

· 临床研究 ·

基于 ICF 的失能评估工具在临床应用中的信度研究

金娟¹ 周莉¹ 毛二莉¹ 刘炜² 李娜² 严娟³ 励建安¹ 刘守国¹¹南京医科大学第一附属医院康复医学科,南京 210029; ²江苏钟山老年康复医院,南京 210000; ³中国社会保障学会医疗保障专业委员会,北京 100000

通信作者:刘守国,Email:liushougou2002@163.com;励建安,Email:lijianan@carm.org.cn

【摘要】 目的 验证基于国际功能、残疾和健康分类(ICF)的功能评定工具在失能评估中的信度。**方法** 选取江苏钟山老年康复医院存在功能障碍的住院患者 371 例参与本研究,以基于 ICF 的功能评定工具对患者进行失能评估,以数字化评定量表(0~10 分)评估条目失能程度。分析失能评定工具的内在一致性、评估者间信度、评估者内信度。**结果** 基于 ICF 的失能评定量表的 Cronbach's α 系数为 0.89,评估者间信度分析显示量表总分的组内相关系数(ICC)为 0.85,量表包含的 20 个条目的 ICC 在 0.78~0.94($P<0.01$);评估者内信度分析显示量表的总分 ICC 为 0.95,除 b230 听力外,量表包含的其它条目 ICC 在 0.72~0.97($P<0.01$)。**结论** 基于 ICF 的失能评定工具与数字评定量表联合在临床失能评估应用中具有良好的内在一致性及评估者间信度和评估者内信度。

【关键词】 国际功能、残疾和健康分类; 功能障碍; 信度**基金项目:**南京市重点临床专科资助课题(2019060002);中国社会保障学会医保专业委员会委托课题(Y2017003)

DOI:10.3760/cma.j.issn.0254-1424.2021.03.007

The reliability of a functioning assessment tool based on the International Classification of Functioning, Disability and Health

Jin Juan¹, Zhou Li¹, Mao Erli¹, Liu Wei², Li Na², Yan Juan³, Li Jianan¹, Liu Shouguo¹¹Department of Rehabilitation Medicine, First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China; ²Jiangsu Zhongshan Geriatric Rehabilitation Hospital, Nanjing 210000, China; ³Medical Security Professional Committee of the Chinese Society for Social Security, Beijing 100000, China

Corresponding authors: Liu Shouguo, Email: liushougou2002@163.com and Li Jianan, Email: lijianan@carm.org.cn

【Abstract】 Objective To investigate the reliability of a tool based on the International Classification of Functioning, Disability and Health (ICF) in the assessment of functioning despite disability. **Methods** A total of 371 patients were assessed using a tool based on the ICF rehabilitation set combined with a numerical rating scale ranging from 0 to 10. The internal consistency, inter-rater reliability and intra-rater reliability were analyzed. **Results** The scale's Cronbach's alpha coefficient was 0.89. The inter-rater correlation coefficient was 0.85, and the correlations among the items ranged from 0.78 to 0.94. The intra-rater correlation coefficient with the scale was 0.95, with the item correlations ranging from 0.72 to 0.97. Only item b230 Hearing function did not correlate well. **Conclusions** The functioning assessment tool based on the ICF rehabilitation set when combined with a numerical rating scale has internal consistency, inter-rater reliability and intra-rater reliability sufficient for use in clinical practice.

【Key words】 International Classification of Functioning, Disability and Health; Dysfunction; Reliability**Funding:** A Nanjing Key Clinical Specialty Funded Project (2019060002); the Chinese Social Security Association Medical Insurance Committee Foundation (Y2017003)

DOI:10.3760/cma.j.issn.0254-1424.2021.03.007

根据国家“健康中国 2030 发展规划纲要”的战略布局,建立以主动健康为抓手的社会保障体系迫在眉睫,该保障体系包括医疗保险、长期照护保险等。在社会保障体系中,老年人群及慢性病患者的功能障碍

(失能)问题是需要关注的重点^[1-3]。失能的来源包括疾病、外伤、先天性疾病、老龄及其它健康问题。由于 92% 的疾病无法以治愈作为结局,故除病死率和发病率之外,功能康复是关键的临床结局指标^[4]。因此,

