Mining Education and
Research Center

Findings from an Audit of Chest Radiograph ILO Classifications performed under the Coal Mine Workers' Health Scheme

Organization Name: University of Illinois Chicago (UIC), School of Public Health, Division of Environmental and Occupational Health Sciences

UIC Investigators	**Audit B Readers**
Robert A Cohen, MD	Robert A Cohen, MD
Leonard HT Go, MD	Robert Tallaksen, MD
Kirsten S Almberg, PhD	Kathleen DePonte, MD
Kathleen Kennedy, PhD	

Report Date: February 5, 2021

Project Sponsor: Resources Safety & Health Queensland

Executive Summary

This audit of the International Labour Office (ILO) classifications of chest radiographs (CXRs) for abnormalities consistent with pneumoconiosis performed by Lungscreen Australia since January 1, 2020 was performed at the request of Resources Safety and Health Queensland (RSHQ) to evaluate the performance of CXR screening. Overall, B readers for Lungscreen Australia are classifying CXRs for pneumoconiosis with high accuracy.

The audit team at the University of Illinois Chicago (UIC) received a sample of 1,857 CXRs obtained through the Coal Mine Workers' Health Scheme that had been classified by Lungscreen Australia's National Institute for Occupational Safety and Health (NIOSH) certified-B readers. These CXRs were independently classified by a team of three U.S.-based NIOSH-certified B readers according to the ILO guidelines for the interpretation of CXRs for pneumoconiosis which includes sections for technical quality; presence and profusion of small opacities; the presence and category of large opacities; presence, width, and extent of pleural changes consistent with pneumoconiosis; as well as the presence of other abnormalities. Lungscreen ILO classifications were abstracted and compared to the audit team classifications to assess any differences in classifications.

Main Findings:

1. **Quality:** There were significant differences in quality grading between the audit team and Lungscreen B readers. The audit team graded 26% of images as quality 1 (highest quality), compared to 55% of images graded by Lungscreen. Notably, the audit team graded images as a quality 3 or 4 (significant deficiency or unreadable) greater than two times more often than Lungscreen (9.5% vs 3.6%). The audit team found the quality of two images to be unacceptable for classification that Lungscreen graded as acceptable. Conversely, Lungscreen reported a small number (2.7%) of images as quality 3 that the audit team graded as quality 1 or 2.

    a. Computed radiography (CR) was associated with worse quality than digital radiography (DR).

    b. Three of the nine radiology clinics had a high rate (>10%) of images graded as poor quality.

    c. There are potential unknown factors that could explain differences in quality grading agreement.

    d. However, there was significantly greater agreement in quality grading between the audit team and Lungscreen for DR images (91%) than CR (87%) ($p = 0.0015$).

2. **Pleural Disease:** The prevalence of pleural abnormalities based on the audit team B readers' classifications was low (n=28, 1.5%). Lungscreen B readers identified three of these 28 cases for a sensitivity of only 11% (**Table 5**). The 25 cases in which only the audit team identified pleural abnormalities have been shared with Lungscreen for review. It is most likely that these pleural changes were the result of non-occupational exposures such as sub-pleural fat versus occupational exposures to mineral fibres, however this would have to be confirmed clinically when the results are reviewed by the subject's supervising physician.

3. **Other Abnormalities:** The audit team reported a higher prevalence of other abnormalities in 34% (n = 628) of cases compared to Lungscreen B readers' 23% (n = 424). A large portion of this difference lay in the reporting of incidental clinical findings that did not require any clinical intervention such as healed rib fractures (n=241). In this comparison, the audit team focused on abnormalities that could be serious including possible cancer (cancer; lung density; or nodular lesion); or an abnormality of cardiac size or shape, or emphysema (including bulla(e)), which is a mine dust lung disease. The audit team found a higher prevalence of symbols and findings that raised the concern for possible cancer including nodular lesion or lung density (n=50) compared to Lungscreen (n=18). There were only nine cases in which the audit team and Lungscreen B readers agreed on this finding. Lungscreen B readers noted emphysema more often (n = 20) than the audit team B readers (n = 8), with five cases in agreement for this finding. Abnormality of cardiac size or shape was a rare finding for either team, with low agreement on this finding.

4. **Referral for physician follow-up:** The audit team indicated that physician follow-up was warranted in 70 cases (3.8%), of which Lungscreen identified five. Physician follow-up was most commonly indicated for findings that raised a concern of possible cancer including nodular lesion, or lung density (n = 42, 60%). Of note again, findings which may have been included in narrative reports separate from the ILO report were not considered in this audit and may account for some of these differences.

5. **Small opacities of pneumoconiosis:** Lungscreen classified a total of eight CXRs (0.4%) compared to three CXRs (0.2%) by the audit team as having small opacities consistent with pneumoconiosis. All three of the audit team's positive CXRs had also been identified as pneumoconiosis by Lungscreen B readers. Due to the low prevalence of pneumoconiosis in this population, conclusive estimates of sensitivity, specificity and agreement of classifications between the audit team and Lungscreen could not be reliably made.

6. **Large opacities of pneumoconiosis or progressive massive fibrosis:** Neither Lungscreen nor the audit team detected large opacities consistent with pneumoconiosis. Consequently, this audit is unable to provide estimates of agreement, sensitivity, or specificity on this feature.

Recommendations:

Conclusions regarding Lungscreen: Areas for improvement include additional attention to the classification of pleural abnormalities, a lower threshold for indicating the presence of other abnormalities, and referring these cases to the primary provider for clinical correlation. There remains potential unknown factors that may explain the difference in reporting quality grading that would benefit further investigation.

Conclusions regarding chest imaging for pneumoconiosis as part of the Coal Mine Workers' Health Scheme: It appears that 53% of radiographs were acquired using CR technology, which is associated with a greater rate (12.3% vs 6.4%, $p < 0.0001$)) of poor quality images (grades 3 or 4). Efforts could be made to encourage the use of DR technology in clinics acquiring images for the Scheme. While based on the NIOSH Guidelines, RSHQ could also consider reviewing its *Standards for Acquiring Digital Chest Radiography Images for Medical Surveillance of Queensland Coal Mine Workers'* to exclude the use of CR. Images from specific clinics had a greater frequency of quality issues. In addition to feedback already provided to clinics by Lungscreen, RSHQ should consider providing feedback from this audit to these clinics to give them an opportunity to improve their imaging techniques.

