

基于 ICF 核心分类集类目开发临床等级评定量表

Eva Grill, Gerold Stucki

王朴¹, 陈迪², 田峻¹译 邱卓英², 张爱氏² 审校

[摘要] 目的 本研究在急性期医院背景下,就上下肢的移动性为例,目的是检验 ICF 类目的临床评定等级是否能够整合成一种多参数的等级评定量表,并且该量表能够根据特定患者的临床问题和康复目标提供有效可靠的临床评估。方法 对急性期医院患有肌肉骨骼系统疾病的患者采用多中心前瞻性队列心理测量方法对数据进行统计分析。结果 234 例肌肉骨骼状况患者(平均年龄 56 岁,50%女性。44%患者诊断涉及下肢,18%患者诊断涉及髋关节,18%患者诊断涉及上肢,16%患者诊断涉及脊柱)。调整不同测量功能项目之后,建立针对上肢或下肢功能的两个独立量表。建立的量表中有 10 个(上肢)和 8 个(下肢)测试项目有很好的拟合度。结论 本研究的结果表明可基于 ICF 的核心分类集类目来开发等级评定量表。这可能是一个有希望的方法,因为在现有的领域里从心理测量方面来讲还没有可靠的测量方法。

[关键词] 结果评价(卫生保健);项目反应理论;分类;身体功能;健康状况测量;活动受限

[中图分类号] N3;R49 **[文献标识码]** A **[文章编号]** 1006-9771(2011)02-0110-05

[本文著录格式] Grill E, Stucki G. 王朴,陈迪,田峻,等译. 基于 ICF 核心分类集类目开发临床等级评定量表[J]. 中国康复理论与实践,2011,17(2):110—114.

1 前言

人类功能是与个体和社会相关的医学核心理论。因此,对人类功能的准确测量为有效的患者保健提供了重要基础^[1]。在康复实践和研究中,人类功能的测量是患者功能评定、康复目标设定、康复服务、干预措施分配、干预管理和结局评价的基础^[2-3]。

在不同的卫生保健系统和服务中,对于卫生专业人员来说,有效的功能测量需要一个简单的并且基于功能的通用语言。应用测量方法根据患者的问题来说明和评估康复的目标也是非常有用的。在康复研究中,测量病因学上应该保持中立,以比较不同康复状况的测量结果;并且测量应该能够准确地有效地反映康复目标和确保康复目标的反应性。

目前的测量方法并不能完全满足以上需求。总体来讲,这些测量方法不简单,因此测量方法只在研究和卫生统计中应用,而不能应用于日常的临床实践。由于缺乏通用语言和功​​能分类,目前测量假定的深层结构,测量项目^[4]的定义、量表和分量表的定义和命名都存在很大差异^[5]。

功能独立性评定(FIM)^[6]或 Barthel 指数^[7]是康复领域广泛认可和使用的测量方法,例如测量资源利用或将患者分配到康复服务领域。然而,它们不是综合性地功能评定^[8],针对不同的健康状况也不能保持不变^[9]。更为重要的是,它们在评定患者功能问题和设定康复目标时都不够具体。此外,现阶段用于康复领域的测量方法是一种分类的测量,因此不适合于参数统计,例如计算平均数、标准差或者甚至变化分数^[10]。

因此,有必要建立一种实用的方法,以便使医疗人员和科研人员研发一种等距量表的临床测量方法,这种测量临床中简单易行,并且易于实施,能够在临床实践中具体地说明不同患者的问题和康复目标。正因为此,我们首先需要统一的架构和分类,这种架构和分类能综合涵盖与功能有关的人类体验。功能有着不同的深层健康状况,并且与个人和环境因素有着广泛

的联系。其次,我们需要显示基于通用分类类目的临床评定等级是否能够转换成参数型的等级评定量表,提供一个特定患者状况和康复目标的可靠评估。

国际功能、残疾与健康分类(ICF)^[11]已成为全球公认的描述和分类功能、残疾和健康的标准语言和康复架构^[12]。ICF 由 4 个成分组成——身体功能、身体结构、活动和参与以及环境因素,分为 1400 多个不同的按等级排列的类目。通过联合应用 ICF 和最近开发的项目反应理论,从原则上讲使得开发与患者病情相适应的临床测量工具进行测量具有可行性,这个测量工具通过简单地整合 ICF 类目以及相关的限定值来开发临床评定等级。ICF 为每个类目提供限定值范围 0~4 来说明问题的严重程度。基于 ICF 限定值给定的定义和明确的界限,专家可用来直接测量人类功能。