失能不仅是长期照护保险需要关注的关键问题,也是医疗保险需要关注的要点^[5]。

国际功能、残疾和健康分类(International Classification of Functioning, Disability and Health, ICF)的康复组合是 ICF 研究中心开发的对患者关键功能进行描述的类目集合^[6],对于临床功能评估有重要价值。前期研究^[5]中,已经基于 ICF 康复组合的框架,构建了结构清晰、符合国情的长期照护保险的失能评估工具,从躯体结构和功能、个体活动与社会交往三个维度分析失能状况。本研究旨在验证基于 ICF 的失能评定工具在临床应用中的量表内一致性以及评估者间与评估者内信度。

资料与方法

一、研究对象

纳入标准:①年龄 ≥ 18 岁的患者;②患者存在功能障碍;③神志清楚、具有认知能力,可进行日常的言语沟通,能配合评估;④自愿参加,并签署知情同意书。

排除标准:①病情危重、生命体征尚未平稳的患者;②存在认知障碍。

选取 2018 年 9 月至 2019 年 9 月江苏钟山老年康复医院收治且符合上述标准的存在功能障碍的患者 371 例,其中男 245 例,女 126 例,年龄 18~87 岁,平均 (45.3 ± 19.6) 岁。

二、研究方法

(一)研究工具开发

课题组在构建初始工具时,基于前期的文献比对,重点考虑老年人失能评估涉及的功能领域:①世界卫生组织进行世界卫生调查^[7]时提出的人类日常活动中涉及最普遍的 8 个功能领域,即睡眠与驱力、情感、疼痛、活动能力、自理能力、人际交往、视力、认知;②美国功能独立性评估(functional independence measurement, FIM)量表^[8]提出的 6 个功能领域,即自我照顾、括约肌控制、移动能力、运动能力、交流、社会;③ ICF 通用组合^[9]6 个功能领域,即能量与驱力、疼痛、情感、自理功能、活动功能、社会功能,由于最后一项 b850 有报酬就业不适合进行老年人评估,故不做考虑;④中国残疾人实用评定标准^[10]包含 5 个领域,即视力残疾、听力言语残疾、智力残疾、肢体残疾、精神病残疾。综合上述特征,课题组最终确定了失能评定的 8 个功能领域,即活动能力、自理能力、睡眠与精神状态、情感、疼痛、认知、感官能力、人际交往或社会功能。

选择功能条目时,通过 8 个功能领域与 ICF 康复组合^[6]、ICF 脑卒中简要核心组合、ICF 肌骨系统疾病简要核心组合、ICF 心肺系统疾病简要核心组合、ICF 老年人非急性期疾病简要核心组合的类目进行比对,

并经多轮专家共识与论证会,最终确定了本研究的 20 个功能条目,20 个条目分别是 b455 活动耐力、d450 步行、d455 到处移动(非步行移动)、b525 排便、b620 排尿、d230 执行日常事务、d510 盥洗自身、d520 护理身体各部、d530 如厕、d540 穿脱、d550 吃、b130 能量与驱力(精气神)、b134 睡眠、b152 情感、b280 痛觉、b114 定向、b144 记忆、d710 基本人际交往、b210 视力、b230 听力。

课题组还编制了基于 ICF 分类的评价标准,采用数字评估量表(numerical rating scale, NRS)对条目进行 0~10 分连续性评分,即 0 分为没有问题、10 分为完全功能障碍;考虑到实操中评价的一致性,课题组又将 0~10 分的连续分值分为 5 个判定区间,即 0 分为没有依赖、1~3 分为轻度依赖、4~6 分为中度依赖、7~9 分为重度依赖、10 分为完全依赖,并制订各区间具体评价细则。评价时在区间选择基础上再进行连续变量打分,确保评价者间一致性。量表总分越高,说明失能程度越严重。

初始工具经过 2 轮专家论证,第一轮专家为康复专家 3 人、护理专家 3 人、保险专家 2 人、政策制订专家 1 人,第二轮专家为康复及临床医学专家 5 人、护理专家 3 人、保险专家 4 人、政策制订专家 6 人,专家对条目名称、含义、条目评价标准进行讨论修改,最终构建了基于 ICF 的失能评估工具。

(二)资料收集方法

课题组根据专家共识会内容制定标准的失能评估手册,以用于评估者培训。江苏钟山老年康复医院的护理人员参与本研究的数据搜集。参与评定的护士为注册护士,有 3 年以上的工作经验。研究开始前,所有参研护士均接受面授培训,培训内容主要包括资料搜集流程、ICF 失能评估工具产生背景、失能评估工具包含的条目含义、条目功能等级划分,评估手册对评估关键点进行了界定。在评估过程中若患者因意识障碍、听力、语言障碍或因其他因素不能回答问题,可由其照护者代为解释老年人各条目相关的身体功能现状,由评估者进行打分,对意识不清者的 b134 睡眠、b152 情感、b280 痛觉、b130 能量与驱力(精气神)四个主观条目计 10 分。

