·综　　述·

# 非特异性腰痛常用评价量表的概况△

孙章胤[1]，毕宇峰[2]，纪昌春[2*]

（1. 陕西中医药大学针灸推拿学院，陕西咸阳 712046；2. 陕西省中医医院，陕西西安 710003）

**摘要**：非特异性腰痛是临床常见病，虽然在临床上针对该病的临床研究中评价量表较多，主要涉及腰部疼痛类量表、功能障碍类量表、情绪障碍类量表以及生活质量类量表等，但存在使用欠准确的情况。故本文评估上述各类量表的内容和特点，总结各量表的优缺点，以期客观准确评价临床治疗非特异性腰痛的疗效。

**关键词**：非特异性腰痛，评价量表，概况

**中图分类号**：R681.5　　**文献标志码**：A　　**文章编号**：1005-8478（2023）05-0424-07

**Overview of commonly used evaluation scales for nonspecific low back pain** *// SUN Zhang-yin[1], BI Yu-feng[2], JI Chang-chun[2]. 1.School of Acupuncture and Massage, Shaanxi University of Traditional Chinese Medicine, Xianyang 712046, China; 2. Shaanxi Hospital of Traditional Chinese Medicine, Xi´an 710003, China*

**Abstract**: Non-specific low back pain is a common clinical condition. Although there are many evaluation scales in clinical studies on this condition, mainly involving low back pain scales, functional impairment scales, emotional disturbance scales, and quality of life scales, there are cases of lack of accuracy in their use. Therefore, in this paper, the content and characteristics of the above-mentioned scales are evaluated and the advantages and disadvantages of each scale are summarized in order to objectively and accurately evaluate the efficacy of clinical treatment of nonspecific low back pain.

**Key words**: nonspecific low back pain, evaluation scale, overview

非特异性腰痛（non-specific low back pain, NLBP）是临床常见的腰痛之一，指无明确病理学因素和解剖学结构异常而出现的腰部疼痛、功能活动障碍等症状[1]。流行病学研究表明，NLBP 的发病率约为腰痛的 90%[2]。由于该病患病率较高、反复发作，不仅严重损害身心健康和生活质量，而且占用大量医疗资源[3]，逐渐受到了广大医患的重视。当前研究涉及 NLBP 模型[4]、探讨疗效[5]以及诊疗指南[6]等方面，然而尚缺乏对该病常用评价量表的评估。在 NLBP 的治疗中需要准确的量表以评价疗效，而目前量表众多，且没有具体的应用范围标准，因此一定程度上影响了疗效的客观评定。根据 NLBP 病症的多因素性，评定该病临床疗效的主要结果为腰部疼痛程度、功能活动状态和情绪健康状态以及生活质量[7]。因此，本文将分别对 NLBP 的 4 种常用评价量表做比较，评估各个量表的适用范围，以期更好地应用于临床疗效的评价。

## 1　NLBP 疼痛的评价

腰部疼痛是 NLBP 的主要症状之一，是该病疗效评价的重要方面。对腰部疼痛的评价量表较多，如疼痛视觉模拟评分法（visual analogue scale, VAS）、疼痛数字评分法（numerical rating scale, NRS）、简明应用疼痛问卷（short form McGill pain questionnaire-2, SF-MPQ-2）和日本骨科协会腰痛评分量表（Japanese Orthopaedic Association Back Pain Evaluation Questionnaire, JOABPEQ）等，因此有必要将上述量表进行介绍并阐述其应用范围（表 1）。

### 1.1　VAS、NRS

VAS 和 NRS 具有适用于大多数疼痛评估工具所需的抽象思维的特点，因效度和反应性较好，现已广泛作为 NLBP 临床疼痛的评价。其中 VAS 量表由 1 条 10 cm 的水平线段表示疼痛分级，线段的一端是

DOI:10.3977/j.issn.1005-8478.2023.05.08

△**基金项目**：国家自然科学基金面上项目（编号：82074560）；2021 年度陕西中医药大学校级科研课题研究生创新项目（编号：2021CX29）

**作者简介**：孙章胤，硕士研究生，研究方向：针灸治疗慢性疾痛的临床疗效评价研究，（电话）15991025778，（电子信箱）1310392079@qq.com

* **通信作者**：纪昌春，（电话）18092005379，（电子信箱）doctor612526@126.com

“无疼痛”，另一端是“剧烈疼痛”或“最严重的疼痛”，故可充分显示微小程度的疼痛，利于统计连续变量。患者需要以书写的方式在量表上划出能代表疼痛程度的印记，因此需要一定抽象思维能力。与 VAS 类似，NRS 由描述疼痛严重程度的术语确定。目前最常用的是由 0~10 的 11 个数字组成的 NRS 数字评估量表，其中疼痛程度分级标准：0 代表无痛、1~3 分表示轻度疼痛、4~6 分表示中度疼痛、7~10 为无法忍受的疼痛。患者被要求在分段量表上口头指出或写出最能描述其疼痛强度的数值，得分越高，说明疼痛强度越大。因此，需要评估者对该量表有一定的抽象理解能力，并且考虑到对量表的理解偏差，故一般不做追踪性的复测。

