

• 临床研究 •

粗大运动功能测试量表 66 项的反应度和精确度研究

史惟 廖元贵 王素娟 施炳培

[摘要] 目的 比较粗大运动功能测试量表 66 项(GMF M 66)与 88 项(GMF M 88)的反应度和精确度。方法 以 173 例(男性 126 例,女性 47 例)接受过两次以上 GMF M 评估的患儿为研究对象,在平均(3.8 ± 4.3)个月后进行第 2 次评估,其中 112 例(64.7%)在平均(3.2 ± 2.2)个月进行第 3 次评估,比较 GMF M 88 项和 GMF M 66 项在前后 3 次评估之间的效应尺度和相对精确度。结果 在各次评估之间,GMF M 66 项的效应尺度与 GMF M 88 项非常接近,差别在 0.03~0.04 之间,且精确度不低于 GMF M 88 项。结论 具有等距特性的 GMF M 66 项有良好的反应度和精确度,比 GMF M 88 项更有利于康复疗效判断。

[关键词] 脑性瘫痪;粗大运动;粗大运动功能测试量表(GMF M);效应尺度;相对精确度

Responsiveness and precision of Gross Motor Function Measure (66 items version) SHI Wei, LIAO Yuan-gui, WANG Su-juan, et al. Rehabilitation Center of Childrens Hospital Affiliated to Fudan University, Shanghai 200032, China

[Abstract] **Objective** To compare the responsiveness and precision between the 66 items version of Gross Motor Function Measure (GMF M 66) and 88 items version (GMF M 88). **Methods** 173 children with cerebral palsy (male 126, female 47) who have received twice assessments were involved. All the second assessment were performed 3.8 ± 4.3 months later. 112 children received the third assessment 3.2 ± 2.2 months after the second assessment. The effect size and relative precision between the GMF M 66 and GMF M 88 were analyzed. **Results** The effect size of GMF M 66 between every two assessments was nearly the same as those of GMF M 88. The precision of GMF M 66 was not lower than that of GMF M 88. **Conclusion** The interval-scale GMF M 66 has satisfying responsiveness and precision. It is more useful than GMF M 88 in measuring the treating effect of children with cerebral palsy.

[Key words] cerebral palsy; gross motor; Gross Motor Function Measure (GMF M); effect size; relative precision

中图分类号:R472.3 文献标识码:A 文章编号:1006-9771(2005)10-0838-03

[本文著录格式] 史惟,廖元贵,王素娟,等.粗大运动功能测试量表 66 项的反应度和精确度研究[J].中国康复理论与实践,2005,11(10):838-840.

粗大运动功能评估是脑瘫康复评估中的重要组成部分,粗大运动功能测试量表(gross motor function measure, GMF M)用于测量脑瘫儿童的粗大运动状况随时间或由于干预而出现的运动功能改变,是目前脑瘫儿童粗大运动评估中使用最广泛的量表^[1,2],共包括 5 区 88 项。2000 年, Russell 等人使用 Rasch 分析法对 GMF M 进行了信度和效度分析,最后确立了 GMF M 66 项^[3]。GMF M 66 项属于等距型量表,提高了总分和变化分数的可理解性,能够合理、客观地反映脑瘫患儿的粗大运动发育变化。经过 Rasch 分析后的 GMF M 66 项删除了能力值较低的 22 个测试项目,全部集中在 A、B、C 区,尤其以能力值较低的 A 区居多,原来的 17 项中被删除了 13 项^[4]。我们已经报道了 GMF M 66 项在婴幼儿脑瘫粗大运动评估中的内在信度和结构效度研究^[5],本研究将进一步探讨该量表的反应度和精确度。

1 对象与方法

1.1 研究对象 2002 年 4 月~2005 年 2 月在本院康

复中心就诊和治疗的脑瘫患儿 173 例,诊断符合 1988 年中国小儿脑瘫座谈会制订的标准^[6],均接受过两次以上 GMF M 88 项评估,其中男性 126 例、女性 47 例,最小 2 个月,最大 94 个月,平均(20.72 ± 11.50)个月;痉挛型 146 例(84.4%),其中四肢瘫 72 例(41.6%)、双瘫 49 例(28.3%)、偏瘫 23 例(13.3%)、三瘫 2 例(1.2%),徐动型 16 例(9.2%),混合型 11 例(6.4%)。

