Establishing assessment criteria for clinical reasoning in orthopedic manual physical therapy: a consensus-building study

Euson Yeung¹, Nicole Woods¹, Adam Dubrowski², Brian Hodges¹, Heather Carnahan¹,2

¹University of Toronto, Toronto, Canada, ²Memorial University of Newfoundland, St John’s, Canada

Objectives: Clinical reasoning (CR) represents one of the core components of clinical competence in Orthopaedic Manual Physical Therapy (OMPT). While education standards have been developed to guide curricular design, assessment of CR has not yet been standardized. Without theory-informed and rigorously developed measures, the certification of OMPTs lacks credibility and is less defensible. The purpose of this study was to use a theory-informed approach to generate assessment criteria for developing new assessment tools to evaluate CR in OMPT.

Methods: A list of assessment criteria was generated based on international education standards and multiple theoretical perspectives. A modified Delphi method was used to gain expert consensus on the importance of these assessment criteria for the assessment of CR in OMPT. The OMPTs from 22 countries with experience in assessing CR were invited to participate in three rounds of online questionnaires to rate their level of agreement with these criteria. Responses were tabulated to analyze degree of consensus and internal consistency.

Results: Representatives from almost half of the OMPT member organizations (MO) participated in three rounds of the Delphi. High levels of agreement were found among respondents regarding the importance and feasibility of most assessment criteria. There was high internal consistency among items within the proposed item subgroupings.

Discussion: A list of assessment criteria has been established that will serve as a framework for developing new assessment tools for CR assessment in OMPT. These criteria will be important for guiding the design of certification processes in OMPT as well as other episodes of CR assessment throughout OMPT training.

Keywords: Manual therapy, Clinical reasoning, Assessment, Consensus

Introduction

Orthopaedic Manual Physical Therapy (OMPT) is a clinical specialty within physical therapy in which professional certification occurs at the post-graduate level. Previous studies have identified clinical reasoning (CR) as a critical behavior, indicative of post-graduate health professional practice in general and specifically within the field of OMPT. Although this is reflected in recent studies aimed at enhancing our understanding of the nature of CR in OMPT, there remains a dearth of literature on how CR competencies can be assessed in a comprehensive manner, particularly for post-graduate OMPT certification.

The challenges associated with assessing CR in the health professions have been previously described and are in part attributed to its complex nature. In OMPT, a key feature of CR that illustrates this complexity is the generation of comprehensive and relevant hypotheses and management plans that account for the dynamic interplay among the range of patients factors, therapist factors, and environmental factors. In light of this, Durning et al. suggest possible recommendations for the assessment of CR skills. In contrast to traditional methods of CR, or assessment that focus on linear processes and a single best path or diagnosis, Durning et al. proposed a theoretical framework for CR assessment that attends to its non-linearity. The authors suggest that CR assessment should not be grounded solely in psychometric theory (reliability and validity), but should additionally account for behaviorist perspectives (assessment of CR through observation of behavior), information processing theories (assessment of one’s mental processes), and situativity theory (CR assessment is inclusive of interactions...
between patient and therapist). By adopting these multiple theoretical perspectives, it is anticipated that an assessment tool will delineate the ‘boundaries’ within which acceptable CR performance can be defined. This will in turn permit a greater range of acceptable CR paths rather than promoting a single ‘correct’ one. This approach is particularly important for CR assessment as there is often no single or optimal path to solve a problem. Moreover, employing these theoretical perspectives in designing certification processes will generate results that are more likely to be generalizable to a number of OMPT training contexts.

In 2008, the International Federation of Orthopaedic Manual Physical Therapy (IFOMPT), comprised of member organizations (MO) from 22 countries in which post-graduate OMPT training occurs, outlined a competency framework encompassing OMPT-specific knowledge, skills, and attributes. The intent of this framework was to outline education standards to inform the design and implementation of post-graduate OMPT curricula and candidate assessment globally. These education standards afford OMPT programs the flexibility to customize curricula and assessment strategies to suit the needs of the learner and of each professional context. However, the onus rests on the education provider to develop certification procedures that are sufficiently comprehensive and rigorous for evaluating the OMPT competencies stipulated by IFOMPT. To date, tools that would assess CR in OMPT certification examinations are non-existent. Without a theory-informed, standardized assessment tool, assessment of OMPT candidates’ competencies may vary widely, leading to inconsistent assessment of candidates and scrutiny by all stakeholders. Given the importance of acquiring CR skills in OMPT training, and the high-stakes nature of the OMPT certification processes, it is essential for the assessment of CR competencies to be theoretically-informed and rigorously developed. This would allow for assessments that reflect a more robust and defensible certification processes that are more likely to be relevant and applicable to a range of IFOMPT MO.

