

# New Composite ('Quadrant') Standard Films for Classifying Radiographs of Pneumoconioses

Michael JACOBSEN\*, William E MILLER and John E PARKER

National Institute for Occupational Safety and Health, Division of Respiratory Disease Studies, Willowdale Road, Morgantown WV 26505, USA.

*Received April 10, 1998 and accepted August 10, 1998*

**Abstract:** 120 chest radiographs of dust exposed persons were classified twice by each of 39 physicians, and three times by 37 of them. All film readers used both the 22 standard films accompanying the current International Labour Office's (1980) Classification of Radiographs of Pneumoconioses, and a modified set that reproduced selected parts from some of them as quadrants of full-size films. Variability within and between readers in small opacity profusion classifications was high, but similar when using the two sets of standards. Some readers recognised irregularly shaped small opacities more frequently, in low profusion categories, when using the quadrant standards. Small rounded opacities were classified more frequently higher on the scale when using the current standards. The effects found were small although the test films included a high proportion with some small opacities. It is concluded that use of the new quadrant standards is unlikely to introduce any important bias into results from occupational health surveys.

**Key words:** ILO, International, Classification, Chest radiograph, Pneumoconiosis

The International Labour Office's Classification of Radiographs of Pneumoconioses<sup>1)</sup> incorporates 22 standard radiographs. Eighteen of them define classifiable levels of profusion of small opacities. We report here briefly<sup>#</sup> on results from a trial designed to determine whether and how film readers' classifications might be altered if only selected parts from some of the standard films are reproduced as quadrant sections of modified standards.

Thirty-nine physicians, in 10 countries<sup>##</sup> and with varying levels of normal film reading commitments, each repeatedly classified 120 chest radiographs of persons known to have been occupationally exposed to dusts. Two sets of standard films were used: (a) the 22 current 'ILO' (1980) standards, and (b) a modified ('QUAD') set consisting of 14 films.

The QUAD set included 10 of the existing (ILO, 1980) full-size standards and four new composite films that reproduce sections from the existing standards that were not included as full-size films (Table 1).

Nineteen of the film readers (from five countries) used the ILO standards on the first occasion that they saw the 120 films ('Round 1'), and the QUAD standards in Round 2 (IQ readers). The other 20 readers used the standards in the reverse (QI) sequence. All (19) IQ readers and 18 of the QI readers reviewed the 120 films a third time (Round 3), so that it was possible also to assess within-reader variability when using the same standards. In Round 3, readers classified 60 of the 120 films using the ILO standards and the other 60 using the QUAD standards. At least six (generally considerably more) weeks elapsed between Rounds 1, 2 and 3. Readers were asked to allow at least 12 hours between viewing the first and second sub-sets of films in Round 3. Results were recorded in a standardised format. Records that were not internally consistent, or that could not be paired with an additional valid classification of the

\* To whom correspondence should be addressed at Institute of Occupational and Social Medicine, University of Cologne, D-50924 Cologne, Germany.

# A detailed report is available; see Acknowledgements, below.

## Canada; China; Czech & Slovak Republics; Finland; France; Germany; Japan; Poland; UK; USA.

**Table 1. Fourteen radiographs comprising the experimental 'QUAD' standards used in this trial**

10 full-size ILO (1980) standard films	4 new composite ('QUAD') films, from sections of existing (1980) standards (showing spatial arrangement of sections)		
0/0 example 1	Composite p/p:	[0/0]	[2/2]
0/0 example 2		[1/1]	[3/3]
1/1 q/q	Composite r/r:	[0/0]	[2/2]
1/1 t/t		[1/1]	[3/3]
2/2 q/q	Composite s/s:	[0/0]	[2/2]
2/2 t/t		[1/1]	[3/3]
3/3 q/q	Composite large opacities:	[A]	[B]
3/3 t/t		[C]	
ILO (1980) pleural shadows quadrant			
ILO (1980) u/u composite (re-arranged)			

**Table 2. Within-reader variability: (weighted) mean percentages (and standard errors) of identical small opacity profusion classifications on 11-point scale**

Rounds involved	Repeated use of		
	ILO standards	QUAD standards	Both standards
R1 & R3	41.5 [19 IQ readers] (2.0)	43.5 [18 QI readers] (2.8)	42.5 (1.7)
R2 & R3	48.1 [18 QI readers] (2.3)	45.3 [19 IQ readers] (2.6)	46.7 (1.7)
All readers	44.7 (1.6)	44.4 (1.9)	44.6 (1.2)

same film by the same reader, were excluded from the analyses. This left 12,929 valid classifications from Rounds 1, 2 and 3, that is, 94% of the 13,800 theoretically possible.