三、统计学方法

采用 Excel 导出数据,导入 SPSS 20.0 版统计软件进行数据分析。采用 $(\bar{x} \pm s)$ 的方式计算失能总分。参与本研究的部分患者在 1 周内接受 2 次同一评估者的失能评定,该部分数据用于评估者内信度(重测信度)分析;参与本研究的部分患者同时接受 2 个不同评估者的评定,该部分数据用于评估者间信度分析。以组内相关系数(interclass correlation coefficient, ICC)进行评估者

间信度和评估者内信度评价^[11];以 Cronbach's α 系数检验失能量表的内在一致性。检验水准 $\alpha=0.05$ 。

结 果

- 一、失能量表的内在一致性
- 量表的内部一致性 Cronbach's α 系数为 0.89,量表内在一致性良好。
- 二、评估者间信度
- 163 例患者接受了 2 位不同评估者的评估,评估者间信度分析发现量表总分的 ICC 为 0.85,量表 20 个条目的 ICC 在 0.78~0.94,提示这 2 位评估者间的信度良好($P<0.01$)。

表 1 163 例患者基于 ICF 的失能评估工具的评估者间信度			
条目	ICC	95%可信区间	P 值
b455 活动耐力	0.92	0.892~0.943	<0.01
d450 步行	0.78	0.709~0.839	<0.01
d455 非步行移动	0.93	0.905~0.950	<0.01
b525 排便	0.88	0.837~0.913	<0.01
b620 排尿	0.88	0.838~0.913	<0.01
d230 执行日常事务	0.87	0.830~0.909	<0.01
d510 盥洗自身	0.86	0.813~0.899	<0.01
d520 护理身体各部	0.87	0.822~0.904	<0.01
d530 如厕	0.87	0.829~0.908	<0.01
d540 穿脱	0.90	0.864~0.928	<0.01
d550 吃	0.88	0.839~0.913	<0.01
b130 精气神	0.91	0.875~0.934	<0.01
b134 睡眠	0.86	0.815~0.902	<0.01
b152 情感	0.83	0.776~0.879	<0.01
b280 痛觉	0.87	0.832~0.910	<0.01
b114 定向	0.82	0.766~0.872	<0.01
b144 记忆	0.93	0.902~0.948	<0.01
d710 基本人际交往	0.93	0.901~0.948	<0.01
b210 视力	0.94	0.917~0.956	<0.01
b230 听力	0.81	0.753~0.865	<0.01
总分	0.85	0.801~0.892	<0.01

- 三、评估者内信度
- 152 例患者在 1 周内接受了 2 次同一评估人员的失能评估,评估者内信度分析发现量表总分的 ICC 为 0.95。除条目 b230 听力的 ICC 为 0.13 外,量表其它条目的 ICC 在 0.72~0.97,提示评估者内的信度良好($P<0.01$)。

讨 论

本研究主要探讨基于 ICF 的失能评定工具在临床失能人群评定中的评估者间信度和评估者内信度,为失能工具的临床应用提供理论依据,研究结果提示该工具具有良好的量表内在一致性及评估者间信度和评估者内信度。

表 2 152 例患者基于 ICF 的失能评估工具的评估者内信度			
条目	ICC	95%可信区间	P 值
b455 活动耐力	0.80	0.738~0.854	<0.01
d450 步行	0.89	0.863~0.926	<0.01
d455 非步行移动	0.82	0.755~0.864	<0.01
b525 排便	0.94	0.924~0.960	<0.01
b620 排尿	0.94	0.920~0.958	<0.01
d230 执行日常事务	0.81	0.748~0.860	<0.01
d510 盥洗自身	0.88	0.835~0.910	<0.01
d520 护理身体各部	0.91	0.872~0.931	<0.01
d530 如厕	0.93	0.898~0.946	<0.01
d540 穿脱	0.91	0.881~0.936	<0.01
d550 吃	0.94	0.915~0.954	<0.01
b130 精气神	0.72	0.629~0.787	<0.01
b134 睡眠	0.72	0.633~0.790	<0.01
b152 情感	0.74	0.663~0.809	<0.01
b280 痛觉	0.83	0.777~0.877	<0.01
b114 定向	0.97	0.961~0.979	<0.01
b144 记忆	0.86	0.810~0.896	<0.01
d710 基本人际交往	0.94	0.920~0.957	<0.01
b210 视力	0.77	0.702~0.832	<0.01
b230 听力	0.13	-0.028~0.288	<0.01
总分	0.95	0.942~0.969	<0.01