可见，VAS 由于能够体现微小的疼痛差异且操作简单而受到评估者青睐，但只能用书写的方式表达疼痛程度故对于有认知障碍的老年患者有一定难度，因此，该量表适宜于有一定认知能力的中青年。而与 VAS 相比，NRS 不仅可通过口头或书面方式描述疼痛程度，而且评分较简单分类明确，因此该疼痛量表的应用更适用于老年患者群体[8]。此外，研究发现 VAS 或 NRS 评分可以预测腰痛患者的残疾状况[9]。

### 1.2 SF-MPQ-2 评分

MPQ 是一份旨在测量成人慢性疼痛患者疼痛性质、强度和心理感受的多维度量表问卷。该评分量表将 102 个疼痛描述词分成 3 类即感觉、情感和评估，内容丰富、涉及面广，然而完成评估所需时间较长。针对原版 MPQ 的局限性，SF-MPQ-2 是在原基础上简化而得，由 22 个疼痛类和 4 个情感类对疼痛的描述词以及疼痛强度（PPI）和 VAS 组成，所有的描述词均由 0~10 分别表示“无痛”和“最剧烈疼痛”，完成时间约为 2~5 min，且保留了原版的科学性和准确性[10]。

SF-MPQ 通过使用准确的描述评估疼痛的性质，可用于流行病学研究和多种原因引起的老年慢性疼痛患者的临床评价以及识别神经病变型疼痛和神经感受型疼痛[11, 12]。然而，地区语言的差异可能会影响疼痛描述的选择[13]。

### 1.3 JOABPEQ 量表

JOABPEQ 量表是一种针对腰椎疾病引起的腰痛患者结构化预后测量问卷。由于 JOABPEQ 中文版是测量腰痛患者多维状态的一种可靠、有效的工具[14]，故在临床腰痛的评估方面应用广泛。该问卷由主观症状、客观临床体征和日常活动受限度以及膀胱功能组成共计 29 分，根据评分判定病情程度，得分越高表示情况越好，其中总分 10 分为轻度，10~20 分为中度，>20 分为重度[15]。但研究发现老年腰痛患者易受问卷中的膀胱功能条目的影响而难以形成客观评估数据，且需要有评估者进行操作。

可见，JOABPEQ 多维度问卷简便易操作，可以对比治疗前后的治疗疗效，但易受膀胱功能影响，且难以独立进行。

**表 1 NLBP 常用腰部疼痛评估量表**

| 问卷 | 评估时间（min） | 推荐使用情况 | 优点 | 缺点 |
|---|---|---|---|---|
| VAS | <1 | 有一定认知能力的中青年 | 能体现微小的疼痛；利于检测统计连续变量，评估疗效 | 只侧重于疼痛，维度单一；只能以书写方式评估病情；患者需要一定抽象思维能力 |
| NRS | <1 | 有一定文化程度老年患者人群 | 评估管理方便（既可口述，亦可书写）；评分简单；分类明确 | 单维度评估疼痛；评估者需要有一定的理解能力和抽象思维能力；不能做追踪性复测；难以体现微小疼痛差异 |
| MPQ | 5~15 | 科学研究 | 内容丰富，多维度评估，涉及面广泛；可评估多种原因引起的疼痛 | 量表耗时较长 |
| SF-MPQ | 2~5 | 普适 | 简捷，耗时短；可多维度评估腰痛患者；可评估多种原因引起的疼痛 | 易受地区语言文化差异影响 |
| JOABPEQ | <5 | 中青年腰痛患者 | 简便易操作；多维度评估；可对预后评估，观察疗效 | 老年患者易受膀胱功能项目的影响；患者难以独立完成问卷 |

## 2 NLBP 腰部功能活动的评价

NLBP 不仅引起腰部疼痛，而且对腰部功能活动障碍具有重要影响，因此需要准确的量表予以评价。目前涉及腰部功能障碍的评定量表包括 ODI 功能障碍指数评分（Oswestry Disability Index, ODI）、RMDQ 功能障碍调查表（Roland-Morris Disability Questionnaire, RMDQ）以及魁北克背痛障碍评分量表（Quebec back pain disability scale, QBPDS）等。现将上述量表的具体情况予以阐述，并对各量表进行比较分析，如表 2。

