

• 骨骼肌肉影像学 •

肌肉脂化率主观评分法的稳定性与可靠性研究

杜婧, 肖江喜, 郑艺明, 于磊, 袁云, 李飞宇

【摘要】 目的:比较 3 位阅片者肌肉脂化率主观评分结果的一致性,分析主观评分法与客观评分法的相关性,评价抗肌萎缩蛋白病中肌肉脂肪浸润主观评分法的稳定性及可靠性。**方法:**搜集本院 2012 年 1 月—2015 年 8 月抗肌萎缩蛋白病患者 194 例,行双下肢横轴面 T₁WI 及 IDEAL-Quant 序列扫描。在横轴面 T₁WI 图像中,由 3 位不同阅片者分别对所有患者下肢肌肉脂肪浸润程度进行两次主观盲法评分;在 IDEAL-Quant 脂肪图中测量得出肌肉脂化率。利用 kappa 检验分析各位阅片者自身及不同阅片者之间主观评分结果的一致性,使用 spearman 相关模型分析主观评分与 IDEAL-Quant 序列脂化率定量结果的相关性。**结果:**在对所有大腿肌肉脂肪浸润的主观半定量评估中,三位阅片者两次阅片评分结果自身比较 Kappa 值分别为 0.75、0.54、0.55;三位阅片者之间阅片评分结果 Kappa 值分别为 0.54、0.23、0.22。主观评分结果与客观定量结果相关($r=0.60\sim0.92, P<0.05$)。但主观评分估计的脂肪浸润程度高于客观定量结果。**结论:**主观评分法的结果相对稳定、可靠,但与临床经验有关;主观评分法会高估肌肉脂肪浸润程度。

【关键词】 磁共振成像; 抗肌萎缩蛋白病; 骨骼肌; 脂肪浸润

【中图分类号】 R445.2; R746.2 **【文献标识码】** A **【文章编号】** 1000-0313(2016)03-0262-05

DOI:10.13609/j.cnki.1000-0313.2016.03.015

The study of stability and reliability of subjective semi-quantitative grading method for leg muscle fatty infiltration in dystrophinopathy DU Jing, XIAO Jiang-xi, ZHENG Yi-ming, et al. Department of Radiology, Peking University First Hospital, 100034 Beijing, China

【Abstract】 Objective: To analyze the intra-observer and inter-observer agreement of subjective semi-quantitative grading scores. Then to compare subjective semi-quantitative grading method with objective fat quantification method in assessing leg muscle fatty infiltration in dystrophinopathy. **Methods:** One hundred and ninety-four boys with genetically and/or pathologically confirmed dystrophinopathy were recruited. The axial T₁-weighted images and IDEAL-Quant fat fraction maps were obtained. The axial T₁-weighted images were visually evaluated by three readers subjectively. The objective muscle fat fraction value could be automatically calculated on GE AW4.6 workstation based on fat fraction maps of IDEAL-Quant sequence. The kappa analysis was used to analyze the intra- and inter-observer agreement. The Spearman correlation was used to compare the two methods. Boxplot diagrams were used to show the differences between them. **Results:** The average intra-observer agreement values for all the muscles were 0.75, 0.54 and 0.55 respectively. The average inter-observer agreement values for all the muscles were 0.54, 0.23 and 0.22 respectively. Significant positive correlation was shown between the semi-quantitative and quantitative methods of all the muscles with the correlation coefficient varied from 0.60 to 0.92 ($P<0.05$), except for the internal obturator muscle and the adductor longus. However, the fat infiltration level evaluated by the semi-quantitative method was higher than the objective method. **Conclusion:** Subjective semi-quantitative grading method was reliable and stable in evaluating leg muscle fatty infiltration in dystrophinopathy. The Stability was correlated with observer's clinical experience. But the subjective semi-quantitative grading method can overestimate the fatty infiltration level.

