A review of general pain measurement tools and instruments for consideration of use in COPD clinical practice

Alisha Maree Johnson¹
Sheree MS Smith¹,²

¹School of Nursing and Midwifery, Western Sydney University, Penrith, ²Respiratory, Sleep, and Environmental Health Research Academic Unit, Liverpool Hospital, Liverpool, NSW, Australia

Background: The experience of pain can have a significant impact on the everyday life of individuals including those with COPD. Recently, pain has emerged as an area in COPD research. When considering pain measurement in COPD studies, it is important to consider the validity, reliability, responsiveness and interpretability of instruments and tools. This review sought to assess these domains of general pain instruments and tools using the consensus-based standards for the selection of health measurement instruments (COSMIN).

Methods: Three separate analyses were used to assess general pain measurement tools and instruments. These comprise COSMIN’s, 1) methodological quality assessment with dichotomous responses, 2) the 4-point rating scale, and 3) overall quality criteria using an assessment scale for clinimetric properties by Terwee.

Results: Overall Pain Sensitivity Questionnaire (PSQ) was found to have the highest rating in all domains of validity, reliability, responsiveness and interpretability. In the first analysis, PSQ and Geriatric Pain Measure (GPM) scored highest in four of the six domains. In the second analysis, using the 4-point rating, the PSQ scored highest in three of four domains. In the third analysis, the GPM scored the highest in all four domains. Overall the PSQ, GPM and Defense and Veterans Pain Rating Scale scores were consistently high in the three separate analyses in this review.

Conclusion: This review found variability in the domains of validity, reliability, responsiveness and interpretability in general pain tools and instruments. The PSQ was found to be the most valid and reliable general pain measurement instrument for adult populations.

Keywords: pain, measurement, validity, reliability, responsiveness, interpretability, COPD

Introduction

COPD causes significant disability and impacts on the lives of people with the disease, their families and health care services.¹ People with COPD often have well-described comorbid conditions, such as heart failure,² anxiety and depression,³ and the most common symptom experienced by people with COPD is breathlessness. More recently, pain has been reported as a frequent symptom in addition to breathlessness.⁴

Pain is known to affect psychological well-being⁵ and has also been linked with a decrease in quality of life.⁶ Reviews of pain and COPD indicate that pain may contribute to an increase in poor clinical outcomes.⁷ COPD patients with pain are prescribed more short-acting beta-agonist therapy and oral steroids than those without pain.⁸ Outpatient prescription data⁹ suggest that many COPD patients use opioid pain medications on
a long-term basis, and it is unclear if this use is related to refractory breathlessness or pain management.\textsuperscript{10}

Pain is a stressful, individual and subjective human experience\textsuperscript{11} and has been associated with feelings of social isolation.\textsuperscript{5} The ability to accurately measure and interpret the symptom of pain through valid and reliable tools or instruments may be clinically important in determining medical therapy and nonpharmaceutical interventions for people with COPD. There are a wide range of general and specific pain tools and instruments available to clinicians. Numerous studies describe the adoption of pain tools and instruments to various acute and chronic conditions in terms of their validity and reliability.\textsuperscript{12,13}

Currently, COPD and pain studies appear to utilize general pain tools or instruments, although none have been developed for chronic respiratory diseases. When differing methods to measure pain are used across studies in COPD patients, this could lead to misinterpretation of the phenotype of pain and fluctuating estimates of pain prevalence being reported. To date, a comparison of general pain measurement instruments or tools’ validity, reliability, responsiveness and interpretability has not been described. The primary aims of this review were to provide an overview of general pain tools and instruments and to evaluate their validity, reliability, responsiveness and interpretability in order to reduce methodological concerns in COPD and pain studies.

**Methods**

The primary aims of this review were to identify general pain measurement tools and instruments used in an adult population and assess their psychometric properties. The review objectives were to provide an overview of the psychometric quality and to compare the assessment of each psychometric property in terms of validity, reliability, responsiveness and interpretability of general pain instruments or tools.

**Search strategy and study selection criteria**

The Preferred Reporting Items for Systematic Reviews and Meta-analyses guideline was used to underpin the development of the review protocol.\textsuperscript{14} An a priori search strategy was devised to identify relevant pain measurement literature, and this search of peer-reviewed literature included CINAHL, Embase, Medline, Pubmed, Proquest, PsychArticles, PsychInfo, Sciedirect and Psychiatry Online electronic databases. The primary search was conducted during January 2016, using combinations of keywords and their abbreviations: “pain”, “scale”, “instrument”, “measurement”, “measure*”, “precision”, “precise”, “validity”, “valid*”, “reliability”, “reliabi*”, “response*”, “responsiveness” and “adults”. The search was not restricted to any publication time-frame although it was restricted to publications in English. The reference lists of selected relevant articles were screened for potential publications. The review protocol outlined the process where abstracts were identified and reviewed for inclusion.