In order to strengthen the existing certification process in OMPT, we sought to use a theoretically-informed approach to identify criteria for assessing CR in post-graduate education in OMPT. This study serves as the first phase toward the development of new tools to assess CR competencies in OMPT. It is anticipated that our study’s results will evolve into assessment tools that can be subsequently validated in a range of OMPT training programs globally.

**Methods**

We used a modified e-Delphi approach to determine the criteria essential for the assessment of the knowledge and skills associated with CR among OMPT candidates (Fig. 1). The Delphi technique is a consensus-seeking method that utilizes a sequential set of questionnaires (rounds) to systematically gather opinions on a topic from a panel of experts. This technique has successfully yielded assessment criteria for professional behaviors, problem-based learning, and procedural skills in the health professions education context. Since assessment of CR in the OMPT certification process is a phenomenon of interest with broad applications in all 22 IFOMPT MOs, the modified Delphi approach enabled us to access the opinions of a geographically diverse group of OMPT experts relevant to our study purpose. We followed Boulkedid et al.’s recommendations to use importance and feasibility as selection criteria for generating and reducing the number of items to achieve the objective of this study.

**Study population and recruitment**

We sought the opinions of OMPTs from the 22 IFOMPT MOs who have experience teaching and/or assessing CR in the OMPT context. Participants were recruited through a delegate from each of the 22 MOs who forwarded the study invitation to eligible OMPTs on our behalf. Owing to issues of privacy, we were unable to obtain personal email addresses of potential participants from MO delegates to contact eligible participants directly. Instead, we obtained email addresses from those who responded to our initial round of questionnaires, which were then used in subsequent rounds. We followed suggestions by Dillman to maximize the response rate through three points of contact with participants: a pre-round invitation email with survey link, a thank you/reminder email, and a final reminder email.

**Item generation**

A total of 80 assessment criteria pertaining to the knowledge, skills, and attributes important for CR in OMPT were generated by one of the investigators (EY) who is familiar with CR assessment in the post-graduate OMPT training context. The investigator constructed the list of criteria based on the IFOMPT education standards document, a review of published literature on CR in the health professions, and in consultation with another OMPT with a high degree of familiarity with CR assessment in the field. These items reflected the theoretical perspectives proposed by Durning et al. For example, some items related to one’s ability to select appropriate clinical procedures that reflects a behaviorist’s priority in assessing CR, while others described the assessment of one’s internal mental processes; an
important principle and concern of information processing theorists. Further, situative theoretical perspectives were accounted for by including items related to the extent to which the environmental and patient-specific social influences are considered and integrated into the reasoning process. Finally,
we included items that were consistent with the
notion of non-linearity. For example, several of
the items described the assessment of one’s ability
to reflect upon and explain the relationship among
subjective examination findings, physical examination
findings and the patient’s context, and stated
priorities and goals.

Furthermore, we sought to ensure that items
generated represented a range of CR theoretical
perspectives important in OMPT including diagnos-
tic, narrative, collaborative, prognostic, and ethical
reasoning.20 The items were not only relevant to the
use of propositional knowledge (‘knowing that’) and
non-propositional knowledge (‘knowing how’),21
but were also related to a particular organization of
knowledge that reflects a broader understanding of the
determinants of health and recovery.20 This included
items concerned with one’s judgment of the relevance
and dominance of particular features of a patient
presentation as well as the social and environmental
contexts in which the patient exists. Further, some of
the items generated reflect Gifford’s22 call for a greater
understanding and consideration of patients’ experi-
ences, beliefs, and feelings toward pain and disability
in a manner that enables the orthopedic manual
therapist to distinguish between adaptive and mal-
daptive mechanisms and responses (behavioral, cog-
nitive, emotional, and physiological).

