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

Euson Yeung¹, Nicole Woods¹, Adam Dubrowski², Brian Hodges¹, Heather Carnahan^{1,2}

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 general¹ and specifically within the field of OMPT.² Although this is reflected in recent studies aimed at enhancing our understanding of the nature of CR in OMPT,³⁻⁶ 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.^{7,8}

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.⁹ In light of this, Durning *et al.*⁷ 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.*⁷ 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

Correspondence to: Euson Yeung, 160-500 University Avenue, Toronto, Ontario, M5G 1V7, Canada. Email: euson.yeung@utoronto.ca

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.¹⁰ 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.¹¹ 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,² 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.^{12,13}

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 for candidates seeking post-graduate certification 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.¹⁴ This technique has successfully yielded assessment criteria for professional behaviors,¹⁵ problem-based learning,¹⁶ and procedural skills¹⁷ 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 MO, 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.*’s¹⁸ 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 Dillman¹⁹ 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,¹¹ 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.*⁷ 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

Figure 1 Modified Delphi technique.

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 diagnostic, 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' experiences, beliefs, and feelings toward pain and disability in a manner that enables the orthopedic manual therapist to distinguish between adaptive and maladaptive mechanisms and responses (behavioral, cognitive, 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 (SurveyMonkey™).

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 importance 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 reduction 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 redundancies, 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 important; 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.^{26,27}

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%).²⁵ 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 reworded 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)

Characteristics	No. of respondents (%)	No. of respondents (%)	No. of respondents (%)
	<i>n=132, unless otherwise stated</i>	<i>n=50, unless otherwise stated</i>	<i>n=34, unless otherwise stated</i>
	Round 1	Round 2	Round 3
Sex	* <i>n=131</i>		
Male	39 (29.8)	13 (26.0)	24 (71.0)
Female	92 (70.2)	37 (74.0)	10 (29.0)
Experience teaching/examining in OMPT (Year: categories)			
5 or less	62 (47.0)	16 (32.0)	12 (35.3)
6–10	21 (15.9)	4 (8.0)	3 (8.8)
11–15	13 (9.8)	7 (14.0)	4 (11.8)
16–20	14 (10.6)	8 (16.0)	5 (14.7)
over 20	22 (16.7)	4 (8.0)	8 (23.5)
Country			
Australia	1 (0.8)	1 (2.0)	0 (0.0)
Austria	1 (0.8)	1 (2.0)	1 (3.0)
Canada	15 (11.4)	10 (20.0)	7 (21.0)
Denmark	1 (0.8)	0 (0.0)	0 (0.0)
Germany	1 (0.8)	1 (2.0)	1 (3.0)
Greece	4 (3.0)	1 (2.0)	1 (3.0)
Hong Kong	1 (0.8)	1 (2.0)	0 (0.0)
Ireland	1 (0.8)	0 (0.0)	0 (0.0)
Italy	1 (0.8)	0 (0.0)	0 (0.0)
New Zealand	1 (0.8)	1 (2.0)	1 (3.0)
Norway	2 (1.5)	0 (0.0)	0 (0.0)
South Africa	53 (40.2)	20 (40.0)	14 (41.0)
Switzerland	4 (3.0)	2 (4.0)	2 (6.0)
United Kingdom	46 (34.8)	12 (24.0)	7 (21.0)

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 post-graduate certification in OMPT. By framing and incorporating the IFOMPT education standards with the theoretical perspectives proposed by Durning *et al.*,⁷ 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.*⁷ 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.*,¹⁸ 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 2 Assessment criteria removed after round 1

Assessment criteria	Mean importance rating out of 4 (SD)	Coefficient of variation, CV (%)
Knowledge of dental and orthodontic dysfunctions related to the neuromusculoskeletal system	2.85 (0.7)	25.6
Knowledge of the nature of surgical interventions for the management of neuromusculoskeletal dysfunctions	2.95 (0.7)	23.2
Knowledge of various manipulative therapy approaches including those in medicine, osteopathy, and chiropractice	2.88 (0.7)	24.6

