The NIOSH B Reader Certification Program

An Update Report

Gregory R. Wagner, MD
Michael D. Attfield, PhD
Richard D. Kennedy, MS
John E. Parker, MD

Physicians trained in the use of the International Labour Office system for classification of radiographs of pneumoconioses who pass a competence test administered by the National Institute for Occupational Safety and Health are designated as B readers. The current certification and recertification examinations for qualification under the NIOSH B reader program are described. Details of the rationale and format of each examination are given, and information on candidates' scores provided for the years 1987–1990.

The National Institute for Occupational Safety and Health (NIOSH) operates a program to train and certify physicians in the use of the International Labour Office (ILO) system for classifying radiographs for the presence of pneumoconioses.[1] To become a B reader, a candidate has to take a standardized certification examination. Thereafter, to remain a B reader, the person must periodically pass a recertification examination. Although the rationale and format of the B reader certification examination have been described previously,[2] changes in the examination together with the lack of published information on the recertification examination have prompted the preparation of this report. This report not only describes in detail the format of both examinations but also summarizes information arising from the results of 1255 certification and recertification examinations that have occurred between 1987 and 1991.

History of the B Reader Program

After studying readings arising from the first few years of a large chest radiographic surveillance program for coal miners, Felson and coworkers concluded that interreader variation was unacceptable.[3] A major reason for the variability between the readers was considered to be lack of familiarity of some readers with the classification scheme for pneumoconiosis. They noted that existing training material, consisting of weekend seminars and home study material prepared by the American College of Radiology (ACR) under contract from NIOSH, had been helpful, but they believed it had not always been effective in rais-

0096-1736/92/3409-0879$03.00/0

ing proficiency to the required level. Another cause of difficulty was thought to lie with readers' unfamiliarity with the basic radiographic characteristics of coalworkers' pneumoconiosis (CWP), and with the wide range of appearances seen in normal chest roentgenographs. In their report, they noted that a certification examination was under development. This examination was described by Morgan.[2]

Training and Certification Procedures for B Readers

The B reader examination originally was developed to identify physicians qualified to serve in national pneumoconiosis programs directed to both epidemiological research and to the compensation of coal miners and others who suffer from dust-related illness.[2] The examination was prepared by Johns Hopkins University under contract to NIOSH. (The B reader examination is administered by NIOSH, and arrangements can be made to take it at the Appalachian Laboratory for Occupational Safety and Health in Morgantown, West Virginia. Additionally, some physicians can take the test after special ACR courses on the pneumoconioses held periodically at major US cities. Persons wishing to participate in the examination process should make application on an Interpreting Physician Certification Document (form CDC/NIOSH (M) 2.12), available from the Morgantown office. Further information on the B reader program can be obtained from Dr. John Parker at the same address.)

Certification Process

Candidates who achieve a passing score on the certification examination are certified as NIOSH B Readers. B readers are certified for a period of 4 years from the date of approval, after which a recertification examination is required.

Temporal Trend in Number of B Readers

Testing of potential B readers began shortly before 1978 and has continued

to the present. Figure 1 shows the temporal trend in numbers of certified readers, based on NIOSH computer records of B readers certified as of December 31 each year. Despite the fact that many B readers choose not to (or failed to) recertify, an increase in the number of B readers has been seen each year until recently. There are some inconsistencies between these numbers and those reported for the earlier years of the program by Felson et al[3] and by Morgan.[2] It has not been possible to reconcile these differences; they probably arise through variations in definition or record-keeping.

The Certification Examination

The test requires classification of 125 radiographs. These were selected principally from the files of the Social Security Administration and NIOSH. The same set is used for each examination. Because the history of the test is very much intertwined with coalworkers' pneumoconiosis, the selection of films is weighted toward that disease. Morgan states, "because of the overriding importance of small opacity profusion, ability to estimate this parameter in chest x-rays is the keystone on which the proficiency examination is based."[2]

Classifications of these roentgeno-

graphs by a panel of 10 expert readers provide the basis for evaluation of each candidate. According to the mean classification of small opacities by these experts, over 50% of the test set has category 1/0 or greater. In addition, 15 test films are deemed to have large opacities. The set also contains 10 roentgenographs believed to demonstrate disease other than pneumoconiosis, including lung cancer and tuberculosis.