【Key words】 Magnetic resonance imaging; Dystrophinopathy; Skeletal muscle; Fat infiltration

在各类神经肌肉病中,随着疾病进展,均会出现不同程度的骨骼肌脂肪浸润^[1-2]。对于不同的疾病种类,肌肉脂肪浸润出现的时间及肌肉的受累顺序存在一定差异,因此,对肌肉脂肪浸润程度的评估,有助于临床判断疾病类型、分析疾病进展程度、观察疗效及预后^[3-5]。对于肌肉脂肪浸润程度的观察方法以骨骼肌

MRI 检查为佳,且近年来得到大力推广^[6-9]。目前,临床上对骨骼肌脂肪浸润程度的判断主要依赖主观半定量评分法,此方法简便快捷,然而,此种判定方法的稳定性及可靠性尚待进一步研究。

材料与方法

1. 一般资料

搜集 2012 年 1 月—2015 年 8 月在本院神经内科就诊,经基因检查和(或)病理学检查证实为抗肌萎缩

作者单位: 100034 北京,北京大学第一医院医学影像科(杜婧、肖江喜、于磊、李飞宇),神经内科(郑艺明、袁云)

作者简介: 杜婧(1988—),女,山西临汾人,博士研究生,住院医师,主要从事 MR 常规序列及功能序列在神经肌肉疾病中的应用研究。

通讯作者: 肖江喜, E-mail: cjr_xiaojiangxi@vip.163.com

蛋白病(Dystrophinopathy)的患者 194 例,均为男性,年龄 2~38 岁,中位年龄 7 岁。其中杜氏肌营养不良(Duchenne's Muscular Dystrophy, DMD)患者 173 例(89.2%),贝克肌营养不良(Becker's Muscular Dystrophy, BMD)患者 21 例(10.8%)。该研究经过伦理审批,所有患者检查前均获得监护人或患者本人的同意并签署知情同意书。

2. 仪器与方法

采用 GE Signa Excite 3.0T MR 及 GE HD750 3.0T MR 扫描仪,使用 8 通道或 32 通道体线圈。进行 MR 常规横轴面 T_1 WI 及 IDEAL-Quant 序列扫描。横轴面 T_1 WI 序列参数:TR 590~740 ms, TE 9~10 ms,ETL 2,层厚 5~7 mm,层间距 1~2 mm,矩阵 512×512 ,视野 28~42 cm, NEX 2。IDEAL-Quant 序列参数:TR 6.3 ms, TE 1 ms, ETL 6,层厚 5~7 mm,矩阵 160×160 ,翻转角 3° ,接收带宽 111.11 kHz,视野 32~40 cm。扫描时长约 5 min。扫描范围从髌前上棘至双膝关节水平。部分患儿检查前需口服水合氯醛镇静。194 例患者均完成横轴面 T_1 WI 扫描,92 例患者完成 IDEAL-Quant 扫描。

3. 图像分析

主观评分:邀请 3 位阅片者,第一位阅片者 D 从事神经肌肉病 MR 研究 5 年,第二位阅片者 X 从事神经肌肉病 MR 研究 2 年,第三位阅片者 Y 从事 MR 工作 10 年,但未参加过神经肌肉病 MR 研究。参照 Mercuri 等^[7]和 Kim 等^[8]制定的肌肉脂化率主观评分标准对 3 位阅片者进行培训:0 分,正常;1 分,肌肉内散在脂肪浸润;2 分,肌肉内小片状脂肪浸润,累及面积 $<30\%$;3 分,肌肉脂肪浸润面积 $30\% \sim 60\%$;4 分,肌肉脂肪浸润面积 $>60\%$ (图 1)。培训完成后,隐去患者全部临床信息,由 3 位阅片者分别对横轴面 T_1 WI 图像中的肌肉脂肪浸润程度进行主观评分。评分均选取右侧盆部及大腿肌肉(共 19 块),于各肌肉肌腹最大层面进行评分。每次阅片时患者顺序均为随机排列。每位阅片者均进行两次阅片,以 D1/D2, X1/X2, Y1/Y2 表示,两次阅片时间间隔大于 1 周。比较各位阅片者两次阅片评分结果的一致性。并比较 3 位