本研究中,所开发的失能评定工具很大程度上参考了 ICF 康复组合,但由于东西方文化背景的巨大差异,ICF 康复组合无法直接作为失能评定工具进行临床应用。ICF 康复组合的部分条目在中国的医疗背景下不适合评估,如 b640 性功能、d640 做家务、d850 获得有报酬的就业、d920 娱乐与休闲;康复组合也有部分条目在维度上存在一定的重叠,如 d410 改变身体的基本姿势、d450 保持一种身体姿势、d420 移动自身;而在失能领域中,有部分非常重要的类目康复组合没有包含,如 b210 听力、b230 视力、b114 定向、b144 记忆,故本研究在 ICF 康复组合的基础上通过专家共识会议的方式与 ICF 脑卒中简要核心组合、ICF 肌骨系统疾病简要核心组合、ICF 心肺系统疾病简要核心组合、ICF 老年人非急性期疾病简要核心组合以及中国残疾人实用评定标准进行综合,开发了目前的失能评定工具。

以往的研究提示评估者间信度受到 ICF 类目内涵和特征、ICF 限定值、评定方式、评估者专业背景和经验的影响^[12],影响 ICF 相关评定工具的信度^[13-14]。Reinhardt 等^[15]关于 ICF 通用组合临床应用的研究发现,中国的评估者对于通用组合条目的理解存在差异,影响评估者间信度,且评估者在应用 ICF 限定值进行评估时无法很好地区分不同限定值之间的界限。李琨等^[16]采用 Rash 分析检验《脊髓损伤护理相关 ICF 组合》中身体功能限定值反应选项,结果发现评估者无法细致和明确地区分 ICF 限定值中的轻、中和重度损伤,提示使用者应更注重 ICF 限定值量化标准的研究;

而评估者对条目定义的理解、评定的严格程度以及评估者本身的专业背景和经验都会对评定结果产生影响。针对上述问题,中国康复医学会与 ICF 研究分中心于 2014 年召开了针对 ICF 康复组合的专家共识会,制订了基于中国国情及文化背景的 ICF 康复组合简洁、描述直观的条目^[17]。在此背景下,2015 年本课题组在全国 20 个省份展开了 ICF 通用组合的多中心大样本队列研究,研究结果提示通用组合与数字评定量表结合有良好的评估者间信度和评估者内信度^[18]。高焱等^[19]研究发现,在 ICF 康复组合条目评定标准量化后进行临床应用有良好的信度和效度。以上的研究结果提示在 ICF 的临床应用中加强条目的评定量化标准、培训评估者、加强评定过程质量管理等措施,能够提高 ICF 评定工具的评估者间信度和评估者内信度。

本研究前期通过多轮专家共识会制订了失能评估工具的条目评定内容标准和条目评定等级标准,并撰写了失能评估使用手册,并在数据采集前进行了统一培训,本研究的评估者间信度分析发现 20 个条目的 ICC 在 0.78~0.94,提示评估者间一致性良好;评估者内信度分析发现 19 个条目的 ICC 在 0.72~0.97,亦提示评估者内一致性良好,但在评估者内信度分析中,b230 听力的 ICC 为 0.13,明显低于其它条目,这主要是因为原始数据中高达 96.7% 的 b230 听力评分为 0,说明轻微的数据扰动便可以大幅度降低 b230 听力的评估者内信度。

综上所述,通过对评估内容和评估方法的标准化、加强对评估者的培训、对评定过程进行质量监控,基于 ICF 康复组合的失能评估工具在临床应用中具有较好的评估者间信度和评估者内信度。