Table of Contents

Glossary

CFR	Code of Federal Regulation (U.S.)
CMWHS	Coal Mine Workers' Health Scheme
CR	Computed radiography
CWHSP	Coal Workers' Health Surveillance Program
CWP	Coal workers' pneumoconiosis
CXR	Chest x-ray or chest radiograph
DR	Digital radiography
HSU	Health Surveillance Unit
ILO	International Labour Office
NIOSH	National Institute for Occupational Safety and Health
RSHQ	Resources Safety and Health Queensland
UIC	University of Illinois Chicago

Background

In response to the re-identification of coal workers' pneumoconiosis in the Queensland coal industry, Monash University and the University of Illinois Chicago (UIC)[1] completed a review (Monash review) of the respiratory component of the Coal Mine Workers' Health Scheme (CMWHS) in July 2016. A range of systemic failures were identified and 18 recommendations were made for improvement. This included additional training for radiologists to classify chest radiographs using the International Labor Office (ILO) International Classification of Radiographs for Pneumoconioses[2] and a guideline for follow-up investigation for workers with abnormal screening results.

While the Queensland Government was implementing the recommendations from the Monash review, as an interim solution, chest radiographs for Queensland's CMWHS were sent to UIC for ILO classification. UIC's team of U.S.-based, United States National Institute for Occupational Safety and Health (NIOSH)-certified B readers classified these chest radiographs following NIOSH's dual reading protocol.[3]

In March 2019, B reading for Queensland's CMWHS transitioned to Lungscreen Australia. Chest radiographs are now sent to Lungscreen Australia, whose team of Australia-based, NIOSH-certified B readers classify the chest radiographs following NIOSH's dual reading protocol.

Aims of Current Audit

The goal of this study was to perform an audit of the performance of the newly implemented Lungscreen Australia screening program. Specifically, the UIC audit team aimed to determine whether the image quality, classification of parenchymal abnormalities, and findings of pleural and other abnormalities on CXRs are performed reliably using the ILO system under the current CMWHS Lungscreen dual reading program.

Methods

Image Selection and Management

Resources Safety and Health Queensland (RSHQ) provided the audit team staff with 1,857 CXRs based on sample size calculations performed by the audit team, assuming a minimum prevalence of parenchymal abnormalities at 0.8%. This prevalence estimate was derived from the audit team's prior work in reviewing 49,978 CXRs from the CMWHS.

The methodology for the selection of CXRs was determined by RSHQ and was performed as follows: The CXRs included in the audit were selected from RSHQ-approved imaging clinics including a geographical spread of those that performed a greater number of examinations for the CMWHS over the previous 12-month period. This covered a range of radiology companies and regional areas including Central Queensland, Mackay Isaac Whitsunday, Wide Bay Burnett, South East Queensland and one clinic from New South Wales. Of note, for one particular company which provided a large volume of CXRs under the CMWHS through several clinics, two clinics were selected from this company for inclusion. For other companies, one representative clinic was selected. This was done to ensure a diversity of companies in the audit. Based on each clinic's previous 12 months of CXR data, sample size for inclusion of CXRs from each clinic was determined using the Australian Bureau of Statistics sample size calculator. Once the minimum CXR audit sample size of 1,857 was reached, no further clinics were selected. In total, RSHQ provided CXRs acquired at nine different locations. The audit team B readers were blinded to the originating clinic of the CXR.

RSHQ included CXRs from the selected clinics if they (1) had Lungscreen International Labour Office (ILO) classifications submitted to RSHQ's Health Surveillance Unit (HSU) from January 1, 2020, ensuring the audit reflected current practices; (2) the coal mine worker had provided consent on their health assessment for this information to be used for audit purposes; and (3) the health assessment form had been completed and submitted to the HSU. Where the number of Lungscreen reports from a selected clinic exceeded the number of reports required for the audit, systematic random selection of images was performed. RSHQ's database of parameters for the review period is limited. Identification of more detailed attributes currently requires manual review of individual records.

Original CXR image files as well as Lungscreen ILO classifications were obtained by audit team staff from RSHQ via secure cloud servers. The audit team staff transmitted the CXR images to participating B readers for full ILO classification via secure cloud servers. Audit team B readers were blind to the Lungscreen classifications of all images.

Chest Radiograph Scoring Protocol

A total of three U.S.-based, NIOSH-certified B readers performed the ILO classifications for the audit team (**See Appendix A**). B readers interpreted each de-identified CXR in accordance with the ILO guidelines for the classification of chest radiographs for pneumoconiosis.[2] The principal findings of interest for this study included those findings of abnormalities consistent with a coal mine dust lung disease.[4] These included image quality; presence of small opacities consistent with pneumoconiosis; the presence of large opacities consistent with pneumoconiosis; the presence of pleural abnormalities; and the presence of other abnormalities including possible diseases.

All CXRs were classified by at least two audit team B readers. If agreement on the presence and profusion of opacities could not be reached through these two classifications, a third B reader classified the image for adjudication. The criteria for assessing agreement is outlined by the U.S. Code of Federal Regulations[3] (CFR) and is the protocol that NIOSH uses for adjudication in the U.S. Coal Workers' Health Surveillance Program. Under this protocol, agreement between two B reads is achieved if both B readers find abnormalities consistent with complicated pneumoconiosis (large opacities in category A, B, or C); or if both B readers find abnormalities consistent with simple pneumoconiosis with profusion scores in the same major category or within one minor category of each other according to the 12-point ILO profusion scale for small opacities. If these criteria are met, the higher of the two classifications is reported. The only exception to this rule is if one reader classifies the image as 0/1 and the other as 1/0. Despite being only one minor category apart on the ILO scale, these reads would not be considered to be in agreement because the former indicates the absence of radiographic pneumoconiosis while the latter indicates its presence. In cases when a third B read is needed for adjudication, agreement is based on the criteria above between any two of the three reads. The other condition that can prompt an adjudicating read is that of one of the two initial B reads classifying the image as unreadable (quality grade 4). Lungscreen employs the same process with up to five B reads if agreement cannot be reached after the first three reads. The U.S. CFR outlines agreement criteria for the presence of opacities consistent with pneumoconiosis only, therefore the criteria employed by the audit team for reporting the final classifications for quality grade, pleural abnormalities, and other abnormalities is outlined below.