2.1 ODI 评分

ODI 是针对急性、亚急性或慢性腰痛患者的疼痛以及相关功能的自我评价量表，因其可信度较高被国内外广泛采用。原始量表共分为 3 类 10 项，包括 1 项疼痛和 9 项日常生活活动，其中每项 1 个问题并都以 0~5 分的顺序尺度进行评分。根据国内文化差异，汉化 ODI 2.0 版量表将性生活一项去掉，且将行走的距离标准的 1 英里换成熟知的 1 km。该量表最终 ODI 评分测算为实际得分与量表总分 45 的比值，比值越高功能障碍越严重，且汉化版的效度也得到认可。ODI 通常由患者使用纸和笔完成，也可以使用通话和计算机管理，完成时间一般在 5 min 左右，操作简单，评估者不需要培训。

可见，ODI 既包含了疼痛评分又有日常功能评分，亦可判定腰痛患者术后疗效[16]。然而，该量表仅涉及疼痛和功能障碍的评估，缺乏对患者其他维度的评定[17]。

2.2 RMDQ 量表

RMDQ 广泛应用于初级保健、受伤工人和多学科康复中心的临床实践，是检测因疼痛引起功能障碍的评分问卷，因其具有较高的信度、效度和灵敏度被广泛使用。该问卷包括 24 项内容，代表了可能受到腰痛影响的日常身体活动和功能的执行情况，如家务活、睡眠、行动能力、穿衣、获得帮助、食欲、易怒和疼痛的严重程度。其中每一项可选是或否，是为 1 分，否为 0 分，分数越高功能障碍越严重。该量表一般 5 min 就可完成，用时较短。虽然它被称为“残疾”量表，但根据国际功能、残疾和健康分类，它包含了损伤、残疾和残疾的要素。一般使用纸质问卷，也可通过电话管理。RMDQ 简洁明了，并且很容易被患者和临床医生理解。

RMDQ 易被患者和评估者准确理解，且便于管理患者的评估，但维度较单一仅涉及因疼痛引起的功能障碍。有研究表明，ODI 对持续性严重残疾的患者更敏感，而 RMDQ 对轻、中度残疾患者的反应更敏感[18]，故更适用于评估治疗前后的短期变化。此外，在信度方面，ODI 要高于 RMDQ[19]。

2.3 QBPDS 量表

QBPDS 是用于评估患者当天功能障碍和睡眠情况的自评问卷，有 20 项条目分为 6 个领域，包括生活自理、腰部功能活动和睡眠等，不涉及社会生活和疼痛强度的问题，以 6 分制进行评分，从 0（根本不困难）到 5（不能做到），总分是由每个项目的分数之和计算而得，分数越高代表功能障碍越严重。评估管理较方便，可以面对面、电子或电话传送，完成填表约 5~10 min。QBPDS 有良好的内部一致性和重测可靠性[20]，但应用于急性腰痛的可靠性较差[21]。此外，由于使用 QBPDS 检测患者病情变化的情况，缺乏其最小的重要差异的经验证据，并且受文化差异影响，跨语言版本的有效性有待验证[22]。

由此，QBPDS 量表简练，用时较少，适用于临床评价。由于计算简便且管理评估便捷，患者可以自主完成评估。另外，相较于 ODI 和 RMDQ，该量表监测腰部功能障碍更加灵敏。然而，该量表无法评价患者社会生活状况和疼痛强度，一般不用于急性腰疼的评估，且易受不同文化差异的影响。此外，ODI，QBPDS 和 RMDQ 在老年腰痛患者中具有相似的测量特性[23]。

表 2 NLBP 常用腰部功能活动评估量表

| 问卷 | 评估时间（min） | 推荐使用情况 | 优点 | 缺点 |
|---|---|---|---|---|
| ODI | <5 | 普适 | 简便易操作；能够涉及患者的腰部疼痛和功能障碍；对于功能障碍程度较严重的群体反应更敏感；评定术后疗效 | 维度较单一，不涉及患者的精神心理状况；对于中低程度的功能障碍不够灵敏 |
| RMDQ | <5 | 快速评估轻、中度患者治疗前后患者短期变化 | 简单明了，不需要对患者和评估者进行培训；可通过书面、电话联系的方式对患者病情管理；对轻、中度患者群体的评价更灵敏 | 维度单一，仅涉及因疼痛引起的功能障碍；对于重度疼痛引起的功能障碍评估的敏感度稍欠缺 |
| QBPDS | 约 5 | 临床评估 | 涉及功能障碍和睡眠情况；管理评估方便；检测患者腰部功能障碍敏感性较高 | 缺乏疼痛强度和社会生活状况的条目；不能用于急性疼痛的评估；受语言文化差异影响较大，跨语言版本的有效性有待验证 |