参 考 文 献

- [1] 陈颖,马丽霞,裴慧丽,等.不同失能程度老年人居家养老照护服务项目需求调查[J].中国实用神经疾病杂志,2016,19(1):38-40. DOI:10.3969/j.issn.1673-5110.2016.01.019.
- [2] 蔡菲菲,张泓,刘晓霞,等.武汉市失能老年人长期照护现状调查[J].护理研究,2017,31(32):4138-4140. DOI:10.3969/j.issn.1009-6493.2017.32.026.
- [3] 蒋华,俞洁毅,刘涛.失能老人长期照护现状与建议[J].全科护理,2019,17(22):2712-2714. DOI:10.12104/j.issn.1674-4748.2019.22.007.
- [4] Stucki G,邱卓英,励建安,等.为中国康复工作中全面应用 ICF 而努力[J].中国康复理论与实践,2011,17(1):5-10. DOI:10.3969/j.issn.1006-9771.2011.01.002.
- [5] 赵元萍,刘守国,张霞,等.长期照护保险老年人失能等级划分研究[J].中国康复医学杂志,2019,34(10):1224-1228. DOI:10.3969/j.issn.1001-1242.2019.10.016.
- [6] Proding B, Cieza A, Oberhauser C, et al. Toward the International Classification of Functioning, Disability and Health (ICF) rehabilitation set: a minimal generic set of domains for rehabilitation as a health

- strategy[J].Arch Phys Med Rehabil,2016,97(6):875-884. DOI:10.1016/j.apmr.2015.12.030.
- [7] World Health Organization. Health statistics and information systems; WHO World Health Survey[R].http://www.who.int/healthinfo/survey/en/.
- [8] Brown AW, Therneau TM, Schultz BA, et al. Measure of functional independence dominates discharge outcome prediction after inpatient rehabilitation for stroke[J]. Stroke,2015,46(4):1038-1044. DOI:10.1161/STROKEAHA.114.007392.
- [9] Cieza A, Oberhauser C, Bickenbach J, et al. Towards a minimal generic set of domains of functioning and health[J].BMC Public Health, 2014,(14):218. DOI:10.1186/1471-2458-14-218.
- [10] 中华人民共和国中央人民政府.中国残疾人实用评定标准[EB/OL].[2006-12-02].http://www.gov.cn/jztzl/gacjr/content_459939.htm.
- [11] Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial[J].Tutor Quant Methods Psychol,2012,8(1):23-34. DOI:10.20982/tqmp.08.1.p023.
- [12] 高焱,燕铁斌,尤黎明,等.《国际功能、残疾和健康分类》相关测量工具信度研究的现状与进展[J].中华物理医学与康复杂志,2016,38(10):789-792. DOI:10.3760/cma.j.issn.0254-1424.2016.10.021.
- [13] Starrost K, Geyh S, Trautwein A, et al. Interrater reliability of the extended ICF core set for stroke applied by physical therapists[J].Phys Ther,2008,88(7):841-851. DOI:10.2522/ptj.20070211.
- [14] Uhlig T, Lillemo S, Moe RH, et al. Reliability of the ICF Core Set for rheumatoid arthritis[J].Ann Rheum Dis,2007,66(8):1078-1084. DOI:10.1136/ard.2006.058693.
- [15] Reinhardt JD, Zhang X, Proding B, et al. Towards the system-wide implementation of the international classification of functioning, disability, and health in routine clinical practice: Empirical findings of a pilot study from mainland china[J].J Rehabil Med,2016,48:515-521. DOI:10.2340/16501977-2102.
- [16] 李琨,燕铁斌,尤黎明,等.《脊髓损伤护理相关 ICF 组合》中的身体功能限定值反应选项的 Rasch 分析[J].中华物理医学与康复杂志,2017,39(2):109-113. DOI:10.3760/cma.j.issn.0254-1424.2017.02.007.
- [17] Proding B, Reinhardt JD, Selb M, et al. Towards system-wide implementation of the International Classification of Functioning, Disability and Health (ICF) in routine practice: developing simple, intuitive descriptions of ICF categories in the ICF generic and rehabilitation set[J].J Rehabil Med,2016,48(6):508-514. DOI:10.2340/16501977-2066.
- [18] Liu S, Reinhardt JD, Zhang X, et al. System-wide clinical assessment of functioning based on the International Classification of Functioning, Disability and Health in China: interrater reliability, convergent, known group, and predictive validity of the ICF generic-6[J].Arch Phys Med Rehabil,2019,100(8):1450-1457. DOI:10.1016/j.apmr.2018.11.014.
- [19] 高焱,燕铁斌,尤黎明,等.《国际功能、残疾和健康分类康复组合》在非急性期患者中使用的信度与效度分析[J].中国康复医学杂志,2019,34(10):1193-1198. DOI:10.3969/j.issn.1001-1242.2019.10.011.

(修回日期:2020-12-28)

(本文编辑:汪 玲)