Continued participation in an asbestos fiber-counting proficiency test with relocatable grid slides†

Martin Harper,*[a] James E. Slaven[b] and Thomas W. S. Pang[c]

Received 11th August 2008, Accepted 18th November 2008
First published as an Advance Article on the web 22nd December 2008
DOI: 10.1039/b813893a

The effect of using relocatable reference slides of chrysotile and amosite in asbestos fiber counting proficiency testing was examined for volunteer analysts from laboratories in the USA. Results of participation in one round have been published; two more rounds are reported here. In the first round, participants were asked to draw what they saw, allowing identification of error type by comparison to the reference. In later rounds only the number of fibers per field was reported since the number of errors per field has been shown to be a reasonable estimate of proficiency. The third round included a training exercise. The total number of participants stayed reasonably constant with some reduction over time. More restricted numbers participated from round to round. Those who dropped out had lower average scores than those that remained in the program; from 2006 to 2007 this difference was significant, but for 2007 to 2008 it was not. The overall results for amosite were generally good compared to an arbitrary proficiency score of 60, and continued to improve further over time. The results for chrysotile were better in rounds 1 and 3 than round 2, so that both attention to detail (drawing the fibers in round 1) and training (round 3) may improve performance, which is consistent with the major type of error being oversight of fine fibers. However, the results are still poor, even by round 3, and no analyst achieved a score of 60 in all three rounds. Further improvement is preferred since chrysotile is the most commonly encountered type of asbestos in the USA. Depending on the adopted score for proficiency many laboratories or analysts may be labeled as poor performers and this may be a deterrent to voluntary participation in this type of exercise, especially for those in most need of assistance. Participants have tested new relocatable reference asbestos proficiency counting slides in three rounds of chrysotile and three rounds of amosite. Performance for amosite was good. Poor performance for chrysotile appears to be improved by greater attention and training.

Introduction

The standard method for assessing exposure to airborne mineral fibers, such as asbestos, is to collect a sample by passing air through a membrane filter, the filter then being cleared and mounted on a slide for counting the fibers under phase-contrast microscopy (PCM) with subsequent calculation of the number concentration.[1–3] Fibers are defined by dimensional characteristics which are measured against a graticule. The limit of detection is defined by the resolution of the microscope which can vary according to the optics and set-up.[4] In addition, the diligence of the microscopist in examining the slide is of paramount importance in assuring an accurate count.[5–7] Rigorous control measures are considered essential to assuring the quality of the analyses, especially for occupational compliance samples. All analysts are required to have attended a week-long training course given by a recognized provider before they can analyze workplace samples. Analysts are required to set-up their microscopes each day according to a specified procedure and the resolution of the microscope is checked. Measurements are made according to a graticule which has had its dimensions calibrated. Analysts are required to analyze reference slides at the beginning of the day and 10% of sample slides in duplicate and be sure their results are within set bounds. Analysts also share slides with others in round-robins. Proficiency testing materials are also available, which have consisted of filters dosed with air-generated fibers or with liquid slurry preparations, or slides made from these filters, or slides made from actual ambient air filter samples.[8–10] In the USA, the supplier of proficiency testing materials is the American Industrial Hygiene Association (AIHA). The AIHA has two programs for fiber-counting of air samples: one for analysts and one for laboratories. Analysts receive proficiency materials through the Asbestos Analysts Testing (AAT) program and through demonstration of

[a]*National Institute for Occupational Safety and Health, Exposure Assessment Branch, Health Effects Laboratory Division, 1095 Willowdale Rd, MS-3030, Morgantown, WV, 26505, USA*
[b]*National Institute for Occupational Safety and Health, Biostatistics and Epidemiology Branch, Health Effects Laboratory Division, 1095 Willowdale Rd, MS-4020, Morgantown, WV, 26505, USA*
[c]*c/o Ryerson University, School of Public and Occupational Health, Room 249A, 350 Victoria St, Toronto, ON, M5B 2K3, Canada*

† Erratum: In our previous paper we had reported that asbestos filters from the Proficiency Analytical Test program of the American Industrial Hygiene Association (AIHA) were produced from air-generated asbestos. This is in fact not the case; these filters have always been generated from deposition of a liquid slurry, and the "REF" slides described above are made from these filters. The materials used in the Asbestos Analytical Test program of the AIHA's Asbestos Analysts Registry (AAR) were made from air-generated samples, but these also have been generated from liquid slurry deposition since 2006.

proficiency they can be registered in the Asbestos Analysts Registry (AAR). Laboratories receive materials through the Industrial Hygiene Proficiency Analytical Testing (IHPAT) program, and through demonstration of proficiency they can be accredited by the AIHA, although for several reasons many laboratories choose to participate in IHPAT but do not apply for accreditation. At the time of writing there are 775 active analysts from 146 organizations in the AAT program and 794 laboratories participating in the IHPAT program (147 of which are accredited for this analysis). Sixty-four of the organizations with AAT analysts also participate in the IHPAT program. While it is not possible to derive the exact number of asbestos PCM analysts in the USA from these figures, an estimate of between 1200 and 1500 appears reasonable. If this many analysts examine two slides per hour for twenty hours a week they may be responsible for more than 2.5 million analyses per year.

