

采用 Kappa 统计量评价腰椎间盘突出 MRI 诊断

李大胜 张芳 高树明 顾金祥 李露 屈辉

【摘要】 目的 应用 Kappa 统计量评价 MRI 诊断腰椎间盘突出的一致性。方法 检查 100 例腰腿痛患者的 300 个椎体。采用 3.0 T MR 扫描仪, 脊柱表面线圈, 应用快速自旋回波序列行腰椎矢状面 T_1 WI 及 T_2 WI、横轴面 T_2 WI。由 2 名放射科医师分别在有及无临床资料的情况下对其中 50 例患者的 L3~4、L4~5、L5~S1 椎间盘的膨出、突出进行 2 次评价; 再由此 2 名医师在无临床资料的情况下共同对其中 52 例患者的 156 个腰椎间盘进行评价。应用 Kappa 统计量评价 2 名医师前后 2 次的诊断结果及其对相同病例诊断结果的一致性。结果 甲乙 2 名放射科医师前后 2 次诊断结果相符合的椎间盘数量分别为 114 和 109 个, 不相符合的椎间盘数量分别为 36 和 41 个, 诊断结果的一致性为中度, Kappa 值分别为 0.60 ± 0.06 和 0.57 ± 0.06 。在有临床资料的情况下, 腰椎间盘膨出的诊断较前明显增加, 分别增加了 10 个和 31 个。在没有临床资料的情况下, 2 名放射科医师之间的诊断结果相符合的椎间盘数量为 77 个, 不相符合者为 79 个, 诊断结果的一致性为弱 (Kappa = 0.24 ± 0.06)。在有无临床资料的情况下, 最大的诊断差异出现在对正常和膨出的腰椎间盘的鉴别上, 2 名医师 2 次诊断膨出的差异分别达 20 和 30 个, 分别占各自总不符合率的 55.6% (20/36) 和 73.2% (30/41); 2 名医师对 156 个椎间盘诊断膨出与正常的差异达 56 个, 为总不符合率的 70.9% (56/79)。结论 对腰椎间盘膨出判断的不一致是造成同一医师 2 次诊断结果和 2 名医师之间诊断结果 Kappa 值低的最主要原因。

【关键词】 椎间盘移位; 腰腿痛; 因素分析, 统计学; 磁共振成像

Application of the Kappa statistic in MRI diagnosis of lumbar disk herniation LI Da-sheng*, ZHANG Fang, GAO Shu-ming, GU Jin-xiang, LI Lu, QU Hui. * Department of Radiology, Beijing Haidian Hospital, Beijing 100080, China

【Abstract】 Objective To assess the concordance of MRI diagnosis for patients suspected of lumbar disk herniation by using Kappa statistic. Methods One hundred patients (48 males and 52 females) with lumbosacral radicular pain, aged from 17 to 86 (average 61). All patients underwent fast spin-echo T_1 and T_2 weighted imaging on a 3.0 T MR scanner and spine surface coil. Two radiologists (doctor A and doctor B) evaluated the lumbar disks from L3—4, L4—5, and L5—S1 in 50 out of the 100 patients independently. The presence of a bulging disk or a herniation was reported. Images were interpreted twice: once before and once after disclosure of clinical information. And disks of 52 patients out of the 100 samples were interpreted by the two radiologists independently without clinical information as well. The Kappa statistics was employed to assess the concordance of each radiologist's diagnoses as well as the observer variation of the two radiologists. Results Diagnoses before and after disclosure to clinical information were concordant in 114 disks for doctor A and in 109 for doctor B, respectively. Diagnoses before and after disclosure to clinical information were not concordant in 36 disks for doctor A and in 41 disks for doctor B, respectively. The Kappa values were 0.60 ± 0.06 and 0.57 ± 0.06 for doctor A and doctor B, respectively. The concordance was moderate. After disclosure to clinical information, the numbers of reported bulging disks increased significantly, by 10 and 31 for doctor A and doctor B, respectively. Without clinical information, the diagnoses of the two radiologists were concordant in 77 disks, while not concordant in 79 disks. The interobserver agreement was poor (Kappa = 0.24 ± 0.06). The difference on diagnoses made between with and without clinical information mainly happened on the differential diagnosis of normal disks and bulging disks. The different diagnoses made between with and without clinical information were on 20 disks and on 30 disks for doctor A and doctor B, respectively; that accounted for 55.6% (20/36) and 73.2% (30/41) of total variation respectively. The diagnostic difference between the 2 doctors happened mainly on differentiation of bulging disks and normal disks, which happened in 56 disks, accounting for 70.9% (56/79) of total variation. Conclusion Variation on diagnoses of the same radiologist or between two