For all included abstracts, full-text articles were retrieved. All included full-text publications were further searched to identify the original psychometric studies of the pain tool or instrument. The protocol restricted the search to psychometric studies that included adult participants and the measurement was pain.

**Data extraction and methodological quality assessment**

The collecting, collating and evaluation of the published data were predetermined by the reviewers before the commencement of the study. The review of abstracts that met the inclusion criteria was performed by two independent reviewers (AMJ, SMSS) and was aided by a predesigned data extraction spreadsheet. This independent review process (AMJ, SMSS) was replicated for retrieved full-text articles. Disagreements over inclusion of full-text articles were settled through discussion and consensus reached.

All included articles were assessed for methodological quality using the consensus-based standards for the selection of health measurement instruments (COSMIN) checklists of domains and properties in each domain.\textsuperscript{15} The COSMIN measurement domains are comprised of validity, reliability, responsiveness and interpretability.\textsuperscript{16} Interpretability of scores was included as an important measurement characteristic that reflects distribution of scores across a population and the reporting of minimal important change or difference.\textsuperscript{15} It should be noted that interpretability could only be assessed using the dichotomous COSMIN checklist and was not a domain in any of the other analyses that used other COSMIN checklists. Each of these domains had prescribed and defined measurement properties (Table 1). For example, validity has the properties of content validity, construct validity and criterion validity.\textsuperscript{16}

COSMIN guidance recommends three approaches in the analysis of instrument measurement properties. Permission was granted to utilize the COSMIN definitions. The first analysis was a quality assessment which employed a specific COSMIN checklist incorporating a dichotomous scale comprising of yes and no with numerical value assignment of 1 and 0, respectively. This checklist incorporates a variety
of components within each aspect of a measurement study’s methodology,\textsuperscript{15} including descriptions of the study design and statistical methods used\textsuperscript{15} (Table 1).

The second analysis of general pain instruments and tools utilized the COSMIN 4-point rating scale,\textsuperscript{17} which calculates a summary of methodological quality for each domain. Each component was assessed using a rating of excellent, good, fair or poor\textsuperscript{17} (Table 2). The lowest rating was considered for each individual domain being assessed, for example, if poor was assigned as a lowest rating in any domain then the overall score was “poor.”\textsuperscript{15}

As per COSMIN guidance an overall score of the general pain instrument or tool was subsequently calculated using a predeveloped assessment scale by Terwee et al\textsuperscript{17} to rate the level of quality. To examine the overall quality of statistical methods used in the development of each general pain instrument or scale, a rating system was adapted from prescribed criteria in accordance with COSMIN recommendations.\textsuperscript{17} A positive (+), negative (−) or unknown (?) was apportioned to each criteria to determine the quality of each element associated with the statistical methods reported in the development of each instrument (Table 3).

The corresponding positive (+), negative (−) or unknown (?) is subsequently linked to a summary of an overall quality rating of strong, moderate, low or unknown (Table 3).

### Results

The database search initially yielded 1,862 potentially relevant studies. Duplicate papers were removed and abstracts screened with 191 full-text versions of the published articles retrieved. Additionally, five pain measurement instrument or tools studies were identified through further searching of reference lists. Eligibility criteria were applied and eight general