In all programs developed for proficiency testing it has not been possible to compare the counts made by microscopists to a reference on an object-by-object basis until recently. It is now possible to obtain "REF" slides made from chrysotile and amosite proficiency test filters with a relocatable grid imprinted on the cover slip.[11] The cover slips have two sets of circular grid openings where each opening can be identified by unique coordinates. These slides have been examined by reference counters of known good performance, who have agreed on what fibers should be counted in each field. There are two ways of using these consensus reference slides. The longer procedure is for the test analyst to draw the objects they see on a template of the grid and label each object to be counted as a fiber. These drawings can then be compared to the agreed reference and exact feedback can be given concerning errors. However, since the kinds of errors are similar from analyst to analyst, it is possible to determine proficiency by the simpler and less time-consuming technique of reporting the number of fibers per field. A proficiency score can then be calculated.[12] Feedback can still be given to the microscopist concerning the general nature of any errors, thereby facilitating continued improvement. These slides have been used in a proficiency test program in Canada for some years.[6] Recently, an improved version of the slides (with round grid openings that are unobstructed by the grid material) has been used in a pilot project in the USA, involving volunteer participants from the AIHA's Laboratory Quality Assurance Programs, including participants in the IHPAT program and the AAR. The results of the first round of testing of amosite and chrysotile slides have been reported.[11] It was shown that, in general, a greater proportion of microscopists may be considered proficient for counting amosite than chrysotile. There are several reasons for this difference: chrysotile fibers are generally thinner than amphibole fibers and their refractive index is closer to that of the mounting medium, so there is less contrast, and they are usually curved and less obvious than the usually much straighter amphibole fibers.

A chrysotile proficiency score of 60 was initially proposed for use in the Canadian program, but it was not actually adopted because of the large number of failures which would have resulted (a score of 50 has been used instead). In our study of volunteer US participants, a similar picture emerged; approximately 2 out of 3 US participants would have been considered non-proficient using a score of 60, but with a score of 50 pass/fail rates were similar to the Canadian situation. Poor performance in new proficiency test programs is commonly encountered, but an improvement is generally seen over time, as participants become used to the differences between the proficiency test materials and the samples they analyze routinely. In this particular case the slides are mounted in a different medium (Euparal in place of triacetin) and are covered by the relocatable grid, while a clear cover-slip would be used on routine samples and other proficiency test materials. However, references suggest no significant difference in fiber resolution between the two mounting media.[13]

Three rounds of this program have been completed (no others are envisioned under this project). In the first round of testing, the objective was to validate the ability to use proficiency scores without drawing. Therefore, participants were asked to draw the fibers they saw and return the drawings for comparison with the reference counts and derivation of their proficiency score. This procedure allowed us to determine the specific types of error and to determine if the errors were similar to those in the Canadian experience, and to determine the validity of the proficiency metric from simply reporting the number of fibers per field. With this latter procedure validated, the second and third rounds of testing required only the reporting of the number of fibers for each relocatable field. However, the second round showed a marked degradation in proficiency, especially for chrysotile, compared to the first round. In an effort to improve performance in the third round, the reference drawings of fibers in one of the two grid areas were provided to the participants to be examined as a training aide before reporting fibers in the other grid area.

Methods

Neither the sponsor of the research, the National Institute for Occupational Safety and Health (NIOSH), nor the AIHA, were made aware of the identity or scores of individual participants. However, in some cases, multiple analysts participated from a single laboratory, and so in order to compare results from round to round, the participants were given a unique identifier. Participation in this program was on a purely voluntary basis, so the number of participants and their identities has varied from round to round. A breakdown of participation is provided in Table 1 for chrysotile and in Table 2 for amosite. Participants in each round received different slides from the pool than those they received in a prior round. Each participant was provided a single slide of chrysotile or a single slide of amosite for each test. Participants were asked to examine designated fields of view on each slide, defined by the relocatable grid. For this project, approximately 100 fibers were intended to be counted on each slide (for some slides there were less), although this information was not conveyed to the participants. The number of fields needing to be examined to reach the total varied from 20 to 100. Results were examined against the "verified fibers" (consensus of two reference analysts) for each field. A proficiency score was calculated based on the number of fibers reported compared to the number verified. The basis of the calculation of the proficiency score is published.[11,12] A proficiency score of 100 is considered optimal. Scores >100 are not possible, since if the optical set-up is better than it should be (a test slide is used in the set-up of microscopes for workplace analyses to ensure a standardized resolution) and