1.2 方法

173 例患儿首次评估后平均(3.8 ± 4.3)个月进行第 2 次评估,其中 112 例(64.7%)在第 2 次评估后平均(3.2 ± 2.2)个月进行第 3 次评估。

评估采用中文译本的 GMF M 88 项。中文译本根据 1993 年 GMF M 修订版^[7]转译,由从事儿童康复的医师和物理治疗师进行翻译。评估由 3 名指定医师或治疗师分别实施,在安静、独立、采光较好的房间内进行,室温控制在 $20^{\circ}\text{C} \sim 30^{\circ}\text{C}$,患儿衣服为 1~2 层。在不违反评估要求的情况下,每次评估都安排相同的家属在场,鼓励患儿发挥出最佳水平。

按照 GMF M 指导手册的要求计算各项评估结果的 GMF M 88 项总分,同时再把各项评估结果输入由 GMF M 66 项所配置的统计软件 Gross Motor Ability Estimator(Version 1.0, 2002),得出 GMF M 66 项分

作者单位:200030 上海市,复旦大学附属儿科医院康复中心。作者简介:史惟(1965-),男,上海市人,医师,主要研究方向:儿童运动发育与运动治疗。

值^[3]。

1.3 统计学处理 重测信度研究包括组内信度研究,对 20 例脑瘫患儿由同一测试者在第 1 次测试后 1 周内进行第 2 次测试。组间信度研究由不同的测试者对 20 例脑瘫患儿同时进行测试,测试过程中测试者之间不进行讨论,以免影响测试结果。以上结果都用等级相关系数 ICC 表达组间和组内信度。

反应度是指在变化状况下测量量表的应变性,常使用效应尺度评价量表的反应度。效应尺度 = (后次得分 - 前次得分) / 前次得分的标准差^[8]。我们比较了 GMFM 88 项和 GMFM 66 项在前后 3 次评估之间的效应尺度。

精确度是指某种评估方法在不同人群或不同测试状况下区分被测者水平的能力,其高低依赖于两方面因素:①区分不同人群和不同测试状况的程度;②测试结果的组内和组间差异水平。相对精确度可以提示某种测试方法与标准方法或原方法相比,精确度的增减情况,是成对 *F* 检测值的比值^[9]。我们测定 GMFM 88 项和 GMFM 66 项在前后 3 次评估之间的相对精确度。以上分析均采用 STATA 7.0 统计软件。

2 结果

表 1 GMFM 88 项与 GMFM 66 项效应尺度和相对精确度比较 (第 1 次与第 2 次评估, n=173)

项目	第 1 次评估			第 2 次评估			两次评估之差			<i>t</i>	效应尺度	<i>F</i>	相对精确度
	分值	<i>SD</i>	<i>SE</i>	分值	<i>SD</i>	<i>SE</i>	分值	<i>SD</i>	<i>SE</i>				
GMFM 88	32.16	22.02	1.67	37.85	23.52	1.79	5.68	7.20	0.55	10.39	0.26	5.39	1.00
GMFM 66	34.90	13.26	1.01	38.74	14.03	1.07	3.84	4.32	0.33	11.70	0.29	6.85	1.27

注:两次评估比较 *t* 检验: *P* < 0.001; GMFM 88 项两次评估比较方差检验: *P* < 0.05; GMFM 66 项两次评估比较方差检验: *P* < 0.001。

表 2 GMFM 88 项与 GMFM 66 项的效应尺度和相对精确度比较 (第 2 次与第 3 次评估, n=112)

项目	第 1 次评估			第 2 次评估			两次评估之差			<i>t</i>	效应尺度	<i>F</i>	相对精确度
	分值	<i>SD</i>	<i>SE</i>	分值	<i>SD</i>	<i>SE</i>	分值	<i>SD</i>	<i>SE</i>				
GMFM 88	34.47	19.76	1.87	39.86	21.97	2.08	5.39	8.20	0.78	6.96	0.27	3.73	-
GMFM 66	36.79	11.49	1.09	39.55	12.13	1.15	2.77	4.75	0.45	6.18	0.24	3.07	-

注:两次评估比较 *t* 检验: *P* < 0.001; 两次评估比较方差检验: *P* > 0.05, 不能计算相对精确度。

表 3 GMFM 88 项与 GMFM 66 项之间的效应尺度和相对精确度比较 (第 1 次与第 3 次评估, n=112)