<table>
<thead>
<tr>
<th>Property</th>
<th>Definition</th>
<th>Rating</th>
<th>Quality criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability</td>
<td></td>
<td>+</td>
<td>Unidimensional (sub)scale AND Cronbach’s alpha between 0.70 and 0.95</td>
</tr>
<tr>
<td></td>
<td>Internal consistency not indicated</td>
<td>?</td>
<td>Dimensionality unknown OR Cronbach’s alpha not determined</td>
</tr>
<tr>
<td></td>
<td>The degree in which measurement</td>
<td>+</td>
<td>ICC weighted kappa &gt; 0.70</td>
</tr>
<tr>
<td></td>
<td>is free from measurement error</td>
<td>?</td>
<td>Time interval not indicated or described</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−</td>
<td>ICC OR weighted kappa &lt; 0.70</td>
</tr>
<tr>
<td>Validity</td>
<td></td>
<td>+</td>
<td>All questionnaires in target relevant AND complete questionnaire</td>
</tr>
<tr>
<td></td>
<td>Content validity</td>
<td>?</td>
<td>No target population involvement</td>
</tr>
<tr>
<td></td>
<td>Criterion validity</td>
<td>+</td>
<td>Target population considers items in questionnaire irrelevant OR considers questionnaire incomplete OR design method flaw</td>
</tr>
<tr>
<td></td>
<td>Responsiveness</td>
<td>?</td>
<td>Correlation with gold standard &lt; 0.70</td>
</tr>
<tr>
<td></td>
<td>Ability of HR-PRO instrument</td>
<td>+</td>
<td>Related constructs is higher than unrelated constructs</td>
</tr>
<tr>
<td></td>
<td>to detect change over time in</td>
<td>−</td>
<td>Related constructs is lower than unrelated constructs</td>
</tr>
<tr>
<td></td>
<td>measured construct</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

Overall quality rating assessment of general pain measurement

<table>
<thead>
<tr>
<th>Level</th>
<th>Criteria</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>Consistently good OR contains 1 or more excellent score</td>
<td>++</td>
</tr>
<tr>
<td>Moderate</td>
<td>Consistently fair OR contains 1 excellent or good score</td>
<td>+</td>
</tr>
<tr>
<td>Low</td>
<td>Consistently fair</td>
<td>OR</td>
</tr>
<tr>
<td>Unknown</td>
<td>Consistently poor</td>
<td>−</td>
</tr>
</tbody>
</table>

**Note:** Data from National Health and Medical Research Council.\textsuperscript{26}

### Table 3

Overall quality level of rating from Terwee’s assessment scale for clinimetric properties

<table>
<thead>
<tr>
<th>Level</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>++ OR</td>
</tr>
<tr>
<td>Moderate</td>
<td>+ OR</td>
</tr>
<tr>
<td>Low</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

**Note:** Reprinted from J Clin Epidemiol, 60(1), Terwee CB, Bot SDM, de Boer MR, et al, Quality criteria were proposed for measurement properties of health status questionnaires, 34–42. Copyright (2007), with permission from Elsevier.\textsuperscript{29}
pain measurement instrument development and validation studies were included for methodological quality assessment (Figure 1). Figure 1 illustrates the selection of general pain measurement studies reported in this review.

An overview of the results from three approaches used in the analysis of the methodological quality of the general pain instruments/tools are reported in Tables 4–6. Detailed results for content validity, criterion validity, reliability, internal consistency, responsiveness and interpretability show several general pain instruments were highly rated in one domain but were found to be poor in other aspects of the psychological properties.

The first analysis employed a dichotomous scale and the Pain Sensitivity Questionnaire (PSQ) scored high in four of the six domains including content validity, internal consistency, responsiveness (without a gold standard comparison) and interpretability. The Geriatric Pain Measure (GPM) also scored highly in four of the six domains including content validity, criterion validity, internal consistency and responsiveness without a gold standard comparison. The highest score for reliability was the Profile of Chronic Pain: Screen (PCP:S).

The second analysis using the COSMIN 4-point rating scale (Table 5), found that the PSQ scored highly in three of the five assessable domains including content validity, internal consistency and responsiveness. The GPM scored highest in two of the five assessable domains including criterion validity and internal consistency. The PainMatcher scored highest in

---

Figure 1 Pain measurement tool search diagram.
Table 4 Summary of methodological quality assessment of measurement tool using COSMIN checklist