Table 1 Participants in the chrysotile rounds

			Participation			
Round	No of labs	Total analysts	All 3 rounds	Rounds 1 + 2 only	Rounds 2 + 3 only	Single round only
1(2006)	31	60	27	4	—	29
2(2007)	29	54	27	4	13	10
3(2008)	29	53	27	—	13	13

Table 2 Participants in the amosite rounds

			Participation			
Round	No of labs	Total analysts	All 3 rounds	Rounds 1 + 2 only	Rounds 2 + 3 only	Single round only
1(2006)	42	85	29	16	—	40
2(2007)	35	75	29	16	8	22
3(2008)	35	62	29	—	8	25

very fine fibers become visible, errors of the type "oversight-extra" will occur and bring down the score. On the other hand, negative proficiency values are possible where there are more than 100 errors per 100 verified fibers.

The participants who dropped out after the first and second rounds were examined as a separate group, to determine whether they could be treated as a homogenous group with those that participated in multiple rounds. Then the participants in multiple rounds were divided as to whether they had achieved a score ≥60 or <60 in one round compared to the subsequent round. Finally, change in performance over time was determined for the group that participated in all three rounds. Differences were examined using t-tests. Based on the estimated population size and the low variance in the results, a sample size of approximately 30 gives the ability to detect a statistically significant difference at an alpha level of 0.05 with a power of 0.93. Doubling the sample size, *i.e.* when all participants in a round could be included, increases the power to 0.99. Using a score of 50 as the divider did not change the significance of the results.

Results

The mean performance for the groups as a whole and for those who stayed from round to round is given in Tables 3 and 4. When the participants who declined to participate in a subsequent round are examined separately, then their average score was worse than those who elected to remain in the study. For drop-outs after the first round this was significant for both chrysotile and amosite, but in the second round the lower average score was not significant.

The round-to-round performance of those who participated in two or more rounds is given in Table 5 (chrysotile) and Table 6 (amosite). For chrysotile, the average score for those who participated in both 2006 and 2007 and who scored <60 in 2006 did not increase significantly in 2007, while the average score for those who scored <60 in 2007 and who participated in 2008 rose significantly. However, those who performed well in 2006 did significantly worse in 2007, while those who performed well in 2007 had a slight, but not significant, decrease in average score in 2008.

For amosite, the average score for those who participated in both 2006 and 2007 who scored <60 in 2006 increased although the change is borderline non-significant ($p = 0.058$). On the other hand, those who scored highly in 2006 tended to remain high in 2007. A similar effect is seen between 2007 and 2008; although for those who scored <60 in 2007 the increase in average score to 2008 is higher than the increase in 2007 for those scoring <60 in 2006, fewer numbers meant having less statistical power to detect differences causing the p-value to be non-significant.

Table 3 Average proficiency scores from round to round: chrysotile

Round	Ave score all participants	% Proficient	Ave score drop-outs (N)	% Proficient	Ave score stayed (N)	% Proficient	Difference (p-value)
1(2006)	48.3	28.3	43.2 (29)	24.1	53.1 (31)	32.3	S 0.041
2(2007)	44.6	20.4	41.8 (14)	28.6	45.6 (40)	17.5	NS 0.50
3(2008)	53.8	41.5	—	—	—	—	—

Table 4 Average proficiency scores from round to round: amosite

Round	Ave score all participants	% Proficient	Ave score drop-outs (N)	% Proficient	Ave score stayed (N)	% Proficient	Difference (p-value)
1(2006)	52.0	56.5	38.5 (40)	45.0	64.0 (45)	66.7	S 0.0039
2(2007)	60.8	65.3	56.5 (38)	63.2	65.3 (37)	67.6	NS 0.29
3(2008)	71.1	83.9	—	—	—	—	—

Table 5 Changes in proficiency score as a function of prior score: chrysotile

	Score <60 in prior round (N 2007 = 21, N 2008 = 33)			Score ≥60 in prior round (N 2007 = 10, N 2008 = 7)		
Round	Ave score in prior round	Average score in current round	Difference (p-value)	Ave score in prior round	Average score in current round	Difference (p-value)
2(2007)	44.9	43.1	NS 0.68	70.2	42.6	S 0.0009
3(2008)	41.4	54.5	S 0.0009	65.6	59.8	NS 0.24