阅片者之间阅片评分结果的一致性。选取两次阅片自身结果一致性最高者的主观评分结果与客观定量结果比较(图 2)。

IDEAL-Quant 图像处理:采用 GE ADW4.6 工作站,于 IDEAL-Quant 序列的脂相图中,选取与肉眼评分相同的层面,手工对所有肌肉分别圈画兴趣区(region of interest, ROI),ROI 沿肌肉边缘进行圈画,由工作站自动计算得出相应 ROI 内的脂肪百分数(fat fraction, FF),见图 3。

4. 统计学分析

以 SPSS 21.0 软件进行统计学分析。使用 Kappa 分析对各位阅片者前后两次阅片结果(D1 与 D2、X1 与 X2、Y1 与 Y2)的一致性进行检验,并对各位阅片者第一次阅片评分结果之间(D1 与 X1、D1 与 Y1、X1 与 Y1)的一致性进行检验。使用 Spearman 相关分析模型研究 D1 评分结果与客观定量结果的相关性,并用箱形图表示两者关系。

结 果

1. 各位阅片者前后两次阅片主观评分的一致性结果

阅片者 D 前后两次评分结果的一致性高(Kappa 值为 0.62~0.86, $P < 0.05$)。阅片者 X 两次评分结果存在中等程度的一致性(Kappa 值为 0.43~0.69, $P < 0.05$;但臀小肌及闭孔内肌 Kappa 值低,分别为 0.37、0.24, $P < 0.05$)。阅片者 Y 两次评分结果存在中等程度的一致性(Kappa 值为 0.39~0.68, $P < 0.05$,表 1)。

2. 各位阅片者之间的主观评分一致性结果

阅片者 D 与 X 评分结果存在中等程度一致性(Kappa 值 0.36~0.75, $P < 0.05$);阅片者 D 与 Y 评分结果的一致性差(Kappa 值 0.12~0.53, $P < 0.05$);阅片者 D 与 Y 评分结果的一致性差(Kappa 值 0.12~0.49, $P < 0.05$,表 2)。

主观评分结果与客观定量结果的差异见图 3、4。阅片者 D 主观评分结果与客观定量结果有显著相关性($r = 0.75 \sim 0.92$, $P < 0.05$,表 3),但长收肌、缝匠肌

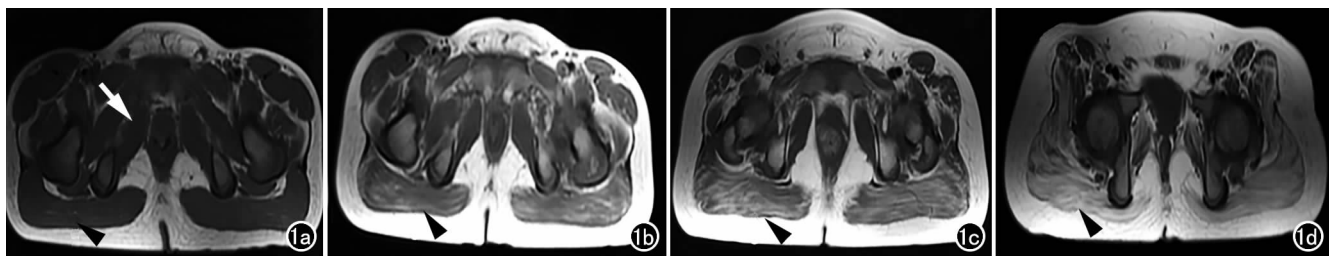


图 1 主观评分参考标准。a) 右侧闭孔内肌 0 分(箭),右侧臀大肌 1 分(箭头); b) 右侧臀大肌 2 分(箭头); c) 右侧臀大肌 3 分(箭头); d) 右侧臀大肌 4 分(箭头)。