The Recertification Examination

The recertification examination was introduced in 1984. The format for this test is similar to the certification examination, although it is based on 50 rather than 125 roentgenograph classifications. During the fourth year of certification, B readers are advised by NIOSH of the need to recertify. Recertification is permitted during the 12 months before the expiration of the current certification.

B readers who either do not take or fail the recertification examination are removed from the NIOSH List of Approved B Readers when the current certification expires.

B readers who fail the recertification examination may take the certification examination before expiration of the current certification. If a passing score is achieved on this ex-
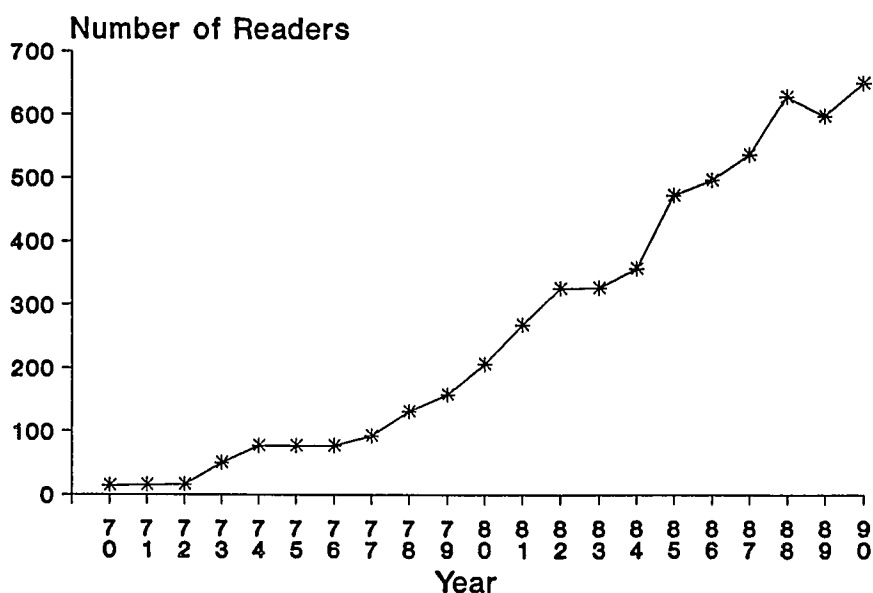

Fig. 1. Number of certified B readers on registry on December 31 each year.

amination, B reader approval is continued. There is no required waiting period for participants who fail the recertification examination before taking the certification examination. However, if a participant fails to achieve a passing score on the certification examination, the candidate must wait a year before retaking it.

The Grading System

Scoring of the B reader certification examination is divided into six sections dealing with small opacities (three parts), large opacities, pleural abnormalities, and other abnormalities (other symbols). Points, totaling a maximum of 100, are awarded separately by section; a candidate's combined score must be at least 50 to pass the examination. Point breakdown by section is shown in Table 1.

The three scoring sections for small opacities are (1) agreement on the presence/absence of small opacities on each film, (2) under-/overreading tendency, and (3) inconsistency index. Agreement on presence/absence is based on the false positive and false negative percentages. The percent false positive (PFP) is the percent of time the candidate said a film was positive (1/0 or greater) when the panel said it was negative (0/1 or less). The percent false negative (PFN) is the percent of time the candidate said a film was negative when the panel said positive. Scoring is accomplished by first summing PFP and PFN, then multiplying by 0.4. The resulting figure is subtracted from 20, with asso-

ciated upper and lower limits of 20 and zero points, respectively. Example: if a candidate obtains 8% false positives and 32% false negatives the score is $20 - [(8 + 32) \times 0.4] = 4$ points. Selection of the factor (0.4) was chosen so that if percent false positive plus percent false negative were more than or equal to 50%, no points are awarded.

Overreading/underreading tendency is concerned with the degree to which a candidate's classifications systematically diverge from those of the panel's. Profusion categories are converted to an integer scale (eg 0/1 = 1, 1/0 = 2, etc) and averaged. The score received is based on subtraction of the absolute difference between the two averages from 1.0 and multiplication by 10, with upper and lower limits of 10 and 0 points respectively. Example: if a candidate's mean score is .2 minor categories below (or above) that of the expert panel, the score received would be $(1.0 - .2) \times 10 = 8$ points.

