

特发性脊柱侧凸 PUMC 分型的可重复性及可信度研究

朱 锋,邱 勇,朱泽章,王 斌,李卫国,王渭军

(南京大学医学院附属鼓楼医院脊柱外科 210008 南京市)

【摘要】目的:探讨特发性脊柱侧凸 PUMC(协和)分型系统的一致性,并对影响分型一致性的因素进行分析。**方法:**随机选取南京鼓楼医院 2004 年~2006 年治疗的 80 例特发性脊柱侧凸患者,其中男性 15 例,女性 65 例,平均年龄 14.6 岁(10~18 岁),均有完整的术前站立位正侧位 X 线片及仰卧位左右 Bending 片共四张。由 4 名脊柱外科专科医生分别根据 PUMC(协和)分型标准进行分型,2 周后此 4 位医生分别对这些患者的 X 线片再次进行分型,收集分型结果分别作可信度和可重复性分析,计算 Kappa 值检验一致性并对影响分型一致性的因素进行统计分析。**结果:**80 例患者均可用 PUMC 分型,分型可信度平均 Kappa 值为 0.801,可重复性 Kappa 值为 0.878。PUMC 分型不一致中包括上胸弯的界定 18 次;胸弯明显时,代偿性与结构性腰弯的界定 15 次;腰弯明显时,代偿性与结构性胸弯的界定 18 次;单弯顶点的判断 20 次;测量角度的差异 18 次。**结论:**PUMC(协和)分型学习曲线相对较短,易于掌握且具有良好的可信度和可重复性。

【关键词】 特发性脊柱侧凸;分型;可信度;可重复性

中图分类号:R682.3 文献标识码:A 文章编号:1004-406X(2008)-12-0915-05

Prospective study of interobserver and intraobserver repeatability and reliability of PUMC classification for idiopathic scoliosis/ZHU Feng, QIU Yong, ZHU Zezhang, et al//Chinese Journal of Spine and Spinal Cord, 2008, 18(12):915-919

【Abstract】 Objective: To determine the interobserver and intraobserver repeatability and reliability of PUMC classification for idiopathic scoliosis and review the factors correlating to this. **Method:** Eighty patients with idiopathic scoliosis treated in Drum Tower hospital from 2004 to 2006 were randomly chosen for study, which included 15 males and 65 females with a mean age of 14.6 years old (range, 10 to 18 years old). Pre-operative standing AP and lateral view X-ray as well as supine lateral bending were taken in all patients. Four spine surgeons independently assign the curve type to each patients following the guidance described by Qiu Guixing et al. The procedure was repeated two weeks later with the patients presented in different order. Kappa value was used to determine the intraobserver and interobserver repeatability and reliability. **Result:** All curves could be assigned into the PUMC classification system. A mean kappa value of 0.801 for interobserver reliability were determined, and a mean kappa of 0.878 for intraobserver repeatability. The disagreement evidenced included the determination of the upper thoracic curve (18 times), distinguishment between structural and non-structural lumbar curve secondary to thoracic major curve (15 times), distinguishment between structural and nonstructural thoracic curve secondary to lumbar major curve (18 times), determination of apic vertebral from single curve (18 times) and angular variation due to each curve. **Conclusion:** PUMC new classification has relative short-time learning curve and is more reliable and repeatability for practical application.

【Key words】 Idiopathic scoliosis; Classification; Interobserver reliability; Intraobserver reliability

【Author's address】 Spinal Surgery, Drum Tower Hospital, Nanjing University Medical School, Nanjing, 210008, China

课题资助:江苏省自然科学基金,项目编号 BK2007003

第一作者简介:男(1977-),主治医师,医学博士,研究方向:脊柱外科

电话(025)83105113 E-mail:cnspine@hotmail.com

青少年特发性脊柱侧凸(adolescent idiopathic scoliosis, AIS)的侧凸类型非常繁杂,其手术策略的制定也是一项系统工程,其中术前对患者的评估及分型显得尤为重要^[1]。一个理想的分型应

当具备以下要素^[1-3]:几乎包含所有的侧凸类型;对脊柱侧凸畸形进行包括矢状面和轴状面的三维评估;有良好的可信度(interobserver reliability)和可重复性(intraobserver reproducibility),易用、易记;最终提出每一种分型的手术策略和融合范围。1983 年出现的 King 分型^[4]是基于第一代 Harrington 和 Luque 矫形装置,属于单平面分型系统,随着第三代脊柱内固定装置的使用,King 分型越来越显出其不足。近年出现的 Coonrad、Lenke 等分型系统^[5,6]又过于复杂。2003 年由邱贵兴报道了自主创新的 AIS PUMC(协和)分型系统^[2],目前除了 PUMC 开发者对其可信度和可重复性研究外,仅有我科的一篇报道^[7]。本文拟在使用 PUMC 分型 2 年后较为熟练的基础上再对这一分型进行可信度和可重复性研究,以分析该系统的学习曲线。