Table 3 Assessment criteria retained from Round 2

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

Table 3 Continued

Assessment criteria	Importance ratings		
	Mean	SD	CV%
Knowledge of the mutual influence between the physical therapist's behavior and the patient's behavior	6.98	1.88	27.07
Ability to reflect and self-evaluate in managing patients	7.65	1.64	21.37
<i>Critical use of knowledge</i>			
Knowledge of the theoretical basis of the assessment of the neuromusculoskeletal system, and its interpretation in determining a physical diagnosis	7.57	1.58	20.84
Knowledge and interpretation of the static, dynamic, and functional postural assessment of the neuromusculoskeletal system	7.61	1.32	17.37
Knowledge of the biomechanics and principles of active and passive movement of the articular system	7.92	1.07	13.46
Knowledge of specific tests for the assessment of functional status of the muscular system	7.96	1.01	13.24
Knowledge of specific tests for the assessment of functional status of the nervous system	7.88	1.04	13.24
Knowledge of specific special or screening tests that enable a safe practice of OMPT	8.29	1.16	14.01
Knowledge and appropriate integration of medical diagnostic tests to determine a physical diagnosis	7.16	1.36	18.99
Knowledge of multimodal physical therapy interventions for the management of neuromusculoskeletal dysfunctions	7.76	1.15	14.86
Knowledge of the theoretical basis for manipulative physical therapy practice in managing neuromusculoskeletal dysfunctions	7.10	1.68	23.67
Knowledge of the indications and contraindications for OMPT interventions	8.35	.94	11.24
Knowledge of appropriate ergonomic strategies to enhance patient function in the work environment	7.35	1.45	19.74
Knowledge of preventative programs or strategies for managing neuromusculoskeletal dysfunctions	7.27	1.44	19.81
Ability to accurately determine diagnoses of neuromusculoskeletal dysfunctions	7.39	1.32	17.89
Ability to integrate principles of mobilization and manipulation as a component of a multimodal management approach	7.96	1.03	12.93
Ability to integrate principles of exercise physiology as it applies to therapeutic rehabilitative exercise programs within a multimodal management approach	7.67	1.25	16.31
Ability to integrate principles of motor-learning as a component of a multimodal management approach	7.41	1.29	17.44
Ability to integrate principles of patient education as a component of a multimodal management approach	7.92	1.10	13.94
Ability to integrate principles of other modalities (i.e.: taping, bracing, electrophysical modalities, acupuncture/needling) as a component of a multimodal management approach	7.14	1.41	19.80
<i>Communication skills</i>			
Ability to use questioning strategies efficiently and effectively to obtain reliable and valid data from the patient	8.12	1.22	15.05
Ability to use active listening skills efficiently and effectively	8.39	1.10	13.15
Ability to collaborate effectively with the patient to inform management decisions	8.29	1.05	12.67

performance-based measures for assessing trainee competence,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 4 Item-total correlation values for the seven subgroups (Round 2)

Subgroup	Cronbach's alpha
1. Reasoning skills (data gathering and hypothesis generation)	0.80
2. Evidence-based practice (critical evaluation of sources of information)	0.90
3. Knowledge and application of biomedical sciences	0.91
4. Knowledge and application of clinical sciences	0.69
5. Knowledge and application of behavioral sciences	0.88
6. Critical use of knowledge	0.91
7. Communication skills	0.80

Table 5 Items that were re-rated for 'importance' (Round 3)

Assessment criteria with negative inter-item correlation values in round 2	Subgroup	Mean importance rating out of 9 (SD)	Coefficient of variation, CV (%)
Ability to accurately and efficiently select inquiry strategies based on early recognition and correct interpretation of relevant clinical cues	Data gathering/hypothesis generation	7.5 (0.92)	12.2
Ability to reflect and self-evaluate in managing patients	Behavioral sciences	7.8 (0.81)	10.8
Knowledge of the specific special tests/screening tests for the safe practice of orthopedic manual therapy	Critical use of knowledge	8.3 (1.07)	12.9
Creativity and innovation in the application of knowledge of orthopedic manual physical therapy (OMPT)	Critical use of knowledge	7.5 (1.04)	13.8
Assessment criteria from 'knowledge and use of the clinical sciences' subgroup (Cronbach's alpha=0.74)			
Knowledge of effectiveness, risks, and efficacy of orthopedic manual therapy interventions		7.9 (0.92)	11.6
Knowledge of the specific diagnostic and evaluative qualities of assessment tools		6.9 (1.07)	15.5
Knowledge of prognostic, risk, and predictive factors of relevant problems and their impact on orthopedic manual therapy management strategies		7.3 (1.09)	15.0
Ability to identify the nature and extent of patient's functional abilities, pain, and multidimensional needs in relation to the International Classification of Function and planned management		6.8 (1.28)	18.8
Ability to determine which assessment and intervention tools are most appropriate		8.2 (0.69)	8.4
Ability to interpret outcomes of assessment and interventions		8.5 (0.70)	8.2
Ability to accurately predict expected changes and progress toward realistic outcomes		7.5 (0.81)	10.8

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 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.