radiologists was mainly caused by disagreement on bulging disks.

【Key words】 Intervertebral disk displacement; Lumbago-leg pain; Factor analysis, statistical; Magnetic resonance imaging

MRI 是诊断腰腿痛患者是否患有腰椎间盘突出
的重要影像检查方法之一, 由于具有较高的空间分
辨率和组织分辨率, 在临床的应用越来越广泛, 但经
常由于观察者经验的不同而导致评价结果的不一
致。笔者通过回顾性分析, 采用 Kappa 统计量评价
同一观察者和不同观察者在有及无临床资料的况
况下诊断腰椎间盘突出的一致性。

资料与方法

1. 病例资料: 回顾性分析北京市海淀医院
2004 至 2005 年 100 例因腰腿痛进行腰椎 MR 检查
的患者, 其中男 48 例, 女 52 例; 年龄 17 ~86 岁, 平
均 61 岁。全部患者均有腰骶部放射性疼痛, 临床怀
疑 L3 ~4、L4 ~5、L5 ~S1 椎间盘病变, 共分析 300 个
椎间盘膨出、突出及神经根受压情况。

2. MR 检查方法: 患者取仰卧位, 采用 GE Signa
3.0 T MR 扫描仪, 使用脊柱表面线圈和快速 SE
(FSE) 序列, 常规扫描腰椎矢状面 T₂WI (TR
3000 ms, TE 109 ms, 4 次采集)、T₁WI (TR 500 ms,
TE 9.3 ms, 2 次采集), 层厚 3 mm, 层间距 1 mm, 矩
阵 448 × 224, 视野 (FOV) 28 cm × 28 cm; 横断面
T₂WI TR 4000 ms, TE 104 ms, 4 次采集, 层厚 5 mm,
层间距 1 mm, 矩阵 384 × 224, FOV 20 cm × 20 cm。

3. 资料分析方法: 由甲、乙 2 名有经验的放射科
医师分别对其中各 50 例患者的 150 个腰椎间盘在有
及无临床资料的情况下进行 2 次评价, 再由该 2 名医
师在无临床资料的情况下共同对其中 52 例患者的
156 个腰椎间盘进行评价。腰椎间盘病变 MRI 诊断
以横断面图像为主并结合矢状面图像。腰椎间盘突
出的诊断分为膨出、突出(包括凸出、脱出及椎间盘游
离)。膨出的诊断标准: 纤维环向四周对称、弥漫地超
出相邻椎体边缘。突出的诊断标准: 椎间盘局限性非
对称超出椎间隙边缘, 与原椎间盘的连接部宽于凸出
部的任一径线长度为凸出; 若椎间盘凸出的程度更
大, 与原椎间盘相连部分的长度小于突出髓核的直径
或与原椎间盘无相连称为脱出; 若移位的椎间盘与原
椎间盘失去联系后称为椎间盘游离^[1-2]。对每个椎间
盘按正常、膨出、突出分别评价并记录, 同时对神经根
受压情况进行评价和记录。