The ILO guidelines for the classification of chest radiographs for pneumoconiosis uses four grades of technical quality. A grade of (1) indicates that the image is of good quality; (2) indicates acceptable quality without technical defects that are likely to affect classification of the CXR; (3) indicates that the image is of acceptable quality with some technical defects, but is still adequate for classification purposes; and (4) indicates that the image is unreadable due to technical quality issues.[2] The audit team final quality grade was derived from the individual quality grades reported by each B reader in their ILO classifications. If all B readers reported the same quality grade, the unanimous grade was considered the final quality grade for that CXR. If there were differences in the quality grade between B readers for the same study, the classification with the worst quality, was chosen as the final quality grade. This process is consistent with how Lungscreen arrives at a final quality grade on their adjudicated ILO forms. If two or more B readers indicated that the study is unreadable (quality grade 4), the final adjudicated quality grade was considered 4 (unreadable). If one out of three B readers deemed a study unreadable, but the other two graded the image as readable (1, 2, 3), the classification with the worst quality grade that was considered readable was chosen as the adjudicated quality grade.

In this audit, we examined the presence or absence of pleural abnormalities in the ILO classifications of all CXRs. If either audit team B reader indicated pleural abnormalities, the image was flagged for additional review. The final classification of the presence/absence of pleural abnormalities was derived from this review by audit team B readers. Similarly, the final classification of other abnormalities was inclusive of all conditions/features noted by any of the audit team B readers. This is the same approach used by Lungscreen.

Lungscreen Data Abstraction

The Lungscreen ILO reports for each CXR were provided by RSHQ to project investigators. Findings from these reports were entered into a database by audit team staff who were blinded to the audit team's classifications of each image. In addition to the unique identifiers needed for linkage with the audit team classifications, all sections of the ILO form, except the free text field in Section 4, were abstracted (**Appendix C**). The Lungscreen classifications were abstracted and then linked to the audit team classifications via unique study identifiers prior to analysis to determine agreement between the audit team and Lungscreen findings.

Statistical Evaluation

Differences in the quality grades reported by the audit team and Lungscreen were evaluated with regards to overall distribution of quality grades by each team, frequency of specific defects, modality (computed radiography, CR; digital radiography, DR), and clinic.

Where feasible, sensitivity and specificity values were calculated separately for the presence of (1) small opacities consistent with pneumoconiosis, (2) large opacities consistent with pneumoconiosis, (3) pleural abnormalities, and (4) selected other abnormalities. We considered the audit team of three B readers as the "gold standard" for presence or absence of abnormalities, and Lungscreen's interpretation was the test under evaluation. In this audit, therefore, sensitivity is a measure of Lungscreen's ability to identify those who have each feature. Specificity measures the ability of Lungscreen to identify those who do not have these features.

Project staff reviewed the agreement statistics and comments from the audit team and Lungscreen B readers to characterize trends in CXR classification seen among both groups and highlight systematic patterns that explain the disagreement observed.

Results

In total, the audit team B readers classified 1,857 CXRs that Lungscreen B readers had previously classified. The mean age of the miners in this audit was 38 years of age (SD 11.8 years, range 15 – 74). This correlates with the average age of 39 across the 42,250 Queensland coal mine workers' CXRs reported by Lungscreen to-date.

Quality Assessment

The distribution of quality grades given by Lungscreen and the audit team for the 1,857 images reviewed in this audit are shown in **Table 1**. Most images (n = 1,855) were considered acceptable for classification purposes by the audit team. The audit team graded 26% of images as quality 1, compared to 55% of images graded by Lungscreen. Notably, the audit team graded images as a quality 3 or 4 over two times more often than Lungscreen (9.5% vs 3.6%). The audit team found the quality of two images to be unacceptable for classification citing overexposure, poor contrast, and improper position. No CXRs graded quality 4 by Lungscreen were provided as part of the sample as these are rejected by Lungscreen back to the originating clinic to be retaken. The retaken image is then reported by Lungscreen and sent to RSHQ. Lungscreen has graded 34 CXRs quality 4 to-date.

Table 1. *Distribution of ILO quality grades across 1,857 chest radiographs by Lungscreen and audit team B Readers*

	Lungscreen		**Audit Team**	
Quality Grade	**n**	**%**	**n**	**%**
1	1,027	55.3	475	25.6
2	764	41.1	1,205	64.9
3	66	3.6	175	9.4
4	0	0	2	0.1

Quality grades from each group, Lungscreen and the audit team, were considered congruent if both grades were either grade 1 or 2, or were either grade 3 or 4. Grades of 1 or 2 compared to a grade 3 or 4 were not considered congruent (**Table 2**). Using this approach, 88.7% of all images had congruent quality grades.

Among those images with non-congruent grades (n = 211), Lungscreen grades were lower, indicating better quality, than the audit team grade in 76% of cases. Two radiographs were graded as a quality 1 by Lungscreen, but graded unreadable (grade 4) by the audit team. An additional 77 cases were graded a quality 1 by Lungscreen and quality 3 by the audit team.

Table 2. *Distribution of the adjudicated quality scores from Lungscreen and audit team classifications of 1,857 chest radiographs.*

Lungscreen Quality Score	Audit Team Quality Score	n	%	Congruent
1/2	1/2	1,630	87.8	Yes
1/2	3/4	161	8.7	No
3/4	1/2	50	2.7	No
3/4	3/4	16	0.9	Yes

Modality

The CXRs in this audit were obtained with computed radiography (CR) systems (n = 982, 53%) and digital radiography (DR) systems (n = 875, 47%). The distribution of these modalities was driven by the participating clinics, as nearly all images from the same clinic were taken using the same modality. Of note, DR images had a significantly higher proportion (94%) of high-quality images, with grades 1 or 2 as compared to 3 or 4, when compared to CR images (88%; $p < 0.0001$). Examining this in a different way, CR images had a significantly higher proportion (12.3%) of quality grades 3 or 4 when compared to DR images (6.4%; $p < 0.0001$) (**Figure 1**). Agreement on quality grades, based on the criteria specified in Table 2, was significantly higher between the audit team and Lungscreen B readers for DR images (n = 798, 91%) compared to CR images (n = 848, 87%) ($p = 0.002$). Both of the CXRs classified as unreadable by the audit team were CR images. CR technology, especially early models has been shown to be inferior to DR imaging in terms of image quality, although it is a cheaper and easier to implement technology.[5,6]