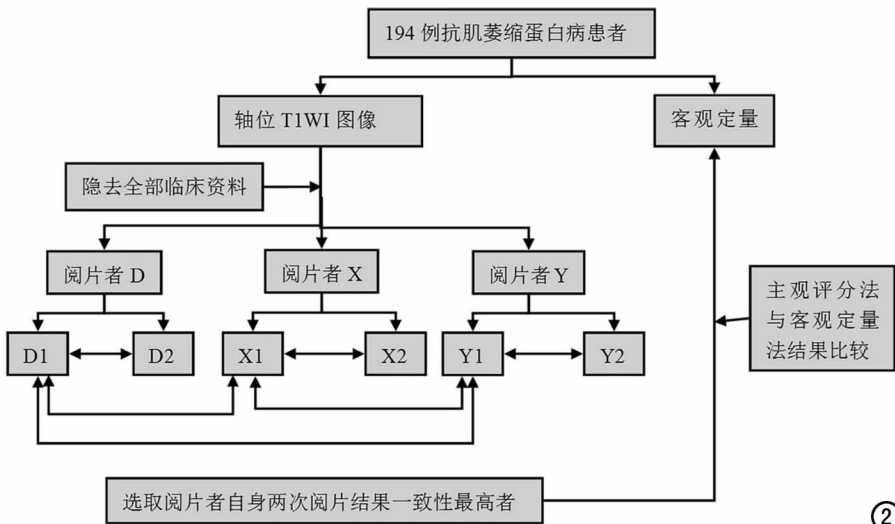


图 2 主观评分法研究流程图。

表 1 各位阅片者自身两次主观评分结果的一致性

名称	GMA	GME	GMI	AM	AL	AB	VL	RF	VM	VI	SM	ST	SA	GR	LBF	IO	OE	PE
肌肉数目	190	93	94	194	194	184	194	194	194	194	193	193	194	194	194	190	165	165
Kappa(D1/D2)	0.71	0.80	0.79	0.86	0.80	0.83	0.75	0.78	0.76	0.74	0.69	0.62	0.74	0.80	0.77	0.65	0.77	0.71
Kappa(X1/X2)	0.54	0.58	0.37	0.69	0.49	0.50	0.67	0.61	0.64	0.63	0.56	0.43	0.56	0.51	0.60	0.65	0.24	0.43
Kappa(Y1/Y2)	0.56	0.56	0.46	0.60	0.43	0.50	0.61	0.65	0.68	0.62	0.61	0.52	0.67	0.27	0.61	0.60	0.47	0.39

注:GMA 臀大肌,GME 臀中肌,GMI 臀小肌,AM 大收肌,AL 长收肌,AB 短收肌,VL 股外侧肌,RF 股直肌,VM 股中间肌,VI 股内侧肌,SM 半膜肌,ST 半腱肌,SA 缝匠肌,GR 股薄肌,LBF 股二头肌长头,IO 股二头肌短头,OE 闭孔内肌,PE 闭孔外肌。表 2、3 中省略词意义同表 1。

表 2 各位阅片者之间主观评分结果的一致性

名称	GMA	GME	GMI	AM	AL	AB	VL	RF	VM	VI	SM	ST	SA	GR	LBF	IO	OE	PE
肌肉数目	190	93	94	194	194	184	194	194	194	194	193	193	194	194	194	190	165	165
Kappa(D1/X1)	0.63	0.64	0.42	0.75	0.56	0.54	0.62	0.69	0.56	0.54	0.42	0.52	0.59	0.36	0.54	0.56	0.58	0.61
Kappa(D1/Y1)	0.35	0.22	0.15	0.53	0.11	0.10	0.49	0.25	0.29	0.30	0.23	0.15	0.12	0.12	0.23	0.16	0.03*	0.11
Kappa(X1/Y1)	0.43	0.42	0.21	0.49	0.09	0.12	0.28	0.24	0.25	0.23	0.15	0.03	0.19	0.16	0.27	0.21	0.02*	0.07

注: * 无统计学意义($P>0.05$),其他结果均有统计学意义。

表 3 主观评分结果与客观定量结果的相关性

名称	GMA	VL	RF	VM	VI	SM	ST	SA	GR	LBF	AM	AL	AB	OE	PE
r 值	0.92	0.91	0.79	0.89	0.89	0.84	0.73	0.75	0.58	0.89	0.84	0.47	0.60	0.37	0.80