4. 统计学方法: 应用 SPSS 10.0 软件。(1) 对
甲、乙 2 名医师在有及无临床资料情况下的诊断结

果一致性分别进行 Kappa 检验; (2) 甲、乙 2 名医师
对相同病例的诊断结果一致性进行 Kappa 检验, 检
验水准为 $\alpha = 0.05$ 。Kappa 值的一致性参考评价原
则为: 小于 0 为极差, 0 ~0.20 为微弱, 0.21 ~
0.40 为弱, 0.41 ~0.60 为中度, 0.61 ~0.80 为显
著, 0.81 ~1.00 为极好^[3-4]。诊断符合率为 2 次诊
断中相符合的椎间盘数除以总椎间盘数, 诊断不符
合率为 2 次诊断不相符合的椎间盘数除以总椎间盘
数, 其中可详细分为 3 种情况, 分别为椎间盘的膨出
与正常、膨出与突出及正常与突出, 分别为 2 次诊断
中不相符合的椎间盘个数相加除以总的 2 次诊断中
不相符合的椎间盘数。

结 果

1. 甲、乙 2 名医师在有及无临床资料情况下的
诊断结果: 甲、乙 2 名放射科医师各自对 150 个椎间
盘在有及无临床资料情况下的诊断结果见表 1, 2。
甲、乙 2 名医师共同对其中 52 例患者的 156 个椎间
盘在无临床资料情况下的诊断结果见表 3。甲乙 2 名
放射科医师分别对 150 个腰椎间盘 2 次诊断结果的
Kappa 值分别为 0.60 ±0.06 和 0.57 ±0.06, 一致性为
中度; 2 名医师在有临床资料的情况下腰椎间盘膨出
的诊断较前明显增加, 分别增加了 10 个和 31 个。
表 3 的结果显示, 2 名医师对 156 个腰椎间盘诊断结
果的 Kappa 值仅为 0.24 ±0.06, 一致性为弱。

表 1 甲医师在有及无临床资料情况下对 50 例
共 150 个腰椎间盘的诊断结果(个)

有临床资料时 的诊断结果	无临床资料时的诊断结果			合计
	正常	膨出	突出	
正常	28	5	0	33
膨出	15	66	7	88
突出	2	7	20	29
合计	45	78	27	150

注: Kappa = 0.60 ±0.06

表 2 乙医师在有及无临床资料情况下对 50 例
共 150 个腰椎间盘的诊断结果(个)

有临床资料时 的诊断结果	无临床资料时的诊断结果			合计
	正常	膨出	突出	
正常	46	3	1	50
膨出	27	45	8	80
突出	1	1	18	20
合计	74	49	27	150

注: Kappa = 0.57 ±0.06

表 3 甲、乙 2 名医师在无临床资料情况下对 52 例
共 156 个腰椎间盘的诊断结果(个)

乙医师诊断结果	甲医师诊断结果			合计
	正常	膨出	突出	
正常	22	54	1	77
膨出	2	38	15	55
突出	0	7	17	24
合计	24	99	33	156

注: Kappa = 0.24 ± 0.06

2. 甲、乙 2 名医师诊断结果一致率: 结果见表 4。在有及无临床资料的情况下, 最大的诊断差异出现在对正常和膨出的腰椎间盘的鉴别上, 2 名医师的 2 次诊断膨出的差异分别达到 20 个和 30 个, 分别占各自总不符合率的 55.6% 和 73.2%。其次是膨出和突出的椎间盘, 2 次诊断膨出的差异分别为 14 个和 9 个, 分别占各自总不符合率的 38.9% 和 21.9%。2 次差异最少的是正常和突出的腰椎间盘, 均为 2 个, 分别占各自总不符合率的 5.6% 和 4.9%。2 名医师对 156 个腰椎间盘诊断结果的最大差异也出现在对膨出和正常的鉴别上, 诊断膨出的差异为 56 个, 为总不符合率的 70.9% (图 1~4), 其次是膨出和突出的差异, 为 22 个, 为总不符合率的 27.8%, 而突出和正常的差异仅为 1 个, 为总不符合率的 1.3%。