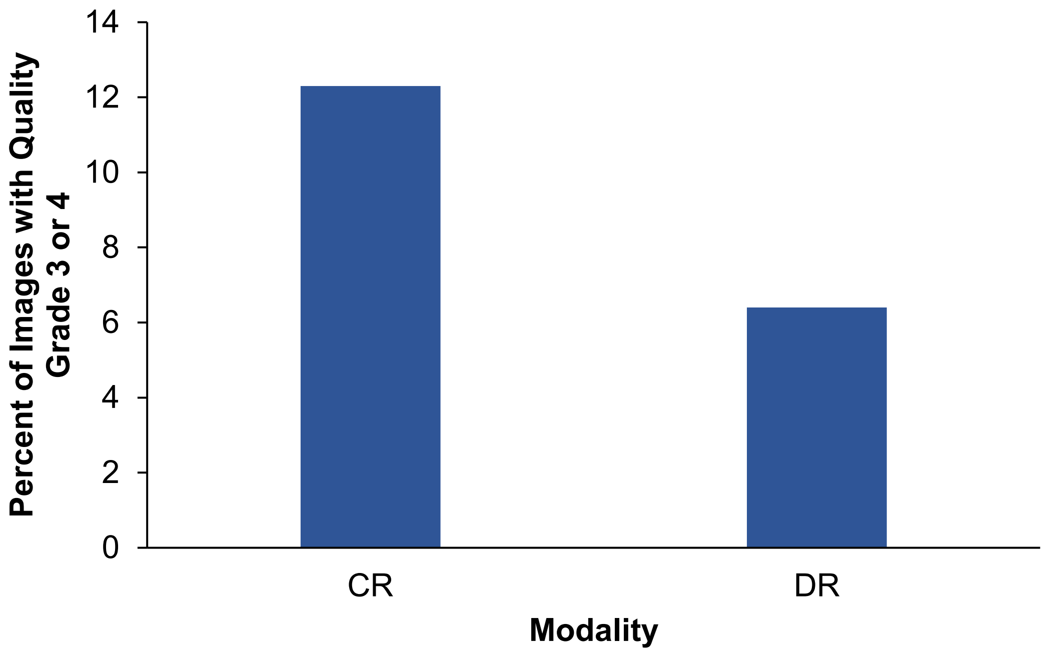

Figure 1. *Proportion of images with technical quality grades of 3 or 4 by modality, either computed radiography (CR) or digital radiography (DR). Quality grades based on the audit team B reader classifications of 1,857 chest x-rays.*

<u>Clinic</u>

The majority of images included in this audit originated from nine clinics (n = 1,793, 96.6%). Results from these nine clinics are discussed in detail below. Of the remaining 3.4% of images (n = 64), either the clinic location could not be identified with sufficient certainty or it originated from another clinic location and the sample size was too small for analysis. This was caused by a processing error during the sample collection later identified by RSHQ.

Of the nine clinics represented in this audit, four used CR systems exclusively (Clinics A, B, C, and F); and five used DR systems exclusively (Clinics D, E, G, H, and I). The images from Clinic D included one acquired by CR modality (<1%), therefore clinic D is categorized as using DR exclusively. The CR image from Clinic D is not used in calculating proportions displayed in Figure 2. The proportion of images with quality grade 3 or 4 differed significantly across clinics (**Figure 2**). Over 10% of images from clinics A, B, and I were graded as a quality 3 or 4. The unreadable images were obtained from clinics A and B. For a list of the distribution of quality defects by clinic, see **Appendix B.**

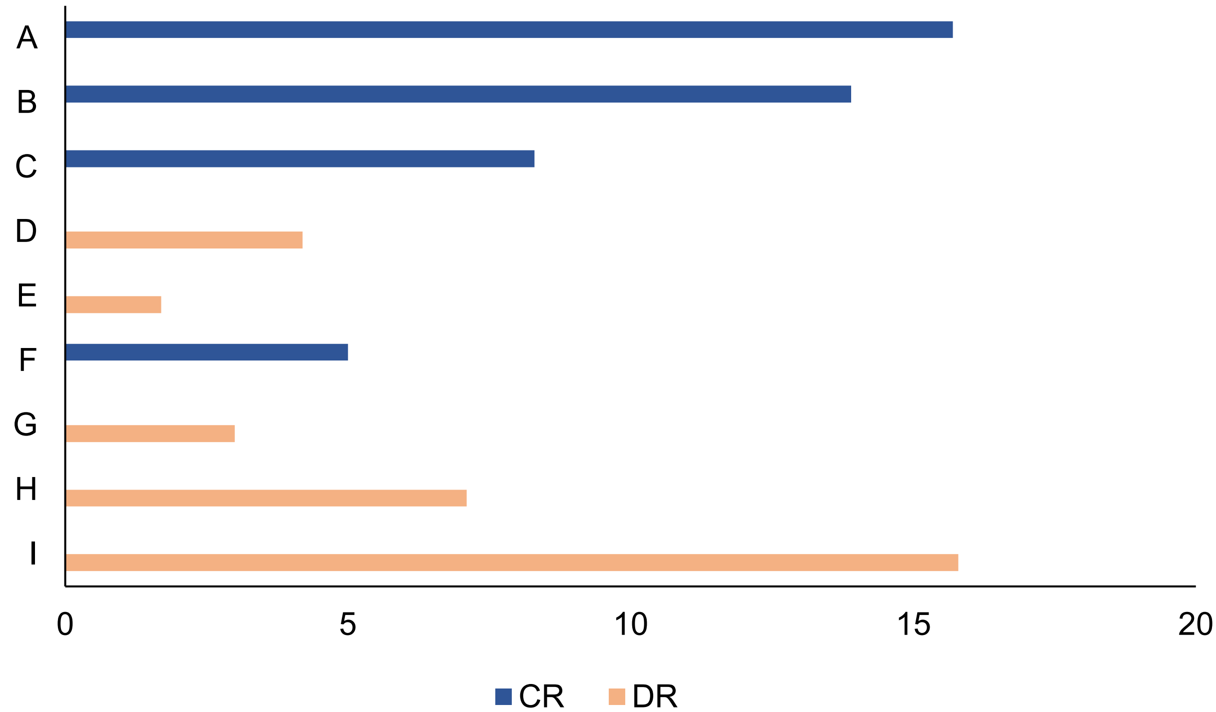

Figure 2. *Proportion of images with International Labor Office technical quality grades of 3 or 4 by clinic and modality, either digital or computed radiography (DR or CR). Quality grades based on the audit team B reader classifications of 1,793 chest x-rays.*

Technical Defects Indicated

By definition, images graded as quality 1 have no technical defects to indicate. As expected, the number of technical defects indicated by the audit team increased significantly as the quality grade increased from grade 2 (mean defects = 1.5) to grades 3/4 (mean defects = 2.2; $p < 0.0001$).