The section weighted most heavily (30 points) is the inconsistency index. This measures the extent to which a candidate's classifications of small opacities differ from those of the panel, after adjustment for any systematic over- or under-reading. The inconsistency index is based on the standard deviation (SD) of differences between the reader's and panel's classifications after the readings have been converted to integers. If the SD lies between 0.8 and 1.8 the score awarded is calculated from 30 times the difference between 1.8 and the SD, with an

upper limit of 30 and a lower limit of zero points. Example: an SD of 1.0 would lead to the awarding of $(1.8 - 1.0) \times 30 = 24$ points.

Points pertinent to the classification of large opacities are based solely on assessment of presence/absence. The method is identical to that for assessment of presence/absence for small opacities noted above.

The pleural diseases section also is based solely on assessment of the presence or absence of pleural abnormalities but uses a weighing factor of 0.2 rather than 0.4, with the result that the maximum score awardable is 10 points compared with 20 as for small and large opacities.

The final section of the certification examination concerns classification of other diseases (symbols). The number of points awarded is set equal to the number out of 10 films for which the disease was correctly identified and named.

Note that the original format of the examination took no account of pleural abnormalities. Instead, a section on technical quality was included that contributed 10 points to the total score. In the late 1970s the current section on pleural disease classification replaced the technical quality evaluation. Like its predecessor, this characteristic received a weight of 10 points.

The recertification examination film test set contains radiographs of similar disease severity and type to those in the certification examination set. Absent from the recertification examination is the section on classification of other symbols. The maximum scores are the same as those of the certification examination, except for a doubling of the pleural disease score from 10 to 20 points.

Examination Participation

NIOSH maintains a record of all persons who ever have taken a B reader examination; however, only those persons who have taken the examination after January 1, 1987 are part of a fully computerized data base. Results presented here are based on the 4-year period beginning on that date.

TABLE 1
B Reader Certification and Recertification Examinations: Weighting Systems for Each Examination Section

Examination Section	Scoring Weights	
	Certification	Recertification
Small opacities		
Agreement on presence or absence	20	20
Overreading or underreading tendency	10	10
Inconsistency index	30	30
Large opacities	20	20
Pleural abnormalities	10	20
Other symbols	10	0
Total	100	100

For the period January 1, 1987 through December 31, 1990, 1255 examinations (927 certification and 328 recertification) have been administered. Of these, 974 persons have taken an examination only once, 138 persons have taken two examinations, and five persons have taken three examinations.

The majority (126/138) of the persons taking two examinations involved those who failed the recertification examination (63) and retook the certification examination or failed the certification examination (63) and retook the certification the next year. Ten participants passed the certification examination (in 1987) and their recertification examination (in 1990); two persons have taken the certification examination and passed the certification examination twice in the 4-year period.

Characteristics of B Reader Candidates

B reader candidates attending the 1990 B reader training course and examination were asked to complete a short questionnaire designed to elicit information on their characteristics and reasons for taking the examination. Just under 200 (about 87%) of the candidates complied with this request.

The medical specialty of almost 75% of the candidates was stated to be radiology and diagnostic roentgenology, with most of the remainder specializing in pulmonary diseases. Specialists in occupational medicine composed only 11% of the total. Over 80% of the candidates were in private practice.

About 70% stated they currently were reading between zero and 10 films per month for pneumoconiosis, and nearly all said they read 50 or fewer films per month. Those who said they were regularly classifying films for occupational lung disease (38% of the total) read most frequently for industry (35%), for patient care (30%), and for attorneys (25%).

When asked why they were taking the B reader course, 75% stated the desire to improve their knowledge of

occupational lung disease. Sixty-five percent said the training would be useful for clinical practice, and the same percentage said it would aid in reading films for surveillance. Thirty-three percent said the training was important for improving credibility as an expert witness. Few said that training for research was at all important (8%).

Pass/Fail Rates

There has been somewhat of a downward trend in the percentage of participants passing the certification examination during the past 4 years.

Compared with 1987, when 57% of those taking the certification examination passed, only 35% passed in 1990. In contrast (apart from 1987 when only 13 examinations were taken), the passing rates for the recertification examination were stable at around 67%. Table 2 shows the number of examinations given by year for the certification and recertification examinations along with the yearly passing rate and the 4-year total rates.