Table 6 Changes in proficiency score as a function of prior score: amosite

	Score <60 in prior round (N 2007 = 15, N 2008 = 12)			Score ≥60 in prior round (N 2007 = 29, N 2008 = 25)		
Round	Ave score in prior round	Average score in current round	Difference (p-value)	Ave score in prior round	Average score in current round	Difference (p-value)
2(2007)	37.8	53.3	NS 0.058	77.6	74.9	NS 0.35
3(2008)	33.3	58.7	NS 0.10	80.6	82.9	NS 0.24

The average scores for those that participated in all three rounds of chrysotile ($N = 27$) rose from 51.7 to 55.5 over the three rounds, which is significant ($p = 0.04$). The average scores for those that participated in all three rounds of amosite ($N = 29$) also rose, from 65.2 to 74.9, but this is not significant ($p = 0.33$).

Discussion and conclusion

The types of errors reported by the US analysts participating in this study are similar to those reported by Canadian analysts in a similar study. In the case of chrysotile, the errors are mostly oversight of fine fibers. This error was referred to as "oversight-missing" in our previous publication and accounted for 97% of the missing fibers and resulted in a 30% average underestimation of the count in the first round.[11] Those that performed poorly in the first round perhaps could do little worse in the second round, and hence their average scores did not change much. However, those that performed well in the first round showed a highly significant drop in performance in the second. The difference between these two rounds was that drawing the fibers was required in the first round but not in the second. Thus it can be hypothesized that the performance change may be the result of less attention to the slides perhaps coupled with complacency over the results from the first round. Interestingly, while the addition of training to the third round significantly increased the performance of those that did poorly in the second, it did not change the performance of those that performed well in the second round, giving some support to the notion of complacency (note that the highest score observed, 93, was still less than a perfect 100 so that further improvement was possible).

For amosite, the performance was overall better than for chrysotile. Those whose score was ≥60 in the first or second round generally maintained their high score in the subsequent round. However, while those who scored <60 in the first or second round increased their average scores in subsequent rounds, the low numbers of such participants resulted in changes of borderline non-significance. While training was added between 2007 and 2008, similar improvement occurred without training between 2006 and 2007. However, after the 2006 round feedback on the nature of the likely errors was provided to the participants. This suggests a possibility that poor performance in amosite counting may be self-rectifiable once recognized. This is not surprising as it has previously been shown that errors for amosite are mostly due to sizing which can be easily corrected through re-calibration.

Overall, group performance for chrysotile did improve significantly in round 3, presumably as a result of the training component. However, not only did no analyst obtain scores ≥60 in all 3 rounds of chrysotile, but from 2006 to 2007, the probability of an analyst scoring ≥60 in 2006 also scoring ≥60 in 2007 was only 10%, and for 2007 and 2008 the probability of an analyst scoring ≥60 in 2007 also scoring ≥60 in 2008 was only 20%. This has repercussions for the use of these slides in proficiency testing programs. Analysts may be unwilling to be subject to a program requiring round-to-round proficiency with so little chance of achieving it. This was the experience in the Canadian program and a lesser pass value of 50 was ultimately adopted. Seventy-two percent of participants in the last Canadian round of chrysotile slides met or exceeded this value, compared to 68% in round 3 of this project. For completeness, all of the statistical analyses in this study were repeated using a score of 50 to divide performance. The averages in Tables 5 and 6 changed but the significance or non-significance of differences did not.

Two observations give rise to concern. The first is that those who declined to participate in the second round were significantly poorer performers than those who elected to remain in the study. Average scores of "drop-outs" were also worse from 2007 to 2008, but not significantly less than those who remained. There are many reasons why an individual analyst or laboratory might not have participated in a second round, including business pressures on commercial laboratories. However, these observations suggest that the analysts that persisted in this study likely exhibited the performance of an above-average population and this raises doubts about the possible performance of those who elected not to participate at all. Nevertheless, as noted, the overall performance is comparable with the Canadian program.

This performance can also be put in context of a higher standard: scores ≥60% for chrysotile had only risen by the third round to just over 40% of participants even with a training component and for analysts that participated in all 3 rounds, the average score was still less than 60, again with no analyst scoring over 60 in all three rounds. Note that if all the errors for chrysotile are "oversight-missing" a score of 50 is equivalent to seeing only half the fibers visible to the reference analysts, experienced microscopists with good equipment properly set-up and calibrated. A score of 60 has only been used as a useful divider of performance for comparative purposes in this study. Should these "REF" slides, or equivalent, be used in other proficiency test programs it will be the decision of the test material providers and/or accrediting agencies as to how proficiency should be determined in their programs.