Given that these assessment criteria are intended
for use within the OMPT context, the items were
categorized and presented within seven subgroups
similar to the groupings that align with the IFOMPT
education standards:
1. reasoning skills;
2. evidence-based practice;
3. knowledge and application of biomedical sciences;
4. knowledge and application of the clinical sciences;
5. knowledge and application of the behavioral
   sciences;
6. critical use of knowledge; and
7. communication skills.

We then administered this questionnaire to expert
OMPTs and sought consensus on the importance,
feasibility, and clarity of each of the assessment criteria
through online survey methods (SurveyMonkey24).

Round 1 – item generation and content validity
The purpose of the first questionnaire was to identify
a comprehensive list of criteria and their level of
perceived importance for CR assessment in OMPT.
Respondents were instructed to rate the level of im-
portance for each of the initial 80 items on a four-point
scale (extremely important, important, minimally
important, not at all important); a scale supported
by Lynn23 as a strategy to eliminate the option of
the ambivalent middle rating in content validity
studies. Items that received a mean rating of less than
3/4 were eliminated from the questionnaire for round
2. Although not used for the purpose of item reduc-
tion in this study, the coefficient of variation (CV), a
measure of response dispersion, was also calculated
to describe agreement among respondents for the
importance ratings. Finally, we invited respondents
to generate additional assessment criteria that were
missing from the list, to report on any item redund-
cancies, and to comment on item clarity.

Round 2 – item reduction
The purpose of round 2 was to reduce the number of
assessment criteria from round 1 to a feasible number
for inclusion in an assessment tool, while preserving
content validity. We also sought to determine the
extent to which items in each of the aforementioned
seven subgroups measured the same general construct
under which they were categorized. In constructing
the questionnaire for round 2, redundant items from
round 1 were combined, or were eliminated where the
item did not meet the aforementioned criteria. Where
necessary, item wording was revised to improve
clarity. The questionnaire in round 2 consisted of
71 assessment criteria and remained categorized
into the same seven subgroups. Participants were
instructed to rate each assessment criteria for clarity
(yes/no), assessment feasibility (yes/no), and level of
importance to CR assessment. In this round, we
used a nine-point scale (1=not at all important, 9-
extremely important) for the importance rating in
an attempt to improve response variability.24

Round 3
Based on the results of round 2, we anticipated that
respondents may encounter respondent fatigue in
subsequent rounds; that they would continue rating
the remaining assessment criteria as highly impor-
tant; and that there would be insufficient variation
in importance ratings that would result in further
item reduction. Thus, we considered the importance
ratings and internal consistency of items within the
seven subgroups in constructing the questionnaire for
round 3. The purpose of round 3 was to finalize the
list of assessment criteria by clarifying the importance
ratings of items with negative inter-item correlation
values and items within the subgroup that received
a Cronbach’s alpha of ≤0.80 from round 2. We
provided participants with the mean importance
rating from round 2 for each of these items and
invited participants to re-rate the importance of these
items on a nine-point scale (1=not at all important, 9=
extremely important).

Data analysis
Consensus was defined a priori as greater than 70%
agreement in all rounds in this study.25 Items that did
not reach 70% agreement regarding importance or
feasibility were eliminated in subsequent rounds. In
addition, respondents were instructed to rate items for
clarity and were encouraged to provide comments for items that were reported as unclear. Items that did not achieve 70% agreement with respect to clarity were revised for subsequent rounds. We also sought to examine the internal consistency of the items under the aforementioned seven subgroups in round 2. Subgroups that yielded Cronbach’s alpha of less than 0.80 were included in round 3 for further investigation.

Results
Round 1
Table 1 summarizes the characteristics of respondents from round 1. Almost all questionnaire items (77/80) received high mean importance ratings (3.2–3.8), low SDs (0.40–0.70), and low CV values (10.1–22.3%). Of the 80 assessment criteria, three of them received a mean importance rating of 3 out of 4 or lower and were subsequently removed for consideration in round 2 (Table 2). Respondents deemed 6 of the 80 assessment criteria to be redundant and were subsequently removed or integrated into other assessment criteria. No additional assessment criteria were generated from round 1. In total, 71 assessment criteria were retained for consideration by participants in round 2.