## 3 NLBP 伴随的情绪障碍评价

NLBP 不仅可使患者腰部疼痛不适，而且可因病情影响患者情绪状况，故有必要对 NLBP 患者情绪障碍进行评估。研究表明慢性疼痛患者的抑郁率高于没有慢性疼痛的患者，且抑郁程度更严重[24]。目前临床多用于 NLBP 患者情绪障碍的量表多为汉化第二版

贝克抑郁量表（Beck Depression Inventor-Ⅱ, BDI-Ⅱ）、汉密尔顿抑郁量表（Hamilton Depression Rating Scale-24, HAMD-24）以及贝克焦虑量表（Beck Anxiety Inventory, BAI）和症状自评量表（symptom checklist-90, SCL-90）。现将 NLBP 伴随的情绪障碍量表做以阐述并进行分析，具体如表 3。

3.1 BDI-Ⅱ量表

由于汉化 BDI-Ⅱ具有较高的可信度和敏感度，受文化差异影响较小，现被广泛应用于国内各种疾病群体和普通人群的自查和临床研究，且研究发现该量表同样适用于医学生抑郁自评[25]。该量表包含 2 个因子共计 21 条项目，每个项目采用 0~3 分评定，根据近 1 周的自我感觉，从每一项条目中选出最适合自己的情况，用笔和纸记录，最后算出总分即为各项总和，其中 0~13 分为正常，14~19 分为轻度，20~28 分为重度，29~63 分为重度，用时约 10~15 min。BDI-Ⅱ是临床最常用的抑郁自评量表之一，适用于成年人各个年龄阶段患者抑郁状态和程度的评估[26]，但对于老年人慎重使用（老年人认知水平下降且项目偏多故敏感度和特异度下降）。

可见，虽然该量表简单，受文化和种族影响较少，但只能自查是否抑郁或抑郁程度，对于继发或者原发性抑郁症的诊断尚需进一步精神科检查。

3.2 HAMD-24 量表

HAMD 量表是临床上常用的抑郁量表之一，可用于临床抑郁状态和程度的评定，以及服用抗抑郁药物后疗效的判定。虽然目前涉及的版本有 2 个，分别是 17 项量表 HAMD-17 和 24 项量表 HAMD-24，但临床使用较广的是最新版本即 HAMD-24。该量表一般采用交谈和观察的方式进行，由经过培训的 2 名评估者对患者进行联合检查，检查结束后分别予以独立评分，评定需要 15~20 min。大部分采用 0~4 分的 5 级评分，0~4 分别代表无、轻度、中度、重度、极重。其中总分为 0~7 分为正常，8~19 分为可能有抑郁，20~34 分为肯定有抑郁，35 分以上为严重抑郁。

可见，该表条目较多，内涵丰富，能够客观判定多种疾病伴随的抑郁情况，尤其适用于抑郁状态的成年人[27]，然而本量表对于抑郁和焦虑状态却不能明确的鉴别[28]。

3.3 BAI 量表

BAI 是一种简短的焦虑测量方法，主要关注焦虑的躯体症状，是一种善于区分焦虑和抑郁的评价方法。BAI 是用于具有焦虑症状的成年人自评焦虑状态的量表，因其信度较高被广泛应用于临床。该量表含有 21 个项目，把患者被多种焦虑症状烦扰的程度作为评定指标，使用 4 级评分，1~4 分别代表无、轻度、中度和重度。评定完成后将各项分数简答相加得到总分，再通过特有公式计算转换成标准分。BAI 是一种相对简短、易于管理、容易得分的焦虑测量方法。它具有良好的心理测量特性，并表现出对变化的敏感性，BAI 的开发是为了减少与抑郁症状的重叠，因此更倾向于关注躯体（例如，心跳加速、头晕）症状，全程需要 5~10 min。由于该量表需要患者根据自身具体状况用纸、笔填写，故需要评估者讲述清楚每项的含义，评定的时间范围是现在或最近一周的自我感觉。

可见该量表内容简单，容易理解，操作分析较方便，且已被证明对精神病人群和医学人群的随时间的变化反应灵敏[29]。然而，需要评估者有一定的理解能力，并且该量表过多关注躯体症状。

