

DOI:10.3724/SP.J.1008.2009.00283

## 我国脑卒中患者临床神经功能缺损评分标准信度、效度及敏感度的评价

陶子荣

中南大学湘雅医院神经外科,长沙 410008

**[摘要]** **目的:**对我国脑卒中患者临床神经功能缺损评分标准(CSS)的信度、效度和敏感度进行研究,评价其临床应用价值。

**方法:**根据纳入标准,选取我院神经内科的126例脑卒中患者,分别用CSS和美国国立卫生研究院卒中量表(NIHSS)对其进行评分,并对CSS进行信度、效度、敏感度评价。信度用相关系数 $r$ 和Cronbach's  $\alpha$ 系数来评价。结构效度采用聚类因子分析方法考核;效标效度以NIHSS量表为效标,应用相关系数分析。各维度敏感度通过标准化效应(SES)来评估。**结果:**共收回有效问卷123份。CSS有良好的信度,各维度的重测信度和评定者间信度均在0.911~1.000;内部一致性信度的Cronbach's  $\alpha$ 系数均在0.8以上。CSS有较好的结构效度;分析CSS与NIHSS间相关系数为0.86,效标效度良好。预后预测CSS准确率为92.4%,稍低于NIHSS的预测准确率94.1%;Logistic回归显示CSS的“凝视功能”和“面瘫”项目未能归入预测方程;除面瘫项目SES为0.38外,其他均大于0.5,大多数维度均显示出了良好的敏感性。**结论:**经初步测试,CSS在脑卒中患者中具有较好的信度、效度、敏感度,但预测预后不及NIHSS,修订时可作相应调整。

**[关键词]** 脑血管意外;临床神经功能缺损评分标准;信度;效度;敏感度

**[中图分类号]** R 743.3 **[文献标志码]** A **[文章编号]** 0258-879X(2009)03-0283-03

### Reliability, validity and sensitivity of Chinese scale for clinical neurological deficit of stroke patients

TAO Zi-rong

Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha 410008, China

**[ABSTRACT]** **Objective:** To investigate the reliability, validity and sensitivity of Chinese scale of clinical neurological deficit of stroke patients (China Stroke Scale, CSS), so as to assess its clinical application value. **Methods:** A total of 126 consecutive inpatients with acute stroke onset were included in our study and they were scored by CSS and the United States National Institutes of Health Stroke Scale (NIHSS) score separately; the reliability, validity, and sensitivity of CSS were evaluated. Reliability was evaluated by correlation coefficient  $r$  and Cronbach's  $\alpha$  coefficient; construct validity was analyzed by factor analysis method of appraisal; criterion validity was analyzed by the correlation coefficient analysis with NIHSS scale as the criterion. Sensitivity in various fields was assessed through standardization of effect (SES). **Results:** Totally 123 valid questionnaires were collected. CSS showed high intrarater reliability, interrater reliability (0.911-1.000) and good internal consistency, with the Cronbach's  $\alpha > 0.8$ . There was concurrent validity between CSS and NIHSS ( $r = 0.86$ ). The prognosis prediction accuracy of CSS was 92.4%, slightly lower than that of NIHSS (94.1%). Logistic regression showed that CSS's "gaze function" and "facial paralysis" were not included in the prediction equation. The facial paralysis had a SES of 0.38, all others had a SES higher than 0.5. Most fields showed a good sensitivity. **Conclusion:** CSS shows an acceptable reliability, validity and sensitivity in patients with stroke, but the predicative validity of CSS is inferior to that of NIHSS, which needs be further revised.