The ability to reproduce small opacity profusion classifications of the same films when using the same standards (in Round 3) varied substantially between the 37 readers who completed all three rounds, from 22% to 70% of paired classifications. But on average, there was little difference between the two sets of standards in the percentages of identical classifications (Table 2). However, the ranges of readers' variability in their duplicate classifications of the same films with the same standards (scoring 1 to 11 for ordered sub-categories) tended to be

less when using the QUAD standards. Twenty readers recorded wider ranges for the paired score differences when using the ILO standards; only four had wider ranges using the QUADs. (The other 13 readers showed no difference between the standards in this respect.) This tendency to register less marked differences on the profusion scale when using the QUAD standards was statistically significant for both IQ and QI readers ( $P < 0.03$ ).

Between reader variability (all 39 readers) was assessed using results from Rounds 1 and 2. Average profusion scores ranged from 2.4 to 5.0 when using the ILO standards ( $SD = 0.6$ ) and from 2.7 to 4.7 when using the QUADs ( $SD = 0.5$ ). The slightly lower between-reader variability

**Table 3.** Summary of small opacity profusion classifications from the first two viewings of 120 films by 39 readers: (a) valid pairs of classifications, and (b) separate distributions into main categories using ILO and QUAD standards, by Reading Sequence and readers' nominations of predominant shapes of small opacities.

Reading Sequence	(a) Paired classifications on 11-point* scale					(b) % Distributions to main categories				
	Predominant shape (both Rounds)	No. of pairs	Percent with profusion higher using -			0	1	2	3	
			QUADs	ILO	same profsn.					using
IQ (ILO in Round 1; 19 readers)	irregular	517	33.3	29.8	36.9	ILO	18.0	55.9	21.1	5.0
						QUAD	14.9	58.4	22.2	4.4
	rounded	679	25.9	35.5	38.6	ILO	9.6	43.4	33.3	13.7
						QUAD	8.2	49.8	31.1	10.9
	remainder	1041	22.4	27.2	50.4	ILO	70.8	22.3	6.2	0.7
						QUAD	72.3	21.5	6.0	0.2
IQ (QUAD in Round 1; 20 readers)	irregular	795	40.1	23.1	36.7	ILO	15.3	63.3	17.4	4.0
						QUAD	10.2	68.0	18.2	3.5
	rounded	517	30.2	29.8	40.0	ILO	6.4	39.4	37.9	16.2
						QUAD	4.8	41.0	39.6	14.5
	remainder	1042	25.9	19.7	54.4	ILO	73.9	20.7	5.3	0.1
						QUAD	69.6	25.3	5.0	0.1
IQ - all shapes		2237	26.0	30.3	43.7	ILO	40.0	36.5	17.9	5.6
						QUAD	39.6	38.6	17.3	4.4
IQ - all shapes		2354	31.7	23.0	45.3	ILO	39.3	39.2	16.5	5.0
						QUAD	35.3	43.2	17.1	4.4
Both Sequences; all shapes		4591	28.9	26.6	44.5	ILO	39.6	37.9	17.2	5.3
						QUAD	37.4	41.0	17.2	4.4

\*sub-categories 0/- and 0/0 were pooled together with unambiguous blanks on the reading sheets that indicated no small opacities.

when using the QUADs could easily have arisen simply by chance ( $P > 0.2$ ).

Three quarters of the 12,929 valid film classifications from all three viewings of the films recorded category 0/1 or higher; 74.4% and 74.7% when using the ILO and QUAD standards, respectively.