The most common defects indicated by quality grades 2 compared to 3/4 are shown in **Table 3**. The most common defects noted for images of quality 3 or 4 were poor contrast (68%), mottle (58%), and overexposed (38%). Overexposure, underexposure, poor contrast, and underinflation were substantially more prevalent among those with quality grade 3 or 4 compared to images with quality grade 2. The “Other” category of technical defect was checked in approximately 13.8% of images with quality grades of 2 or worse (n =191), with even distribution between grade 2 vs 3/4. The other defects noted were overwhelmingly related to improper position (172/191; 90%). Scapula overlay was the most common position defect noted (n = 150). There were two images which the audit team characterized the quality as grade 4, and Lungscreen did not. The first unreadable image was considered by the audit team to be overexposed with poor contrast and upper zones that were too dark for interpretation. The second unreadable image was considered to be improperly positioned, with the lung apices not included in the image. Both images were acquired using CR technology.

The types of quality defects indicated by the audit team varied markedly by modality. The most common defects listed for DR images with a quality grade of 2 or 3 (n = 621) were mottle (80%), improper position (31%), and poor contrast (13%). Among CR images with a quality grade of 2, 3, or 4 (n = 761) the most common defects noted were poor contrast (57 %), overexposure (41%), mottle (33%), and improper position (25%) (See **Table 4**).

Table 3. *Distribution of specific technical quality defects indicated by the audit team for the 1,382 chest x-rays with a quality grade of 2, 3, or 4.*

	Grade 2		**Grade 3/4**	
Technical Quality Defect	**n**	**%**	**n**	**%**
Overexposed	246	20.4	67	37.9
Underexposed	15	1.2	20	11.3
Artifacts	7	0.6	1	0.6
Improper Position	348	28.9	39	22.0
Poor Contrast	398	33.0	121	68.4
Poor Processing	0	0.0	0	0.0
Underinflation	103	8.6	27	15.3
Mottle	646	53.6	103	58.2
Excessive Edge Enhancement	51	4.2	8	4.5
Other Defect (write in)	167	13.9	24	13.6

Table 4. *Distribution of specific technical quality defects indicated by the audit team by image modality. The percentages presented in the table are based on the number of chest radiographs acquired using digital radiography (n = 621) and computed radiography (n = 761) with an audit team quality grade of 2, 3, or 4.*

Technical Quality Defect	**Digital Radiography (DR) n (%)**	**Computed Radiography (CR) n (%)**
Artifacts	8 (1.3)	0 (0)
Excessive edge enhancement	23 (3.7)	36 (4.7)
Improper position	194 (31.2)	193 (25.4)
Mottle	495 (79.7)	254 (33.4)
Overexposed	5 (0.8)	308 (40.5)
Poor contrast	83 (13.4)	436 (57.3)
Underexposed	22 (3.5)	13 (1.7)
Underinflation	63 (10.1)	67 (8.8)

Small Opacities Consistent with Pneumoconiosis

The analysis of small opacities consistent with pneumoconiosis excluded the two cases graded as unreadable by the audit team, leaving 1,855 CXRs for analysis. The excluded cases were classified as negative for small opacities (0/0 or 0/1 profusion score) by Lungscreen.

Lungscreen classified a total of eight CXRs (0.4%) as having small opacities consistent with pneumoconiosis (small opacity profusion score of 1/0 or higher). The audit team classified only three CXRs (0.2%) as 'positive' for small opacities consistent with pneumoconiosis. All three of the CXRs classified as positive for small opacities consistent with pneumoconiosis were also identified as such by Lungscreen B readers. Five additional cases were classified as positive by Lungscreen and negative by the audit team (**Table 5**). These results indicate a 99.7% agreement between the audit team and Lungscreen B readers. Due to the low prevalence of pneumoconiosis in this population, conclusive estimates of sensitivity and specificity could not be reliably made.

Table 5. *Comparison of audit team findings and Lungscreen findings for the presence of small opacities consistent with pneumoconiosis.*

Small Opacities Consistent with Pneumoconiosis		**Audit Team Findings**	
		Positive	**Negative**
Lungscreen Findings	**Positive**	3	5
	Negative	0	1,847

Large Opacities Consistent with Pneumoconiosis

Neither Lungscreen nor the audit team indicated the presence of large opacities consistent with pneumoconiosis on any of the CXRs included in this audit. Consequently, we are unable to provide estimates of agreement, sensitivity, or specificity on this feature.

Pleural Abnormalities

The prevalence of pleural abnormalities based on the audit team B readers' classifications was low (n=28, 1.5%). Lungscreen B readers identified three of these 28 cases for a sensitivity of only 11% (**Table 6**). The 25 cases in which only the audit team identified pleural abnormalities were shared with Lungscreen for review.

Table 6. *Comparison of audit team and Lungscreen findings for the presence or absence of classifiable pleural abnormalities. Sensitivity and specificity were calculated using the audit team findings as the gold standard.*

Classifiable Pleural Abnormalities		**Audit Team Findings**		**Total**
		Positive	**Negative**	
Lungscreen Findings	**Positive**	3	1	4
	Negative	25	1,826	1,851
	Total	28	1,826	1,855

Sensitivity (95% CI) = 10.7% (2.3%, 28.2%)

Specificity (95% CI) = 99.95% (99.7%, 100.0%)

Other Abnormalities

Section 4 (A, B, C, D, and E) of the ILO CXR classification form is where other abnormalities must be recorded through the use of 28 obligatory symbols or a list of 18 other diseases. A free text field is available for abnormalities not covered by the symbols or diseases list, or for additional comments. The audit team reported a higher prevalence of other abnormalities in 34% (n = 628) of cases compared to Lungscreen B readers' 23% (n = 424). The five most common abnormalities reported by the audit team were rib (9.9%) and non-rib fractures (7.1%); post-surgical changes (4.0%); scoliosis (3.2%); and density/nodule, nodular lesion (2.7%). The five most common abnormalities indicated by Lungscreen were heavy or increased bronchovascular markings (3.7%); plate atelectasis (3.1%); eventration (2.3%); post-surgical changes (2.2%); and fractured ribs (2.2%) (See **Table 7**).