Acknowledgements

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

References

- Rushton A, Lindsay G. Defining the construct of masters level clinical practice in healthcare based on the UK experience. *Med Teach*. 2008;30:e100–7.
- Rushton A, Lindsay G. Defining the construct of masters level clinical practice in manipulative physiotherapy. *Man Ther*. 2010;15:93–9.
- Kerry R, Taylor AJ. Cervical arterial dysfunction assessment and manual therapy. *Man Ther*. 2006;11:243–53.
- Smart K, Doody C. The clinical reasoning of pain by experience musculoskeletal physiotherapists. *Man Ther*. 2007;12:40–9.
- Kerry R, Taylor AJ. Cervical arterial dysfunction: knowledge and reasoning for manual physical therapists. *J Orthop Sports Phy Ther*. 2009;39:378–87.
- Sweeney A, Doody C. The clinical reasoning of musculoskeletal physiotherapists in relation to the assessment of vertebrobasilar insufficiency: a qualitative study. *Man Ther*. 2010;15:394–9.
- Durning SJ, Artino AR, Schuwirth L, van der Vleuten. Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. *Acad Med*. 2013;88:442–8.
- Schuwirth L. Is assessment of clinical reasoning still the Holy Grail? *Med Educ*. 2009;43:298–9.
- Higgs J, Jones M, Loftus S, Christensen N. Clinical reasoning in the health professions, 3rd edn. Oxford: Butterworth-Heinemann; 2008.
- Norman G. Research in clinical reasoning: past history and current trends. *Med Educ*. 2005;39:418–27.
- Educational Standards in Orthopaedic Manipulative Physical Therapy. International federation of orthopaedic manipulative physical therapy. 2008 [cited 2013 Jul 29]. Available from: http://www.ifompt.com/site/ifompt/files/pdf/IFOMT_Education_Standards_and_International_Monitoring_20080611.pdf
- Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356:387–96.
- Rees CE, Monrouxe LV. Theory in medical education research: how do we get there? *Med Educ*. 2010;44:334–9.
- Kalaian SA, Kasim RM. Terminating sequential Delphi survey data collection. *Pract Assess Res Eval*. 2012;17:1–10.
- Bartlett D, Lucy SD, Bisbee L. Item generation and pilot testing of the Comprehensive Professional Behaviours Development Log. *J Allied Health*. 2006;35:89–93.
- Des Marchais JE. A Delphi technique to identify and evaluate criteria for construction of PBL problems. *Med Educ*. 1999;33:504–8.
- McKinley RK, Strand J, Gray T, Schuwirth L, Alun-Jones T, Miller H. Development of a tool to support holistic generic assessment of clinical procedure skills. *Med Educ*. 2008;42:619–27.
- Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One*. 2011;6:e20476.
- Dillman DA. Mail and internet surveys: the tailored design method, 2nd edn. New York, NY, USA: John Wiley & Sons Inc.; 2000.
- Jones MA, Rivett DA. Clinical reasoning for manual therapists. Edinburgh: Butterworth-Heinemann; 2004.
- Higgs J, Titchen A. Propositional, professional and personal knowledge in clinical reasoning. In clinical reasoning in the health professions, 2nd edn. Oxford: Butterworth-Heinemann; 2008.
- Gifford L. Pain, the tissues and the nervous system: a conceptual model. *Physiotherapy*. 1998;84:27–36.
- Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986;35:382–5.
- Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. New York: Oxford University Press; 2008.
- Dajani JS, Sincoff MZ, Talley WK. Stability and agreement criteria for the termination of Delphi studies. *Technol Forecast Soc*. 1979;13:83–90.
- Ferketich S. Focus on psychometrics: internal consistency estimates of reliability. *Res Nurs Health*. 1990;13:437–40.
- Nunnally JC, Bernstein IH. Psychometric theory. Sydney: McGraw-Hill; 1994.
- Harden R, Stevenson M, Downie W, Wilson G. Assessment of clinical competence using an objective structured examination. *Br Med J*. 1975;1:447–51.
- Hodges B, Mellroy J. Analytic global OSCE ratings are sensitive to level of training. *Med Educ*. 2003;37:1021–16.
- Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med*. 1998;73:993–7.
- DeVellis RF. Scale development: theory and applications (applied social research methods series, Vol. 26). Newbury Park: Sage; 1991.
- DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007;39:155–64.
- Kwak N, Radler B. A comparison between mail and web surveys: response pattern, respondent profile, and data quality. *Journal Off Stat*. 2002;18:257–73.
- Chau J, Chadbourn P, Hamel R, Mok S, Robles B, Chan L, et al. Continuing education for advanced manual and manipulative physiotherapists in Canada: a survey of perceived needs. *Physiother Can*. 2012;64:20–30.