及闭孔内肌的主观评分结果与客观定量结果呈中等程度相关(r 分别为 0.47、0.58、0.37, $P<0.05$)。

讨 论

抗肌萎缩蛋白病是最常见的神经肌肉病类型,由抗肌萎缩蛋白(Dystrophin)基因突变导致的一种 X-染色体连锁隐性遗传性骨骼肌疾病,包括 DMD 及 BMD 两种类型^[2,10]。DMD 及 BMD 中,不同肌肉受累严重程度存在差异,且肌肉受累严重程度与年龄相关^[4,6]。因此选择大样本抗肌萎缩蛋白病作为主观评分法的研究资料,可覆盖到各肌肉组的各种不同受累程度,在不同情况下全面评估主观评分法的可靠性及稳定性。抗肌萎缩蛋白病中双侧大腿肌肉均可见受累,双侧肌肉受累程度对称^[7],因此均选取右侧肌肉作为观察对象,保持各位评分者的观察对象一致,避免偏倚,且既往也有文献只选择右侧肌肉作为研究对象^[8]。抗肌萎缩蛋

白病中肌肉为弥漫受累,肌腹及肌腱均可见受累,受累程度较为均匀。因为肌腹最大层面的面积较大,利于观察;且各肌肉的肌腹最大层面均在扫描野内,部分肌肉的肌腱并不一定都在扫描范围内,无法评价;客观定量中,肌腹最大层面的面积较大,各肌肉解剖结构容易区分,利于 ROI 圈画,可以较好的避免皮下脂肪的干扰;因此选取各肌肉肌腹最大层面进行主观评分及客观定量。

② 该研究中,阅片者 D 前后两次评分结果的一致性高,阅片者 X、Y 两次评分的自身结果比较,

结果存在中等程度的一致性,提示阅片者自身主观评分结果较稳定,且神经肌肉病诊断经验丰富的阅片者主观评分结果更稳定。通过对不同阅片者之间的主观评分结果的比较,发现评分一致性差,说明不同阅片者之间主观评分的结果不稳定。Alizai 等^[11]在绝经后妇女小腿肌肉脂肪浸润程度的主观评分研究中得出,阅片者自身、阅片者之间的一致性均较高,Kappa 值分别为 0.83、0.81,与本研究结果存在差异。可能是因为 Alizai 等^[11]选择的研究对象肌肉受累程度轻,大部分为 1 分或 2 分,肌肉脂肪浸润程度差异小,且病例数较少,因此一致性高于本研究。

本研究在评价主观评分结果的可靠性中,使用了一种新的脂肪定量技术(IDEAL-Quant)的脂肪定量结果作为参考标准。IDEAL-Quant 是一种基于化学位移原理来进行水-脂分离的脂肪定量技术^[12],目前多用于脂肪肝、椎体的脂肪定量研究。既往文献报道,