**[KEY WORDS]** cerebrovascular accident; Chinese Scale of Clinical Neurologic Deficit; reliability; validity; sensitivity

[Acad J Sec Mil Med Univ, 2009, 30(3): 283-285]

我国现行的脑卒中患者临床神经功能缺损评分标准(China Stroke Scale, CSS)是在1995年全国第四次脑血管病学术会议上修订而成的<sup>[1]</sup>。十余年来,CSS在临床脑卒中患者神经功能缺损的评价中得到了较为广泛的应用<sup>[2]</sup>。王新等<sup>[3-4]</sup>曾于1999年

对CSS进行量表的质量研究,初步证实CSS具有良好的信度和效度,同时提出CSS存在缺陷:预测效度不如美国国立卫生研究院卒中量表(NIH Stroke Scale, NIHSS)全面;有一个无预测值的项目“面瘫”;“步态”项目对急性期脑出血患者使用困难,认

**[收稿日期]** 2008-07-20 **[接受日期]** 2008-09-18

**[作者简介]** 陶子荣,硕士生. E-mail:lxw514@126.com

为 CSS 需要进一步修订。其后未再见有相关研究报道,而 CSS 在临床也取得了较为满意的效果,并且王新等<sup>[3]</sup>的研究样本含量偏小,以此作为 CSS 较为重要的修订依据,证据不足。本研究设计对 CSS 的信度、效度及其敏感度进行进一步评价,为进一步修订更符合国人特点的 CSS 提供依据。

### 1 对象和方法

1.1 研究对象 选取 2008 年 3~6 月在我院神经内科入院治疗的急性脑卒中患者。纳入标准为 1982 年国际神经系关联病及卒中协会的诊断标准,并行头颅 CT 和(或)MRI 确诊。排除标准为:(1)患者入院距发病时间均大于 72 h;(2)既往曾有脑卒中病史,入院前就存在一定程度的神经功能障碍定位表现者;(3)短暂性脑缺血发作(TIA)及病灶不明确者;(4)有严重心、肝、肾等脏器疾病者;(5)昏迷和痴呆者。最后共入选 126 例。

1.2 数据收集方法 采用问卷调查法进行数据收集。调查前均向入选对象说明本次研究的目的、意义及方法,取得对方的合作。量表采用自评法,由调查对象本人填写,若被调查者填表有困难可由其家属根据被调查者本人的理解代为填写。

初次评分:分别应用 CSS 和 NIHSS 对患者进行调查,2 份调查表填写间隔 30 min。全部 126 例患者均进行了初次评分,在入院 24 h 内进行。调查均在同一名医师的指导下进行。

重测信度与评定者间信度:评定量表为 CSS。从中随机选取 50 例患者,随机分成两组:一组 25 例由同一检查者于其第 1 次记录后 1 d 进行再次评定;另一组 25 例患者由另一名检查者于第 1 次记录后 1 d 进行再次评定,评定步骤与前一次一致。

1.3 CSS 质量评价方法 所有资料均采用 Excel 建立数据库。重测信度与评定者间信度用相关系数 *r* 表示,内容一致信度用 Cronbach's  $\alpha$  值表示。结构效度采用聚类因子分析进行评价,效标效度采用 CSS 与 NIHSS 两者间的相关系数 *r* 表示,预测效度用 Logistic 回归模型进行分析。各维度敏感度通过标准化效应 (SES) 大小来评估。用 Stata 9.0 软件进行统计分析。

### 2 结果

2.1 一般资料 根据病例纳入标准,本研究共发出问卷 126 份,实际收回 123 份,接受率为 97.6%。在收回的 123 份问卷中,核查条目均全部完成,完成率为 100%。该 123 例患者中,年龄 42~92(66.1±

12.5)岁;男 69 例,女 54 例;脑梗死 92 例,脑出血 31 例;病灶部位左侧 46 例,右侧 40 例,双侧 31 例,脑桥小脑 6 例;71 例既往有高血压病史,52 例既往有心脏病史,39 例既往有糖尿病史。

#### 2.2 CSS 量表的信度

2.2.1 重测信度 25 例患者在相隔 1 d 时间内,两次评定,各维度的 *r* 值为 0.903~0.972 ( $P$  均 < 0.001)。重测信度最好的项目是手肌力,最差的项目是水平凝视功能(表 1)。这表明 CSS 受环境、时间等随机因素干扰小,重测信度好。

