

《1986尘肺诊断标准》读片重复性的研究

中国预防医学科学院劳动卫生与职业病研究所 丁茂柏 张翠娟 卢世璋

摘要 为了验证《1986尘肺诊断标准》的读片重复性,作者用集体读片方式分析读片差异、分析原因。

两次读片诊断符合程度75%;小阴影形态70%;密集度75%;分布范围65%;大阴影、胸膜斑、小阴影聚集及弥漫胸膜增厚则完全相符。

影响诊断结果相符程度的主要因素是密集度判定差异。产生差异的原因可能与标准片对比方法、胸片示局部小阴影密集度较高及混合形态阴影对判定的影响等因素有关。

关键词 尘肺 诊断标准 X线

《1986尘肺诊断标准》^{〔1〕} (以下简称《标准》)应用价值的高低、能获得公认的程度,取决于确实性强、重复性高及实用性好三项条件。《标准》颁布执行后,全国各地均对以上三项内容进行验证,大多反映较积极。惜无正式总结材料发表。原因可能与缺乏统一、简便的验证方法有关。本文利用1988年6月沈阳国际尘肺讨论会上与外国专家交流读片的机会,用交流活动中中方独立完成的部分资料总结而成。旨在探讨《标准》读片结果的重复性,分析读片差异的特点。

读片按中国传统方式进行,即3人组成诊断组,共同读片,讨论决定结果。所指读片重复性乃同一读片组在不同时间读同一胸片的符合程度,不强调在单个人之间相比。不同读片组间的差异将另文讨论。

材料与方 法

片组来源:以参加1988年沈阳国际尘肺讨论会某外国专家准备的一套胸片共200张(称全组)为基准,从中随机取40张(称试验组)作重复读片。

诊断组3人讨论读片,意见分歧时以两票为准作记录。无论诊断结果、密集度、阴影形态均未出现经讨论仍有三种不同意见的实例。记录采用《标准》附录D表格略加修改。读片顺序由供片人随意排定,排列无规律。重读40张穿插在正常读片顺序中进行。胸片乱排及随机抽取重读片目的在于尽可能消除或减少主观性。

所得结果总结如表1~7。

表1 胸片质量

质量等级	全 组		试 验 组	
	片 数	%	片 数	%
1	58	29.0	10	25.0
2	95	47.5	18	45.0
3	34	17.0	8	20.0
4*	13	6.5	4	10.0

* 等级4定义为废片(Unacceptable),按《标准》规定不作读片记录。本次读片前约定,尽可能作形态观察。

表2 尘肺诊断分期分布比较

分 期	全 组		试 验 组	
	片 数	%	片 数	%
0	57	28.5	11	27.5
0+	12	6.0	1	2.5
I	40	20.0	6	15.0
I+	23	11.5	4	10.0
II	23	11.5	7	17.5
II+	19	9.5	6	15.0
III	16	8.0	4	10.0
III+	5	2.5	0	0
不 定*	5	2.5	1	2.5
总 数	200	100.0	40	100.0

* 不定:①部分4级片经努力仍不能作明确结论者。②根据《标准》要求需结合临床资料才能作诊断者。

表1说明试验组胸片质量与原片组相比,大致相似,有随意性。

表2表明试验组期别分布与原片组相比大致相同,两组胸片均以小阴影片为主,占87%左右。

表3中实线框中数字为两次读片相符片

表3 诊断结果差异

第一次读片	0	0'	I	I'	II	II'	III	III'	IV	IV'	不定	总数
0	10											10
0'		1										1
I			6	2	1							9
I'			2	1								3
II			2	1	3							6
II'					2	3						5
III						4						4
III'							0					0
IV								0				0
IV'									2			2
总数	10	1	10	4	6	3	4	0	2			40

表4 小阴影形态观察差异

第一次读片	0	p	s	q	t	r	u	不定	总数
0	10								10
p		1	2	1					4
s			2						2
q				7	1				8
t				5	2				7
r				2		4	1		7
u							1		1
不定								1	1
总数	10	1	2	16	4	4	1	2	40

表5 小阴影密集度差异

第一次读片	0	1	2	3	总数
0	13				13
1		12	2		14
2		4	3	1	8
3			3	2	5
总数	13	16	8	3	40

数,共30片。大方框表示差异集中的区段,共23片,此区段两次读片相符者13片。

表4实线框中总数28,形态差异特点为:大小差异p与q为2,r与q为2;圆与不规则间差异p

表6 小阴影分布范围观察差异

第一次读片	0	1	2	3	4	5	6	总数
0	11							12
1		0						0
2			1	1		1	1	4
3			1	1			2	4
4			1	1	2	1	1	6
5				1		1	1	3
6						1	10	11
总数	11	0	3	3	4	4	15	40