项目	第 1 次评估			第 2 次评估			两次评估之差			<i>t</i>	效应尺度	<i>F</i>	相对精确度
	分值	<i>SD</i>	<i>SE</i>	分值	<i>SD</i>	<i>SE</i>	分值	<i>SD</i>	<i>SE</i>				
GMFM 88	28.63	17.61	1.66	39.86	21.97	2.08	11.23	11.63	1.10	10.21	0.64	17.82	1.00
GMFM 66	33.02	10.95	1.04	39.55	12.13	1.15	6.53	6.20	0.59	11.16	0.60	17.89	1.00

注:两次评估比较 *t* 检验和方差检验均为 *P* < 0.001。

3 讨论

康复医学量表除了可以评估功能障碍,区分康复人群与正常人群之间的差别,体现临床亚型之间的区别,更重要的还必须有效地反映康复疗效,而且儿童康复医学量表与成人量表之间的最大差别在于必须考虑自然生长发育因素。自然生长发育因素给康复疗效判

2.1 GMFM 量表的重测信度 组内信度:对 20 例脑瘫患儿,同一测试者在间隔 1 周内的两次测试结果之间具有非常高度相关性, GMFM 88 项的 ICC 为 0.9898(95 %可信区间 0.9721 ~ 0.9956), GMFM 66 项的 ICC 为 0.9666(95 %可信区间 0.9157 ~ 0.9868)。

组间信度:对 20 例脑瘫患儿,两位测试者同时进行的测试结果之间具有非常高度相关性, GMFM 88 项的 ICC 为 0.9875(95 %可信区间 0.9684 ~ 0.9951), GMFM 66 项的 ICC 为 0.9782(95 %可信区间 0.9450 ~ 0.9914)。

2.2 GMFM 88 项与 GMFM 66 项效应尺度和相对精确度比较 3 次评估结果显示, GMFM 88 项和 GMFM 66 项的效应尺度非常接近,差异在 0.03 ~ 0.04 之间,表明 GMFM 66 项的反应度与 GMFM 88 项相似。由于 GMFM 88 项与 GMFM 66 项得分在第 2 次与第 3 次评估之间的配对方差检验均无显著性差异,所以不能计算相对精确度;而第 1 次评估的精确度与第 2 次评估的精确度之间, GMFM 66 项的精确度略高,但第 1 次评估的精确度与第 3 次评估之间无显著性差异,表明 GMFM 66 项的精确度不低于原来的 GMFM 88 项(见表 1 ~ 表 3)。

断带来了较大的困难,特别是在发育相对快速的 0 ~ 3 岁阶段,而等距型量表得分间隔的稳定性有利于临床医学研究中设定良好的对照组,在一定程度上避免儿童自然发育因素对疗效判断的影响。

GMFM 量表制定者在 GMFM 88 项基础上运用 Rasch 分析法确立了 GMFM 66 项。Rasch 分析属于

项目反应理论(item response theory, IRT)中的一种分析模型。传统的项目分析理论是从原始数据中直接推断出项目的难度、信度和效度,确定原始数据与项目难度之间的关系;而 Rasch 是通过将样本总体进行评价后,在被测试者的全部测试结果中确定被测试者的能力和项目的难易度以及他们之间的关系,所以需要较多的测试项目和测试结果以确保分析结果的稳定性。与传统的项目分析理论相比,Rasch 分析有以下 3 个特征:①能够在共同尺度的基础上实现不同项目的能力测定(test-free person measurement);②在不同样本群中均可以测定出共同的项目特性测定值(sample-free item calibration);③能够精确地测定项目难易度和被测试者的能力以及他们之间的关系(multiple reliability estimation)。由于 Rasch 分析属于项目反应理论中的单参数模型,项目鉴别力被假定为无差异性,所以经 Rasch 分析后得出的能力分值具有等距性。近年来,康复医疗界也开始使用 Rasch 分析进行各种量表的建立和效度分析^[10],确定项目的等级,从而更好地反映功能状态的恢复过程,同时可以删除不适合项目,提高等级量表的准确性。

经 Rasch 分析后的 GMFM 66 项具有以下特点:①能力分值的等距性可以提高能力分值和改变分值的可理解性;②确立了测试项目的难度顺序;③删除了部分不适合项目,增加了评估的单维性;④符合心理学测定意义上的信度、效度^[4]。GMFM 66 项的等距特性能够合理、客观地反映脑瘫患儿的粗大运动发育变化。GMFM 66 项删除了原来量表中的 22 个项目,虽然项目的减少可节省评估时间,但可能会影响测定脑瘫儿童粗大运动发育水平变化状况的有效程度。为此,我们检验了 GMFM 66 项与 GMFM 88 项之间的效应尺度和相对精确度,以确定 GMFM 66 项的反应度和精确度。