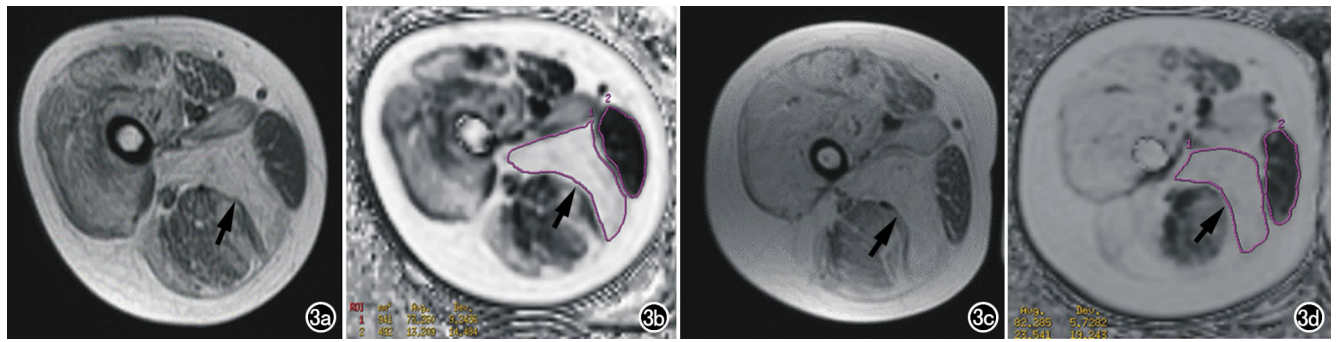


图 3 IDEAL-Quant 图像处理方法。图 a、b 为同一患者,图 c、d 为另一患者。a) T₁ WI 横轴面图像,右侧大收肌主观评分为 4 分(箭); b) 为同一层面 IDEAL-Quant 图像处理方法,圈画 ROI 后机器自动计算出 ROI 内 FF 值,右侧大收肌 FF 为 73.26% (箭); c) T₁ WI 横轴面图像,右侧大收肌主观评分为 4 分(箭); d) 为同一层面 IDEAL-Quant 图像处理方法,圈画 ROI 后机器自动计算出 ROI 内 FF 值,右侧大收肌 FF 为 83.29%(箭)。图 b 与 d 中右侧大收肌脂化率相差约 10%,但肉眼主观评分均为 4 分,无法区分二者差异,提示主观评分结果无法反映肌肉内脂肪浸润程度间较小的差异。

在体外研究中 IDEAL-Quant 技术脂肪定量结果与样本真实的脂肪百分比接近。在肝脏、脊柱脂肪定量研究中,IDEAL-Quant 技术也可以客观、准确定量脂肪百分比,其结果与 MRS 结果高度相关^[13-14]。因此,本研究以 IDEAL-Quant 技术测量的脂肪百分比作为参考标准。

该研究发现经验丰富的阅片者主观评分结果与客观定量结果存在明显相关性,说明经验丰富的阅片者主观评分的结果比较可靠。但主观评分结果高于客观定量值,说明主观评分会高估肌肉的脂肪浸润程度,该结论与 Alizai 等^[11]及 Wokke 等^[15]的结论一致,证明主观评分会高估肌肉脂肪浸润程度。推测可能的主要

原因为:主观评分时肉眼无法分别出增生的结缔组织和脂肪,二者均为高信号,而肌肉组织为低信号,因此肉眼观察时将较高信号的各种结构都判断为脂肪组织,使得肉眼评分结果偏高;其次,不同的肌肉走行存在差异,斜切面或平行切面时会高估脂肪浸润程度。

文献报道,健康人群下肢骨骼肌脂肪百分比约为 $\leq 5\%$ ^[11],依据 Mercuri 等^[7]和 Kim 等^[8]的脂肪浸润标准:0 分,正常;1 分,散在小片状脂肪浸润;2 分,脂肪浸润 $< 30\%$;3 分,脂肪浸润 $30\% \sim 60\%$;4 分,脂肪浸润 $> 60\%$ 。为区分各评分等级差异,本研究通过对上述标准进行修订,设定脂肪浸润 $< 5\%$ 为 0 分, $5\% \sim 10\%$ 为 1 分, $10\% \sim 30\%$ 为 2 分,3 分及 4 分标准同