在开发与患者问题和康复目标相适应的临床测量时,人们要依据 ICF 的某些功能类目,这些功能类目标针对特定健康状况^[13],能综合性地反映功能的各个方面,或涵盖卫生保健情境如急性期医院或早期急性康复机构^[14]中功能的各个方面。那些类目即所谓的 ICF 核心分类集(ICF Core Sets)是通过一个世界卫生组织所开展的国际性合作而开发的。

例如开发量表用于评定在急性期医院中患者上下肢运动功能,可以从急性期医院适用的 ICF 核心分类集中选取与上下肢相关的 ICF 类目。然后应用“项目反应理论”来分析挑选出来的类目是否确实包括了共同的根本特质(例如移动),可以编制一个量表,其中挑选出来的类目适合于假定的特质,并且包括了典型患者所表现的所有能力。

本文的目的是检验 ICF 类目的临床等级评定是否能够转换为开发的参数型的等级评定量表,并且使用一个急性期医院情境中上下肢移动项目的例子来说明和评估特定患者情况和康复目标。

本研究的具体目标为:①检查一个与特定状况相关的以及从肌肉骨骼系统急性期 ICF 核心分类集挑选出来的类目子集是否能编制成一个等级评定量表,该量表具有 Rasch 模型不随个人因素变化的属性;②确定该子集的类目能够并且应该包括在一个临床等级量表中,使其能够简便和有效地在临床实践中进行测量。

译者单位:1. 武汉大学中南医院康复医学科,湖北武汉市 430071; 2. 中国康复研究中心康复信息研究所,北京市 100068。译者简介:王朴(1981-),男,湖北武汉市人,博士研究生,主要研究方向:神经系统疾病康复。

2 方法

2.1 研究设计和患者 本心理测量研究的数据采集来自在 2005 年 1 月到 10 月就诊的肌肉骨骼系统状况的患者,作为急性医院的一个更大的前瞻性队列研究的一部分。在瑞士的 20 个急性医院的骨科、风湿科和外科来征集患者。所有患者均提供书面的知情同意书。该研究经每个参与医院的伦理委员会批准。

2.2 测量 ICF 是属于 WHO 国际分类家族的多目的的分类系统用于报告和组织健康和健康相关领域的广泛信息。ICF 分为两个部分,每一部分有两种成分:第一部分包括身体功能(b)、身体结构(s)以及活动与参与(d)。第二部分为背景性因素,包括环境因素(e)和个人因素。在 ICF 分类中,字母 b、s、d 和 e 分别指 ICF 的各个组成成分,紧跟字母的是编码数字,开始是章数(1 位数字),接着是二级水平(2 位数字)和第三和四级水平(各为 1 位数字)。举一个例子,步行是一个二级类目,编码为 d450。这表明步行是活动与参与成分(d),并且属于活动和参与成分的第四章 d4 活动。第四章的其他类目如 d455 到处移动和 d465 利用设备到处移动。

我们在肌肉骨骼系统急性期 ICF 核心分类集^[15]和附加的二级水平类目中选择信息,这些额外的类目被包括进来是因为它们与物理治疗师相关^[16]。

如果患者没有限制,类目的评定使用限定值 0,如果患者有一定限制,类目的评定使用限定值 1。如果信息不足以评定等级,则观察值被视为缺失。该方法已经显示了可接纳的信度^[17]。分级的 ICF 类目形成了一个针对特定患者有关活动受限的类目分布表。然而,这些等级并不能形成一个单一的等级评定量表。

2.3 数据收集 本研究对急性事件或手术后 24~48 h 的患者进行评定。由经过 ICF 应用和原则培训的物理治疗师来征集和评定患者。

2.4 数据分析 本研究的统计分析应用 Rasch 单维测量模型的理论^[18]。该理论解释了等级测量量表的项目表现如何以及测试者的能力和量表属性之间有何种联系。Rasch 模型可以对