Round 2
Fifty of the 132 respondents from round 1 completed the questionnaire in round 2 resulting in a response rate of 37.9%. Table 1 summarizes the characteristics of the respondents in this round. Although some respondents commented on item ambiguity or the use of incomprehensible terminology, over 70% of respondents rated all 71 assessment criteria as clear and feasible for assessment. Problematic items were worded or edited for clarity for round 3. Regarding the importance ratings, 53 of the 71 items received mean ratings of at least 7 out of 9 from over 70% of respondents, with CV values ranging from 11.2 to 27.1% (Table 3). In other words, for 18 of the 71 items, less than 70% of respondents gave scores of greater than 7/9 on the importance rating. These items were subsequently removed from the list of assessment criteria.

We also examined the internal consistency among items within each of the aforementioned seven subgroups. Table 4 provides a summary of the inter-item correlations within these seven subgroups. This analysis yielded Cronbach’s alphas greater than 0.80 for all subgroup except for subgroup four (‘Knowledge and application of the clinical sciences’), and negative inter-item correlation values for 4 of the 71 assessment criteria (Table 5).

Owing to the disproportionately large number of participants from South Africa, a secondary analysis was conducted to examine the proportion of South African respondents who provided an importance rating of ≥7 compared to other respondents, using a Fisher’s exact test. The results indicate that there is

Table 1 Characteristics of survey respondents (Rounds 1, 2, and 3)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. of respondents (%)</th>
<th>No. of respondents (%)</th>
<th>No. of respondents (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=132, unless otherwise stated</td>
<td>n=50, unless otherwise stated</td>
<td>n=34, unless otherwise stated</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>39 (29.8)</td>
<td>13 (26.0)</td>
<td>24 (71.0)</td>
</tr>
<tr>
<td>Female</td>
<td>92 (70.2)</td>
<td>37 (74.0)</td>
<td>10 (29.0)</td>
</tr>
<tr>
<td>Experience teaching/examining in OMPT (Year: categories)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 or less</td>
<td>62 (47.0)</td>
<td>16 (32.0)</td>
<td>12 (35.3)</td>
</tr>
<tr>
<td>6–10</td>
<td>21 (15.9)</td>
<td>4 (8.0)</td>
<td>3 (8.8)</td>
</tr>
<tr>
<td>11–15</td>
<td>13 (9.8)</td>
<td>7 (14.0)</td>
<td>4 (11.8)</td>
</tr>
<tr>
<td>16–20</td>
<td>14 (10.6)</td>
<td>8 (16.0)</td>
<td>5 (14.7)</td>
</tr>
<tr>
<td>over 20</td>
<td>22 (16.7)</td>
<td>4 (8.0)</td>
<td>8 (23.5)</td>
</tr>
<tr>
<td>Country</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>1 (0.8)</td>
<td>1 (2.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Austria</td>
<td>1 (0.8)</td>
<td>1 (2.0)</td>
<td>1 (3.0)</td>
</tr>
<tr>
<td>Canada</td>
<td>15 (11.4)</td>
<td>10 (20.0)</td>
<td>7 (21.0)</td>
</tr>
<tr>
<td>Denmark</td>
<td>1 (0.8)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Germany</td>
<td>1 (0.8)</td>
<td>1 (2.0)</td>
<td>1 (3.0)</td>
</tr>
<tr>
<td>Greece</td>
<td>4 (3.0)</td>
<td>1 (2.0)</td>
<td>1 (3.0)</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>1 (0.8)</td>
<td>1 (2.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Ireland</td>
<td>1 (0.8)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Italy</td>
<td>1 (0.8)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1 (0.8)</td>
<td>1 (2.0)</td>
<td>1 (3.0)</td>
</tr>
<tr>
<td>Norway</td>
<td>2 (1.5)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>South Africa</td>
<td>53 (40.2)</td>
<td>20 (40.0)</td>
<td>14 (41.0)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>4 (3.0)</td>
<td>2 (4.0)</td>
<td>2 (6.0)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>46 (34.8)</td>
<td>12 (24.0)</td>
<td>7 (21.0)</td>
</tr>
</tbody>
</table>
no significant difference among participants’ importance ratings except for two items.