In Rounds 1 and 2, eight of the IQ readers and 13 of the QI readers read more profusion when using the QUADs, respectively, than when they used the ILO standards. The other 18 readers read more profusion when using the ILO standards. Table 3 summarises the 4591 pairs of valid classifications from Rounds 1 and 2. In aggregate, 2043 (44.5%) were identical on the 11-point profusion scale when using the two sets of standards. The slightly higher proportion with more profusion when using the QUADs (28.9%) as compared with the ILO standards (26.6%) is reflected in the separate aggregate marginal distributions to the four major profusion categories (the bottom-right section of Table 3). This shows 2.2% more classifications in categories 1, 2 and

3 combined when using the QUADs (i.e., 62.6% versus 60.4%).

The other sections of Table 3 draw attention to two factors that influenced the aggregate results: (i) the sequence in which the standards were used in Rounds 1 and 2, and (ii) readers' perceptions of the predominant shapes of the small opacities that they detected. First, IQ readers ("all shapes") tended to read more profusion when using the ILO standards, in Round 1; but QI readers tended to read more profusion when using the QUADs—also in Round 1. However, section (b) of Table 3 shows that, when using the QUADs, both IQ and QI readers generated higher proportions of classifications into category 1, and lower proportions into category 3.

The latter tendencies, to read more profusion when using the QUADs for films with lower profusion categories, but less profusion when using the QUADs for films with higher profusion categories, were related to readers' judgements about the shapes of small opacities. In particular, films repeatedly classified as showing small irregular opacities

tended to attract higher profusion classifications when using the QUAD standards. This trend was reflected in more classifications into categories 1 and 2 when using the QUADs; was more marked among QI than IQ readers; was shown to be statistically significant among both IQ and QI readers ( $P < 0.0001$  and  $P < 0.03$ , respectively); and was evident not only for films showing predominantly 's'-sized shadows, but also for those categorised as predominantly size 't'. (The latter finding is remarkable because all three 't/t' standards in the QUAD set were full-size reproductions of those in the ILO set—see Table 1.)

Additional analyses showed that IQ readers read more profusion when using the QUAD standards in Round 3 than when they used them in Round 2 (4.6% more frequent classifications to category 1 or higher;  $P < 0.03$ ). There was no analogous tendency among QI readers to read more profusion in Round 3 than in Round 2 when they used the ILO standards repeatedly with the same films.

Despite wide variation between readers regarding the presence of large opacities (ranging from 2.5% to 12.5% of the duplicated viewings in Rounds 1 and 2) a tendency to recognise them less frequently when using the QUAD standards was statistically significant ( $P < 0.0005$ ).

Analyses of covariance, based on results from all three rounds, verified that small opacity profusion classifications when using the two sets of standards varied, depending on (a) the sequence in which readers used the standards; (b) readers' views about the shapes of the small opacities that they saw; and (c) other reader-specific factors. All these factors interacted to yield a small net effect, amounting to about three more profusion sub-categories per 100 film classifications when using the QUADs, or 2% more frequent classifications with the QUADs to categories 1, 2 and 3 combined. Some readers recognised the presence of irregularly shaped small opacities more frequently when using the QUAD standards, and tended to classify them in

higher profusion categories, in the lower part of the profusion scale, when using the QUADs.

The 120 films used for this study were selected deliberately to include an unusually high proportion (about 60%) with signs likely to be classified as abnormalities. A wide range of appearances associated with the pneumoconioses was thus included, as intended, but the resulting high prevalence of small opacities is atypical of real occupational health survey situations. Therefore we conclude that use of the new quadrant standards as alternatives or supplements to the present set is unlikely to introduce any important bias that might interfere with the interpretation of results from future studies. If the availability of these new composite radiographs encourages readers to use standard films more frequently, then this may help ultimately to reduce the high variability between readers in their assessments of the same films.

### Acknowledgements

This work was planned on the initiative of various experts advising the ILO. It was co-ordinated by the Division of Respiratory Disease Studies of the US National Institute for Occupational Safety and Health, and it would not have been possible without generous donation of expertise and time from the physicians who read the films. Their names and affiliations are recorded in a detailed report (HETA 93-0340, July 1997) available from NIOSH Publications Office, 4676 Columbia Parkway, Cincinnati, Ohio 45226, USA.

### Reference

- 1) Guidelines for the use of ILO International Classification of Radiographs of Pneumoconioses, Revised edition 1980. Occupational Safety and Health Series No. 22 (Rev.) Geneva: International Labour Office, 1980.