表 3 NLBP 伴随的情绪障碍评价

| 问卷 | 评估时间（min） | 推荐使用情况 | 优点 | 缺点 |
|---|---|---|---|---|
| BDI-Ⅱ | 10~15 | 普适 | 各种疾病伴随抑郁状态群体和普通人群自查；条目简单，易于理解；受文化和种族影响较少；有效评价抑郁程度 | 不适用于老年患者群体使用；继发或者原发性抑郁症的诊断尚需进一步精神科检查 |
| HAMD-24 | 15~20 | 抑郁状态的成年人 | 他评量表，评分客观，准确性较高；条目内涵丰富，可有效判定多种疾病伴随的抑郁情况；临床抑郁状态和程度的评定以及疗效检测 | 用时较长；评估者需要培训；不能鉴别抑郁和焦虑状态 |
| BAI | 5~10 | 焦虑症状的成年人 | 能够有效区分焦虑和抑郁状态；具有良好的心理测量特性，对焦虑的反应有较好的敏感性；时效性较佳 | 需要评估者有一定的理解能力；更多关注躯体症状 |
| SCL-90 | 15~20 | 心理卫生的诊断和精神病学研究 | 包含较宽泛的精神病症状学内容；适用群体宽泛 | 条目较多，用时较长；受语言文化差异较大 |

3.4 SCL-90 量表

SCL-90 常用于临床精神健康的筛查和精神疾病的诊断和评估[30]，目前已广泛应用于不同职业群体的心理卫生问题，如学生群体和企事业单位以及精神障碍患者[31-33]。该量表共有 90 个项目，包含较宽泛的精神病症状学内容，涉及感觉、情感、思维、意识、人际关系、和饮食睡眠等方面，采用 10 个因子分别反应 10 个方面的心理症状情况。其中每 1 个项目进行 5 级评分，从 0（完全没有）到 4（严重），说明过去 1 周患者的困扰程度。总分为 90 个项目单项分相加之和，分值越高病情越严重，总均分为总分/90，表示从总体情况看，受试者的自我感觉位于 0~4 级间的哪个分值上，一般用时 15~20 min。

可见，SCL-90 具有容量大、反应症状丰富的特点，更能准确反映患者自觉症状，可以用于心理健康状况的诊断和精神病学研究。然而，由于语言文化差异，有些项目文本容易产生歧义，且条目较多，用时较长。

## 4 NLBP 生活质量评价

NLBP 因其反复发作的腰部疼痛和功能活动障碍严重影响生活质量[34]，因此有必要探究关于该病的生活质量评价量表。目前在该病的临床对照实验中常用的生活量表主要涉及简明生活质量量表（short form 36 health survey questionnaire, SF-36）、简明健康质量量表（12-item short-form health survey, SF-12）、世界卫生组织与健康有关生存质量问卷（WHO Quality of Life-BREF, WHOQOL-BREF）。此外，研究发现睡眠质量与生活质量关系密切[35]，故对 NLBP 中常用的匹兹堡睡眠质量指数（Pittsburgh Sleep Quality Index, PSQI）予以评估。现评估 NLBP 生活质量量表和睡眠障碍量表，如表 4。

4.1 SF-36 与 SF-12 量表

SF-36 是目前广泛使用的简化生活质量评估量表，现代研究证实了该量表在评估患者生活质量和病情状况以及评价干预疗效方面的重要作用。该量表从生理机能、生理职能、一般健康状况、躯体疼痛和精神健康等 8 个方面共计 36 项全面概括生存质量，计分方法为分别计算各个方面下包含项目的分数再予以换算，换算公式为换算得分=（实际得分-该方面可能的最低分）/该方面的最高得分与最低得分之差×100，并且各组所得分数可以互相比较。量表内容较丰富，项目简单，一般用时 5~10 min。由于是中文简缩版，故受到语言文化习惯的影响，对患者的认知有一定的要求。而 SF-12 是在 SF-36 的基础上进一步简化而来，包含与 SF-36 相同的 8 个维度共计 12 项，两者在患者生理和心理健康方面的评估具有高度相关性[36]。因其重点突出，用时较短（约 3 min），故在临床上尤其适用于老年患者的评估。在管理评估方面，该量表可自主填写、计算机系统完成或者电话联系填写，且最近的研究发现 SF-36 和 SF-12 的纸质和电子版管理模式下具有等效性[37]。

可见 SF-36 与 SF-12 虽然条目精简、效度和信度较佳且两者相关性较强，在临床广泛应用，然而两者都易受文化差异的影响，此外，其中 SF-36 尚需被评估者有一定的认知能力，而 SF-12 的应用范围相对较小。

