the period and round within period, with weights inversely proportional to the estimated variance of the predicted value. The model allowed for differences among periods and changes within each of the periods by including linear and quadratic components for change by round within period. Since the fullest model allowed for many parameters, Mallow's C_p was used to select models with a reduced number of parameters.[10] This statistic allows the user to find models that explain as much as possible of the variability in the data, but without including so many explanatory variables as to inflate the variance of the model predictions. Simultaneous 96.7% confidence limits were constructed for each period's predicted total RSDs under the models at each of the three loadings. All comparisons among periods and over all three levels were made with 90% confidence, since the error probabilities (3.33% at each loading) added to 10%.

INTRACOUNTER RSD

Regression models were constructed in which the intracounter RSD was expressed as a function of the period, the average of the medians for the matched samples, and the round within period. Also, terms representing products of these various factors were included. As with the total variability, the natural logs of the intracounter RSDs were used in the analysis. For this model, weights were inversely proportional to the degrees of freedom associated with the RSD estimates. Since the fullest model involves many factors, Mallow's C_p procedure, as for total variability models, was used to choose a group of acceptable models.[8]

The loading range for the paired samples changed somewhat from period to period. In Periods 6 and 7 there were few loadings in excess of 700 fibers/mm^2. If interest were restricted to two loadings, 300 and 500 fibers/mm^2, results on those loadings could be obtained from a single model, which models RSD_{WF} as a function of round within period and loading.

INTERCOUNTER RSD

Models for intercounter variability based on the differences of estimated intracounter RSD were determined by procedures analogous to intracounter models, and were obtained at 300 to 500 fibers/mm^2. The weight used on each round was inversely proportional to the degrees of freedom associated with the intracounter estimate for that round. The intracounter and total RSD estimates had approximately equal degrees of freedom, and as indicated in the previous section, the intercounter (and intracounter) RSD appeared to be relatively constant fractions of the total RSD for most of the rounds. Thus, the degrees of freedom associated with the RSD_{BC} estimates should be approximately proportional to the degrees of freedom of the intracounter (or total) RSD estimates.