In addition to this broad comparison between the audit team and Lungscreen B readers, the audit team focused on abnormalities that could be serious, or were an indication of mine dust lung disease. Therefore a comparison was carried out reviewing the presence of emphysema (symbols EM or BU), concern for possible lung cancer (CA; lung density; or nodular lesion); or an abnormality of cardiac size or shape (CO). Lungscreen and audit team B readers indicated emphysema in 20 and 7 cases, respectively, with only four cases in common. The audit team indicated possible cancer, nodular lesions, or lung densities in 50 cases compared to 18 cases identified by Lungscreen. Importantly, there were only nine cases in which the audit team and Lungscreen B readers agreed on these findings. Abnormality of cardiac shape and size was a rare finding by either team, but of the 10 cases identified by the audit team, only two were noted as such by Lungscreen.

As noted above, the audit team could not account for abnormalities that may have been recorded in a narrative report by the clinic's radiologist and not included in Lungscreen's ILO classification form.

Table 7. *Distribution of other abnormalities as indicated by the audit team and Lungscreen B readers. The number of cases in which the audit team and Lungscreen B readers both reported this listed condition is noted.*

ILO Other Abnormalities/Diseases	**Audit Team n (%)**	**Lungscreen n (%)**	**Number of Common Cases**
Abnormality of cardiac size or shape (CO)*	10 (0.5)	7 (0.4)	2
Aorta, anomaly of	1 (0.0)	0 (0.0)	0
Atherosclerotic Aorta (AA)	50 (2.7)	0 (0.0)	0
Azygos lobe	11 (0.6)	8 (0.4)	8
Bony chest cage abnormality	25 (1.4)	9 (0.5)	6
Bronchovascular markings, heavy or increased	22 (1.2)	68 (3.7)	2
Calcification in small pneumoconiotic opacities (CN)	1 (0.0)	0 (0.0)	0
Calcified non-pneumoconiotic nodules (e.g. granuloma) or nodes (CG)	14 (0.8)	12 (0.7)	4
Cancer (CA); Density, lung; Nodule, nodular lesion*	50 (2.7)	18 (1.0)	9
Emphysema (EM) or Bulla(e) (BU)*	8 (0.4)	20 (1.1)	5
Enlargement of non-calcified hilar or mediastinal lymph nodes (HI)	6 (0.3)	3 (0.2)	0
Eventration	15 (0.8)	42 (2.3)	9
Foreign body	45 (2.4)	1 (0.0)	0
Fracture, healed (non-rib)	131 (7.1)	37 (2.0)	33
Fracture, not healed (non-rib)	7 (0.4)	3 (0.2)	2
Fractured rib(s) (acute or healed) (FR)	183 (9.9)	40 (2.2)	35
Hiatal hernia	8 (0.4)	3 (0.2)	3
Hyperinflation	8 (0.4)	37 (2.0)	4
Ill-defined diaphragm border (ID)	0 (0.0)	1 (0.0)	0
Inflitrate	0 (0.0)	2 (0.1)	0
Marked distortion of an intrathoracic structure (DI)	1 (0.0)	1 (0.0)	0
Parenchymal bands (PB)	1 (0.0)	4 (0.2)	0
Plate atelectasis (PA)	37 (2.0)	57 (3.1)	20
Pleural effusion (EF)	5 (0.3)	4 (0.2)	1
Pleural thickening of an interlobular fissure (PI)	0 (0.0)	6 (0.3)	0
Post-surgical changes/sternal wire	75 (4.0)	41 (2.2)	38
Scoliosis	59 (3.2)	30 (1.6)	20
Significant apical pleural thickening (AT)	1 (0.0)	4 (0.2)	1
Tuberculosis (TB)	0 (0.0)	1 (0.0)	0
Vertebral column abnormality	3 (0.2)	11 (0.6)	0

* Indicates condition of concern

Note: Cases may have mad more than one abnormality reported, therefore the total numbers in this table will exceed the number of cases with abnormalities reported.

B readers are required to check a box, ILO form (Section 4D), to "See Doctor" when they find an abnormality that warrants follow-up by their personal physician. The audit team indicated physician follow-up was warranted in 70 cases (3.8%), of which Lungscreen identified five. Physician follow-up was most commonly indicated for findings that raised the concern for possible cancer, including the symbol ca, lung density, or nodular lesion (n = 42, 60%) (See **Table 8**). In total, Lungscreen suggested physician follow-up in seven cases, two because of nodular lesions, and five cases because of findings noted in free text (ILO Form Section 4E). The free-text comments included findings of cardiomegaly and opacities warranting further investigation.

It should be noted that radiology reports by radiologists at the clinic taking the image were not included in the scope of this review (these are completed in addition to the ILO reports by Lungscreen). These reports should also identify any significant other abnormalities requiring further review. RSHQ has forwarded the abnormalities identified by the audit team to the appropriate physician for assessment (e.g. Appointed Medical Adviser under the CMWHS).

***Table 8**. Distribution of other abnormalities for cases in which the audit team recommended physician follow-up.*

ILO Other Abnormalities/Diseases	Audit Team (n)	Lung Screen (n)
Abnormality of cardiac size or shape (CO)	3	1
Atherosclerotic Aorta (AA)	8	
Bronchovascular markings	1	
Bulla(e) (BU)	1	
Calcification in small pneumoconiotic opacities (CN)	1	
Calcified non-pneumoconiotic nodules (e.g. granuloma) or nodes (CG)	4	1
Cancer (CA); Density, lung; Nodule, nodular lesion	42	3
Emphysema (EM)	2	
Enlargement of non-calcified hilar or mediastinal lymph nodes (HI)	4	
Eventration	4	
Fracture healed (non-rib)	5	
Fractured rib(s) (acute or healed) (FR)	8	
Hiatal hernia	7	
Hyperinflation	2	
parenchymal bands (PB)	1	
Plate atelectasis (PA)	5	
Pleural effusion (EF)	1	
Post-surgical changes/sternal wire	3	
Scoliosis	2	
Significant apical pleural thickening (AT)	1	

Note: Cases may have mad more than one abnormality reported, therefore the total numbers in this table will exceed the number of cases referred to physician follow-up.

Discussion

This review of Lungscreen's classification of chest images for the presence of abnormalities consistent with pneumoconiosis showed excellent agreement overall. Of note, while representative of the average age of Queensland coal mine workers screened since 2019, the images selected for this review came from a population that was quite young (mean age 38 years), and therefore likely had low work tenures with low cumulative exposure to mine dust. The prevalence of abnormalities consistent with disease seen in this group was low compared to the broader population screened since 2016. There were only eight images that were positive for indications of pneumoconiosis detected by Lungscreen, three of which were also classified as such by the audit team. This means that the study was underpowered to detect a difference in classifications between the audit team and Lungscreen. Therefore, we cannot reliably comment on the sensitivity of Lungscreen for the detection of opacities consistent with simple or complicated pneumoconiosis compared to the audit team.