表7 其它尘肺阴影观察差异

尘肺阴影	全组		试验组		观察重复数
	片数	%	片数	%	
胸膜斑	19	9.5	3	7.5	3
弥漫胸膜增厚	14	7.0	2	5.0	2
肋膈角闭锁	7	8.5	3	7.5	3
胸膜钙化	7	3.5	0	0	0
小阴影聚集	11	5.5	2	5.0	2
大阴影	16	8.0	4	10.0	4

与t为1,q与t为6。即圆窗差与圆条差之比为4比7。其中以q与t间差异为主,为6/12,占半数。

表5实线框中总数30片,两次符合75%。

表中数字表明,0与1间界线分明。

总数40片中,0占13,容易掩盖问题性质。若单以小阴影即双线框数字计总数27,符合17,占62.9%。

表6实线框中总数26片。框外数字表示差异分散状况。虚线框内数字表示一次读片分布范围不到4个肺区,而另一次读片小阴影分布范围超过4个肺区者。凡在虚线框外的差异对最后诊断不发生实质影响。

表7表明试验组胸片在所列尘肺阴影与原片组相比也有随意性。观察表明这几类尘肺阴影两次观察完全相符。

讨论

1. 由于尘肺胸片分类目的不同,中外读片方式有差别。但不论用何方式读片,学术界一致认为尘肺胸片分类的读片差异是客观存在的。国外有关读片差异的研究以个人读片方式

为主, 研究不同个人读同一胸片结果间的差异或同一读片人不同时间读同一胸片结果间的差异。前者称人际差异 (inter-variation), 后者称自身差异 (intra-variation)^[3]。

分类方法的重复性与读片差异有关, 但并不等同。易言之, 差异过大说明重复性差, 但无差异是不可想象的。至于差异小到什么程度算是可接受的且又能表达分类方法的重复性好, 则没有任何明确标准, 只是同行之间大致界限, 即符合程度达60~80%是可接受的。较为困难的单项观察结果或读片生手使用《标准》时符合率约60%左右; 经验较为丰富的人, 总的符合程度能达80%左右。汪绍训教授生前曾说过, 若将他的尘肺读片记录加以整理, 符合率大约不到80%。这与 Dr JC McDonald 1981年在我所访问时发表的意见是一致的。

我国《职业病诊断管理条例》要求尘肺读片诊断应在集体讨论下完成, 因此我国读片实际工作情况与国外有差别。不仅所有的常规诊断读片活动是集体读片, 一些研究项目^[4]也不例外。国外报道读片差异时虽也谈到个人读片方式较能表达结果的客观性、独立性。但并未排斥集体讨论方式或直指集体读片结果无独立性。事实上国外也有用集体读片作研究的^[5,6]。有些大型研究项目虽分别读片, 但结果采用三票两定制^[7]。若能作好安排, 尽可能排除主观干扰, 前后读片在相似环境下进行, 能客观表达读片差异状况。

本试验用集体读片方式, 是从适应我国实际情况出发, 易于推广, 便于各地结果相比较等情况所定。工作安排的特点是供片人负责打乱片序, 重复读片随机抽取, 穿插进行正常读片程序中, 供片人不参加读片等。以减少主观干扰, 免除某种固定片型引起的读片倾向性。两次读片环境相似, 所用读片时间略同等, 可避免重读片多想多忆所造成的主观干扰。

2. 表3~7表明, 两次读片诊断相符程度可达75%; 形态观察相符70%; 小阴影密集度相符75%; 分布范围相符65%; 其它尘肺阴影包括胸膜斑、小阴影聚集、弥漫胸膜增厚及大阴影两次读片结果完全相同。

诊断差异集中发生在 I 到 II⁺ 区段上。表3表明借助标准片的作用, 解决 0 与 I 的分界较

为成功; III 以大阴影达 1×2cm 作指标, 一致性很好。同时也表明诊断差异较集中的区段符合程度不甚理想, 表3大方框中数字符合率56.5%。

除大阴影容易观察外, 小阴影聚集亦不易忽略; 我国《标准》对胸膜斑的记录要求较国外简而明, 其要点在“有”或“无”, 不强调分辨其所在部位、形态、长度及宽度, 故观察重复性高。