1 资料与方法

1.1 临床资料

病例全部来源于南京鼓楼医院脊柱外科,为 2005 年~2006 年连续收治的 80 例 AIS 患者,入选标准:年龄 10~18 岁,Cobb 角 40°~75°,临床诊断为特发性脊柱侧凸。排除标准为:影像学有明显的先天性脊柱发育畸形,MRI 提示神经系统发育异常,临床查体有任何感觉、运动、其他神经系统或结缔组织异常,肿瘤、结缔组织或代谢性疾病导致的脊柱侧凸畸形。抽取结果为男性 15 例,女性 65 例,平均年龄 14.6 岁(10~18 岁)。术前主弯 Cobb 角 40°~75°,平均 54°,Risser 征 1~4 级,距月经初潮 0~72 个月。手术方法包括单纯后路矫形内固定植骨融合术、前路矫形内固定植骨融合术。每例患者均有完整的术前 X 线片资料,包括术前三站立位全脊柱正侧位及仰卧位左右 Bending 片。

1.2 方法

根据北京协和医院骨科发表的 PUMC(协和)分型标准^[2,3]对 80 例患者进行分型。PUMC 分型:该分型系统根据顶点多少将侧凸分为三型,1 个顶点为 I 型,2 个顶点为 II 型,3 个顶点为 III 型。同时,每型中再分不同的亚型,共计 13 个亚型:I a(单胸弯)、I b(胸腰弯)、I c(腰弯)、II a(双胸弯)、II b1 和 II b2(胸弯>胸腰弯/腰弯 10°以上)、II c1 和 II c2(胸弯≈胸腰弯/腰弯,即二者 Cobb 角差小于 10°)、II d1 和 II d2(胸弯<胸腰弯/腰弯

10°以上)、III a 和 III b。

手术前由 4 名脊柱外科专科医生分别对 80 例患者进行分型,该 4 名医生对侧凸的 SRS(脊柱侧凸研究协会)定义以及 PUMC 各类型已有详细了解,并且使用 PUMC 分型系统 2 年以上,每位医生阅片分型超过 200 例。4 名医生对 80 例 AIS 患者分别分型,收集结果作观察者间可信度分析。2 周后将患者顺序打乱,再由这 4 位医生对相同的 80 例患者进行分型,收集结果作分型的可重复性分析。数据分析采用一致性检验,计算 Kappa 值。在做一致性比较时,共需要对比 800 次,其中可信度比较 480 次,可重复性比较 320 次,为了详细分析何种因素影响分型一致性,特列出各种因素出现不一致的次数。

2 结果

80 例患者均可用 PUMC 分型,分型可信度平均 82.9%,平均 Kappa 值为 0.801,可重复性为平均 90.2%,Kappa 值平均 0.878(表 1、2)。

影响分型一致性的主要因素依次为:(1)上胸弯的界定(I a 还是 II a, III a 还是 II),出现 18 次不一致;(2)胸弯明显时,代偿性与结构性腰弯(I a 还是 II b1),出现 15 次不一致(图 1);(3)腰弯明显时,代偿性与结构性胸弯(II d1 还是 I c 或 I

表 1 4 位医师对 80 例特发性脊柱侧凸患者 PUMC 分型的可信度分析

医师	分型一致的例数	分型一致的百分比(%)	Kappa值
1-2	61	77.5	0.726
1-3	68	85.0	0.827
1-4	75	93.8	0.928
2-3	60	75.0	0.712
2-4	60	75.0	0.712
3-4	73	91.3	0.899
平均	66.1	82.9	0.801

表 2 4 位医师对 80 例特发性脊柱侧凸患者 PUMC 分型的可重复性分析(第一次与第二次)

医师	分型一致的例数	分型一致的百分比(%)	Kappa值
1	71	88.1	0.871
2	72	90.0	0.884
3	73	91.3	0.899
4	73	91.3	0.899
平均	72.25	90.2	0.878

b), 出现 18 次不一致; (4) 单弯顶点的判断, 胸腰弯与腰弯, 胸腰弯与长胸弯 (影响 I 型 a、b、c 亚型分类), 出现 20 次不一致; (5) 测量角度的差异, 主要影响 II 型 b 和 c 亚型出现 18 次不一致 (表 3, 图 2)。