2.2.2 评定者间信度 25 例患者在相隔 1 d 由不同评定者进行调查,各维度的 *r* 值为 0.982~1.000 ( $P$  均 < 0.001,表 1)。表明评定者之间的评分误差变异可以忽略,CSS 评定者间信度良好。

2.2.3 内部一致性信度 CSS 各维度的内容一致性都很好,各维度的 Cronbach's  $\alpha$  值均在 0.8 以上(表 1)。

表 1 CSS 量表信度评价

Tab 1 Reliability of CSS

Item	Intrater reliability ( <i>n</i> = 25)	Interrater reliability ( <i>n</i> = 25)	Cronbach's $\alpha$ ( <i>n</i> = 123)
Consciousness	0.921	0.995	0.832
Horizontal gaze	0.903	0.982	0.826
Facial paralysis	0.911	0.991	0.831
Language	0.923	0.993	0.846
Arm movement	0.953	1.000	0.876
Hand movement	0.972	1.000	0.878
Crura movement	0.934	1.000	0.861
Impaired walking	0.928	1.000	0.854
Total score	0.937	0.996	0.859

#### 2.3 CSS 量表的效度

2.3.1 结构效度 对 CSS 量表的测试结果进行因子分析,8 个维度采用主成分分析法选出一个因素,贡献率达 69.5%,高于结构效度检验的最低标准 40%,且量表的每个维度都在这个因子上有较高的负荷值(0.719~0.823),其中意识水平 0.731、凝视功能 0.721、面瘫 0.719、语言 0.785、上肢肌力 0.823、手肌力 0.798、下肢肌力 0.801、步行能力 0.768),皆大于 0.7,表明 CSS 量表有良好的结构效度。

2.3.2 效标效度 以 NIHSS 量表的检测结果为效标,CSS 的检测结果与之进行相关分析,发现 CSS 的实证效度与 NIHSS 量表相近,总评分的相关系数 *r* 值为 0.86 ( $P$  < 0.001)。

2.3.3 预测效度 所有 123 例患者出院后均随访 3 个月,记录生存预后。住院期间死亡 2 例,随访期间

死亡4例(3例因脑卒中复发,1例因合并肺部感染),本组卒中后病死率为4.9%。用 Logistic 回归分别分析 CSS、NIHSS 初评总分对死亡结局的预测性,发现 CSS 预测的准确率为 92.4%,NIHSS 的预测准确率 94.1%。将两个量表各项目的评分与随访结局进一步做 Logistic 回归发现:NIHSS 的各项目均归入方程,而 CSS 的“凝视功能”和“面瘫”项目未能归入方程。

2.4 敏感度 采用 SES 表示敏感度。SES $\leq$ 0.2 为无敏感;0.2<SES $\leq$ 0.5 为轻度敏感;0.51<SES<0.7 为中度敏感;SES $\geq$ 0.7 为高度敏感。本组除面瘫项目外,其他 SES 均大于 0.5(意识水平 0.62、凝视功能 0.53、面瘫 0.38、语言 0.57、上肢肌力 0.64、手肌力 0.65、下肢肌力 0.59、步行能力 0.54),表明 CSS 量表有良好的敏感度。

### 3 讨论

脑卒中患者临床神经功能缺损的评估对急性期治疗和预后的判断均具有重要意义。我国现行的评分标准(CSS)是以斯堪的纳维亚神经卒中量表(Scandinavian Stroke Scale, SSS)<sup>[5]</sup>为基础,在1995年全国第四次脑血管病学术会议上修订而成的。该量表共包括8个维度<sup>[1]</sup>,分别是意识、水平凝视功能、面瘫、言语、上肢肌力、手肌力、下肢肌力和步行能力。其中意识维度分为两项提问、两项指令、强烈局部刺激3个条目,其余每个维度均为1个条目。每个条目根据实际情况赋值0~6分。合计总分最高45分,最低0分;据此分为轻型(0~15分)、中型(16~30分)、重型(31~45分)。本研究对 CSS 量表本身的质量进行研究,并评价其临床应用价值。