量表进行标准化,量表的测量项目可以是两选或多选,可以将单个项目的测量值相加形成一个总分,因此可以按照统计学的方法计算变化分数和影响大小。Rasch 模型的基本假设是:与困难项目相比,人们更可能通过简单项目;并且较高能力的人比较低能力的人更可能通过一个项目。通过一个项目的概率被定义为个人能力 θ 和项目难度 b 之间的一个 Logistic 函数。

$$P_m = \frac{\exp(\theta_n - b_i)}{1 + \exp(\theta_n - b_i)}$$

P_m 为个体 n 通过项目 i 的概率; θ 表示个人能力; b 表示在一个等级量表连续范围内的项目难度,理论上是从 $-\infty$ 到 $+\infty$ 。

如果一个项目集合的观察反应值(the observed responses)与 Rasch 模型相拟合,也就是说简单项目可以被更多的患者通过并且能力越强的人通过的项目越多,然后这个项目集是单维的并且构成了一个 Rasch 量表。假说可通过专门的拟合度统计来验证。然后可以假设用这个量表的项目来测量单个特质。这个假设通过分析项目的局部独立性来确定。如果深层的单个特质(即所谓的 Rasch 因素)被删除,那就应该没有关联存在。此外,另一个测量的重要特质是,某一量表针对不同类别的患者,所起的作用一致,也就是说对于男性和女性,年轻人和老年人来说,一个项目的难易程度相当。因此,可在不同个人因素水平之间检查每个项目的区分项目功能(DIF)。

本研究举例说明如果 ICF 核心分类集的子集是单维的,我们可以进行一系列的分析。因为 ICF 的维度非常复杂并且本研究的目的是调查是否有可能建立有意义的代表康复目标的等级评定量表,因此我们将对 ICF 类目的进一步分析限制在“活动与参与”成分内。

在建立量表时,首先根据项目内在的通用概念来合理选择项目。然后我们关注两个独立的类目集,一个集合包括与上肢移动相关的类目,另一个类目集与下肢移动相关。表 1 显示有关的 ICF 类目以及通过统一标准建立的两个子集。此外,个人因素的诊断、年龄和性别作为协变量用于 DIF 进一步的分析。

表 1 首先被考虑收入两个量表的 17 个 ICF 类目

ICF 编码	类目描述	集合 1;上肢	集合 2;下肢	人数	受限百分率(95%置信区间)
d240	控制应激和其他心理需求	×	×	178	19.7(14.5;26.1)
d410	改变身体的基本姿势	×	×	206	75.2(68.9;80.6)
d415	保持一种身体姿势	×	×	196	69.9(63.1;75.9)
d420	移动自身	×	×	203	42.9(36.2;49.7)
d430	举起和搬运物体	×	×	174	87.4(81.6;91.5)
d440	精巧手的使用	×	×	229	16.6(12.3;22)
d445	手和手臂的使用	×	×	226	23.5(18.4;29.4)
d450	步行	×	×	199	78.4(72.2;83.5)
d4551	攀登	×	×	155	71.6(64.1;78.1)
d460	在不同地点到处移动	×	×	168	64.9(57.4;72.7)
d465	利用设备到处移动	×	×	132	65.2(56.7;72.7)
d510	盥洗自身	×	×	230	80.4(74.8;85)
d520	护理身体各部	×	×	230	77.4(71.6;82.3)
d530	如厕	×	×	219	37.0(30.9;43.6)
d550	吃	×	×	233	22.3(17.4;28.1)
d560	喝	×	×	234	15.8(11.7;21)
d760	家庭人际关系	×	×	228	5.7(3.4;9.5)

然后分别分析这两个类目集是否与 Rasch 模型相拟合。拟合度通过整体拟合度统计和每个测量项目拟合度统计进行分析。我们使用 3 个整体拟合度统计,其中 2 个测量项目-个人相

互作用的统计(测量项目和个人拟合)是 Z 分布即平均数为 0,标准差为 1,表明非常拟合;测量项目-特质相互作用的统计为 χ^2 分布。通过测量项目-特质相互作用统计的一个非显著水平