Distribution of Scores

Figure 2 shows the distribution of total scores for the certification and

TABLE 2
B Reader Passing Rates for Certification and Recertification Examinations, 1987 to 1990

Year	Certification		Recertification	
	Number*	Percent	Number*	Percent
1987	124/217	57	11/13	85
1988	98/222	44	25/36	69
1989	136/256	53	113/173	65
1990	84/240	35	71/106	66
Total	442/935	47	220/328	67

* First number is candidates who passed, second number is total number of persons who took examination.
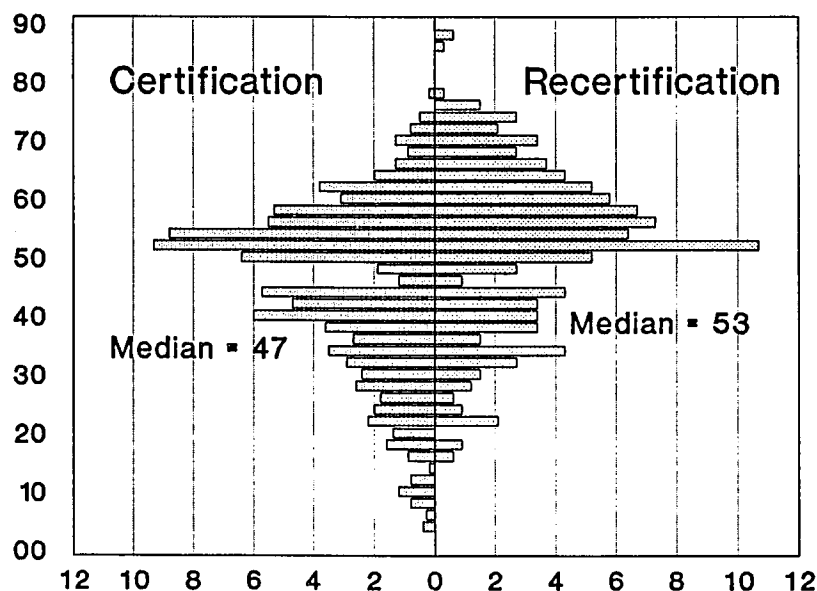

Fig. 2. B reader final scores, 1987 to 1990. Percentage distribution, certification versus recertification examination.

TABLE 3
Mean and Standard Deviation (SD) for Sectional Scores by Pass/Fail and Certification versus Recertification Examination, B Readers 1987 to 1990

	Certification					Recertification				
	Maximum Possible	Pass Observed		Fail Observed		Maximum Possible	Pass Observed		Fail Observed	
		Mean	(SD)	Mean	(SD)		Mean	(SD)	Mean	(SD)
Small opacity										
Agreement on presence/absence	20	10.7	(2.6)	5.2	(4.6)	20	9.5	(3.4)	4.4	(3.6)
Over/under reading	10	7.4	(1.8)	4.4	(3.1)	10	6.7	(2.4)	2.9	(2.9)
Inconsistency index	30	20.0	(3.8)	11.0	(5.6)	30	22.8	(4.3)	14.4	(5.5)
Large opacity	20	9.3	(3.8)	5.2	(4.6)	20	13.5	(4.3)	9.5	(6.3)
Pleural diseases	10	2.4	(2.0)	1.7	(1.9)	20	8.8	(2.9)	6.4	(3.5)
Other symbols	10	7.3	(1.7)	5.8	(2.2)	0	–	–	–	–
Total	100	57.1				100	61.3			

recertification examinations taken between 1987 and 1990. The trough in each distribution occurring just below 50 points is probably due to the current NIOSH practice of manually scoring the readings of candidates with borderline scores. Scores can only be increased in this review process.

Examination scores by section for persons passing and failing the certification examination are given on the left side of Table 3. Those failing the examinations obtained about half the total points of those that passed. Furthermore, this trend was evident for each section. For example, the mean score on the inconsistency index for those who failed the certification examination was 11.0, or just over half the 20.0 points obtained by those passing. Sectional scores were more variable for failing candidates compared with those passing.