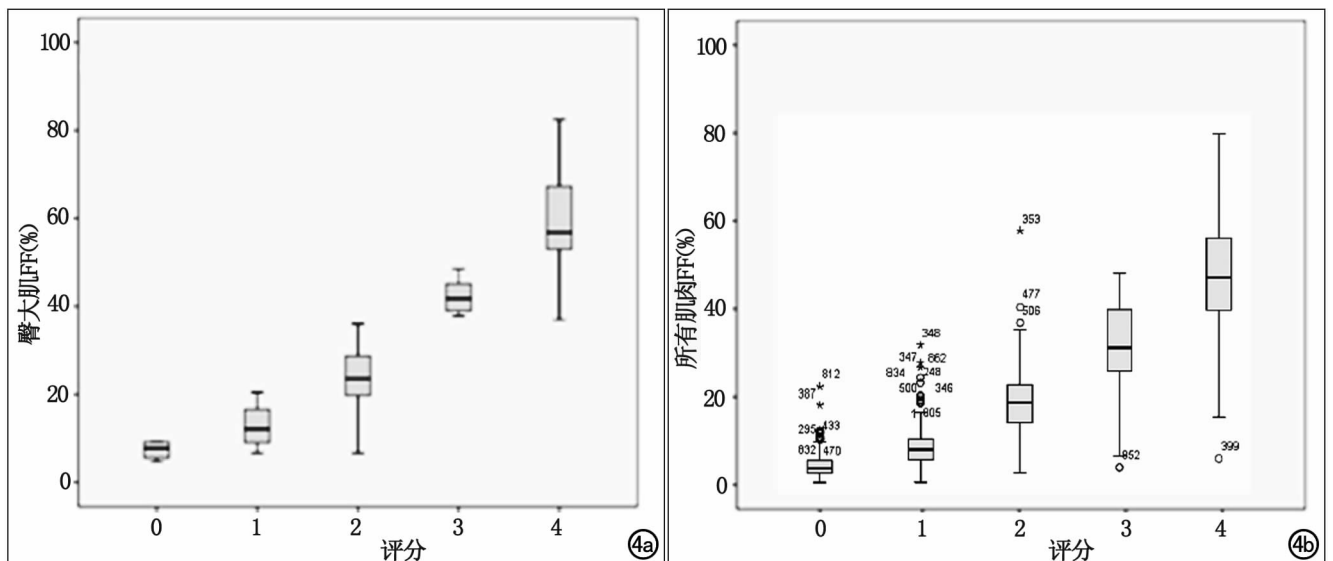


图 4 主观评分结果与客观评分结果的差异。a) 臀大肌主观评分与客观评分值的差异,显示主观评分值高于客观评分值。如对于评 4 分的所有臀大肌,其客观定量平均 FF 为 55%,小于 4 分(肌肉脂肪浸润程度 $> 60\%$)标准。且对于评 4 分的所有臀大肌,其客观 FF 为 38%~85%,说明客观结果可以更精确显示脂肪浸润程度; b) 所有肌肉的主观评分与客观评分结果差异,也同时显示主观评分的平均结果高于客观定量评分平均值,尤其以评分为 4 分时为著。箱形图以外的点为异常点,显示主观评分在某些评分时,其结果可能低于或高于客观评分结果。图 a 及 b 中,部分箱形图的区域可见重叠,以 0 分及 1 分时箱形图重叠面积为最大,提示肉眼区分 0 分及 1 分存在一定困难。

上。本研究发现在同一评分时,不同肌肉脂肪浸润程度也存在很大差异,以臀大肌为例,4 分时肌肉脂肪浸润 37%~83%。且不同评分级别时,肌肉脂肪浸润程度出现重叠,在 0 分及 1 分中出现重叠情况较多。且主观评分法只能大致评估肌肉脂肪浸润程度,无法精细表示出不同肌肉脂肪浸润程度之间的差异,且对于随访或治疗后出现的轻度改变不敏感。因此,在需要精确评估脂肪浸润程度或观察细微变化时,需要使用敏感性更高的方法。

本研究存在一定局限性。首先仅选择抗肌萎缩蛋白病作为研究对象,不能完全代表所有类型的神经肌肉病,但本研究采用盲法评分,且样本量大,肌肉受累跨度较大,代表性较好。其次,本研究无组织学水平脂肪定量结果作为金标准,但在实践中无法获得所有肌肉的组织学定量结果,因此以 IDEAL-Quant 技术的脂肪定量结果作为参考标准。再次,本文中并无同年龄段健康儿童肌肉 MRI 定量结果作为参考标准,但儿童检查部分需要镇静,且磁共振存在一定噪声,无法募集同年龄段的健康儿童作为对照,对正常肌肉的脂肪定量标准参考了文献中提供的成人的数据。