来表示拟合度,表明测量项目等级与项目深层特质的水平无关,是恒定的。个体项目拟合统计表明一个测量项目是否能够收入等级量表。我们报告了拟合残差作为来自模型的个体和测量项目差异总体,残差范围应该在 ± 2.5 内;并且在经过多重测试校正之后, χ^2 统计应该是非显著的。

如果要分析等级量表分值和个人因素相独立,我们可以考察在某人既定的能力水平上,在个人因素层面上通过某一项目的概率是否相等。深入思考发现,当调查类目步行(项目)时,如果 Rasch 假设成立,那么一个下肢移动(个人能力)有一定水平限制的女性,与有相同水平限制的男性相比,具有相同的步行(通过这个项目)能力的概率。因此,在年龄和性别层面对每个测量项目的 DIF 进行调查。每个项目的个人-反应残差表明在既定能力水平上,某人能力反应和在该能力水平上期望的能力反应的差异程度。在经过多重测试的适当校正之后,将个人因素和组间因素作为自变量的 ANOVA 分析得出了一个非显著的结果。如果有些项目上有一致的 DIF,那么不论个体的能力水平如何,在任意层面检测到的反应差别是相同的,那些测量项目可以根据个人因素的改变而作出调整。相反,对于不一致的 DIF 项目,即在任意层面的反应差别很大,应该不予考虑。本研究我们通过诊断把样本分成有上下肢、脊柱或髋关节状况,通过年龄把样本分成两个年龄组还有性别分组,考察了 DIF。

通过主成分分析对最后一个单维测试即个人-项目反应残差进行了分析^[19]。

Person 分离指数可被用来表示测量项目定义的患者在量表中的分布状况,因此表明了量表可以很好的区分不同患者。当所有个体在相同水平使用同一方式时,Person 分离指数接近 0,而个体在项目连续过程中分布得越多,Person 分离指数越接近 1。

显著性水平被设置为 0.05。Bonferroni 方法被用来进行多重测试校正,产生一个 $0.05/k$ 的显著水平,这里的 k 是指同时进行的测试数目。

表 2 上肢量表类目和 Rasch 模型的拟合度(按位置排序)

ICF 编码	类目描述	位置	拟合误差	χ^2	P	受限%
d420	移动自身	-4.769	-0.335	0.742	0.863	12.2
d530	如厕	-3.307	-0.280	0.667	0.881	22.5
d240	控制应激和其他心理需求	-3.183	0.032	5.754	0.124	20.7
d410	改变身体的基本姿势	-2.206	-0.769	4.915	0.178	30.0
d560	喝	0.089	0.412	2.185	0.535	59.5
d440	精巧手的使用	0.735	-0.918	4.655	0.199	68.3
d430	举起和搬运物体	2.462	0.047	3.410	0.333	84.2
d550	吃	2.517	-0.465	0.466	0.926	85.7
d520	护理身体各部	3.256	-0.378	1.324	0.724	90.5
d510	盥洗自身	4.407	-0.399	0.936	0.817	92.9

这些类目代表了量表的最终版,以项目的难度排序(移动自身是最小可能受限的类目,盥洗自身是最可能受限的类目)。

为了检测 DIF,我们对年龄组和性别组差异进行双向分析。对于年龄或性别,既没有一致也没有不一致的 DIF。根据前后分析,该结果表明测量项目的难度在年龄和性别中比较稳定。

主成分分析显示无残差联系,即在 Rasch 因素排除之后没有第二个因素能够提供拟合度的证据。个人-项目位置如图 1 所示,表明在个人能力整体范围内类目显示了一个良好的分布,没有较大的地板或天花板效应。Person 分离指数为 0.83,为优等。该图表明个体和能力与测量项目难度的关系。测量项目难度越小,该测量项目通过的可能性越大。信息曲线表明测量项目提供的信息的比例。在这里,曲线的下降深度是由两个测量

使用 RUMM 2020 软件包(Perth, RUMM 实验室)进行 Rasch 分析,使用 SAS V9.1 软件(Cary, NC)进行二元统计分析。

3 结果

本研究共纳入 234 例患者,平均年龄为 56 岁(中位数 58 岁,年龄范围 18~91 岁),女性 50%。44% 的患者诊断主要涉及下肢,如膝关节置换术、肌腱断裂或骨折;18% 的患者诊断涉及髋关节,主要为髋关节置换术;18% 的患者诊断主要涉及上肢;16% 的患者诊断涉及脊柱。活动和参与成分的 17 个 ICF 类目普遍受限如表 1 所示。