讨 论

自从 1960 年 Kappa 统计量作为评价判断一致性程度的指标, 在疾病流行病学调查和临床试验等方面都已经得到广泛的应用, 在国外被普遍应用于临床研究, 其中对影像科医师诊断结果一致性的评价受到特别的重视, Kappa 检验在国内的临床应用较少, 但近年来也受到越来越多临床工作者的关注^[4-5]。Kappa 统计量是比较同一放射科医师对同一观察对象, 进行 2 次以上的观察, 作出各次判断的结果是否一致; 2 个或多个放射科医师对同一观察对象进行某指征的观察判断的一致性。本研究应用 Kappa 统计量对同一放射科医师在有及无临床资料的情况下 2 次诊断结果的一致性和 2 名放射科医师

图 1, 2 女, 47 岁, 左下肢放射性疼痛。甲医师在有临床资料情况下诊断为 L5~S1 椎间盘膨出, 在没有临床资料情况下, 甲、乙医师均诊断为正常椎间盘(箭) 图 3, 4 男, 71 岁, 右侧腰骶部放射性疼痛。乙医师在有临床资料情况下诊断为 L5~S1 椎间盘膨出, 在没有临床资料的情况下, 乙医师诊断为正常腰椎间盘, 甲医师诊断为腰椎间盘膨出(箭)

在没有临床资料的情况下对相同病例诊断结果的一致性进行评价。结果显示, 应用 MRI 进行腰椎间盘突出诊断的一致性不高, 分别为中度和弱; 而造成诊断结果一致性不高的最主要因素是对腰椎间盘膨出的判断。

MRI 普遍被认为是诊断腰椎间盘突出较为准确的方法^[6-7], 但从本研究的统计结果来看, 在对怀疑有腰椎间盘突出病变的患者进行诊断时, 即使是同一放射科医师对相同病例在有及无临床资料的情况下 2 次诊断结果的符合率分别只有 76.0% 和 72.7%, 而 2 名放射科医师在没有临床资料的情况下对相同病例的诊断符合率只有 49.4%。造成同一医师 2 次诊断和 2 名医师对同一患者诊断结果不一致

表 4 甲、乙 2 名医师诊断结果的一致率(个)

诊断结果	符合			不符合					
	总椎体数	符合椎体数	符合率(%)	膨出与正常	不符合率(%)	膨出与突出	不符合率(%)	突出与正常	不符合率(%)
甲医师 2 次诊断结果	150	114	76.0	20	55.6	14	38.9	2	5.6
乙医师 2 次诊断结果	150	109	72.7	30	73.2	9	21.9	2	4.9
甲、乙医师诊断结果	156	77	49.4	56	70.9	22	27.8	1	1.3

的原因是对腰椎间盘突出认识,而其中最主要的原因是对正常和膨出腰椎间盘的鉴别,本组研究结果与 Van Rijn 等^[8]的研究结果基本一致。但是在本研究中发现同一放射科医师在有临床资料的情况下对腰椎间盘突出诊断明显增多,分别增加了 10 和 31 个,这表明放射科医师在有临床资料的情况下更多地将在没有临床资料情况下认为正常的腰椎间盘突出诊断为膨出,从而更多地将在临床症状的原因归结为腰椎间盘突出,也可以认为在有临床资料的情况下,放射科医师对腰椎间盘突出诊断标准降低了,而在腰椎间盘突出诊断上没有这种明显的倾向。

腰椎间盘突出可以压迫神经根而引起腰腿疼的临床症状,也可以见于没有临床症状的退行性改变的人群中;同时在临床治疗中没有症状的腰椎间盘突出也通常被认为非手术适应证。腰椎 MRI 的观察指标主要包括椎间隙高度的变化、髓核信号的减低、椎间盘形态的改变、纤维环信号的变化和纤维环后方高信号带;其中椎间盘的形态改变是判断椎间盘正常、膨出和突出最主要的指标。在临床工作中判断椎间盘膨出的标准非常容易导致诊断结果的不一致,虽然在本研究中观察的放射科医师和病例数量较少,但也能够反映出在实际工作中放射科医师在判断椎间盘形态所出现的观察者间的不一致性较大。