**Round 3**

Thirty-four of the 50 respondents from round 2 completed the round 3 questionnaire resulting in a response rate of 68.0%. Table 1 summarizes the characteristics of the respondents in this round. In round 3, all four items from round 2 with negative inter-item correlation values received an importance rating of 7.5 out of 9 or higher, with CV ranging from 10.8 to 13.8. Analysis of internal consistency of items in the ‘knowledge and application of the clinical sciences’ subgroup yielded a Cronbach’s alpha of 0.74.

**Discussion**

The purpose of this study was to generate assessment criteria for developing new tools to assess CR competencies among candidates seeking postgraduate certification in OMPT. By framing and incorporating the IFOMPT education standards with the theoretical perspectives proposed by Durning et al., the resulting items are particularly relevant to the theory and practice of CR in OMPT. This study represents the critical first steps toward developing a theory-informed assessment tool that should serve to improve the rigor of post-graduate OMPT certification.

Results from round 1 demonstrate that a comprehensive list of assessment criteria was generated in the initial round based on the literature and the IFOMPT education standards. This is evidenced by the lack of additional items generated by respondents in the initial round, and the minimal redundancies noted (6/80). These results support the use of the theoretical framework proposed by Durning et al. and show promise for its application in future CR research. Furthermore, round 1 results indicate that respondents were in strong agreement that almost all of the assessment criteria presented were deemed important for CR assessment in OMPT certification. Notably, respondents represented 14 of the 22 IFOMPT MO in this round. As noted by Boulkedid et al., participation from a wide range of relevant stakeholders enhances the credibility and acceptance of the eventual list of assessment criteria; thus, our study results may be a suitable starting point for designing MO-specific assessment tools for OMPT certification.

In round 2, our analysis of the internal consistency amongst assessment criteria yielded high Cronbach’s alpha values for six of the seven subgroups. This demonstrates that the categorization proposed in this study may be a promising method for grouping the assessment criteria as global constructs of CR in OMPT. Three of the seven subgroups in round 2 yielded alphas of 0.90 or higher in round 2 (Table 3), which suggest a higher level of item redundancy in these subgroups as compared to the other subgroups. Future studies are required to investigate the extent to which items are redundant and whether or not there are unnecessary items that need to be removed. Moreover, the adequacy and applicability of these global constructs (subgroups) will require further validity testing in OMPT training programs within each of the IFOMPT MO.

Our analysis also yielded Cronbach’s alpha values ≥0.80 for all subgroups except for the ‘knowledge and application of the clinical sciences’ subgroup (Table 5). This may be explained by the broad range of clinical approaches employed in OMPT, as well as the diverse philosophical approaches associated with the recognition and selection of specific clinical procedures. Different perspectives regarding the role and value placed on clinical procedures (and their diagnostic properties) in OMPT practice may have an impact on respondents’ opinions regarding the requisite knowledge and skills for CR. In light of this, targeted feasibility and acceptability evaluations of the assessment criteria generated in this study are needed in order to account for any between-country, or within-country, differences in philosophical approaches that may threaten the validity of the subsequent assessment tools.

Studies are currently underway examining the subgroups proposed in this study by examining actual assessment data from OMPT certification examinations in the Canadian context. This will help to justify the use of global rating scales in addition to, or in place of, individual assessment criteria. Although medical educators have historically employed a checklist of

<table>
<thead>
<tr>
<th>Assessment criteria</th>
<th>Mean importance rating out of 4 (SD)</th>
<th>Coefficient of variation, CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of dental and orthodontic dysfunctions related to the neuromusculoskeletal system</td>
<td>2.85 (0.7)</td>
<td>25.6</td>
</tr>
<tr>
<td>Knowledge of the nature of surgical interventions for the management of neuromusculoskeletal dysfunctions</td>
<td>2.95 (0.7)</td>
<td>23.2</td>
</tr>
<tr>
<td>Knowledge of various manipulative therapy approaches including those in medicine, osteopathy, and chiropractice</td>
<td>2.88 (0.7)</td>
<td>24.6</td>
</tr>
</tbody>
</table>
## Table 3  Assessment criteria retained from Round 2