表 3 不同影响因素在 PUMC 分型不一致中出现的频率 (次)

影响因素	不同医生分型不一致次数	同一医生分型不一致次数
上胸弯界定	13	5
腰弯 (结构性或代偿性)	13	2
胸弯 (结构性或代偿性)	9	9
单弯顶点	18	2
角度差异	11	7

3 讨论

AIS 的类型纷繁复杂, 科学的术前评估和合理的治疗策略制定显得尤为重要, 而脊柱侧凸分

型系统正是对 AIS 侧凸规律的归纳与总结, 优秀的分型系统应当具备下列五个特征: (1) 通用性, 包含所有的 AIS 侧凸类型; (2) 合理性, 简单、易记、临床可操作性强; (3) 临床指导性, 为每一分型提出治疗方案以及融合节段选择; (4) 先进性, 能够三维评估, 符合第三代三维矫形理念; (5) 一致性, 必须有良好的可信度与可重复性。

1983 年出现的 King 分型是特发性脊柱侧凸的经典分型, 其分型依据是侧凸的部位、顶椎、侧凸严重程度和柔韧性, 重要的是 King 提出的稳定椎的概念和选择性融合^[4]曾经是特发性脊柱侧凸的分型“金标准”, 但它仅适用于胸椎侧凸, 为单平面分型系统, 没有考虑矢状面形态和横断面旋转; 对新型三维矫形系统的融合节段选择指导意义不大。而且 King 分型的一致性并不理想, Hannes 等^[8]的研究发现 King 分型可信度 Kappa 值为 0.46, 一致性百分比 57%, 可重复性 Kappa 值 0.79, 一致性百分比 80%。Kappa 值为真正的观察值一致减



图 1 女, 16 岁, 特发性脊柱侧凸 a 正位 X 线片示胸弯 48°, 腰弯 35° b 侧位 X 线片示胸椎后凸 (T5~T12) 0° c 胸弯 Bending 相为 20° d 腰弯 Bending 相为 10° (正确分型应该为 PUMC II b1, 但如果忽略代偿性腰弯, 容易分型为单胸弯 PUMC I a) 图 2 女, 14 岁, 特发性脊柱侧凸 a 正位 X 线片示胸弯 50°, 腰弯 45° b 侧位 X 线片示胸椎后凸 (T5~T12) 15° c 胸弯 Bending 相为 15° d 腰弯 Bending 相为 4° (因为对主胸弯和主腰弯 Cobb 角测量的差异, 分型结果分别为 PUMC II c3 和 PUMC II b1)

去偶然因素造成的结果一致,因此 Kappa 值在 $[-1, +1]$ 之间,当 Kappa 系数为 0 时,表明结果的一致纯属偶然, Kappa 值为 1 时表明完全一致,因此 Kappa 值越大,一致性越好。根据 Landis 和 Koch 的推荐^[9], Kappa 值在 0.61~0.80 表明一致性较高 (substantial agreement), Kappa 值在 0.81~1.00 表明几乎完全一致 (almost perfect agreement)。而 Svanholm 定义 Kappa 值 ≥ 0.75 时一致性很好 (excellent)^[10]。

2001 年 Lenke 提出了新的分型系统^[6]包括 6 种侧凸分型 (I~VI), 3 个腰椎修正 (ABC) 和 3 个矢状面修正 (+N-), 因而共有 42 种亚型, 虽然比较全面, 继续包括了侧凸所有类型, 区分了结构性与非结构性弯曲, 并考虑了胸椎矢状面畸形, 但亚型过多记忆较困难, 也不利于临床医生之间交流。Lenke 本人报道的分型总体可信度与可重复性分别为 92% 和 83%^[11]。但 Ogon 等 (非 Lenke 分型开发者) 的研究发现, Lenke 分型总体可信度 Kappa 值为 0.62, 总体一致率仅为 41%, 可重复性 Kappa 值为 0.75^[12]。不同研究者对 Lenke 分型一致性评估差异的原因有: 观察者对分型掌握程度不一, 部分作者没有经过系统培训, 对分型理解存在差异; 腰椎修正的判定相对主观性较大, 容易引发争议; 患者的选取不同, Lenke 报道的 29 例患者是作者经过预先筛选, 多数为胸弯患者, 而 Ogon 的 51 例患者经过随机筛选。另外 Lenke 分型包括三个参数, 相对复杂, 虽然每个参数的 Kappa 值并不低, 侧凸分型、腰椎修正、矢状面修正的可信度 Kappa 值分别为 0.75, 0.69, 0.91; 可重复性 Kappa 值为 0.81, 0.81, 0.93, 但因为 Lenke 分型需要将三个参数综合在一起, 导致其一致性明显降低。