As might be expected because of the weighting applied to it, the inconsistency index provided the greatest contribution to the total score. On the other hand, the pleural abnormality section provided the least, which was partly a result of its lower weighting and partly due to generally poor classification by the candidates. The other sections provided roughly equal contributions. In general, candidates obtained the greatest scores relative to the total available for the inconsistency index, over-/underreading tendency, and other symbols, and least for
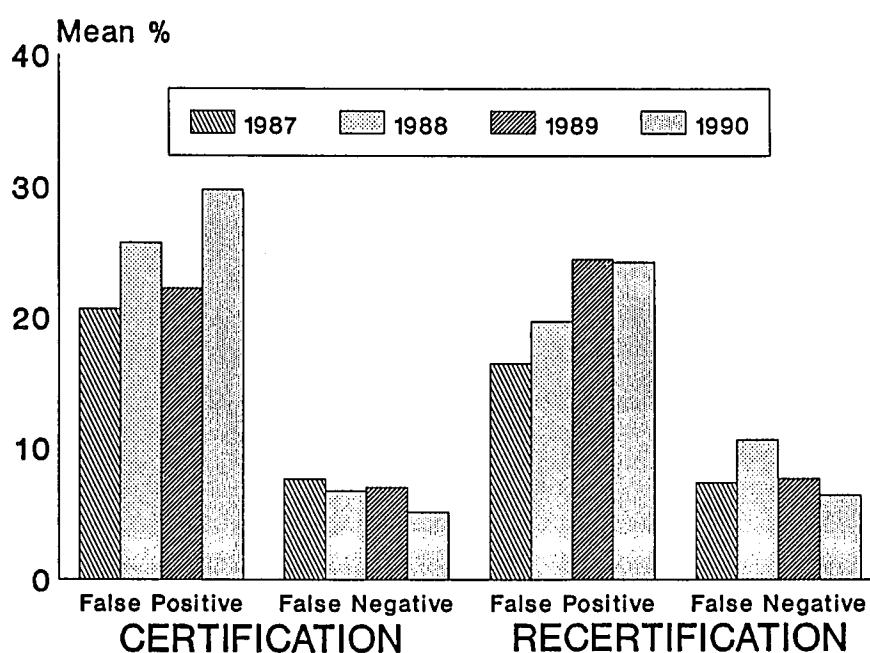

Fig. 3. B reader certification versus recertification examination, mean percent of false positives and false negatives for small opacities, 1987 to 1990.

pleural abnormalities. Results for the recertification examination were similar to those for the certification test.

False Positives and False Negatives

Figure 3 displays the mean number of small opacity percent false positives and percent false negatives (as defined by reference to the determinations by the expert panel) by year separately for all certification and recertification examination candidates. The findings, generally similar for both examinations, show that candidates tended to find two to three times as many false positives as false negatives. In addition, there are suggestions that the percentage of false positives rose over the 4 years, although the percentage of false negatives showed slight temporal falls.

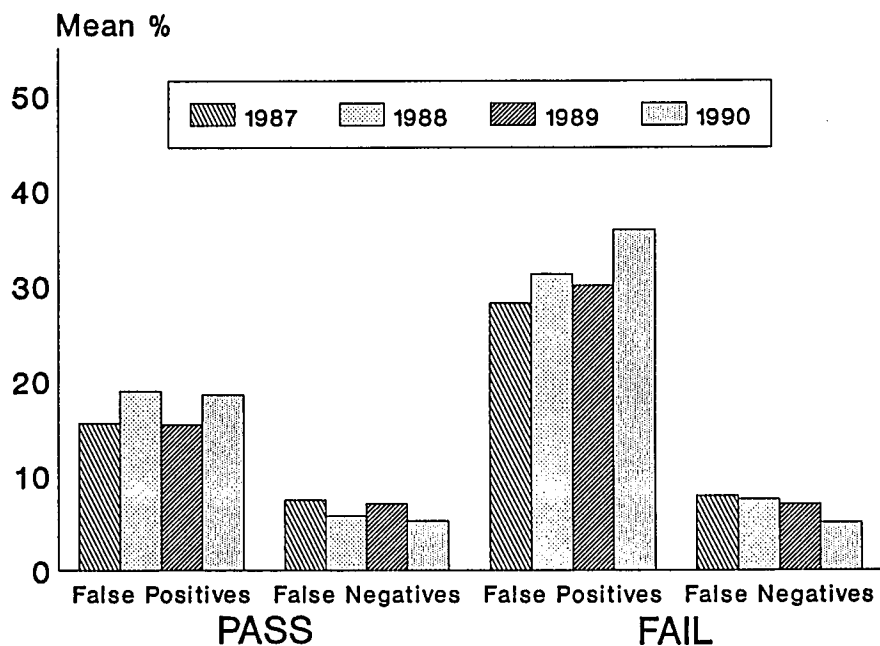Results for the certification examination candidates alone, divided into those who did and did not pass are