The second observation of concern is the decline in the performance in round 2 of chrysotile of analysts scoring ≥60 in round 1. This suggests a need for vigilance through continued testing. For the most part, participants have not been provided with specific feedback on their performance on a fiber-by-fiber basis, although this can be done, with the possible outcome of further improvements in performance, but with the corollary of greater expenditure of effort on the part of both the analyst and the proficiency test provider.

The possibility of being considered non-proficient may make analysts and laboratories wary or reluctant to participate in a scheme involving REF slides. Nevertheless, it is suggested that these slides be used in proficiency testing. The option of performance feedback also makes this procedure more useful than that used in current proficiency programs and this could lead to improved participant performance over time. Improved performance should be a goal. The data presented here indicate that many analysts see as little as 50% of chrysotile fibers seen by reference analysts and some analysts make more than 100 errors per 100 fibers. The observation that no analyst scored higher than 60 in all 3 rounds indicates a need for additional training. These REF slides could usefully be added to the required analyst training programs.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

Acknowledgements

This research was supported by the National Institute for Occupational Safety and Health through National Occupational Research Agenda (NORA) pilot project funding.

References

1 J. Carter, D. Taylor and P. A. Baron: Fibers, Method 7400, Revision #3: 4/15/89, in *Manual of Analytical Methods*, ed. P. M. Eller, DHHS (NIOSH) Pub. No. 84–100, Cincinnati, OH, 3rd edn, 1984.
2 Health and Safety Executive (HSE): *Asbestos Fibres in Air: Light Microscope Methods for Use with Control of Asbestos at Work Regulations*, (HSE MDHS 39/3), London, HSE, 1990.
3 World Health Organization (WHO): *Determination of Airborne Fiber Number Concentrations, A Recommended Method, by Phase-Contrast Optical Microscopy (Membrane Filter Method)*, Geneva,WHO, 1997.
4 S. J. Rooker, N. P. Vaughan and J. M. LeGuen, On the Visibility of Fibres by Phase Contrast Microscopy, *Am. Ind. Hyg. Assoc. J.*, 1982, **43**, 505–515.
5 P. A. Baron, Measurement of Airborne Fibers: A Review, *Ind. Health*, 2001, **39**, 39–50.
6 T. W. S. Pang, Precision and Accuracy of Asbestos Fiber Counting by Phase Contrast Microscopy, *Am. Ind. Hyg. Assoc. J.*, 2000, **61**, 529–538.
7 M. Harper and A. Bartolucci, Preparation and Examination of Proposed Consensus Reference Standards for Fiber-Counting, *Am. Ind. Hyg. Assoc. J.*, 2003, **64**, 283–287.
8 C. A. Esche, J. H. Groff, P. C. Schlecht and S. A. Shulman: *Laboratory Evaluation and Performance Reports for the Proficiency Analytical Testing (PAT) and Environmental Lead Proficiency Analytical Testing (ELPAT) Programs* (DHHS[NIOSH] Publication no. 95–104), Cincinnati, OH, US Department of Health and Human Services, Public Health Service, National Institute for Occupational Safety and Health, 1994.
9 N. P. Crawford, P. Brown and A. J. Cowie, The RICE and AFRICA Schemes for Asbestos Fiber Counting, *Ann. Occup. Hyg.*, 1992, **36**, 59–69.
10 M. C. Arroyo and J. M. Rojo, National vs. International Asbestos Fiber Counting Schemes: Comparison Between the Spanish International Quality Control Programme (PICC-FA) and the Asbestos Fiber Regular Informal Counting Arrangement (AFRICA), *Ann. Occup. Hyg.*, 1998, **42**, 97–104.
11 T. W. S. Pang and M. Harper, The quality of fiber counts using improved slides with relocatable fields, *J. Environ. Monit.*, 2008, **10**, 89–95.
12 T. W. S. Pang, A New Parameter to Evaluate the Quality of Fiber Count Data of Slides of Relocatable Fields, *J. Occup. Environ. Hyg.*, 2007, **4**, 129–144.
13 J. M. M. LeGuen and S. Galvin, Clearing and Mounting Techniques for the Evaluation of Asbestos Fibers by the Membrane Filter Method, *Ann. Occup. Hyg.*, 1981, **28**, 273–280.