总之,MR T₁WI 横轴面主观半定量评分法简便、快捷,神经肌肉病诊断经验丰富的阅片者评分结果较为稳定、可靠。且主观评分结果与客观定量结果存在显著相关性,但主观评分法会高估肌肉脂肪浸润程度,且较为粗略,无法反应细微差别。因此,在应用主观评分法对图像进行快速阅读后,需要应用更敏感的定量技术对骨骼肌脂肪浸润程度进行准确评估。

参考文献:

- [1] Hilton TN, Tuttle LJ, Bohnert KL, et al. Excessive adipose tissue infiltration in skeletal muscle in individuals with obesity, diabetes mellitus, and peripheral neuropathy: association with performance and function[J]. *Physical Therapy*, 2008, 88(11): 1336-1344.
- [2] 王新德, 沈定国, 吴士文, 等. 神经病学: 肌肉疾病[M]. 北京: 人民军医出版社, 2007: 150-154, 214-222.
- [3] Ishu A, Willcocks RJ, Forbes SC, et al. Examination of effects of corticosteroids on skeletal muscles of boys with DMD using MRI and MRS[J]. *Neurology*, 2014, 83(11): 974-980.
- [4] Fischmann A, Hafner P, Gloor M, et al. Quantitative MRI and loss of free ambulation in Duchenne muscular dystrophy[J]. *J Neurol*, 2013, 260(4): 969-974.
- [5] Li W, Zheng Y, Wei Z, et al. Progression and variation of fatty infiltration of the thigh muscles in Duchenne muscular dystrophy, a muscle magnetic resonance imaging study[J]. *NMD*, 2015, 25(5): 375-380.
- [6] 李文竹, 郑艺明, 杜婧, 等. 迪谢内肌营养不良骨骼肌磁共振成像研究[J]. *中华神经科杂志*, 2014, 47(1): 16-20.
- [7] Mercuri E, Pichiecchio A, Allsop J, et al. Muscle MRI in inherited neuromuscular disorders: past, present, and future[J]. *J Magn Reson Imaging*, 2007, 25(2): 433-440.
- [8] Kim HK, Laor T, Horn PS, et al. T₂ mapping in Duchenne muscular dystrophy: distribution of disease activity and correlation with clinical assessments[J]. *Radiology*, 2010, 255(3): 899-908.
- [9] 杜婧, 程晓悦, 肖江喜, 等. 常规 MRI 及功能 MRI 诊断抗肌萎缩蛋白病[J]. *中国医学影像技术*, 2014, 30(9): 1399-1402.
- [10] Bushby K, Finkel R, Birnkrant DJ, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management[J]. *Lancet Neurol*, 2010, 9(1): 77-93.
- [11] Alizai H, Nardo L, Karampinos D, et al. Comparison of clinical semi-quantitative assessment of muscle fat infiltration with quantitative assessment using chemical shift-based water/fat separation in MR studies of the calf of post-menopausal women[J]. *Euro Radiol*, 2012, 22(7): 1592-1600.
- [12] Liu CY, McKenzie CA, Yu H, et al. Fat quantification with IDEAL gradient echo imaging: correction of bias from T₁ and noise[J]. *Magn Reson Med*, 2007, 58(2): 354-364.
- [13] Bernard CP, Liney GP, Manton DJ, et al. Comparison of fat quantification methods: a phantom study at 3.0T[J]. *J Magn Reson Imaging*, 2008, 27(1): 192-197.
- [14] Sina M, Catherine DG, Gavin H, et al. Quantification of hepatic steatosis with T₁-independent, T₂*-corrected MR imaging with spectral modeling of fat: blinded comparison with MR spectroscopy[J]. *Radiology*, 2011, 258(3): 767-775.
- [15] Wokke BH, Bos C, Monique Reijnerse MD, et al. Comparison of dixon and T₁-weighted MR methods to assess the degree of fat infiltration in duchenne muscular dystrophy patients[J]. *J Magn Reson Imaging*, 2013, 38(3): 619-624.

(收稿日期: 2015-10-15)