Mean %

Fig. 4. B reader certification examination, mean percent of false positives and false negatives, small opacities by pass/fail, 1987 to 1990.
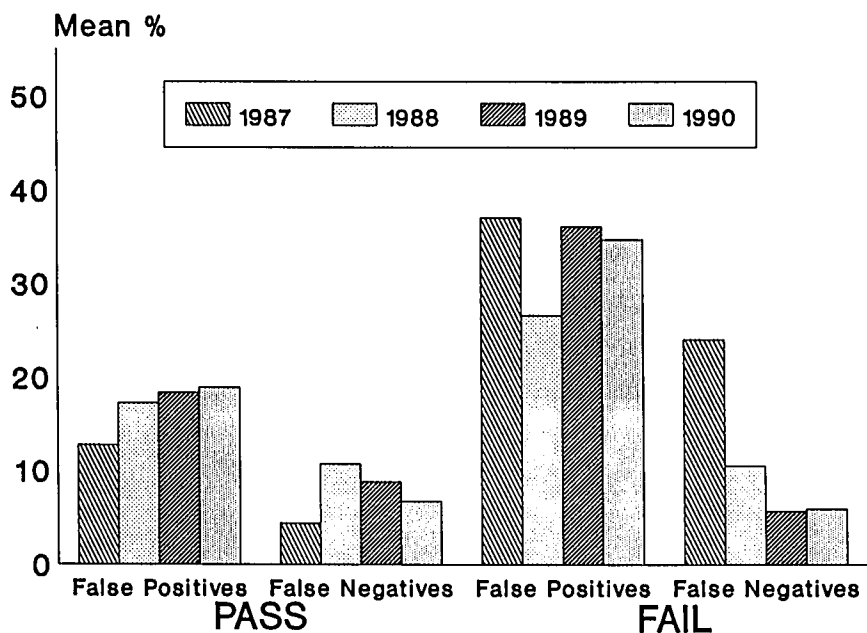
Mean %

Fig. 5. B reader recertification examination, mean percent of false positives and false negatives, small opacities by pass/fail, 1987 to 1990.

shown in Figure 4. This shows that those that failed the examination had about twice as many false positives as those who passed the examination. However, the percent of false negatives was about the same in the two groups. Similar findings are evident

for the recertification candidates (Figure 5).

Discussion

Although there is little doubt that the extra training and materials asso-

ciated with the B reader test have led to better understanding of the ILO classification system and of normal and abnormal radiographic appearances, there is evidence that substantial variability remains among certified B readers.[4-6] Some variation is to be expected, because even the expert panel disagreed among themselves on the classification of the films used in the B reader examinations, but critics note that interreader variation is excessive.[7,8] In an attempt to uncover reasons for this variability and to recommend improvements in training and quality control, a workshop was held in Chicago in 1990. A report on this is the subject of a companion article in this issue of *JOM*.[9]

References

1. International Labour Office. *International Classification of Radiographs of Pneumoconiosis*, rev ed. Occupational Safety and Health Series 22, Rev 80. Geneva: ILO; 1980.
2. Morgan RH. Proficiency examination of physicians for classifying pneumoconiosis chest films. *Am J Radiol.* 1979; 132:803–808.
3. Felson B, Morgan WKC, Bristol LJ, et al. Observations on the results of multiple readings of chest films in coal miners' pneumoconiosis. *Radiology.* 1973;109: 19–23.
4. Attfield MD, Althouse RB, Reger RB. An investigation of inter-reader variability among x-ray readers employed in the underground coal miner surveillance program. *Ann Am Conf Gov Ind Hyg.* 1986;14:401–409.
5. Ducatman AM, Yang WN, Forman SA. 'B-Readers' and asbestos medical surveillance. *J Occup Med.* 1988;30: 644–647.
6. Bourbeau J, Ernst P. Between- and within-reader variability in the assessment of pleural abnormality using the ILO 1980 international classification of pneumoconiosis. *Am J Ind Med.* 1988; 14:537–543.
7. Parker DL, Bender AP, Hankinson S, Aeppli D. Public health implications of the variability in the interpretation of 'B' readings for pleural changes. *J Occup Med.* 1989;31:775–780.
8. Rom W. Should NIOSH scrap the "B" reader program? *Occup Environ Med Rep.* 1989;3:89–90.
9. Attfield MD, Wagner GR. A report on a workshop on the NIOSH B reader certification program. *J Occup Med.* 1992; 34:875–878.