<table>
<thead>
<tr>
<th>Assessment criteria</th>
<th>Importance ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reasoning skills</strong></td>
<td></td>
</tr>
<tr>
<td>Understanding of the process of hypothetico-deductive clinical reasoning, CR</td>
<td>7.67, 1.67, 21.77</td>
</tr>
<tr>
<td>Ability to apply hypotheses generated in relation to diagnosis, treatment and prognosis</td>
<td>7.18, 1.81, 25.26</td>
</tr>
<tr>
<td>Appropriate use of the pattern recognition process</td>
<td>7.27, 1.41, 19.42</td>
</tr>
<tr>
<td>Recognition of dysfunction(s) that require further investigation and/or referral to another health care professional</td>
<td>8.10, 1.45, 17.85</td>
</tr>
<tr>
<td>Ability to evaluate common CR errors</td>
<td>6.98, 1.72, 24.98</td>
</tr>
<tr>
<td>Ability to integrate the following in clinical decision-making: scientific evidence, clinical data, patient’s perceptions and goals, and all factors related to the clinical context and patient circumstance</td>
<td>7.71, 1.41, 18.33</td>
</tr>
<tr>
<td>Ability to demonstrate effective prioritization in the assessment and management of patients</td>
<td>7.80, 0.93, 11.87</td>
</tr>
<tr>
<td>Ability to critically reflect during and after a clinical encounter</td>
<td>7.80, 1.28, 16.39</td>
</tr>
<tr>
<td><strong>Evidence-based practice</strong></td>
<td></td>
</tr>
<tr>
<td>Ability to demonstrate an evidence-based assessment and management approach</td>
<td>7.63, 1.27, 16.68</td>
</tr>
<tr>
<td>Ability to demonstrate an evidence-based approach to modify or progress treatment based on an accurate interpretation of the results of treatment</td>
<td>7.86, 1.34, 17.06</td>
</tr>
<tr>
<td>Appropriate use of outcome measures to evaluate the effectiveness of orthopedic manual physical therapy (OMPT)</td>
<td>7.49, 1.49, 19.84</td>
</tr>
<tr>
<td><strong>Knowledge and application of biomedical sciences</strong></td>
<td></td>
</tr>
<tr>
<td>Knowledge of anatomy, physiology, and biomechanics of the musculoskeletal, neurological, vascular, and lymphatic systems that enables evaluation of normal and abnormal function</td>
<td>8.02, 1.13, 14.14</td>
</tr>
<tr>
<td>Knowledge of biomechanical properties of viscoelastic tissues to enable evaluation of normal and abnormal function</td>
<td>7.29, 1.62, 22.18</td>
</tr>
<tr>
<td>Knowledge of pathology and pathogenesis of mechanical dysfunctions of the neuromusculoskeletal system</td>
<td>7.96, 1.03, 12.93</td>
</tr>
<tr>
<td>Knowledge of non-mechanical dysfunctions of the neuromusculoskeletal system</td>
<td>7.67, 1.27, 16.52</td>
</tr>
<tr>
<td>Knowledge of neurological dysfunctions of the neuromusculoskeletal system</td>
<td>7.76, 1.17, 15.09</td>
</tr>
<tr>
<td>Knowledge of cardiovascular dysfunction to enable differentiation from neuromuscular dysfunctions</td>
<td>6.98, 1.75, 25.13</td>
</tr>
<tr>
<td>Knowledge of pain sciences as related to the neuromusculoskeletal system</td>
<td>8.14, 1.03, 12.65</td>
</tr>
<tr>
<td>Knowledge of examination procedures to enable differential diagnosis of neuromusculoskeletal, neurological, vascular, and lymphatic dysfunctions</td>
<td>7.73, 1.63, 21.02</td>
</tr>
<tr>
<td>Ability to apply knowledge of anatomy, physiology, pathology, and biomechanics to the examination and management of patients with neuromusculoskeletal dysfunctions</td>
<td>8.12, 1.00, 12.34</td>
</tr>
<tr>
<td>Ability to critically evaluate the contribution of the biomedical sciences (i.e.: anatomy, physiology, pathology, biomechanics) to the patient’s presentation</td>
<td>7.55, 1.67, 22.07</td>
</tr>
<tr>
<td><strong>Knowledge and application of the clinical sciences</strong></td>
<td></td>
</tr>
<tr>
<td>Knowledge of effectiveness, risks, and efficacy of orthopedic manual therapy interventions</td>
<td>7.92, 1.21, 15.27</td>
</tr>
<tr>
<td>Knowledge of specific diagnostic and evaluative qualities of assessment tools (i.e.: reliability, validity, specificity, sensitivity etc)</td>
<td>6.96, 1.45, 21.60</td>
</tr>
<tr>
<td>Knowledge of prognostic, risk, and predictive factors of relevant health problems and their impact on OMPT interventions</td>
<td>7.04, 1.54, 21.85</td>
</tr>
<tr>
<td>Ability to select the most appropriate assessment and intervention tools/methods</td>
<td>7.94, 1.20, 15.14</td>
</tr>
<tr>
<td>Ability to interpret the outcomes of assessment and interventions</td>
<td>8.20, 1.23, 14.98</td>
</tr>
<tr>
<td>Ability to accurately predict expected changes and progress toward realistic patient outcomes</td>
<td>7.31, 1.15, 15.69</td>
</tr>
<tr>
<td><strong>Knowledge and application of the behavioral sciences</strong></td>
<td></td>
</tr>
<tr>
<td>Knowledge of relevant theories of behavior and behavior change (i.e.: behavioral reactions to pain and limitations, coping strategies, etc)</td>
<td>7.02, 1.52, 21.63</td>
</tr>
<tr>
<td>Knowledge of the role of the biopsychosocial model in the assessment and management of patients</td>
<td>7.59, 1.23, 16.17</td>
</tr>
<tr>
<td>Ability to apply the biopsychosocial model in OMPT interventions</td>
<td>7.71, 1.44, 18.70</td>
</tr>
</tbody>
</table>
28 researchers have criticized the use of such performance-based measures in favor of global rating scales over the past two decades. The criticism was largely a result of the incongruence among the use of discrete performance-based behaviors for complex, non-linear behaviors such as analytical or communication skills of which CR is partially comprised.29,30