- The audit team and Lungscreen agreed on quality 1 or 2 88% of the time. However in 8.7% of the cases the audit team was more stringent in the evaluation of quality finding quality 3 or 4 when Lungscreen found quality 1 or 2. The converse was also true in a smaller number of cases where Lungscreen graded more stringently in some cases than the audit team. The two cases deemed unreadable by the audit team were given a quality grade of 1 by Lungscreen. Of note, in the past, B-readers were not specifically tested on the classification of quality. This has changed and the new NIOSH B-reader syllabus and exam have focused extensively on quality because of the impact of changes in quality on the classification of pneumoconiosis. The differences in quality grading may also be influenced by differences in display software and algorithms. Further investigation into this area is warranted.

- The audit team noted pleural and other abnormalities more frequently than Lungscreen. The ILO classification system requires classifying pleural abnormalities in addition to parenchymal abnormalities. As noted above, some cases that were not classified by Lungscreen may have been the result of sub-pleural fat. ILO requires classification of all pleural abnormalities and this was clarified during consensus calls when results were discussed with Lungscreen.

- The audit team also noted other abnormalities that warranted referral to the subject's physician more frequently than Lungscreen. This may be due in part to the audit team taking a more conservative approach and reporting all findings, even though they were not necessarily relevant to the presence of an occupational exposure. In addition, comments that may have been noted on a narrative report and not recorded on the ILO form were not considered by the audit team. In general, B-readers are encouraged to err on the side of increased referral to be sure that any possible abnormality is brought to the notice of the subject's primary provider. This was also discussed with Lungscreen and RSHQ provided the audit results related to other abnormalities to the relevant physician.

Conclusions and Recommendations

1) **Study power:** Overall, Lungscreen is classifying chest images for abnormalities consistent with pneumoconiosis with high accuracy. Of note, this study was underpowered for an evaluation of the detection of pneumoconiosis due to the very low prevalence of disease in the images of the population referred for evaluation. The audit team recommends future studies focus on a population of miners with more than 20 years of mining tenure and at least 40 years of age in order to be able to evaluate the sensitivity and specificity of Lungscreen's classifications. We understand RSHQ's new online digital health records system 'ResHealth' will make this possible for future audits. The use of control images could also be considered.

2) **Increased attention to the classification of pleural abnormalities:** Lungscreen should consider encouraging their readers to classify any pleural abnormalities present on the image and note other considerations in the comments.

3) **Increased attention to the identification of other abnormalities:** Lungscreen should consider encouraging their readers to have a lower threshold for identifying other abnormalities and referring these cases to the primary provider for clinical correlation.

4) **Increase the use of Digital Radiography:** Of note, four out of the nine major facilities participating in this audit are still using Computed Radiography technology which is associated with a greater rate of poor quality images (grades 3 or 4). Efforts could be made to encourage the use of Digital Radiography technology in clinics acquiring images for the Scheme.

5) **Feedback to providers with high rate of poor quality images:** RSHQ should consider providing feedback from this audit to these clinics to give them an opportunity to improve their imaging techniques.

6) **Further investigation into causes of poor quality images and differences in quality ratings:** RSHQ should consider further analysis of images with quality 3 and 4 ratings to determine if equipment, technique, image display, or interpretation differences explain some of the differences noted in this report.

References

1. Review of Respiratory Component of the Coal Mine Workers' Health Scheme - monash-qcwp-final-report-2016.pdf. Accessed September 1, 2016. https://www.dnrm.qld.gov.au/__data/assets/pdf_file/0009/383940/monash-qcwp-final-report-2016.pdf

2. International Labour Office, ed. *Guidelines for the Use of the ILO International Classification of Radiographs of Pneumoconioses*. Revised edition 2011. International Labour Office; 2011.

3. US Code of Federal Regulations. Method of obtaining definitive chest radiograph classifications. 42 CFR 37.53. Electronic Code of Federal Regulations (eCFR). Accessed January 28, 2021. https://ecfr.federalregister.gov/on/2021-01-28/title-42/chapter-I/subchapter-C/part-37/subpart-ECFRef4963eeed1af04/subject-group-ECFR7e2909b44451903/section-37.53

4. Petsonk EL, Rose C, Cohen R. Coal mine dust lung disease. New lessons from an old exposure. *Am J Respir Crit Care Med*. 2013;187(11):1178-1185. doi:10.1164/rccm.201301-0042CI

5. Seibert JA. Digital radiography: The bottom line comparison of CR and DR technology. Published online 2009:8.

6. Digital Radiography versus Computed Radiography. News-Medical.net. Published November 27, 2018. Accessed January 29, 2021. https://www.news-medical.net/health/Digital-Radiography-versus-Computed-Radiography.aspx

Appendix A: Biosketches of participating B Readers

Robert A. Cohen

Dr. Robert Cohen is Professor of Medicine and Director of the Occupational Lung Disease Program at Northwestern University. He is also Clinical Professor of Environmental and Occupational Health Sciences at the University of Illinois Chicago, School of Public Health. His major research interests are occupational lung disease, particularly mineral dust exposed workers. He has served as a consultant to several agencies of the United States government in areas of mining related health issues including the Respiratory Health Division of the National Institute of Occupational Safety and Health, MSHA, the Division of Coal Mine Workers Compensation of the US DOL. He is the principal investigator on the Black Lung Center of Excellence as well as co-Investigator for the Black Lung Clinics Program funded by the Office of Rural Health Policy of the Health Resources and Services Administration. He has served on several government oversight committees including the Mine Safety Research Advisory Committee and the US National Academies of Science and Institute of Medicine committee to review Personal Protective Technologies. He is a NIOSH certified B-reader since 1998 and has worked closely with the US National Institute for Occupational Safety and Health to develop a new teaching syllabus and exam for B readers in the United States as well as the development of new standard images for the International Labour Office. He has worked extensively internationally in the area of medical surveillance for coal mine dust and silica exposed workers. He has worked on projects in Ukraine, Colombia, Argentina, and most recently has been working with Queensland State Government agencies to review and improve their Coal Mine Workers' Health Scheme.