3. 符合程度最差的是小阴影分布范围观察结果。表6统计为65%, 但考虑到 0 与小阴影满布 6 个肺区容易观察, 其数冲淡了结果的真实差别。若试将第一次读片结果为 0 或 6 的共23片除去不计, 计算符合程度仅29.4%。另一种计算即两次读片结果为 0 或 6 的全除不计, 符合程度也仅41.7%。

虽云小阴影占 6 个肺区容易观察, 读片差异并不小, 表6示第二次读片15张小阴影分布占 6 个肺区的胸片, 在第一次读片时分散度大, 分散为 5 栏, 最大差异达 4 个肺区梯度。

4. 影响尘肺诊断差异的诸因素中起决定作用的是小阴影密集度判定。

表5示密集度判定前后一致者30片。考虑到“0”较为特殊。除去13片“0”不计, 余密集度前后相符的17片中, 查其前后诊断亦相符者15例, 占88.2%。而这15例中有1例有形态差异, 8例有小阴影分布范围差异, 即60%含有其它差异而不影响最后诊断。反之, 密集度判定两次不符的10片中查其诊断, 除3例因同时出现大阴影或小阴影聚集对诊断起决定性作用外, 其余7例全部出现诊断差异, 其中差一整期(即I与II之差)者3例, 近一半。

小阴影分布范围对诊断差异影响较小, 表6虚线中数字表示理应对诊断发生影响的片数, 仅占全部有差异数之半, 即有半数差异居虚线框外, 其差异无诊断影响。虚线框内数7片深入查对, 实际发生诊断差异的只有2例。在全部诊断差异10片中所占比重为20%。

影响范围差异的主要原因是胸片质量差, 有范围差异的14片中3级片占70%, 其次原因是大阴影掩盖或有肺气肿影响观察。

5. 影响密集度判定因素有二: ①对标准片认识; ②形态差异的影响。

有了标准片, 多数情况下密集度判定无困

难,但少数情况下会因认识不一致产生差异。

按《标准》的规定,密集度判定可分解为形态认定、选择相应的标准片、观察密集度时分区判定、最后作全片评估等几个步骤。其中分区判定时允许和任何一张标准片中所标定的区域密集度相对比,而全片评估时要点在以两个较高级别的密集度肺区作为判定依据。

根据以上程序,一般情况下判定较顺利,发生争议时可从各标准片同级别密集度肺区中找最低者。实践证明,标准片对1级密集度与不够1级间分界的问题解决得很好,重复性较好。但1级和2级间的分界往往发生问题。

表8 标准片中同密集度的肺区数

	p	q	r	s	t	u
1	13	10	14	14	10	1
2	4	16	12	6	10	1
3	6	0	6	2	4	1

上表可见同形态同密集度的样本,不仅在特定标准片中出现,也可在其它标准片中出现。 q_2 共16个肺区, t_2 有10个肺区。这些肺区同是2级密集度,但高低有差别。因此出现用 I_q 标准片比,密集度不够2,但用另一张标准片比就能判定为2。这是2级密集度差异大的主要原因,也提示2级密集度标准片应在修订时加以调整,特别是 I_q 与 I_t 。因为它在全套32张标准片中不是最低者,没有体现低限片的原则。

局部小阴影密集度较高,特别是3级密集度时,常因判定仅一个肺区为3或有两个肺区够3而发生差异。若是局部小阴影密集度高的部位并不整齐落在标准肺区区划上时更易出现上述差异,其结果致密集度评估差一级。

小阴影形态判定差异也会对密集度判定产生影响,实践证明q与t的2级密集度差异最大。用q作观察够2级密集度的,再用t作比较不够2。 r 与 u 也有差异, r 形2级密集度改用 u 作对比,似够3级,而q与t及r与u两种形态阴影混合存在的情况较为多见。

6. 经验证明,全片小阴影只是单一形态的情况较为少见。同是类圆形小阴影或不规则形小阴影,可混合存在不同直径或不同宽度影,如q与r混合、s与t混合;类圆形小阴影与不规

则形混合存在同样普遍,如q与t混合、r与u混合。其实,类圆形小阴影显示并非规整,因此在不同对比度的衬托下,判定为不规则影从而产生差异是很常见的。提示修订标准片时,选明确的混合阴影是值得探讨的。