The value of using a global rating scale in certification needs to be explored within different OMPT contexts,

Table 3 Continued

<table>
<thead>
<tr>
<th>Assessment criteria</th>
<th>Importance ratings</th>
<th>Mean</th>
<th>SD</th>
<th>CV%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of the mutual influence between the physical therapist’s behavior and the patient’s behavior</td>
<td></td>
<td>6.98</td>
<td>1.88</td>
<td>27.07</td>
</tr>
<tr>
<td>Ability to reflect and self-evaluate in managing patients</td>
<td></td>
<td>7.65</td>
<td>1.64</td>
<td>21.37</td>
</tr>
</tbody>
</table>

Critical use of knowledge

Knowledge of the theoretical basis of the assessment of the neuromusculoskeletal system, and its interpretation in determining a physical diagnosis
Knowledge of interpretation of the static, dynamic, and functional postural assessment of the neuromusculoskeletal system
Knowledge of the biomechanics and principles of active and passive movement of the articular system
Knowledge of specific tests for the assessment of functional status of the muscular system
Knowledge of specific tests for the assessment of functional status of the nervous system
Knowledge of specific special or screening tests that enable a safe practice of OMPT
Knowledge and appropriate integration of medical diagnostic tests to determine a physical diagnosis
Knowledge of multimodal physical therapy interventions for the management of neuromusculoskeletal dysfunctions
Knowledge of the theoretical basis for manipulative physical therapy practice in managing neuromusculoskeletal dysfunctions
Knowledge of the indications and contraindications for OMPT interventions
Knowledge of appropriate ergonomic strategies to enhance patient function in the work environment
Knowledge of preventative programs or strategies for managing neuromusculoskeletal dysfunctions
Ability to accurately determine diagnoses of neuromusculoskeletal dysfunctions
Ability to integrate principles of mobilization and manipulation as a component of a multimodal management approach
Ability to integrate principles of exercise physiology as it applies to therapeutic rehabilitative exercise programs within a multimodal management approach
Ability to integrate principles of motor-learning as a component of a multimodal management approach
Ability to integrate principles of patient education as a component of a multimodal management approach
Ability to integrate principles of other modalities (i.e.: taping, bracing, electrophysical modalities, acupuncture/needling) as a component of a multimodal management approach