4.2 WHOQOL-BREF 问卷

WHOQOL-BREF 问卷是由世界卫生组织研制，以 WHOQOL-100 为基础简化而来的，旨在评估一般人群和患者近 2 周与健康相关的生活质量状况。因其可信度和效度较好，被广泛应用于临床。该量表包含 4 个维度即生理、心理和环境以及社会关系，共计 26 项，将 26 项得分相加即为总分值，分值越高代表生存质量越好。因其简单易操作且占用时间较短，故多用于大型流行病学调查和临床试验疗效评定。

可见，WHOQOL-BREF 维度多样，内容简短准确，然而由于该表跨国家的特点，易受文化差异的影响，在“社会关系”项中有较高缺失比[38]。

4.3 PSQI 量表

PSQI 是目前广泛应用的睡眠障碍评估指数量表，适用于多维度的评价睡眠障碍患者和精神障碍患者以及健康人近一月的睡眠质量。研究发现，睡眠质量是 NLBP 发病的重要危险因素[39]，因此有必要使用 PSQI 量表评价该病患者的睡眠质量。该量表由 19 个自评和 5 个他评条目组成，其中前 18 个条目分成主观睡眠质量、入睡时间、睡眠效率、睡眠障碍、等 7 个因子，患者需要 5~10 min 完成填表。在计分方法上，每个因子以 0~3 分来算，累计 7 个因子总分范围是 0~21 分，总分越高，睡眠质量越差，其中 0~5 分表示睡眠质量正常，6~10 分代表轻度睡眠障碍，11~15 分表示中度睡眠障碍，8~21 分提示重度睡眠障碍。较早研究已发现 PSQI 可用于在所有卫生保健环境中与老年人进行初步评估和持续的比较测量。PSQI 是对睡眠的一种主观测量方法，患者的自我报告能够较准确的反应睡眠情况，然而性别差异会导致对睡眠质量的理解有所偏差[40]。

可见 PSQI 具有精神障碍和睡眠质量的评估，适用于精神科和睡眠医学的研究。然而对睡眠障碍的程度区分度不够明显。

表 4 NLBP 生活质量评价

| 问卷 | 评估时间（min） | 推荐使用情况 | 优点 | 缺点 |
|---|---|---|---|---|
| SF-36 | 5~10 | 普适 | 角度多维，涉及面广泛，可用于多种疾病生活质量的评价；可评价干预疗效 | 易受语言文化和生活习惯的影响；对患者的认知有一定的要求；缺乏对 NLBP 的特异性 |
| SF-12 | 3 | 老年患者的评估 | 评估管理方式多样；维度多样，涉及生理、心理状况；重点突出，用时较少 | 易受语言文化和生活习惯的影响 |
| WHOQOL-BREF | — | 普适性 | 维度多样；内容简短准确 | 易受文化差异的影响；相关项目不适应我国具体情况而有较高缺失比 |
| PSQI | 5~10 | 适用于我国精神科和睡眠医学的临床研究 | 多维度的评价睡眠障碍患者和精神障碍；能够有效评估睡眠质量 | 睡眠障碍的程度区分度不够明显；性别差异可能影响结果的客观性 |

## 5 小 结

NLBP 反复发作的腰部疼痛和功能障碍以及伴随的情绪障碍严重影响生活质量，虽然目前针对该病的临床研究较多，但对评定疗效的量表尚欠缺应用方面的比较分析，而能够客观判定临床疗效是提高临床诊疗效率的关键环节。故本文对 NLBP 临床研究中疼痛量表、功能障碍量表、情绪障碍量表以及生活质量量表进行分析比较，以期能够高效地选择临床疗效的评估量表。

上述量表的可信度和效度都已得到验证，并在临床中广泛应用。然而尚存在一下问题亟待解决：（1）有些量表近些年欠缺常规模型的建立，故需要在原始量表的基础上进行研究，将量表本土化，更新建立常规模以提高量表的客观性；（2）上述量表多由其他版本汉化而来，有些条目与我国生活习惯和语言文化差异较大，需要进一步研究优化以适应我国患者群体。总之，上述评价指标在 NLBP 临床研究的疗效评定中具有重要作用，在今后临床研究中应根据各量表的特点选择适当的量表进行有效的临床评估，并且需要进一步研究及时更新常模和适应我国文化和语言差异，以提高评估效率，进而能够客观评价临床治疗效果。

## 参考文献

[1] 施玉博, 郭卫春, 余铃 . 非特异性下腰痛:北美脊柱协会(NASS)循证医学指南解读 [J] . 中国修复重建外科杂志, 2021, 35 (10) : 1336-1340.

[2] Maher C, Underwood M, Buchbinder R. Non-specific low back pain [J] . Lancet, 2017, 389 (10070) : 736-747.