综上所述,虽然 MRI 给临床工作者诊断腰椎间盘突出病变提供了很多信息,但观察者间的诊断结果不一致现象客观存在,而这种不一致很大程度上是由于腰椎间盘突出定义和诊断标准所造成的,探讨一个更恰当和准确的腰椎间盘突出命名和诊断标准是临床工作者所面临的一个课题。

参 考 文 献

- [1] Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med*, 1994, 331: 69-73.
- [2] Milette PC, Fontaine S, Lepanto L, et al. Differentiating lumbar disc protrusions, disk bulges, and discs with normal normal contour but abnormal signal intensity. *Spine*, 1999, 24: 44-53.
- [3] 李春波, 何燕玲, 张明园. 一致性检验方法的合理应用. *上海精神医学*, 2000, 12: 228-232.
- [4] 方积乾, 陆盈. *现代医学统计学*. 北京: 人民卫生出版社, 2002, 80-86.
- [5] 郑卓肇, 孔忠强, 范家栋, 等. Kappa 统计量评价半月板的 MRI 诊断. *中国医学影像技术*, 2002, 18: 587-588.
- [6] 张燕, 余卫. 腰椎间盘突出形态学分型及 MRI 的诊断价值. *国外医学临床放射学分册*, 2002, 25: 161-164.
- [7] 屈辉, 孙楠, 程晓光. 腰椎间盘突出症的影像学诊断. *中国临床医学影像杂志*, 2002, 13: 418-422.
- [8] Van Rijn JC, Klemetso N, Reitsma JB, et al. Observer variation in MRI evaluation of patients suspected of lumbar disk herniation. *AJR*, 2005, 184: 299-303.

(收稿日期: 2007-12-24)

(本文编辑: 隋行芳)

第三届全国泌尿系统影像诊断与技术应用研讨会征文通知

由中华放射学杂志编辑委员会、山东省医学影像学研究所、山东省医学影像学研究会共同举办的“第三届全国泌尿系统影像诊断与技术应用研讨会”, 将于 2008 年 7 月 31 日至 8 月 4 日在湖北省宜昌市召开。大会将邀请知名医学影像专家作学术报告。欢迎大家积极撰写学术论文, 参会交流, 参加大会者将获得国家级继续教育学分。

1. 征文范围: 凡未正式公开发表过的关于泌尿系统影像诊断(包括普通放射诊断、CT、MRI、DSA、超声和核医学), 影像技术学, 护理学等方面的论著、实验研究、综述、短篇报道和经验介绍, 文章内容应体现真实性、实用性和科学性。

2. 征文要求: 来稿请寄 4000 字左右全文及 800 字的摘要各一份, 摘要请按“目的、方法、结果、结论”四部分撰写。文章务必 Word 格式打印, (请附磁盘和打印稿, 也可发 Email)。文章需附作者简介, 写于首页左下方, 内容包括: 姓

名、出生年月、性别、民族、籍贯、学历学位、职称和研究方向, 联系电话。作者请自留底稿, 概不退稿。来稿请注明“第三届全国泌尿系统影像诊断与技术应用研讨会”字样, 寄至: 100710 北京市东四西大街 42 号中华放射学杂志编辑部张晓冬收, 电话: 010-85158292, Email: zhangxd@cma.org.cn 或 cjr.zhangxd@vip.163.com。优秀论文经有关专家审阅通过后, 可优先在《中华放射学杂志》或《医学影像学杂志》上发表。所有投稿, 经专家评审通过的将收录在论文汇编中。

3. 截稿日期: 2008 年 6 月 20 日, 以当地邮戳为准。会议具体时间、地点见第 2 轮通知。

(中华放射学杂志编辑委员会
山东省医学影像学研究所
山东省医学影像学研究会)