由北京协和医院骨科提出的 PUMC 分型系统根据顶点多少将侧凸分为三型, 1 个顶点为 I 型, 2 个顶点为 II 型, 3 个顶点为 III 型。PUMC 分型根据顶点多少来分型既符合临床上特发性脊柱侧凸的特点, 又便于记忆。同时, 每型中再分不同的亚型, 共计 13 个亚型。该分型系统还兼顾了脊柱侧凸三平面内的畸形特点, 与 Lenke 分型相比, 亚型数大大减少, 容易掌握且不失全面。本文所得出的 PUMC 分型可信度与可重复性 (82.9% 和 90.2%) 与该分型系统的开发者邱贵兴等^[12, 3]报道的 85% 与 91% 基本一致, 邱贵兴报道的可信度与可重复性 Kappa 值分别为 0.832 和 0.898, 本研究

得出的可信度与可重复性的 Kappa 值分别为 0.801 和 0.878。根据 Kappa 值评估, 两次该分型系统的可信度与可重复性均为优秀 (Kappa 值 ≥ 0.75)。并没有发现分型的可信度和可重复性在开发者与非开发者之间存在明显差异的现象。

邱勇等^[7]在使用 PUMC 分型的初期也曾对随机抽取的 100 例 AIS 患者做过可信度和可重复性的比较, 结果 PUMC 分型的可信度为 84.3%, Kappa 值为 0.819, 可重复性为 92%, 平均 Kappa 值为 0.907, 并且明显高于 King 分型和 Lenke 分型。经过 2 年以上对 PUMC 分型系统的使用和深入体会后, 我们对 80 例 AIS 患者进行分型一致性比较, 结果可信度和可重复性基本一致, 没有太大变化。参与本研究的 4 位观察者均为专业脊柱外科医师, 在接触 PUMC 分型前对特发性脊柱侧凸已有深入了解。在开展本前瞻性研究前 2 年已经开始使用 PUMC 分型, 并经过该分型开发者的专门培训, 因此四位观察者对 PUMC 分型的定义和每项指标均有深刻了解, 在本次研究中各自的观察一致性均较高 88.1%、90.0%、91.3%、91.3%。作者体会该分型根据侧弯的顶点数目分为 I、II、III 型, 一目了然便于记忆, 然后根据站立位各弯曲的部位、大小及卧位 Bending 片显示的胸弯和腰弯的柔韧性差别归属不同亚型, 针对每一亚型 PUMC 开发者设定了严格的界定标准, 因而很容易划分各侧弯归属的类型。通过本中心的前后两次分型可信度和可重复性比较, 可以发现使用者可在较短时间内掌握 PUMC 分型, 并在后续使用中始终保持了较高的一致性。设计者在开发这一分型系统时不仅考虑了冠状面畸形, 也考虑了矢状面与轴状面的畸形, 如胸腰段的交界性后凸、腰弯的旋转度等, 这些均是现代三维矫形融合节段选择必须考虑的因素, 因此可以满足三维矫形术前评估的要求。同时开发者根据每一亚型提出相应手术策略和融合节段, 具有很高的临床指导意义。

本研究发现因 PUMC 分型上胸弯界定导致的的不一致出现 18 次。Ogon 等^[12]发现 Lenke 对上胸弯的判断有 30% 存在争议, 是影响一致性的最大因素。而上胸弯的界定对融合节段选择非常重要。Winter^[13]建议从临床观察和影像学判断结构性胸弯的存在, 主要指标是肩部是否水平、两侧斜方肌是否对称, X 线片上 T1 是否向右侧倾斜。Lee

等^[14]建议左肩抬高的患者不考虑 T1 是否倾斜均作上胸弯融合。Lenke 等^[15]认为如测量角度大于 30°, Bending 片矫正大于 20°, 大于 I 度旋转, 左肩抬高、T1 倾斜入上胸弯凹侧要界定为结构性上胸弯并融合到 T2。Suk 等^[16]则将结构性上胸弯角度界定为 25°。笔者认为上胸弯的界定要综合考虑, 即使在 T1 水平或向左侧倾斜, 双肩水平或右肩抬高的患者也不能排除结构性上胸弯的存在, 临床外观的评估以及 Cobb 角的大小有时更为重要。