患者受限最多的 ICF 类目是举起和搬运物体,有 87% 的患者该类目受限;受限最少的是家庭人际关系,有 6% 的患者该类目受限。

3.1 上肢活动量表 表 1 显示最初考虑收入上肢活动量表中的 12 个类目,其中 4 个类目也被考虑收入下肢活动量表中。

几乎所有类目均显示不完全拟合 Rasch 模型,导致高度显著性的项目-特质相互作用。根据诊断有一致和不一致的 DIF,即在不同的健康状况下,量表所起的作用不同。因此,我们将以上数据集拆分为 3 组:诊断涉及上肢的患者,诊断涉及下肢的患者,诊断涉及脊柱的患者。3 组的患者没有重叠。然后我们重新分析了诊断仅涉及上肢状况的患者组的量表。将类目 d445“手和手臂的使用”从分析中排除,因为所有患者均在该类目上受限。接着将类目 d760“家庭人际关系”从分析中排除,因为该类目的项目拟合度不足。剩余 10 个类目的平均项目拟合度为 -0.305 (标准差 0.392),个人拟合度为 -0.268 (标准差 0.318)。项目特质相互作用表明对于不同残疾水平的患者该量表是恒定的(自由度 30, $\chi^2 = 25.053$, $P = 0.722$)。没有类目显示显著的项目特质拟合度统计($P = 0.003$),也即没有任何失配并且所有项目的拟合残差均在 -2.5 到 2.5 之间。患者最可能受限的类目是 d510“盥洗自身”,最不可能受限的类目是 d420“移动自身”(项目统计和位置见表 2)。

项目之间相关的较大差距造成的。见图 1。

3.2 下肢活动量表 表 1 显示最初考虑收入下肢活动量表中的 9 个类目,其中 4 个类目也被考虑收入上肢活动量表中。

几乎所有类目均显示不完全拟合 Rasch 模型,导致高度显著性的项目-特质相互作用。和上肢活动类目一样,由于诊断而出现了一致和不一致的 DIF。因此我们将数据集进行拆分,只对诊断涉及下肢和髋关节的患者进行数据分析。类目 d460“在不同地点到处移动”因为不充分拟合而被排除。

剩余的 8 个类目显示出一个显著的项目-特质拟合统计。类目 d420“移动自身”在年龄上显示出显著一致的 DIF。随后我

们就年龄对该类目进行拆分。

9 个类目的项目拟合平均值为 -0.587(标准差 0.782), 个人拟合平均值为 -0.383(标准差 0.311)。项目特质相互作用表明对不同残疾水平的患者量表比较稳定(自由度 27, $\chi^2 = 29.754, P = 0.325$)。没有类目显示显著的项目特质拟合统计

($P = 0.003$), 即没有任何的失配。所有项目的拟合残差均在 -2.5 到 2.5 之间。患者最可能的受限类目是 d450“步行”, 最不可能受限的类目是 d420“家庭人际关系”(项目统计和位置见表 3)。

表 3 下肢/腕关节量表和 Rasch 模型的类目拟合度(按位置排序)

ICF 编码	类目描述	位置	拟合误差	χ^2	P	受限(%)
d760	家庭人际关系	-4.737	-0.063	5.459	0.141	5.0
d240	控制应激和其他心理需求	-3.662	0.963	6.507	0.089	17.9
d420	自我转移(年龄≤55)	-2.143	-0.582	1.320	0.724	30.6 ^a
d420	自我转移(年龄>55)	-0.534	-1.461	3.933	0.269	71.1 ^b
d465	利用设备到处移动	0.112	-1.291	4.665	0.198	71.8
d415	保持一种身体姿势	1.242	-0.773	4.329	0.228	86.2
d4551	攀登	2.718	-0.327	1.568	0.667	94.4
d410	改变身体的基本姿势	3.182	-1.004	1.119	0.773	93.2
d450	步行	3.822	0.747	0.855	0.836	96.7

这些类目代表了量表的最终版, 以项目的难度排序(家庭人际关系是最小可能受限的类目, 步行是最可能受限的类目)。a: 年龄在 55 岁和 55 岁以下的患者; b: 年龄在 55 岁以上的患者。