<table>
<thead>
<tr>
<th>Measure</th>
<th>Reliability</th>
<th>Criterion validity</th>
<th>Content validity</th>
<th>Responsiveness without gold standard</th>
<th>Responsiveness with gold standard</th>
<th>Statistical validity</th>
<th>Design score total (D+G+S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain Catastrophizing Scale</td>
<td>D=3</td>
<td>S=1</td>
<td>S=1</td>
<td>D=1</td>
<td>S=1</td>
<td>D=1</td>
<td>D=3</td>
</tr>
<tr>
<td>Pain Catastrophizing Questionnaire</td>
<td>D=3</td>
<td>D=4</td>
<td>D=4</td>
<td>D=4</td>
<td>D=4</td>
<td>S=0</td>
<td>D=3</td>
</tr>
<tr>
<td>Defense and Veterans Pain Rating Scale</td>
<td>D=3</td>
<td>D=2</td>
<td>D=4</td>
<td>D=4</td>
<td>D=4</td>
<td>S=0</td>
<td>D=4</td>
</tr>
<tr>
<td>Pain Evaluation Scale</td>
<td>D=3</td>
<td>D=0</td>
<td>S=0</td>
<td>S=0</td>
<td>D=0</td>
<td>S=0</td>
<td>S=0</td>
</tr>
<tr>
<td>Geriatric Pain Measure</td>
<td>D=3</td>
<td>S=1</td>
<td>S=1</td>
<td>D=1</td>
<td>S=1</td>
<td>D=1</td>
<td>S=2</td>
</tr>
<tr>
<td>Painmatcher</td>
<td>D=3</td>
<td>D=0</td>
<td>D=0</td>
<td>D=0</td>
<td>D=0</td>
<td>S=0</td>
<td>S=0</td>
</tr>
<tr>
<td>Chronic Pain-cat</td>
<td>D=5</td>
<td>D=4</td>
<td>D=4</td>
<td>D=4</td>
<td>D=4</td>
<td>S=0</td>
<td>S=0</td>
</tr>
<tr>
<td>Profile of Chronic Pain: Screen</td>
<td>D=7</td>
<td>S=0</td>
<td>S=0</td>
<td>S=0</td>
<td>S=0</td>
<td>S=0</td>
<td>S=0</td>
</tr>
</tbody>
</table>

Abbreviations: COSMIN, common standards for the selection of health measurement instruments. D, Design requirement; G, General requirement; S, Statistical methods.

Overall the PSQ and GPM scored consistently highest in all domains. The PSQ scored overall the highest when being assessed by COSMIN, with the GPM scoring second highest.

Discussion

Our review of validation studies revealed a wide variety of pain measurement instruments and tools have been studied and included general pain and disease-specific pain measurement. This review utilized three approaches to assess selected general pain instruments or tools and found that several pain measurement instruments scored significantly higher on the COSMIN criteria in validity, reliability, responsiveness and interpretability. From this review, the findings suggest that there is wide variability in quality of pain measurement instruments and tools across all domains. This variability may be due to increased rigor and changing expectations in validation studies across time resulting in the potential for varying methodologies and quality of publication.

The results from our analyses associated with content and criterion validity demonstrated the PSQ, DVPRS and GPM scored the highest in both of these domains. These findings were consistent in each separate analysis. Reliability was assessed in terms of reliability and internal consistency. Reliability was found highest in PCP:S, whereas the PSQ and GPM scored the highest for internal consistency. The findings for the PSQ and GPM were consistent in all three analyses. Although the PCP:S scored high in the COSMIN, it did not have reportable findings in overall quality criteria for pain measurement tools using Terwee’s assessment scale of clinimetric properties from Terwee et al. In assessing responsiveness the new instrument was required to be evaluated against a gold standard and if a gold standard was not used then different criteria were utilized. The PSQ scored highest without a gold standard comparison, whereas the GPM scored the highest in reliability with a gold standard comparison. In this review, interpretability was assessed with the highest scores found in the Pain Catastrophizing Scale (PCS) and PSQ.

Overall the PSQ and GPM scored consistently highest in all domains. The PSQ scored overall the highest when being assessed by COSMIN, with the GPM scoring second highest.
It should be noted that some pain measurement tools/instruments were originally developed in specific disease populations, such as the Brief Pain Inventory, which was developed for the measurement of pain in cancer patients and has been utilized in COPD studies.

Therefore, when considering a pain measurement instrument for either clinical practice or research it is important to examine validity, reliability, responsiveness and interpretability to ensure that the best evidence-based pain measurement tool is selected.

Limitations
There are several limitations associated with this review. Although we undertook a systematic and extensive search for published articles associated with the original development of general pain measurement tool or instruments, some publications may have been missed. Secondly, new pain measurement tools and instruments may have been published after the completion of the extensive search in January 2016. Finally, articles published in other languages may report the development of general pain measurement instruments or tools and these were not included in this review.

Conclusion
This review of general pain instruments or tools revealed there was variability in the validity, reliability, responsiveness and interpretability when using a standardized approach. The PSQ was found to be the most valid and reliable general pain measurement instrument for adult populations.

Author contributions
AMJ and SMSS conceived the study; both authors contributed to defining and refining the review question and scope of review. Both authors reviewed the findings and contributed to the content, writing, drafting and revision of the manuscript, and agreed to the final version. AMJ is guarantor of the manuscript.

Disclosure
The authors report no conflicts of interest in this work.

References