Kathleen DePonte

Kathleen A. DePonte, M.D. is a board-certified diagnostic radiologist and NIOSH-certified B-reader currently in private practice as a partner with Mountain Empire Radiology, P.C. She completed her residency in diagnostic radiology at Wake Forest University. As president of Diagnostic Imaging Associates, P.C, she specializes in imaging of pneumoconiosis and occupational lung diseases. She serves as a consultant for multiple clinics including Washington and Lee Black Lung Clinic. As a panel surveillance reader for NIOSH she served on the ACR/NIOSH task force for the development of the new digital NIOSH B-reader certification examination. Dr. DePonte also serves as a member of the Adjunct Clinical Faculty at DeBusk College of Osteopathic Medicine.

Robert Tallaksen

Robert J. Tallaksen, MD, is a graduate of the University of North Carolina School of Medicine. He trained in Diagnostic Radiology at the U.S. Naval Hospital in Bethesda, MD, and did further fellowship training in Thoracic Radiology at the Armed Forces Institute of Pathology in Washington, DC. After completing his military service, he worked for several years in private practice. He joined the faculty of the West Virginia University School of Medicine in 2001 as Chief of Cardiothoracic Radiology. He is a Senior Consultant to the Respiratory Health Division of NIOSH/CDC.

Appendix B: Technical Quality Defects by Clinic

Table B1. Distribution of technical quality defects indicated on chest x-rays with a quality grade of 2, 3, or 4 as classified by the audit team B readers by clinic.

			Technical Quality Defect															
			Overexposed		Underexposed		Artifacts		Improper Position		Poor Contrast		Underinflation		Mottle		Excessive Edge Enhancement	
Clinic[a]	N	X-rays with Quality Grade 2, 3, or 4 n (%)	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
A	312	213 (68)	129	60.6	2	0.9	0	0.0	38	17.8	178	83.6	15	7.0	2	0.9	1	0.5
B	295	245 (83)	9	3.7	4	1.6	0	0.0	57	23.3	40	16.3	15	6.1	197	80.4	34	13.9
C	276	231 (84)	146	63.2	4	1.7	0	0.0	68	29.4	188	81.4	26	11.3	38	16.5	0	0.0
D	239	158 (66)	1	0.6	0	0.0	0	0.0	72	45.6	13	8.2	24	15.2	93	58.9	0	0.0
E	174	99 (57)	1	1.0	2	2.0	1	1.0	16	16.2	12	12.1	5	5.1	80	80.8	0	0.0
F	40	28 (70)	17	60.7	0	0.0	0	0.0	13	46.4	15	53.6	6	21.4	2	7.1	1	3.6
G	132	97 (73)	0	0.0	0	0.0	1	1.0	10	10.3	6	6.2	9	9.3	90	92.8	0	0.0
H	154	136 (88)	4	2.9	0	0.0	4	2.9	51	37.5	20	14.7	7	5.1	118	86.8	23	16.9
I	171	127 (74)	0	0.0	20	15.7	1	0.8	46	36.2	30	23.6	18	14.2	111	87.4	0	0.0

[a] The table excludes 64 images taken from eight additional clinics. The nine clinics listed in this table account for 96.6% of images reviewed in the audit.

Appendix C: ILO Form Data Abstraction Tool

LungScreen ILO Report

Record ID	______________________
First Name	______________________ (Miner's First Name)
Last Name	______________________ (Miner's Last Name)
Date of Birth	______________________
Date of Radiograph	______________________

IMAGE QUALITY

Film Quality (1)	◯ 1 ◯ 2 ◯ 3 ◯ U/R ◯ No Entry (B read film quality)
Image Quality Issues	☐ Overexposed (dark) ☐ Underexposed (light) ☐ Artifacts ☐ Improper position ☐ Poor contrast ☐ Poor processing ☐ Underinflation ☐ Mottle ☐ Excessive edge enhancement ☐ Other (Check all boxes that are marked in section 1)
Image Quality Other	______________________

ABNORMALITIES

NORMAL STUDY? No pneumoconiosis or significant abnormalities	◯ Yes ◯ No

PARENCHYMAL ABNORMALITIES

Parenchymal abnormalities consistent with pneumoconiosis? (2A)	◯ Yes ◯ No (Mark yes if positive findings for pneumoconiosis)
Primary small opacities, shape/size (2B)	◯ p ◯ q ◯ r ◯ s ◯ t ◯ u ◯ No Entry (Small opacities shape/size primary)
Secondary small opacities, shape/size (2B)	◯ p ◯ q ◯ r ◯ s ◯ t ◯ u ◯ No Entry (Small opacities shape/size secondary)

Lung Zones

	Right	Left
Upper	☐	☐
Middle	☐	☐
Lower	☐	☐

Profusion score (2B)	◯ 0/- ◯ 0/0 ◯ 0/1 ◯ 1/0 ◯ 1/1 ◯ 1/2 ◯ 2/1 ◯ 2/2 ◯ 2/3 ◯ 3/2 ◯ 3/3 ◯ 3/+ ◯ No Entry
Large opacities (2C) consistent with pneumoconiosis?	◯ O ◯ A ◯ B ◯ C ◯ No Entry

PLEURAL ABNORMALITIES	
Any pleural abnormalities consistent with pneumoconiosis? (3A)	◯ Yes ◯ No (Mark yes if positive findings for pleural abnormalities)

OTHER ABNORMALITIES	
Any other abnormalities? (4A)	◯ Yes ◯ No
Other Symbols (4B)	☐ aa ☐ at ☐ ax ☐ bu ☐ ca ☐ cg ☐ cn ☐ co ☐ cp ☐ cv ☐ di ☐ ef ☐ em ☐ es ☐ fr ☐ hi ☐ ho ☐ id ☐ ih ☐ kl ☐ me ☐ pa ☐ pb ☐ pi ☐ px ☐ ra ☐ rp ☐ tb (Select all that are marked in section 4B)
Other Abnormalities (4C)	☐ Eventration ☐ Hiatal hernia ☐ Bronchovascular markings, heavy or increased ☐ Hyperinflation ☐ Bony chest cage abnormality ☐ Fracture, healed (non-rib) ☐ Fracture, not healed (non-rib) ☐ Scoliosis ☐ Vertebral column abnormality ☐ Azygos lobe ☐ Density, lung ☐ Inflitrate ☐ Nodule, nodular lesion ☐ Foreign body ☐ Post-surgical changes/sternal wire ☐ Cyst ☐ Aorta, anomaly of ☐ Vascular abnormality
Should the worker see a personal physician because of findings? (4E)	◯ Yes ◯ No ◯ No Entry
Comments	______________________
Date of B Read	______________________