主成分分析显示无残差联系, 即在 Rasch 因素排除之后没有第二个因素能够提供拟合度的证据。个人-项目位置见图 2。该图显示, 表明在个人能力整体范围内类目显示了一个良好的

分布, 没有较大的地板或天花板效应。Person 分离指数为 0.51, 为中等。解释见图 1。

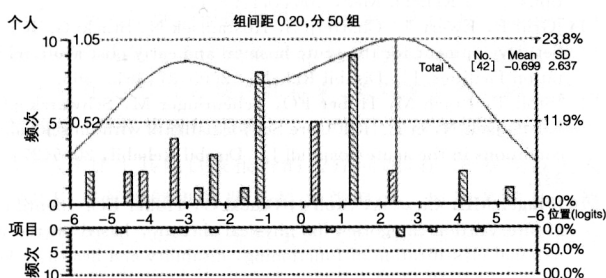


图 1 上肢量表个人-项目位置分布图

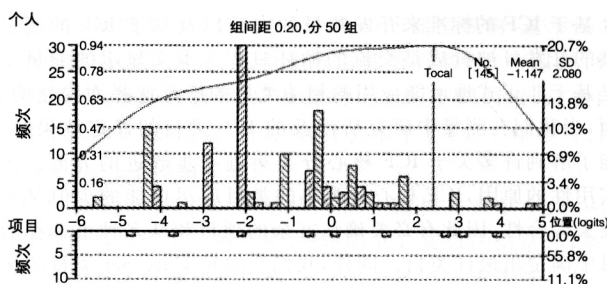


图 2 下肢量表个人-项目位置分布图

4 讨论

本研究通过使用 ICF 限定值简单地整合了与 ICF 类目相关的临床评分等级, 研究结果证明了为开发与临床测量相关的问题的理论。图示过程也表明了开发上下肢活动的临床测量方法可能作为一种实用的方法, 使得卫生工作人员和研究人员开发与患者和服务需求相符的参数型临床测量方法。

我们发现, 这两个量表具有内在一致性, 覆盖了测量项目难度的全部范围并且在人口学研究上也不存在任何相关的天花板或地板效应。两个指标测量值之间没有任何关联进一步说明两个量表可测量不同的概念。

尝试开发一些不同的上肢和下肢活动量表^[20-22], 可能由于不一致, 与一个单个的功能量表相比, 那些独立量表的进行没有他们所希望的那么好^[23]。借助于现代测量理论来开发基于 ICF 核心分类集的测量方法有很多优势。首先 ICF 核心分类集确定了测量的内容, 他们提供了一个完备的和理论上合理的变量选择。该选择可根据特定的目的进行修改和扩大, 在急性期医院满足评定需求和确定康复需求。如果数据拟合 Rasch 模型, 则那些有序的类目可添加到测量量表或指标当中, 这些类目不随患者和观察者特征的改变而改变, 并且提供对患者能力的评估, 评估不随使用项目的改变而改变。通过直接应用

ICF 类目和 ICF 类目的一些简单说明来开发基于 ICF 的测量方法似乎是可行的。ICF 类目的特征是最具有代表性并且世界范围内可应用的方法, 只需要全球通用的 ICF 分类, 不考虑其他任何信息。在发展中国家更精确的和自我管理的测量方法通常都没有, 因此, ICF 在研究和临床实践中的应用才显得尤为重要。基于我们的结果, 类目分布表不仅应用在国际大型调查, 而且可通过人类功能各个方面的测量做出更多差异性分析。

虽然本文的目的不是考察 ICF 成分的维度, 但我们发现建立一个包含 ICF 活动和参与成分所有类目的单独量表是不可能的。全套类目的 Rasch 分析显示严重的不一致表明不止存在一种深层次的结构(数据没有显示)。然而, 两个独立的量表被证明具有一致性。

跨越不同指标中类目的定位表现出较高的表面效率。在上肢活动量表中“盥洗自身”是最有可能受限的项目, 其次是“护理身体各部”。在下肢活动量表中, “步行”和“改变身体姿势”最有可能受限。