引起 GMFM 测试结果变化的主要因素是被测量者的粗大运动功能状况。但测量环境、被测量者的情绪和依从性、年龄和背景、测量者的经验以及测量者与被测量者之间的熟悉程度、测量的方法等都有可能影响结果。在本研究中,我们首先测定了 GMFM 88 项和 GMFM 66 项的重测信度,结果表明,GMFM 66 项依然保持着良好的重测信度。在此基础上我们检测了 GMFM 88 项和 GMFM 66 项对脑瘫儿童粗大运动能力前后 3 次评估之间的效应尺度和相对精确度,结果显示,GMFM 66 项的反应度和精确度均不逊色于原来的 GMFM 88 项。

GMFM 88 项经过 Rasch 分析,删除了 22 个项目,重新建立了具有等距特性的 GMFM 66 项。项目

的减少可节省评估时间,而且与 GMFM 88 项相比保持了相似的效应尺度,精确度也不低于 GMFM 88 项,提高了 GMFM 量表在脑瘫儿童粗大运动功能评估中的检测水平。GMFM 66 项的等距性提高了能力分值和改变分值的可理解性,在临床上通过合理的对照分析可以判断治疗前后粗大运动能力的变化,也可以比较各种脑瘫治疗方法之间的效果,在康复疗效判断上优于 GMFM 88 项。由于 GMFM 66 项具有良好的反应度和精确度,所以在对脑瘫儿童进行康复治疗过程中,为了更好地掌握患儿的发育变化状况,建议每隔 3 个月左右进行 1 次 GMFM 66 项评估,或采用每月平均改变值评估康复疗效^[11]。

[参考文献]

- [1]史惟,廖元贵,杨红,等.粗大运动功能测试量表与 Peabody 粗大运动发育量表在脑性瘫痪康复疗效评估中的应用[J].中国康复理论与实践,2004,10:423—424.
- [2]Kondo I, Hosokawa K, Iwata M, et al. Effectiveness of selective muscle release surgery for children with cerebral palsy: longitudinal and stratified analysis[J]. Developmental Medicine and Child Neurology, 2004, 46: 540—547.
- [3]Russell D, Rosenbaum PL, Avery LM, et al. Gross Motor Function Measure (GMFM-66 & GMFM-88) Users Manual [M]. London: Mac Keith, 2002. 30—123.
- [4]Russell D, Avery LM, Rosenbaum PL, et al. Improved scaling of the gross motor function measure for children with cerebral palsy: evidence of reliability and validity[J]. Physical Therapy, 2000, 80: 873—885.
- [5]史惟,廖元贵,王素娟,等. GMFM66 项在婴幼儿脑性瘫痪粗大运动评估中的信度和效度研究[A]. 2004 年全国小儿脑性瘫痪专题研讨会论文汇编, 昆明, 2004. 北京:中华医学会中华儿科杂志编辑委员会, 2004. 21—27.
- [6]林庆. 小儿脑性瘫痪的定义、诊断条件及分型[J]. 中华儿科杂志, 1989, 27: 162.
- [7]Russell D, Rosenbaum PL, Gowland C, et al. Gross motor function measure manual [M]. 2nd ed. Ontario: Institute for Applied Health Sciences, McMaster University, 1993. 1—112.
- [8]方积乾. 医学统计与电脑实验 [M]. 第 2 版. 上海:上海科学技术出版社, 2001. 238—246.
- [9]Fitzpatrick R, Norquist JM, Dawson J, et al. Rasch scoring of outcomes of total hip replacement[J]. J Clin Epidemiol, 2003, 56: 68—74.
- [10]Strong DR, Breen RB, Lesieur HR, et al. Using the Rasch model to evaluate the south oaks gambling screen for use with nonpathological gamblers [J]. Addictive Behaviors, 2003, 28: 1465—1472.
- [11]史惟,施炳培,廖元贵,等. 运动发育推拿法治疗小儿脑瘫 [J]. 中国康复, 2004, 19: 351—352. (收稿日期: 2005-05-12)