[3] Blanchette MA, Stochkendahl MJ, Borges DSR, et al. Effectiveness and economic evaluation of chiropractic care for the treatment of low back pain: a systematic review of pragmatic studies [J] . PLoS One, 2016, 11 (8) : e160037.

[4] 柯嵩, 徐钢, 李忠海 . 慢性非特异性腰痛的动物模型相关研究进展 [J] . 中国矫形外科杂志, 2020, 28 (7) : 629-633.

[5] 刘爱峰, 王平, 张超, 等 . 射频等离子针刺不同功率时间组合蛋清实验及治疗非特异性腰痛疗效分析 [J] . 中国矫形外科杂志, 2017, 25 (13) : 1153-1157.

[6] 顾蕊, 王岩, 陈伯华 . 中国非特异性腰背痛临床诊疗指南 [J] . 中国脊柱脊髓杂志, 2022, 32 (3) : 258-268.

[7] Benz T, Lehmann S, Elfering A, et al. Comprehensiveness and validity of a multidimensional assessment in patients with chronic low back pain: a prospective cohort study [J] . BMC Musculoskelet Disord, 2021, 22 (1) : 291.

[8] 徐城, 杨晓秋, 刘丹彦 . 常用的疼痛评估方法在临床疼痛评估中的作用 [J] . 中国疼痛医学杂志, 2015, 21 (3) : 210-212.

[9] Shafshak TS, Elnemr R. The Visual Analogue Scale versus Numerical Rating Scale in measuring pain severity and predicting disability in low back pain [J] . J Clin Rheumatol, 2021, 27 (7) : 282-285.

[10] 万丽, 赵晴, 陈军, 等 . 疼痛评估量表应用的中国专家共识 (2020 版) [J] . 中华疼痛学杂志, 2020, 16 (3) : 177-187.

[11] Jumbo SU, MacDermid JC, Packham TL, et al. Reproducibility: reliability and agreement parameters of the Revised Short McGill Pain Questionnaire Version-2 for use in patients with musculoskeletal shoulder pain [J] . Health Qual Life Outcomes, 2020, 18 (1) : 365.

[12] Dworkin RH, Turk DC, Trudeau JJ, et al. Validation of the Short-form McGill Pain Questionnaire-2 (SF-MPQ-2) in acute low back pain [J] . J Pain, 2015, 16 (4) : 357-366.

[13] Wang JL, Zhang WJ, Gao M, et al. A cross-cultural adaptation and validation of the short-form McGill Pain Questionnaire-2: Chinese version in patients with chronic visceral pain [J] . J Pain Res, 2017, 10 : 121-128.

[14] Yao M, Li ZJ, Zhu S, et al. Simplified Chinese version of the Japanese Orthopaedic Association Back Pain Evaluation Questionnaire: cross-cultural adaptation, reliability, and validity for patients with low back pain [J] . Spine (Phila Pa 1976) , 2018, 43 (6) : E357-E364.

[15] Tominaga R, Sekiguchi M, Yonemoto K, et al. Establishment of reference scores and interquartile ranges for the Japanese Orthopaedic Association Back Pain Evaluation Questionnaire (JOABPEQ) in patients with low back pain [J] . J Orthop Sci, 2018, 23 (4) : 643-648.

[16] Werner D, Grotle M, Gulati S, et al. Can a successful outcome after surgery for lumbar disc herniation be defined by the Oswestry Disability Index Raw Score [J] . Global Spine J, 2020, 10 (1) : 47-54.

[17] Lee CP, Fu TS, Liu CY, et al. Psychometric evaluation of the Oswestry Disability Index in patients with chronic low back pain: factor and Mokken analyses [J] . Health Qual Life Outcomes, 2017, 15 (1) : 192.

[18] Smeets R, Koke A, Lin CW, et al. Measures of function in low back pain/disorders: Low Back Pain Rating Scale (LBPRS) , Oswestry Disability Index (ODI) , Progressive Isoinertial Lifting Evaluation (PILE) , Quebec Back Pain Disability Scale (QBPDS) , and Roland- Morris Disability Questionnaire (RDQ) [J] . Arthritis Care Res (Hoboken) , 2011, 63 (Suppl 11) : S158-S173.

[19] Chiarotto A, Maxwell LJ, Terwee CB, et al. Roland-Morris Disability Questionnaire and Oswestry Disability Index: Which has better measurement properties for measuring physical functioning in nonspecific low back pain? Systematic review and meta-analysis [J] . Phys Ther, 2016, 96 (10) : 1620-1637.

[20] Speksnijder CM, Koppenaal T, Knottnerus JA, et al. Measurement properties of the Quebec Back Pain Disability Scale in patients with nonspecific low back pain: systematic review [J] . Phys Ther, 2016, 96 (11) : 1816-1831.