存在的几个限制值得考虑。针对不同的健康状况上下肢两个说明性的指标是不一致的。因此在测量有上下肢状况的患者时, 必须分开测量。患有下肢疾病的患者样本大小是足够

的^[24]。患有上肢疾病的患者样本相对较小限制了拟合度测试的说服力。

虽然这个结果可能有悖于常理,“步行”比“爬楼梯”更困难,但必须注意到的是有 97% 和 94% 的患者分别在“步行”和“爬楼梯”上受限,这两个类目非常接近,因而样本的特质并不能将两者真正区别开来。

此外,量表的稳定性即测试项目的难度随着时间推移需要进一步评估,通过检验入院和出院时的项目拟合度进行评估。

尽管理论上改善功能是康复的主要议题,但该分析提出的两个量表并没有考虑到患者的偏好。目前这项研究并不能回答患者是否达到目标以及如何达到目标。然而,当前有一些方法来说明量表在目标达成方面存在方法学挑战^[25],这应该是下一步研究议程。

该指数在观察者进行患者严重程度评级时是否恒定仍然需要被证实。一个最主要的难题是使用 ICF 限定值直接测量 ICF 类目时评定者之间的信度问题。我们需要做更多的工作以达到对信度决定因素的更好了解,例如,在不同的环境、不同的卫生专业和培训水平上。

我们的研究是一个很好的例子。需要进一步证据以得到直接的基于 ICF 的测量方法,例如,那些 ICF 核心类目集类目是否适用于不同的健康状况^[13,26]以及从急性期医院到社区机构的不同情况^[14,27]。

基于 ICF 的标准来开发测量的内容以及基于 ICF 的测量工具的开发可被看成是交流的循环过程。本文显示的测量工具是基于非正式地可能应用的相关类目的广泛选择而构建的。然而,当我们在测量中研究所涉及的 ICF 类目的评定效果时,也能了解到许多关于 ICF 核心分类集进一步改进的方面。出于实用性的原因,ICF 核心分类集的类目应尽可能的少但又不能丧失综合性,因此有必要确认冗余的项目和从多种可选择的类目中确定出最佳类目。同样,我们在缩短 Core Set 类目的长度时,重要的不能去除测量范围两个端点的任何一个或作为某一领域惟一的一个类目,因为这些类目有利于区分能力差异性。因此,Rasch 分析是 ICF 核心分类集进一步发展的重要工具。

本研究结果显示可能开发基于 ICF 核心分类集类目的等级评定量表。这可能是一个有希望的方法,因为在现有的领域里从心理测量方面来讲还没有可靠的测量方法。

致谢:本研究得到了慕尼黑 Ludwig Maximilia 大学 Friedrich-Baur 基金会和瑞士物理疗法学会的赞助支持。我们感谢起到突出贡献的医院以及瑞士苏黎世大学医学院物理医学学院研究合作中心对数据的收集和帮助。

[参考文献]

- [1] Greenfield S. The state of outcome research: are we on target? [J]. *N Engl J Med*, 1989, 320: 1142-1143.
- [2] Stucki G, Melvin J. The International Classification of Functioning, Disability and Health: a unifying model for the conceptual description of physical and rehabilitation medicine[J]. *J Rehabil Med*, 2007, 39: 286-292.
- [3] Stucki G, Cieza A, Melvin J. The International Classification of Functioning, Disability and Health (ICF): a unifying model for the conceptual description of the rehabilitation strategy[J]. *J Rehabil Med*, 2007, 39: 279-285.
- [4] Stuck AE, Walthert JM, Nikolaus T, Bula CJ, Hohmann C, Beck JC. Risk factors for functional status decline in community-living elderly people: a systematic literature review[J]. *Soc Sci Med*, 1999, 48: 445-469.
- [5] Cieza A, Stucki G. Content comparison of health-related quality of life (HRQOL) instruments based on the international classification