7. 本次研究还注意到阴影形态记录方式也应作研究,以避免或减少混合型影对诊断差异的影响。

附加代号中pc和pt与胸膜阴影的正式描述重复。还应增加ax一项以解决有明确小阴影融合趋势,而又不夠标准片所示小阴影聚集 I_+ 的程度时作记录的问题。

小 结

根据本文试验,应用标准片作分类时,显示《1986尘肺诊断标准》诊断重复性良好。结合实践经验说明I期起点诊断的重复性问题解决得最好。

重复性较差的环节是I与II诊断分界、小阴影2级密集度的上、下限。提示标准片中2级密集度特别是 I_q 与 I_t 需作调整以体现低限片原则。

本文结论认为标准片应用价值好,但应重视混合形态阴影的密集度标准等技术难点,以求解决,在修订标准(已列入国家七五攻关计划)时提出新片。

此外,《标准》记录表格、附加代号也需改进。胸膜阴影分类宜简不宜繁的原则是对的。

参 考 文 献

1. 国家标准.《尘肺X线诊断标准》:GB5906-86.北京:国家标准出版社,1986.
2. Liddell FDK, Lindars DC. An elaboration of ILO Classification of simple pneumoconiosis. Brit J Ind Med 1969; 26:89.
3. Rossiter CE. Initial repeatability trials of U/C Classification of the radiographic appearances of pneumoconiosis. Brit J Ind Med 1972; 29:409.
4. Huang JQ, Hong YZ. A comparison of Chinese diagnostic standards of silicotic radiographs and the ILO International Classification of radiographs of pneumoconiosis. Ann Occup Hlth 1984; 28:13.
5. Shapiro HA, ed. The reduction of observer variation in categorizing CWP. Pneumoconiosis, London: Oxford Press, 1970; 268~273.
6. Cochrane AC, Garland CD. Observer error in the interpretation of chest films. An international investigation. Lancet 1952; 2:505.
7. Morgan WKC, Brygess DB, et al. The prevalence of CWP in US coal miners. Arch Environ Hlth 1973; 27:221.

Abstracts of Original Articles

**Intra-variation within a observer group
in the application of 1986 Chinese
Classification of Pneumoconioses**

Ding Mao-bo, et al.

Pneumoconiosis is recognized by certain fairly distinctive roentgenographic appearances and a dust-exposure history. Film reading can easily be done under the guidance of a well designed classification system and a set of standard films. However, the variation of results does exist among the observers within the expert group. This study is an attempt to determine the degree and characteristics of intra-variation in filmreading within a certain group of observers.

Key words, pneumoconiosis X-Ray

**Evaluation on diagnostic indices
in lead poisoning**

Zhang Jimei.

Biological evaluation was studied in a total of 688 subjects including groups of exposed, polluted and controlled under lead concentration in air of 0.44~1.92, 0.013~0.24 and 0~0.0007mg/m³ respectively.

The results showed a close dose-response relationship between the levels of PbU, PbB, ZPP and the environmental Pb concentrations.

Key words, diagnostic indices lead poisoning ZPP

**Observations on Renal Function
in Occupational Chronic Plumbism**

Hu Disheng, et al.

The results of PSP excretory test, blood uric acid, β_2 MS, β_2 MU and physical examination in 30 cases of mild lead poisoning showed

(1) The abnormal occurrence rates in lead poisoning group were higher than those of normal group ($p < 0.01$) (2) The level of lead in blood and urine were associated with the increase of β_2 MU. (3) The level of β_2 MU in patients with plumbism decreased after the treatment with chelating agent. The results indicate that it seems to have a subclinical and reversible renal impairment in some chronic plumbism.

Key word, β_2 M PSP excretory test blood uric acid lead poisoning**Clinical study of oral DMSA
treatment for lead poisoning**

Ni Wei-min, et al.

The therapeutic effects of oral administration of DMSA in 70 cases of lead poisoning, 47 cases of lead-absorption, and 20 lead-exposed subjects were studied. Urine Pb excretion was significantly elevated while blood Pb level decreased and δ -ALAD activity returned to normal. No significant adverse effects were noticed.

Key word, DMSA lead poisoning

**Decompression sickness in Hong
Kong railway construction**

Wu Zhe-En, et al.

The symptomatology and treatment of 1534 cases of decompression sickness in a mass transit railway construction site in Hong Kong was reported. Among the 1519 case of type I decompression sickness, 928 cases had limb pain in one site only and 2 cases in as many as eight sites. The knees were the predominant sites of limb pain.

Key words, decompression sickness Recompression minimum effective pressure method