在 PUMC 分型过程中, 当出现结构性胸主弯或腰主弯时, 对代偿弯的判断有时很困难, 正如 King 分型中 II 型和 III 型的鉴别, 有时对代偿腰弯是否超过中线很难取得统一^[17]。但这种判定困难并没有太多临床意义, 因为 PUMC 分型 I a 与 II b1, II d1 与 I c 虽然分属不同亚型, 但手术策略基本一致, 如 I a 与 II b1 均可仅融合胸弯, II d1 与 I c 均仅融合腰弯。可见 PUMC 分型已经充分认识到这一问题。

分析本研究几个主要影响分型一致性的因素, 如上胸弯的界定, 代偿弯与结构弯的判定等, 这也是影响 King 分型或 Lenke 分型等所有脊柱侧凸分型系统必须面对的问题。但 PUMC 分型系统是一开放的分型系统, 随着这一分型系统的推广应用, 可以对某些指标进行更为明确的界定, 必然会进一步提高 PUMC 分型系统的科学性。

4 参考文献

- 邱贵兴. 特发性脊柱侧凸分型研究的回顾与展望[J]. 中华骨科杂志, 2003, 23(1): 10-13.
- 邱贵兴, 仇建国, 王以朋, 等. 特发性脊柱侧凸的 PUMC(协和)分型系统[J]. 中华骨科杂志, 2003, 23(1): 1-9.
- Qiu G, Zhang J, Wang Y, et al. A new operative classification of idiopathic scoliosis: a Peking Union Medical College Method [J]. Spine, 2005, 30(12): 1419-1426.
- King HA, Moe JH, Bradford DS, et al. The selection of fusion levels in thoracic idiopathic scoliosis [J]. J Bone Joint Surg Am, 1983, 65(9): 1302-1313.
- Coonrad RW, Murrell GAC, Motley G, et al. A logical coronal pattern classification of 2000 consecutive idiopathic scoliosis cases based on the Scoliosis Research Society-defined apical vertebra[J]. Spine, 1998, 23(12): 1380-1391.
- Lenke LG, Betz RR, Harms J, et al. Adolescent idiopathic scoliosis: a new classification to determine extent of spinal arthrodesis [J]. J Bone Joint Surg Am, 2001, 83-A (8): 1169-1181.
- 邱勇, 朱泽章, 朱锋, 等. 青少年特发性脊柱侧凸 King, Lenke 和 PUMC(协和)分型的可信度和可重复性比较及意义[J]. 中华骨科杂志, 2007, 27(10): 748-751.
- Hannes B, Karlmeier G, Michael O, et al. Multisurgeon assessment of coronal pattern classification systems for adolescent idiopathic scoliosis: reliability and error analysis [J]. Spine, 2002, 27(7): 762-767.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data[J]. Biometrics, 1977, 33(1): 159-174.
- Svanholm H, Starklint H, Gundersen HJG, et al. Reproducibility of histomorphologic diagnosis with special reference to the kappa statistic[J]. APMIS, 1989, 97(8): 689-698.
- Lenke LG, Betz RR, Bridwell KH, et al. Intraobserver and interobserver reliability of the classification of thoracic adolescent idiopathic scoliosis [J]. J Bone Joint Surg Am, 1998, 80(8): 1097-1106.
- Ogon M, Giesinger K, Behensky H, et al. Interobserver and intraobserver reliability of Lenke's New Scoliosis Classification System[J]. Spine, 2002, 27(8): 858-862.
- Winter RB. The idiopathic double thoracic curve pattern: its recognition and surgical management [J]. Spine, 1989, 14(12): 1287-1292.
- Lee CK, Denis F, Winter RB, et al. Analysis of the upper thoracic curve in surgically treated idiopathic scoliosis: a new concept of the double thoracic curve pattern[J]. Spine, 1993, 18(12): 1599-1608.
- Lenke LG, Bridwell KH, Baldus C, et al. Preventing decompensation in King type II curves treated with Cotrel-Dubousset instrumentation: strict guidelines for selective thoracic fusion [J]. Spine, 1992, 17(Suppl 8): 274-281.
- Suk SI, Kim WJ, Lee CS, et al. Indications of proximal thoracic curve fusion in thoracic adolescent idiopathic scoliosis: recognition and treatment of double thoracic curve pattern in adolescent idiopathic scoliosis treated with segmental instrumentation[J]. Spine, 2000, 25(18): 2342-2349.
- 邱勇. King II 特发性脊柱侧凸: 选择性胸椎融合还是非选择性融合[J]. 中国脊柱脊髓杂志, 2003, 13(5): 266.

(收稿日期: 2008-03-01 修回日期: 2008-11-03)

(英文编审 蒋欣)

(本文编辑 彭向峰)