[21] Garg A, Pathak H, Churyukanov MV, et al. Low back pain: critical assessment of various scales [J] . Eur Spine J, 2020, 29 (3) : 503-518.

[22] Wewege MA, Jones MD, McAuley JH. Clinimetrics: Quebec Back Pain Disability Scale [J] . J Physiother, 2020, 66 (4) : 270.

[23] Jenks A, Hoekstra T, van Tulder M, Ostelo RW, et al. Roland-morris disability questionnaire, oswestry disability index, and quebec back pain disability scale: which has superior measurement properties in older adults with low back pain [J] . J Orthop Sports Phys Ther, 2022, 52 (7) : 457-469.

[24] Velly AM, Mohit S. Epidemiology of pain and relation to psychiatric disorders [J] . Prog Neuropsychopharmacol Biol Psychiatr, 2018, 87 (Pt B) : 159-167.

[25] 王金燕, 曹文君, 管振钧, 等 . 贝克抑郁量表在医学生中的信度与效度评价和应用 [J] . 中国卫生统计, 2018, 35 (2) : 253-255.

[26] 孙晓艳, 李怡雪, 余灿清, 等 . 中文版抑郁量表信效度研究的系统综述 [J] . 中华流行病学杂志, 2017, 38 (1) : 110-116.

[27] 郑磊磊, 李惠春 . 常用焦虑及抑郁评估量表 [J] . 中华全科医师杂志, 2016, 15 (5) : 334-336.

[28] 姚树桥, 杨彦春 . 医学心理学 [M] . 6 版 . 北京: 人民卫生出版社, 2013.

[29] Lee YW, Park EJ, Kwon IH, et al. Impact of Psoriasis on Quality of Life: relationship between clinical response to therapy and change in health-related quality of life [J] . Ann Dermatol, 2010, 22 (4) : 389-396.

[30] Carrozzino D, Vassend O, Bjorndal F, et al. A clinimetric analysis of the Hopkins Symptom Checklist (SCL-90-R) in general population studies (Denmark, Norway, and Italy) [J] . Nord J Psychiatr, 2016, 70 (5) : 374-379.

[31] 王伟 . 少数民族大学生症状自评量表 (SCL-90) 测量研究定性及定量分析 [J] . 中国健康心理学杂志, 2012, 20 (3) : 435-437.

[32] 王姿欢, 俞文兰, 沈壮, 等 . 症状自评量表 (SCL-90) 应用于中国职业女性心理测评的信效度评价 [J] . 中国工业医学杂志, 2017, 30 (4) : 247-250.

[33] Kostaras P, Martinaki S, Asimopoulos C, et al. The use of the Symptom Checklist 90-R in exploring the factor structure of mental disorders and the neglected fact of comorbidity [J] . Psychiatr Res, 2020, 294 : 113522.

[34] 王微晶, 石含英, 颜如冰, 等 . 慢性腰痛患者躯体症状负担与健康相关生活质量的关系 [J] . 颈腰痛杂志, 2019, 40 (4) : 501-504.

[35] Darchia N, Oniani N, Sakhelashvili I, et al. Relationship between Sleep Disorders and Health Related Quality of Life-Results from the Georgia SOMNUS Study [J] . Int J Environ Res Public Health, 2018, 15 (8) : 1588.

[36] 何燕, 赵龙超, 刘丹萍, 等 . SF-36 和 SF-12 在人群生命质量调查中的性能比较研究 [J] . 现代预防医学, 2017, 44 (5) : 852-862.

[37] White MK, Maher SM, Rizio AA, et al. A meta-analytic review of measurement equivalence study findings of the SF-36 (R) and SF-12 (R) Health Surveys across electronic modes compared to paper administration [J] . Qual Life Res, 2018, 27 (7) : 1757-1767.

[38] 李佳玲, 赵艳霞, 顾菁, 等 . 基于项目反应理论修订生存质量测定量表简表 WHOQOL-BREF [J] . 中山大学学报 (医学科学版) , 2015, 36 (2) : 284-293.

[39] 李浩曦, 谢湘涛, 陈源, 等 . 壮族中青年办公室人员非特异性腰痛调查 [J] . 中国矫形外科杂志, 2020, 28 (7) : 613-618.

[40] Morris JL, Rohay J, Chasens ER. Sex differences in the psychometric properties of the Pittsburgh Sleep Quality Index [J] . J Womens Health (Larchmt) , 2018, 27 (3) : 278-282.

（收稿：2022-03-07 修回：2022-09-08）

（同行评议专家：王守国 张喜善 刘 敏 王 渊）

（本文编辑：宁 桦）