量表的质量一般可以通过其信度、效度及敏感度评价。信度是指所测试结果的可靠程度或可重复程度,通常采用重测信度、评定者间信度、分半信度和 Cronbach's  $\alpha$  系数来分析<sup>[6]</sup>。效度是反映一个测定工具能否有效地测定到了它所预测定内容,或测定工具的测定结果与预想结果的符合程度,包括结构效度、效标效度、预测效度和内容效度等<sup>[6]</sup>。敏感度是指在内外环境变化时,测量结果对此变化做出反应的敏感程度。常用的研究方法包括使用效应尺度统计量以及治疗前后的统计学分析。

我们的研究表明:CSS 重测信度最好的项目是手肌力,最差的项目是水平凝视功能。推测这可能与神经科医师通常熟练掌握肌力评定,而对于水平凝视功能的检查容易忽视有关。评定者间信度各维度的  $r$  值为 0.982~1.000,表明 CSS 评定者间信度良好,

CSS 的评定者之间的评分误差变异可以忽略。本研究中 Cronbach's  $\alpha$  值均在 0.8 以上,表明量表的内部一致性信度较好<sup>[6]</sup>。这与 NIHSS 中文版的 Cronbach's  $\alpha$  值 0.92<sup>[7]</sup> 基本相当,也与王新等<sup>[3]</sup>的研究结果一致。

另外,本研究通过主成分分析发现:CSS 量表每个维度都有较高的负荷值,具有良好的结构效度。而将 NIHSS 量表作为效标,发现 CSS 的实证效度与 NIHSS 量表相近,相关系数  $r$  值为 0.86,CSS 也具有较好的效标效度。但预测效度的研究发现,CSS 量表对死亡结局的预测不及 NIHSS 量表,尤其是 CSS 量表的“凝视功能”和“面瘫”两项未能归入随访预后 Logistic 方程,存在两个无预测值的项目,而王新等<sup>[3]</sup>的研究结果中仅“面瘫”一项未能列入回归方程,可能与两项研究所列入的患者病变类型的构成比不一致有关,对此仍需大样本的更有代表性的研究进一步证实。

对于 CSS 敏感度的研究,本组大部分项目 SES 值均大于 0.5,表明 CSS 量表有良好的敏感度。

综上,我们初步评价了 CSS 量表,证实该量表在重测信度、评定者间信度、内容一致性、结构效度、效标效度、敏感度上均有较好的表现,但对预后的预测不及 NIHSS,尤其是“凝视功能”和“面瘫”两项未发现和预后明显相关,修订时应作相应调整。当然,本研究样本量仅为 123 例,因样本量的不足等可能给研究结果造成一定的偏倚。因此,研究结果仍有待于在更大人群中进行验证。

### [参考文献]

- [1] 全国第四次脑血管病学术会议(1995). 脑卒中患者临床神经功能缺损程度评分标准[J]. 中华神经科杂志,1996,29:381.
- [2] 曾宇,付林,朱焰,李顺池,李雅萍. 依达拉奉辅助治疗急性脑梗死对其早期神经功能缺损康复的影响[J]. 医学临床研究,2006,23:1782-1784.
- [3] 王新,王拥军,颜振瀛,刘峥. 脑卒中患者临床神经功能缺损评分标准的信度和效度研究[J]. 卒中与神经疾病,1999,6:93-95.
- [4] 王新,王拥军,刘峥,曲镭. 四个脑卒中量表信度与效度的对比研究[J]. 中华物理医学与康复杂志,1999,21:140-143.
- [5] Scandinavian Stroke Study Group. Multicenter trial of hemodilution in ischemic stroke-background and study protocol[J]. Stroke, 1985,16:885-890.
- [6] 刘朝杰. 问卷的信度与效度评价[J]. 中国慢性病预防与控制,1997,5:174-177.
- [7] Sun T K, Chiu S C, Yeh S H, Chang K C. Assessing reliability and validity of the Chinese version of the stroke scale: scale development[J]. Int J Nurs Stud,2006,43:457-463.