Communication skills

Ability to use questioning strategies efficiently and effectively to obtain reliable and valid data from the patient
Ability to use active listening skills efficiently and effectively
Ability to collaborate effectively with the patient to inform management decisions

Table 4 Item-total correlation values for the seven subgroups (Round 2)

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reasoning skills (data gathering and hypothesis generation)</td>
<td>0.80</td>
</tr>
<tr>
<td>2. Evidence-based practice (critical evaluation of sources of information)</td>
<td>0.90</td>
</tr>
<tr>
<td>3. Knowledge and application of biomedical sciences</td>
<td>0.91</td>
</tr>
<tr>
<td>4. Knowledge and application of clinical sciences</td>
<td>0.69</td>
</tr>
<tr>
<td>5. Knowledge and application of behavioral sciences</td>
<td>0.88</td>
</tr>
<tr>
<td>6. Critical use of knowledge</td>
<td>0.91</td>
</tr>
<tr>
<td>7. Communication skills</td>
<td>0.80</td>
</tr>
</tbody>
</table>
and may additionally serve as a valuable tool for other episodes of CR assessment throughout the OMPT training process.

Finally, our results from the analysis of internal consistency will require further clarification. In round 2, our analysis yielded negative inter-item correlation values for 4 of the 71 items in this round. Although a negative inter-item correlation may suggest that the item is not measuring the same construct to which the item is grouped, further analysis of internal consistency of items in these subgroups needs to be undertaken using assessment data from actual OMPT certification examinations. Similarly, in round 3, we sought to clarify the degree of homogeneity of items within the ‘knowledge and application of the clinical sciences’ subgroup. Although the internal consistency estimate for this subgroup was <0.80 (0.74), it still exceeds the acceptable alpha of 0.70 recommended by other researchers for new instruments.31,32

Limitations
While there may be inherent biases associated with the use of one author (EY) in the initial item generating process, our literature search and use of a wide variety of experts were a way of ensuring broad-based input. Notably, we received responses from several IFOMPT MOs; however, it is important to note that there remain opinions of OMPTs from other MOs which are not represented in our study sample. Thus, the application of the assessment criteria generated from this study must undergo further content validity in specific OMPT contexts. Moreover, we recognize that a response rate of 37% in the initial round represents the low end of acceptable response rates for surveys in general;19 however, our response rates are very similar to those found in a study comparing mail surveys and internet-based surveys33 and in other studies using web-based methods to survey OMPTs.34

Conclusion
In this study, the 53 CR assessment criteria generated were grounded in multiple theoretical perspectives and approximated the OMPT-specific education standards set out by IFOMPT. The results of the current study indicate that participants strongly agreed with the CR assessment criteria generated, and that they are sufficiently clear and feasible for assessment. At present, there is a dearth of theory-informed, rigorously developed assessment tools for OMPT certification. Thus, our study’s results will be important for the development of new assessment tools for OMPT certification. Furthermore, there is great potential to use the assessment criteria from this study in evaluating and advancing our understanding of the effectiveness of educational interventions aimed at development of CR skills in OMPT. In addition to improving certification processes, assessment tools that are developed based on these assessment criteria could also serve as means for evaluating change in learning outcomes throughout OMPT training. Although the aforementioned assessment criteria have the potential to be used as noted above, further studies are warranted in order to determine their validity and applicability in various contexts.
to apply these assessment criteria to specific OMPT contexts globally.

**Conflict of Interest**
There were no conflicts of interest issues in this study.

**Disclosure Statement**
This study was approved by the University of Toronto Research Ethics Board Research Ethics Board.

**Acknowledgements**
The authors would like to thank the IFOMPT for its assistance with participant recruitment for this study.

**References**