- of functioning, disability and health (ICF)[J]. *Qual Life Res*, 2005, 14: 1225-1237.
- [6] Hamilton BB, Granger CV, Sherwin FS, et al. A uniform national data system for medical rehabilitation[M]. // Fuhrer MJ. *Rehabilitation outcomes: Analysis and measurement*. Baltimore, MD: Brookes, 1987;137-147.
- [7] Mahoney FI, Barthel DW. Functional evaluation: the Barthel index [J]. *Md State Med J*, 1965, 14: 61-65.
- [8] Grill E, Stucki G, Scheuringer M, Melvin J. Validation of International Classification of Functioning, Disability, and Health (ICF) Core Sets for early postacute rehabilitation facilities: comparisons with three other functional measures[J]. *Am J Phys Med Rehabil*, 2006, 85: 640-649.
- [9] Lundgren-Nilsson A, Tennant A, Grimby G, Sunnerhagen KS. Cross-diagnostic validity in a generic instrument: an example from the Functional Independence Measure in Scandinavia[J]. *Health Qual Life Outcomes*, 2006, 4: 55.
- [10] Tennant A. Principles and practice of measuring outcome[M]. // Barat M, Franchignoni F. *Advances in rehabilitation Assessment in physical medicine and rehabilitation*. Pavia: Maugeri Foundation, 2004.
- [11] World Health Organization. International classification of functioning, disability and health (ICF)[M]. Geneva: WHO, 2001.
- [12] Stucki G. International Classification of Functioning, Disability, and Health (ICF): a promising framework and classification for rehabilitation medicine[J]. *Am J Phys Med Rehabil*, 2005, 84: 733-740.
- [13] Cieza A, Ewert T, Üstün B, Chatterji S, Kostanjsek N, Stucki G. Development of ICF Core Sets for patients with chronic conditions[J]. *J Rehabil Med*, 2004, 44: 9-11.
- [14] Grill E, Ewert T, Chatterji S, Kostanjsek N, Stucki G. ICF Core Set development for the acute hospital and early post-acute rehabilitation facilities[J]. *Disabil Rehabil*, 2005, 27: 361-366.
- [15] Stoll T, Brach M, Huber EO, Scheuringer M, Schwarzkopf SR, Kostanjsek N, et al. ICF Core Set for patients with musculoskeletal conditions in the acute hospital[J]. *Disabil Rehabil*, 2005, 27: 381-387.
- [16] Finger ME, Cieza A, Stoll J, Stucki G, Huber EO. Identification of intervention categories for physical therapy, based on the international classification of functioning, disability and health: a Delphi exercise[J]. *Phys Ther*, 2006, 86: 1203-1220.
- [17] Gloor T, Grill E, Huber EO, Stucki G. Assessment of functioning in the acute hospital: operationalization and reliability testing of ICF categories relevant for physical therapists' interventions[J]. *Physiotherapy*, 2007, 93: S326.
- [18] Rasch G. Probabilistic models for some intelligence and attainment tests[M]. Chicago, IL: University of Chicago Press, 1980.
- [19] Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals[J]. *J Appl Meas*, 2002, 3: 205-231.
- [20] Lawrence RH, Jette AM. Disentangling the disablement process [J]. *J Gerontol B Psychol Sci Soc Sci*, 1996, 51: S173-S182.
- [21] Simonsick EM, Kasper JD, Guralnik JM, Bandeen-Roche K, Ferrucci L, Hirsch R, et al. Severity of upper and lower extremity functional limitation: scale development and validation with self-report and performance-based measures of physical function[J]. *J Gerontol B Psychol Sci Soc Sci*, 2001, 56: S10-S19.
- [22] Wolinsky FD, Armbrecht ES, Wyrwich KW. Rethinking functional limitation pathways[J]. *Gerontologist*, 2000, 40: 137-146.
- [23] Long JS, Pavalko E. Comparing alternative measures of functional limitation[J]. *Med Care*, 2004, 42: 19-27.
- [24] Linacre JM. Sample size and item calibration stability[J]. *Rasch Measure Tran*, 1994, 7: 328.
- [25] Tennant A. Goal attainment scaling: current methodological challenges[J]. *Disabil Rehabil*, 2007, 29(20-21): 1583-1588.
- [26] Stucki G, Grimby G. Applying the ICF in medicine[J]. *J Rehabil Med*, 2004; 5-6.
- [27] Stucki G, Üstün B, Melvin J. Applying the ICF for the acute hospital and early postacute facilities[J]. *Disabil Rehabil*, 2005, 27: 349-352.

(收稿日期: 2011-01-17)

(原文载: *J Clin Epidemiol*, 2009, 62(9): 891-898.)